



Instituto de Economía

Facultad de Ciencias Económicas y de Administración
Universidad de la República - Uruguay

Un enfoque de modelos mixtos de clases latentes para
analizar la trayectoria nutricional y el desempeño
escolar de niños y niñas

Alejandra Marroig

INSTITUTO DE ECONOMÍA

Serie Documentos de Trabajo

Febrero, 2018

DT 02/2018

ISSN: 1510-9305
ISSN: 1688-5090

(en papel)
(en línea)

El trabajo contó con apoyo financiero de la Comisión Académica de Posgrado (UdelaR), Instituto de Economía e Instituto de Estadística (FCEyA – UdelaR).

Forma de citación sugerida para este documento: Marroig, A. (2018) “Un enfoque de modelos mixtos de clases latentes para analizar la trayectoria nutricional y el desempeño escolar de niños y niñas.”. Serie Documentos de Trabajo, DT 02/2018. Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Uruguay.

Un enfoque de modelos mixtos de clases latentes para analizar la trayectoria nutricional y el desempeño escolar de niños y niñas

Alejandra Marroig[§]

Resumen

En el trabajo se analiza la trayectoria nutricional de una cohorte de niños y niñas de Uruguay en edad escolar y su vínculo con el desempeño educativo mediante la estimación de Modelos Mixtos de Clases Latentes y Modelos Mixtos Conjuntos de Clases Latentes. La metodología se enmarca en los modelos estructurales y permite analizar el cambio con la edad del estado nutricional y su relación con el desempeño escolar. Además, supone heterogeneidad entre individuos no capturada por variables observadas que se representa mediante la pertenencia a grupos latentes. Se identificaron tres grupos de niños y niñas según trayectorias nutricionales en el período escolar. Un grupo de niñas tiene trayectoria nutricional normal (50%) para la edad, otro normal supera el umbral de sobrepeso para ciertas edades (40%) y un grupo tiene problemas de sobrepeso (10%) en edad escolar. Por su parte, un grupo de niños tiene trayectoria normal (70%), otro supera el sobrepeso en ciertas edades (20%) y un grupo minoritario (10%) presenta obesidad. Las niñas con sobrepeso tienen mayor peso y talla al nacer pero no hay diferencias en el nivel educativo de sus madres. El grupo de niños con obesidad tiene mayor peso al nacer pero la talla no presentó diferencias significativas. Además, hay mayor proporción de madres con nivel educativo terciario en el grupo de niños con obesidad. En cuanto a la relación entre trayectoria nutricional y desempeño escolar, los niños abandonan antes que las niñas. Sin embargo, los grupos según trayectorias nutricionales no presentan diferencias en la edad de abandono escolar. Ninguno de los grupos identificados muestra déficit nutricional o decrecimiento durante el período escolar, pero sí problemas de sobrepeso u obesidad lo que podría ser objeto de estudio en futuras investigaciones.

Palabras clave: Clase latente, Modelo conjunto, Trayectoria nutricional, Desempeño escolar, Uruguay

Código JEL: C33, I12, J13

[§] Instituto de Economía, Facultad de Ciencias Económicas y de Administración, UdelarR.

Latent Classes Mixed Models approach to analyze child's nutritional trajectory and school performance

Alejandra Marroig

Abstract

The aim is to analyze the nutritional trajectory of a cohort of uruguayan boys and girls in school age and its association with school performance. With this purpose, Latent Class Mixed Models and Joint Latent Class Mixed Models were fitted. This methodology is framed in the structural models and allows an analysis of nutritional status change with age of the cohort. In addition, the models assume that it is not possible to capture heterogeneity among individuals by any observed variable and, therefore, this is represented by latent groups. The model identified three groups of boys and girls according to its nutritional trajectories during school. A group of girls have normal nutritional trajectory (50%) for age, another exceed the overweight threshold for certain ages (40%) and last group have overweight (10%) in school age. Besides, a group of boys have normal nutritional trajectory (70%) for age, another have overweight for certain ages (20%) and, finally, a minority group have obesity (10%). The group of overweight girls is characterized by greater weight and height at birth, however there are no differences in mothers educational attainment. Amongst boys the obesity group is heavier at birth but size at birth did not show significant differences. In addition, there is a higher proportion of mothers with superior educational attainment in the obese group of boys. Regarding the relationship between nutritional trajectory and school performance, results indicates that boys drop out at younger ages than girls. However, groups according to its nutritional trajectories do not differ in terms of school dropout age. None of the groups exhibited a deficit or decrease in nutritional status during school period, although the group with overweight or obesity could be of concern and the object of future research.

Keywords: Latent class, Joint model, Nutritional trajectory, School performance, Uruguay

JEL Classification: C33, I12, J13

1. Introducción

El estudio del estado nutricional de las personas es relevante en sí mismo pero, en particular, durante la infancia debido a que los problemas nutricionales en esta etapa de la vida pueden tener efectos adversos en otras dimensiones como, por ejemplo, los logros cognitivos de los niños o la asistencia escolar (ACC/SCN, 2000). Para el diseño de políticas públicas eficientes resulta de interés la asociación que muestra la literatura entre el déficit nutricional y los menores logros educativos alcanzados (Glewwe, 2005; Levinger, 1992; Victora *et al.*, 2008). A su vez, las enfermedades crónicas asociadas al estado nutricional pueden conllevar problemas en lo laboral y consumen recursos de los servicios primarios de salud (ACC/SCN, 2000).

La transición epidemiológica¹ ha llevado a que en los países en desarrollo el foco de atención sea la desnutrición infantil, mientras que en los países desarrollados lo han sido la obesidad y el sobrepeso. Estos últimos problemas nutricionales se han reportado como factores de riesgo para la salud tanto de corto como largo plazo (deOnis & Lobstein, 2010) y recientemente han cobrado mayor atención. La prevalencia de obesidad en el mundo se ha más que duplicado entre 1980 y 2014 según la Organización Mundial de la Salud (OMS), por tanto, entender su trayectoria y las características de los individuos que transitán hacia problemas de obesidad y sobrepeso puede dar un mayor entendimiento del proceso y, por tanto, de las acciones a tomar (Stuart & Panico, 2016).

Si bien en Uruguay, existen trabajos que han analizado la situación nutricional en la niñez y sus determinantes (ver, por ejemplo, Bérgolo *et al.*, 2006) no se encuentran estudios específicos sobre trayectorias nutricionales durante el período escolar. Por tanto, uno de los objetivos de este trabajo consiste en analizar la evolución temporal de las trayectorias para niños y niñas de Uruguay en edad escolar.

Por su parte, Failache *et al.* (2015) estudiaron los determinantes del desempeño escolar en Uruguay, específicamente analizando la probabilidad de repetición y abandono escolar. Los autores concluyen que el déficit nutricional afecta positivamente la probabilidad de repetición. Sin embargo, no reportan efectos directos de este déficit sobre la probabilidad de abandono del sistema educativo.

Estos resultados surgen de la estimación de un *probit* para la probabilidad de repetir y otro para la de abandonar, donde el estado nutricional se introduce para explicar dichas probabilidades. Además, los autores reportan que el nivel educativo y el ingreso del hogar así como características de los niños vinculadas a sus habilidades no cognitivas fueron significativos para explicar los desempeños considerados.

Los resultados reportados por este trabajo motivan el análisis de la asociación entre la trayectoria nutricional y el desempeño escolar que se aborda mediante una metodología diferente que permite analizar dicha asociación considerando a estos procesos dinámicamente en el tiempo.

El tema central de este trabajo consiste en analizar la trayectoria nutricional de una cohorte de niños y niñas de Uruguay y su vínculo con el desempeño escolar mediante la estimación de Modelos Mixtos de Clases Latentes y Modelos Mixtos Conjuntos de Clases Latentes.

Para este análisis se aplica un Modelo Mixto de Clases Latentes para el Índice de Masa Corporal (IMC)² de niños y niñas, que permite identificar grupos o clases a partir de la trayectoria del índice con la edad (Muthén, 2004; Proust-Lima *et al.*, 2014). Este modelo no requiere

¹ La transición epidemiológica consiste en un proceso de cambio dinámico de la morbilidad y mortalidad de la población. Esta transición, acompañada por la demográfica, se caracteriza por disminución de la prevalencia de enfermedades transmisibles y aumento de las no transmisibles, desplazamiento de la morbilidad y mortalidad de los más jóvenes hacia los grupos de mayor edad y aumento de los problemas asociados a la morbilidad en detrimento de la mortalidad. Los países en desarrollo en general se encuentran en una etapa previa de la transición epidemiológica respecto a los desarrollados (Arredondo *et al.* (2003), Wikipedia).

² El IMC es un cociente que relaciona el peso y la estatura de un individuo y se utiliza para evaluar el estado nutricional. La OMS cuenta con estándares de crecimiento por sexo según el IMC de 5 a 19 años.

contar con medidas longitudinales para las mismas edades o en los mismos momentos del tiempo para los niños y niñas, sino que considera la disparidad de medidas o momentos en los que se miden¹. A su vez, se trata de un caso especial de Modelos de Ecuaciones Estructurales donde los parámetros que gobiernan la curva para cada niño o niña se estiman como variables latentes (Johnson, 2015).

El abandono del sistema educativo resulta un problema importante para Uruguay. Según información de la Encuesta Continua de Hogares (ECH) el porcentaje que no asiste al sistema educativo asciende a 3.36 %, 7.41 % y 12.96 % para 13, 14 y 15 años respectivamente en 2012. Esto muestra un problema importante vinculado a la no asistencia en este tramo de edad que empeora notablemente. Si bien para 2016 estas tasas disminuyen el problema persiste. La no asistencia según la información de la ECH 2016 se sitúa en 2.79 %, 5.8 % y 9.54 % para 13, 14 y 15 años respectivamente.

A su vez, los Modelos Conjuntos han cobrado mayor relevancia recientemente en el análisis biométrico. En la literatura internacional existen varias aplicaciones fundamentalmente en el área médica (Lin *et al.*, 2002; Proust-Lima *et al.*, 2009, 2015; Rizopoulos, 2012; Tsiatis & Davidian, 2004; Zajacova & Ailshire, 2012). Existe un único trabajo hecho en Uruguay que aplica esta metodología para el análisis del deterioro cognitivo en adultos mayores para un estudio longitudinal de Suecia (Massa, 2015).

En el presente trabajo, como desempeño escolar se considera el abandono o desvinculación del sistema educativo, que se estudia en el marco de un Modelo de Sobrevivencia donde la variable de interés es la edad en la que el niño deja de estudiar en el sistema educativo formal. A la vez, se considera un Modelo Mixto de Clases Latentes para la trayectoria nutricional. Entonces, se estima un Modelo Mixto Conjunto de Clases Latentes que permite evaluar el vínculo entre la trayectoria nutricional y la edad de abandono escolar.

La información utilizada corresponde al Estudio Longitudinal del Bienestar en Uruguay (EL-BU) realizado por el Instituto de Economía, con apoyo de la Comisión Sectorial de Investigación Científica (CSIC), Unicef y el Fondo Clemente Estable. El estudio consiste en una muestra de niños y niñas que se encontraba cursando primer año en escuelas públicas de capitales departamentales del país en 2004. La encuesta se repite en años ulteriores (2004, 2006 para Montevideo y área metropolitana y 2011/12). En 2015/16 con financiamiento de CSIC se realizó una nueva ola, en la que se entrevistó a los jóvenes con aproximadamente 17 y 18 años de edad. En este trabajo, se utiliza la información disponible hasta 2012².

2. Estado nutricional y desempeño escolar

La privación durante la infancia reporta particular interés en la discusión sobre el bienestar a lo largo del ciclo de vida, puesto que las capacidades accesibles durante esta etapa condicionan el espacio de funcionamientos a los que una persona podrá acceder en su vida adulta (Sen, 1999). La nutrición de las personas es uno de los funcionamientos elementales para sobrevivir y su privación puede generar incapacidades para alcanzar otros funcionamientos. Por ser una capacidad elemental del ser humano, resulta relevante el estudio en sí mismo del estado nutricional y su evolución a lo largo del tiempo.

A su vez, los logros nutricionales pueden ser un medio para alcanzar otras capacidades. Por tanto, la presencia de problemas en la nutrición puede tener efectos, incluso de muy largo plazo, sobre otras dimensiones (ACC/SCN, 2000). En particular, la literatura indica que el déficit nutricional se asocia con menores logros educativos (Georgiadis *et al.*, 2016; Glewwe, 2005; Levinger, 1992; Victora *et al.*, 2008).

¹ Se supone un proceso “*missing at random*”. Ver, por ejemplo, Rizopoulos (2012).

² Los datos definitivos de última ola no estaban disponibles cuando se realizó el análisis para este trabajo.

En los estudios de crecimiento y nutrición en la infancia, se pone de manifiesto la interacción de varios factores -hereditarios, referentes al nivel socio económico, al acceso a servicios y características de la comunidad en las que viven las familias- para determinar el estado y evolución nutricional (ACC/SCN, 2000; Alderman *et al.*, 1997; Behrman & Wolfe, 1987; Behrman, 1990). Asimismo, en la literatura biomédica se tiende a dar mayor relevancia a factores del entorno para condicionar las trayectorias de desarrollo y crecimiento de los niños (Barker *et al.*, 1989; deOnis *et al.*, 2004; Gluckman *et al.*, 2005). La interacción de genética con medio ambiente, da como resultado el crecimiento y nutrición alcanzados, afectando o favoreciendo, la salud a corto, mediano y largo plazo. Además, entender cómo se transita hacia los distintos problemas nutricionales puede dar un mejor entendimiento del proceso y, por tanto, de las acciones a tomar. En este sentido, la comprensión de esta evolución puede guiar la práctica clínica y poner de relieve los mecanismos que se producen antes de que las trayectorias queden “establecidas” (Engle *et al.*, 1997; Regnault & Gillman, 2014; Stuart & Panico, 2016; Victora *et al.*, 2010).

En cuanto a la vinculación entre trayectoria nutricional y desempeño escolar, Glewwe (2005) plantea tres períodos relevantes para el desarrollo del niño en donde su estado nutricional juega un rol importante en su desempeño educativo: desde el nacimiento hasta los dos años, a partir de entonces y hasta el momento de ingresar a la escuela (5 o 6 años) y el período de edad escolar hasta que se evalúan los desempeños. A su vez, indica que un conjunto de decisiones que toman los padres durante la primer infancia del niño, constituyen elementos centrales para comprender los desempeños posteriores de los hijos. En el trabajo de Zhao & Glewwe (2010) se resalta también la importancia de aspectos tales como la motivación del niño, sus gustos por la educación y las habilidades innatas. Además, Cueto & Chinen (2000) indican que un estado nutricional y de salud adecuados no son suficientes para que los niños obtengan logros en términos de aprendizaje.

Antecedentes

El estudio del estado nutricional puede realizarse a través de medidas antropométricas que representan adecuadamente el estado de salud y nutrición en una población. Estas medidas se utilizan en general para identificar problemas nutricionales o de salud de una población y realizar intervenciones correspondientes, así como para monitorear estos problemas y luego aplicar otras herramientas más específicas (deOnis & Habicht, 1996).

En Uruguay, dos estudios contienen medidas antropométricas a lo largo de la infancia para muestras representativas de una población: el ELBU y la Encuesta de Nutrición, Desarrollo Infantil y Salud (ENDIS). La ENDIS es una encuesta longitudinal llevada adelante por el Instituto Nacional de Estadística (INE), en conjunto con Uruguay Crece Contigo (UCC) y la Universidad de la República (UdelaR). Se realiza a hogares donde residen niños de 0 a 3 años y 11 meses de localidades urbanas (más de 5 mil habitantes), habiéndose llevado a cabo dos relevamientos: la primer entrevista en 2014 y otra en 2016.

El ELBU, fuente de información utilizada en el presente trabajo, es un estudio representativo de niños y niñas que cursaban primer año de educación primaria en escuelas públicas de capitales departamentales y el área metropolitana en 2004 y que fueron entrevistados posteriormente en otras tres instancias: 2006, 2011/12 y 2015/16.

A partir de información contenida en el ELBU, trabajos previos han contribuido a analizar el estado nutricional de los niños uruguayos y sus determinantes (ver, por ejemplo, Bérgolo *et al.*, 2006). Sin embargo, los trabajos realizados para Uruguay utilizaron técnicas que se aplican a datos de tipo sección cruzada para el análisis del estado nutricional y sus determinantes.

Identificaron el peso al nacer, la historia reproductiva de la madre, la talla de la madre, el nivel educativo del hogar y en algunos casos la dotación de bienes públicos como principales factores

que afectan la nutrición de los niños y niñas de Uruguay (Bérgolo *et al.*, 2006). Además, la morbilidad del niño y la situación nutricional de la madre resultaron significativamente asociadas con sus logros nutricionales.

Por su parte, varios trabajos internacionales analizan la trayectoria nutricional en la infancia utilizando métodos para datos longitudinales. En particular, Stuart & Panico (2016) analizan la trayectoria nutricional de una cohorte nacida entre 2000 y 2002 en Reino Unido. En el trabajo se estima la trayectoria latente de crecimiento para el IMC utilizando un modelo mixto semi paramétrico, de donde se desprende la existencia de grupos en función de la trayectoria nutricional de los niños. Además, se presentan modelos multinomiales para caracterizar estos grupos en función de características socio económicas y del comportamiento de las familias durante el embarazo y la infancia. Los autores identifican cuatro grupos de trayectorias: dos transitán en niveles del IMC dentro de un rango normal y representan en conjunto alrededor del 85 % de los niños de la cohorte; un segundo grupo representa el 14 % de los niños y se caracteriza por una trayectoria de sobrepeso; el último grupo con el 3 % de los niños mantiene una trayectoria en niveles de obesidad. A su vez, plantean que los perfiles socio económicos de los grupos sugieren que las diferencias pueden provenir de la infancia temprana.

Por otro lado, el trabajo de Georgiadis *et al.* (2016) vincula el estado nutricional en diferentes períodos de la infancia con los logros cognitivos alcanzados por los niños. Los autores indican que la evidencia respecto a la relación entre el estado nutricional, en diferentes períodos de la niñez, y los logros cognitivos es ambigua en la literatura y lo atribuyen en parte a las diferentes formas de modelar las trayectorias de crecimiento. Por tanto, utilizan un modelo general para aproximar la trayectoria nutricional con el fin de comparar distintas metodologías y explorar canales a través de los cuales se relacionan el estado nutricional y los logros cognitivos. Utilizando información proveniente del estudio *Young Lives*¹ de cuatro países (Etiopía, India, Perú y Vietnam), sus resultados sugieren que existe asociación entre los logros cognitivos de los niños a los 8 años y sus trayectorias de crecimiento en distintos períodos de la primera infancia.

Respecto al análisis del desempeño educativo en Uruguay Failache *et al.* (2015) analizan los factores asociados mediante la estimación de la probabilidad de repetir y abandonar, utilizando información contenida en el ELBU. En los modelos para estas probabilidades incluyen como variable explicativa la situación nutricional del individuo en el período previo. Esta variable se construye a partir del valor de la puntuación del *z-score* sobre el IMC, que se obtiene con los programas de cálculo que provee la OMS². Esta variable permite separar entonces a los niños en tres grupos según su estado nutricional: *z-score* por debajo de -1 con problemas de desnutrición, entre -1 y 1 rango de valores normales, es decir, sin problemas nutricionales serios, y por encima de 1 con sobrepeso u obesidad.

Estos autores, en el modelo para la repetición reportan que el déficit nutricional observado en el período previo (*z-score* debajo de -1) se asocia positivamente con una mayor probabilidad de repetir. Esto es, aquéllos niños que presentan déficit nutricional tienen, en promedio, mayor probabilidad de repetir algún año del ciclo educativo que aquéllos que tienen un estado nutricional en el rango normal (*z-score* entre -1 y 1).

Luego, reportan que la repetición afecta significativamente el abandono y, además, plantean una posible endogeneidad de la repetición como regresor para explicar el abandono por lo que estiman mediante variables instrumentales. Sin embargo, plantean que no habría un efecto significativo directo del déficit nutricional sobre la probabilidad de abandono del sistema educativo. Asimismo, haber asistido a enseñanza preescolar, el ingreso del hogar, la educación de la ma-

¹ *Young Lives* es un estudio longitudinal de la pobreza infantil que realiza la Universidad de Oxford que sigue a niños de Etiopía, India (en los estados de Andhra Pradesh y Telangana), Perú y Vietnam, durante aproximadamente 15 años.

² Estos programas de cálculo se pueden encontrar en la web de la OMS.

dre y haber obtenido un puntaje en el rango de normalidad de pruebas psicométricas aparece negativamente relacionado con el abandono (Failache *et al.*, 2015).

Por tanto, esto constituye una motivación del presente trabajo en donde uno de los objetivos consiste entonces en aportar al conocimiento de la relación entre la trayectoria nutricional y el desempeño educativo utilizando una metodología diferente a los trabajos previos. Si bien estos trabajos dan un buen indicio de la asociación entre estado nutricional y desempeño escolar de los niños en Uruguay, la metodología allí utilizada no explota el carácter longitudinal de los datos del ELBU. En este sentido, el análisis realizado hasta el momento ha sido con técnicas para datos de sección cruzada donde no se considera la evolución temporal sino el estado nutricional de los niños en un momento dado.

Otros trabajos para Uruguay, han argumentado que las condiciones de los hogares donde crecen los niños son determinantes de la desvinculación del sistema educativo de los adolescentes. Estas condiciones, en general, se asocian con carencias de ingreso o niveles educativos bajos de los padres (Boado & Fernández, 2010; Bucheli & Casacuberta, 2000).

El problema de desvinculación del sistema educativo es relevante y su estudio constituye un desafío para el caso de educación media ya que allí se encuentran tasas de desvinculación muy elevadas, en tanto, para primaria el problema más importante es la repetición. Según se observa en el Cuadro 1 el porcentaje de varones y mujeres que no asisten al sistema educativo aumenta de forma muy importante con la edad, según datos del Censo 2011. Además, la no asistencia crece de forma más marcada con la edad para los varones que para las mujeres.

Cuadro 1: Porcentaje de no asistencia al sistema educativo por edad y sexo

| Edad | Mujer | Varón | Total |
|-------|-------|-------|-------|
| 12 | 1.6 | 2.1 | 1.8 |
| 13 | 4.1 | 5.6 | 4.9 |
| 14 | 8.5 | 13.1 | 10.8 |
| 15 | 14.6 | 22.0 | 18.4 |
| 16 | 21.8 | 30.5 | 26.2 |
| 17 | 29.4 | 40.9 | 35.3 |
| 18 | 43.6 | 56.2 | 50.0 |
| 19 | 53.2 | 65.1 | 59.2 |
| Total | 21.9 | 29.0 | 25.5 |

Fuente: deMelo *et al.* (2015) en base al Censo 2011.

Para los niños que integran el ELBU, la tasa de abandono escolar en la tercera ola, realizada en 2011/12 cuando tenían entre 13 y 14 años aproximadamente, es del entorno de 9%. A su vez, según la ECH 2011 la tasa desciende hasta 6.3% para el tramo etario de 13 y 14 años que cursaron primaria pública (Failache *et al.*, 2015). Por tanto, si bien la cobertura en educación primaria es prácticamente universal el abandono se vuelve un problema serio en el tramo etario que cursa educación media en Uruguay.

Por otro lado, deMelo *et al.* (2015) muestran tres momentos destacados de desvinculación en educación media: primer año (13 y 14 años), culminación de ciclo medio básico (14 y 15 años) y culminación de ciclo medio superior (17 y 18 años), donde este último tramo sería el de mayor probabilidad de abandono. Por su parte, Aristimuño & deArmas (2012) hacen referencia a dos períodos críticos de abandono, las franjas etarias donde se presentan las tasas más altas de abandono son 14-15 y 18-19 años.

Relevancia del trabajo

La relevancia de la investigación se fundamenta, por un lado, en entender los factores asociados a la problemática nutricional de los niños y su vinculación con el desempeño escolar. Entender estos factores puede contribuir al diseño de políticas y programas públicos que atiendan dichos aspectos. El desarrollo y aplicación de Modelos Conjuntos no se ha hecho aún con este fin, lo cual puede aportar nueva evidencia. En la literatura internacional se encuentran varias aplicaciones de estos modelos, fundamentalmente en el área médica (Proust-Lima *et al.*, 2009, 2015; Rizopoulos, 2012; Tsiatis & Davidian, 2004; Wu *et al.*, 2012). Por su parte, existe un único trabajo en Uruguay que aplica esta metodología para el análisis del deterioro cognitivo en adultos mayores para un estudio longitudinal de Suecia (Massa, 2015).

Como ya fue mencionado, en Uruguay existen otras bases de datos longitudinales que pueden ser fuente para distintos estudios sociales, como la ENDIS, el ELBU y la Encuesta sobre Situaciones Familiares y Desempeños Sociales (ESF)¹. Por lo tanto, existe información que permite el estudio de distintos temas de importancia social a los que se podrían aplicar los modelos estudiados en este trabajo. Tener conocimiento sustantivo de estos modelos, sus ventajas y desventajas, sus propiedades, diferentes extensiones y especificaciones puede extenderse a otros problemas de investigación.

Se propone analizar la evolución temporal del estado nutricional de los niños y niñas mediante la estimación de Modelos Mixtos de Clases Latentes e indagar en la relación de esta trayectoria nutricional con la desvinculación escolar a través de la estimación de Modelos Mixtos Conjuntos de Clases Latentes.

3. Modelos longitudinales y de sobrevivencia

En esta sección se presenta una reseña de los modelos estadísticos que se aplican en el trabajo, los cálculos que se deben realizar luego de la estimación de los modelos y las consideraciones metodológicas referentes a la aplicación de estos modelos al análisis de la trayectoria nutricional y el desempeño escolar.

3.1. Modelos para datos longitudinales

Los Modelos Lineales Mixtos (LMM) se han convertido en una herramienta estándar de análisis de la trayectoria de una medida longitudinal gaussiana a lo largo del tiempo y son utilizados para analizar el efecto de covariables sobre dichas trayectorias (Rizopoulos, 2012; Verbeke & Lesaffre, 1996; Verbeke & Molenberghs, 2000). Sin embargo, debido a que la estructura de datos longitudinales puede ser compleja, que puede llevar al incumplimiento de los supuestos acerca de las distribuciones en que se basan los LMM, se ha planteado la utilización de modelos más generales (Proust-Lima & Jacqmin-Gadda, 2005; Proust-Lima *et al.*, 2014, 2015, 2016b).

Los LMM, además, asumen que el efecto de las covariables es constante a lo largo del rango de valores de la medida longitudinal, por lo que surgen los Modelos Lineales Mixtos Generalizados (GLMM). Estos últimos son una extensión de los LMM para medidas binarias, ordinales o Poisson.

Los Modelos Mixtos de Clases Latentes (LCMM) resultan de utilidad para solucionar algunos problemas que presentan los LMM aplicados a medidas longitudinales. En particular, los LCMM

¹ La ESF es un panel de mujeres de dos olas: la primera en 2001 encuestó una muestra de 1806 mujeres de entre 25 y 54 años residentes en el Área Metropolitana, la segunda en 2008 logró entrevistar a 828 mujeres de la primera ola.

resuelven los problemas generados por el supuesto de continuidad de la medida longitudinal y distribución gaussiana de las desviaciones aleatorias¹ de los LMM (Proust-Lima *et al.*, 2016b).

Los LCMM se enmarcan en la literatura de Modelos de Ecuaciones Estructurales (SEM). En ellos se separa el modelo estructural, que describe la trayectoria del proceso de interés (latente) a lo largo del tiempo en función de covariables, del modelo de medición que vincula este proceso con las observaciones. En este contexto, los Modelos Mixtos para Procesos Latentes (LPMM) suponen un LMM para el proceso latente y separan el modelo de medición (observable) del estructural (latente). Además, los LCMM incorporan la heterogeneidad de la población a través del supuesto de pertenencia a grupos que capturan la evolución diferencial del proceso latente, donde estos grupos se suponen internamente homogéneos en cuanto a la evolución de dicho proceso.

La aplicación de los LCMM, como los GLMM, constituye entonces una alternativa que flexibiliza los supuestos restrictivos de los LMM.

En este contexto, se detallan los tres modelos antes mencionados que se aplican en el contexto de datos longitudinales: LMM, LPMM y LCMM. Se reseñan los principales supuestos, los métodos de estimación usualmente utilizados y los cálculos post-estimación relevantes para este trabajo.

3.1.1. Modelo lineal mixto

Se considera la variable de interés Y_{ij} , correspondiente al individuo i en la ocasión j , medido en el momento t_{ij} , con $i = 1, \dots, N$ y $j = 1, \dots, n_i$. La ocasión j no coincide con el tiempo t_{ij} para permitir diferenciar el momento así como la cantidad de ocasiones de medición entre individuos.

Laird & Ware (1982) plantean el **Modelo Lineal Mixto** que puede expresarse como:

$$Y_{ij} = \mathbf{X}_{Li}(t_{ij})^\top \boldsymbol{\beta} + \mathbf{Z}_i(t_{ij})^\top \mathbf{b}_i + \omega_i(t_{ij}) + \varepsilon_{ij} \quad (1)$$

donde $\mathbf{X}_{Li}(t_{ij})$ y $\mathbf{Z}_i(t_{ij})$ son dos vectores de covariables en el momento t_{ij} de tamaño p y q respectivamente, $\boldsymbol{\beta}$ es un vector en \mathbb{R}^p que contiene los efectos fijos², \mathbf{b}_i es un vector en \mathbb{R}^q que contiene los efectos aleatorios tal que $\mathbf{b}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$. El componente ε_{ij} es la perturbación en cada ocasión, independiente y con varianza constante σ_ε^2 tal que $\varepsilon_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_i = \sigma_\varepsilon^2 \mathbf{I}_{n_i})$. Finalmente, $\omega_i(t_{ij})$ es un proceso estocástico tal que $\omega_i \sim N(\mathbf{0}, \mathbf{R}_i = \sigma_\omega^2 \boldsymbol{\Omega}_i)$, donde los parámetros que definen la matriz $\boldsymbol{\Omega}_i$ se denotan como ρ .

3.1.2. Modelo mixto para procesos latentes

En ciertos análisis de datos longitudinales, puede que el proceso de interés no sea directamente observable y sea necesario aproximarse mediante indicadores o medidas. En estos casos, no es posible obtener el proceso de interés directamente de los individuos sino medidas repetidas de un indicador del mismo y, por tanto, puede resultar pertinente la utilización de un Modelo Mixto para Procesos Latentes (Proust-Lima *et al.*, 2015, 2016b). La idea básica consiste en plantear un **modelo estructural** para el proceso latente y un **modelo de medición** para los indicadores que aproximan dicho proceso de interés. A continuación, se presenta el caso de un único indicador

¹Las desviaciones aleatorias refieren de forma general a los efectos aleatorios, los errores correlacionados y los errores de medida.

² En este caso los efectos fijos corresponden a los parámetros de las covariables del modelo. La literatura económica de datos longitudinales o de panel considera a los efectos fijos con otro significado. Ver, por ejemplo, Cameron & Trivedi (2010); Frees (2004); Wooldridge (2002, 2010).

para aproximar el proceso latente y puede consultarse Proust-Lima *et al.* (2016b) para el caso de múltiples indicadores.

El **modelo estructural** describe la evolución en el tiempo del proceso latente en función de covariables. Éste se dice latente ya que no es observable directamente de los individuos. Por su parte, el **modelo de medición** representa la evolución longitudinal de la medida o indicador observable que intenta capturar dicho proceso latente (Proust-Lima *et al.*, 2016b). El interés se encuentra entonces, en modelar la trayectoria temporal del proceso latente y entender los efectos de covariables, a partir de un indicador medido a los individuos repetidamente en el tiempo.

En el modelo estructural, el proceso latente puede aproximarse correctamente por un LMM estándar sin error de medida, que en tiempo continuo se denota $\Lambda_i(t)$ y se representa como:

$$\Lambda_i(t) = \mathbf{X}_{Li}(t)^\top \boldsymbol{\beta} + \mathbf{Z}_i(t)^\top \mathbf{b}_i + \omega_i(t) \quad \forall t \in \mathbb{R}^+ \quad (2)$$

donde $\mathbf{X}_{Li}(t)$ es un vector de p covariables en t asociado a los efectos fijos $\boldsymbol{\beta}$, $\mathbf{Z}_i(t)$ es un vector de q covariables en t asociado a los efectos aleatorios \mathbf{b}_i (típicamente $\mathbf{Z}_i(t)$ incluye funciones del tiempo). El vector \mathbf{b}_i de q efectos aleatorios satisface los mismos supuestos que en la sección 3.1.1, al igual que el proceso estocástico ω_i pero ahora en tiempo continuo.

Por su parte el modelo de medida, toma en cuenta la diferencia con el proceso latente y vincula $\Lambda_i(t_{ij})$ con Y_{ij} , en los momentos t_{ij} , mediante la **función de enlace**, como sigue:

$$Y_{ij} = H(\tilde{Y}_{ij}; \boldsymbol{\eta}) = H(\Lambda_i(t_{ij}) + \varepsilon_{ij}; \boldsymbol{\eta}) \quad (3)$$

donde ε_{ij} son los errores gaussianos independientes con varianza σ_ε^2 , la función de enlace (H) tiene parámetros $\boldsymbol{\eta}$ y \tilde{Y}_{ij} denota el proceso latente con error en el momento t_{ij} .

Si el indicador es una medida cuantitativa, H^{-1} es una función continua monótona creciente. Pero, también es posible especificar una función discreta. En Proust-Lima *et al.* (2016b) se presentan alternativas para la función H .

3.1.3. Modelo mixto de clases latentes

Tanto el LMM como el LPMM asumen que la población es homogénea y queda representada por un único perfil medio del proceso longitudinal. Esto puede no aproximarse a la realidad. El LCMM asume que la población es heterogénea y está compuesta por G grupos latentes de individuos caracterizados por un perfil medio de la trayectoria del proceso. La existencia de grupos latentes que capturan la heterogeneidad entre individuos asume que no es posible hacerlo a través de ninguna covariable observada (Proust-Lima *et al.*, 2016b).

El LCMM parte de una estructura de G clases latentes donde cada individuo pertenece a un único grupo latente. La pertenencia a estos grupos se define mediante una variable c_i que es igual a g si el individuo i pertenece a g (donde $g = 1, \dots, G$). Cada una de estas G clases genera un perfil medio homogéneo del proceso. A su vez, la variable de medida para el individuo i , en la ocasión j y en el momento t_{ij} se denota como Y_{ij} .

La variable latente (no observable) c_i sigue una distribución de probabilidad que se describe mediante un modelo logit multinomial en función de ciertas covariables \mathbf{X}_{Ci} :

$$\pi_{ig} = P(c_i = g | \mathbf{X}_{Ci}) = \frac{e^{\mathbf{X}_{Ci}^\top \boldsymbol{\lambda}_g}}{\sum_{l=1}^G e^{\mathbf{X}_{Ci}^\top \boldsymbol{\lambda}_l}} \quad (4)$$

donde $\boldsymbol{\lambda}_g$ es un vector de parámetros específicos del grupo g asociados a las covariables \mathbf{X}_{Ci} invariantes temporalmente de dimensión q_1 . Cuando no se introducen covariables este modelo se reduce a una probabilidad específica de cada grupo (solo incluye intercepto).

Ahora se tienen G perfiles longitudinales medios acordes al tiempo y las covariables en un LMM específico, por lo que, los efectos fijos y la distribución de los aleatorios podría diferir entre grupos. Por último, la incorporación de un LPMM disocia el modelo estructural del de medición según lo presentado en la sección 3.1.2.

Entonces, el **modelo estructural específico** queda representado en la siguiente ecuación en tiempo continuo:

$$\Lambda_i(t)|_{c_i=g} = \mathbf{X}_{Li}(t)^\top \boldsymbol{\beta}_g + \mathbf{Z}_i(t)^\top \mathbf{b}_{ig} + \omega_i(t) \quad \forall t \in \mathbb{R}^+ \quad (5)$$

Luego, el **modelo de medición específico**, suponiendo una función de enlace lineal, se presenta en la siguiente ecuación:

$$Y_{ij}|_{c_i=g} = \mathbf{X}_{Li}(t_{ij})^\top \boldsymbol{\beta}_g + \mathbf{Z}_i(t_{ij})^\top \mathbf{b}_{ig} + \omega_i(t_{ij}) + \varepsilon_{ij} \quad (6)$$

que difiere de la ecuación (1) en el vector de efectos fijos $\boldsymbol{\beta}_g$ y aleatorios \mathbf{b}_{ig} que ahora pueden diferir entre grupos. Entonces, la matriz de covarianzas de los efectos aleatorios también puede diferir entre grupos incorporando un coeficiente de proporcionalidad σ_g^2 . Por consiguiente, su distribución es $\mathbf{b}_{ig} \sim N(\mathbf{0}, \mathbf{D}_g = \sigma_g^2 \mathbf{D})$ donde \mathbf{D} es la misma matriz definida en la sección 3.1.1. Para incorporar la separación entre el proceso latente y la evolución de la medida longitudinal en la ecuación (6) se puede introducir una función de enlace H con parámetros $\boldsymbol{\eta}$ como en la sección 3.1.2.

3.2. Modelo de sobrevivencia

En los modelos conjuntos que se aplican en el presente trabajo se modela el tiempo hasta la ocurrencia de un evento mediante la formulación de un modelo de sobrevivencia. En este la variable T_i^* corresponde al tiempo hasta que ocurre cierto evento de interés. Para describir su distribución se utiliza, en general, la función de sobrevivencia (S) que expresa la probabilidad de que el evento ocurra después de t , es decir, la de sobrevivir hasta t :

$$S(t) = P(T^* > t) = \int_t^\infty f(s) \, ds$$

donde f es la función de densidad de T^* que se asume continua. Además, se define la función de riesgo instantáneo de ocurrencia del evento en un intervalo $[t, t + dt]$ dado que se ha sobrevivido hasta t :

$$h(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T^* < t + dt \mid T^* \geq t)}{dt}, \quad t > 0$$

Dadas estas definiciones, es posible establecer una relación entre las funciones anteriores:

$$S(t) = \exp\{-\mathcal{H}(t)\} = \exp\left\{-\int_0^t h(s) \, ds\right\} \quad (7)$$

donde \mathcal{H} es la función de riesgo de ocurrencia del evento acumulado hasta el momento t . Esta función también se interpreta como el número esperado de eventos observados hasta t .

Al estimar modelos de sobrevivencia es necesario tomar en cuenta la existencia de censura. Por tanto, si C_i es el momento donde se presenta la censura, se define el tiempo observado de ocurrencia del evento T_i como el mínimo entre el verdadero tiempo y el momento de censura: $T_i = \min(T_i^*, C_i)$.

Además, se define $D_i = \mathbb{1}_{T_i^* \leq C_i}$, una variable indicadora de la ocurrencia del evento, que vale uno si al individuo i le ocurrió el evento y cero en caso contrario. El análisis de sobrevivencia se interesa en la distribución de T_i^* utilizando la información disponible T_i y D_i .

Una especificación de estos modelos corresponde a la regresión de riesgo relativo (Rizopoulos, 2012; Therneau *et al.*, 1990). En esta especificación se asume que los efectos multiplicativos de las covariables sobre el riesgo instantáneo, por tanto:

$$h_i(t|\mathbf{X}_{Si}) = h_0(t)e^{\mathbf{X}_{Si}^T\boldsymbol{\delta}}$$

donde \mathbf{X}_{Si} es el vector de covariables que se suponen asociadas al riesgo de ocurrencia del evento de cada individuo y $\boldsymbol{\delta}$ el vector de coeficientes. La función $h_0(t)$ corresponde a la de riesgo de línea de base, que es la correspondiente a un individuo con $\mathbf{X}_{Si}^T\boldsymbol{\delta} = 0$.

3.3. Modelos conjuntos para datos longitudinales y de sobrevivencia

En los estudios de seguimiento se suele tener medidas de diferentes resultados longitudinales de los individuos así como del tiempo que transcurre hasta la ocurrencia de un evento de interés. Las preguntas de investigación en base a estos estudios pueden requerir el análisis de la estructura de asociación entre la medida longitudinal y tiempo hasta el evento (Rizopoulos, 2012).

En general, en estos estudios se encuentra una naturaleza dinámica de los resultados. Es decir, la tasa de cambio en el tiempo puede ser diferente entre individuos y además ésta puede cambiar a lo largo del tiempo para el mismo individuo. Por tanto, el potencial de una medida longitudinal para predecir el tiempo del evento y entender su asociación con la sobrevivencia del individuo solamente se alcanza al evaluar medidas repetidas en el tiempo para los individuos (Rizopoulos, 2012).

Para responder preguntas acerca de la asociación entre una medida longitudinal y el tiempo hasta el evento, se han desarrollado modelos estadísticos conocidos como Modelos Conjuntos para Datos Longitudinales y de Sobrevida que han cobrado mayor relevancia recientemente, particularmente en el análisis biométrico (Faucett & Thomas, 1996; Rizopoulos, 2012; Wulfsohn & Tsiatis, 1997).

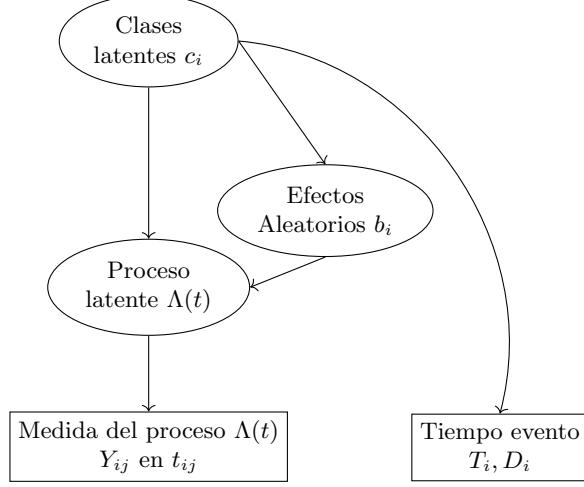
Una aproximación de los modelos conjuntos lo constituye una extensión del modelo presentado en 3.1.3 conocido como Modelo Mixto Conjunto de Clases Latentes (JLCMM). El JLCMM, tiene como supuesto fundamental que la estructura de grupos latentes captura la asociación entre la trayectoria longitudinal de un proceso y el riesgo de ocurrencia del evento.

La existencia de grupos latentes implica una trayectoria específica del proceso longitudinal, que se puede aproximar por un LMM o LPMM, y un modelo de sobrevivencia específico¹ para cada uno de los grupos. La hipótesis clave del modelo, conocida como **independencia condicional**, es que la evolución de la medida longitudinal y el riesgo del evento son independientes condicional a los grupos latentes (Proust-Lima *et al.*, 2009, 2014).

La Figura 1 muestra un grafo que representa el esquema del JLCMM donde la medida longitudinal Y_{ij} disponible para los momentos t_{ij} con $j = 1, \dots, n_i$ está generada por el proceso latente $\Lambda(t)$, medido en tiempo continuo $t \in \mathbb{R}^+$.

¹Esto refiere a un riesgo de ocurrencia del evento específico para cada grupo.

Figura 1: Grafo dirigido que define el JLCMM



Fuente: Proust-Lima *et al.* (2015)

3.3.1. Modelo mixto conjunto de clases latentes

Como fue mencionado, el JLCMM se compone de un modelo de pertenencia a las clases, LMM o LPMM presentados en 3.1 y un modelo de sobrevivencia como el de 3.2. Sin embargo, este último será específico para cada grupo. Por tanto, en el grupo latente g el riesgo del evento queda:

$$h_i(t|c_i = g) = h_{0g}(t)e^{\mathbf{X}_{Si}^T \boldsymbol{\delta}_g} \quad (8)$$

donde \mathbf{X}_{Si} es un vector de covariables asociadas al vector de parámetros $\boldsymbol{\delta}_g$ que pueden tener efectos específicos en cada grupo g . La función $h_{0g}(t)$ corresponde a la función de riesgo de línea de base, que depende de ciertos parámetros ξ_g , que puede diferir entre grupos.

En definitiva, el JLCMM se representa de la siguiente forma:

$$\left\{ \begin{array}{l} \pi_{ig} = P(c_i = g | \mathbf{X}_{Ci}) = \frac{e^{\mathbf{X}_{Ci}^\top \boldsymbol{\lambda}_g}}{\sum_{l=1}^G e^{\mathbf{X}_{Ci}^\top \boldsymbol{\lambda}_l}} \\ Y_{ij}|_{c_i=g} = \mathbf{X}_{Li}(t_{ij})^\top \boldsymbol{\beta}_g + \mathbf{Z}_i(t_{ij})^\top \mathbf{b}_{ig} + \omega_i(t_{ij}) + \varepsilon_{ij} \\ h_i(t|c_i = g) = h_{0g}(t)e^{\mathbf{X}_{Si}^T \boldsymbol{\delta}_g} \end{array} \right. \quad (9)$$

donde $\boldsymbol{\lambda}_g$ es un vector de parámetros específicos del grupo g asociados a las covariables \mathbf{X}_{Ci} invariantes temporalmente, $\boldsymbol{\beta}_g$ y \mathbf{b}_{ig} corresponden al vector de efectos fijos y aleatorios para cada grupo g y la matriz de covarianzas de los efectos aleatorios puede diferir entre grupos incorporando un coeficiente de proporcionalidad σ_g^2 . Por tanto, $\mathbf{b}_{ig} \sim N(\mathbf{0}, \mathbf{D}_g = \sigma_g^2 \mathbf{D})$ donde \mathbf{D} es la misma matriz definida en 3.1.1. Se puede vincular el proceso latente con la evolución de la medida longitudinal agregando una función H con parámetros $\boldsymbol{\eta}$ como en la sección 3.1.3. Además, se incorpora una función de riesgo específico para cada grupo g .

3.4. Cálculos posteriores a la estimación

Una vez estimados los modelos se realizan una serie de cálculos posteriores que permiten validarlos y, para los modelos de clases latentes, permiten elegir el número de grupos que mejor

ajusta. En general, esta elección debe basarse en un conjunto de consideraciones y no solamente en minimizar algún criterio de información (Proust-Lima *et al.*, 2014).

Para realizar la clasificación *a posteriori* se computan las probabilidades de pertenencia a las diferentes clases para cada individuo, basándose en la probabilidad *a posteriori* de la fórmula de Bayes. Esta probabilidad se utiliza para la clasificación de los sujetos así como para evaluar la bondad de ajuste del modelo mediante tablas de probabilidades medias de pertenencia a los grupos (Proust-Lima *et al.*, 2016b).

Otras medidas utilizadas para elegir el número de grupos son el **Criterio de Información Bayesiano** (BIC), que resulta usualmente utilizado en el contexto de modelos mixtos (Proust-Lima *et al.*, 2014, 2016b), y el **Índice de Entropía** (E), que muestra una medida de la calidad de discriminación que se logra con ese número de grupos. Un valor del BIC más pequeño indica un mejor ajuste mientras que un valor de E más cercano a 1 representa una mejor discriminación de los individuos en los grupos considerados.

3.5. Aspectos metodológicos: trayectoria nutricional y desempeño escolar

A continuación se explicitan las consideraciones metodológicas realizadas al aplicar los modelos antes descritos al análisis de la trayectoria nutricional y su relación con el desempeño escolar.

En primer lugar, se elige el IMC como indicador del estado nutricional de los niños y niñas. De las medidas antropométricas disponibles, el IMC para la edad se utiliza generalmente y presenta estándares de crecimiento dados por la OMS para niños y niñas en el rango de edad considerado en este trabajo (Butte *et al.*, 2007; deOnis & Habicht, 1996; deOnis *et al.*, 2004, 2007; deOnis & Lobstein, 2010).

En segundo lugar, el desempeño escolar considerado es la edad de desvinculación del sistema educativo. Esto se basa en que el problema del abandono es importante y presenta tasas altas en Uruguay que comienzan a ser elevadas particularmente a partir de los 13-14 años, edad que tienen los niños del ELBU en la entrevista del 2011/12. Además, se considera la edad de abandono y no otros desempeños, como la repetición por ejemplo, ya que es la información que está disponible en el ELBU para aplicar los modelos. En los mismos se analiza el momento en que se produce el evento, lo que está disponible únicamente para el abandono, si bien se preguntó sobre la repetición no se cuenta con información del momento en que se produce el episodio y, por tanto, no es posible analizar este desempeño mediante la metodología propuesta¹.

Por último, el análisis se hace por separado para niñas y niños dado que las tasas de crecimiento y desarrollo en el tramo de edad considerado difieren según género (deOnis & Habicht, 1996; deOnis *et al.*, 2007; Stuart & Panico, 2016).

3.5.1. Base de datos

La información pertenece a las primeras rondas del ELBU. Éste constituye una muestra de niños que cursaba primer año en escuelas públicas en 2004 representativa de las capitales departamentales que fue seguida en años ulteriores (2004, 2006, 2011/12 y 2015/2016, aproximadamente a los 6, 8, 14 y 18 años). Dado que las tasas de asistencia en primer año escolar son universales, la muestra no tiene los problemas de sesgo de selección que habitualmente presentan los relevamientos realizados a estudiantes en la adolescencia. La muestra fue seleccionada en base al Tercer Censo de Talla Escolar de 2002 (Amarante *et al.*, 2007). Es representativa de

¹ En el ELBU tampoco se tiene información de la ocurrencia de otros desempeños relevantes, como podría ser un test de aprendizaje por ejemplo, que puedan ser objeto de análisis mediante la metodología propuesta en este trabajo.

los niños que concurrían a primer año de la educación primaria pública -aproximadamente un 85 % de los niños que estaban en primer año en 2004- en capitales departamentales y el área metropolitana, si bien presenta un truncamiento en los estratos socio económicos más altos, ya que solo el 43 % de los niños del quintil más alto concurrían a escuela pública ese año (Failache *et al.*, 2015). Las escuelas fueron seleccionadas de los departamentos Artigas, Canelones, Colonia, Florida, Montevideo, Paysandú, Rivera y Soriano.

En la primera ola, se realizaron más de 3000 encuestas a hogares y para 2811 se dispuso de información socio económica y antropométrica. En 2011/12 se re entrevistaron 2152 hogares, lo cual implica que se recuperó el 65 % del panel inicial (desgranamiento del 35 %)¹. Combinando ambas olas se cuenta con 1563 casos con información antropométrica y socio económica para ambas. Por su parte, en 2006 fueron entrevistados 1160 hogares de Montevideo y Canelones, de los cuales 900 se recuperaron en 2011/12 (Arim *et al.*, 2014).

En el marco de un proyecto CSIC, se obtuvo financiamiento para la realización de una nueva ola de este estudio. Esto permitió obtener información tanto socio económica como antropométrica para cuando tienen entre 17 y 18 años. Aún no se cuenta con la base de datos final, pero con ella se podrán analizar aspectos poco explorados en este trabajo lo que constituye una línea de investigación a futuro.

El ELBU contiene la información requerida para este trabajo. En primer lugar, se tomaron medidas antropométricas en el primer año escolar, al año siguiente y, por último, en 2009 las medidas se tomaron al finalizar el ciclo primaria. En definitiva, se cuenta con una base de datos longitudinal con tres momentos de medición de peso y talla de los niños que darán cuenta de su trayectoria nutricional. Cabe mencionar que la segunda ola fue realizada únicamente para los niños de Montevideo y Área Metropolitana y, además, el estudio presenta problemas usuales de datos faltantes. Con las medidas de peso y talla de niños se construye el IMC, que aproxima la situación nutricional, además, se calculan los *z-scores* para el IMC según la metodología de la OMS.

Se cuenta con la fecha de nacimiento y de toma de las medidas que permiten calcular el *z-score* de los individuos a lo largo de las distintas mediciones en el estudio. Estos *z-score* se utilizan para tomar valores que se encuentran dentro de los rangos considerados posibles según la OMS². En los modelos donde se analiza la evolución de la trayectoria nutricional las variables utilizadas son la edad y edad al cuadrado (en meses) y el IMC.

Por otro lado, el ELBU contiene información socio económica de los hogares de los niños, así como información referente al nacimiento del niño, recogida de forma retrospectiva, características de los integrantes del hogar, particularmente la madre del niño. En las últimas olas se cuenta, además, con resultados de test psicológicos que intentan captar características asociadas a las habilidades cognitivas y no cognitivas de los niños³.

Para analizar la asociación entre trayectoria nutricional y desempeño escolar en el JLCMM, se cuenta con las medidas del IMC en las distintas olas, en 2011/12 se preguntó si el niño o niña continuaba estudiando y, en caso que no lo haga, la edad que dejó de estudiar. El desempeño escolar considerado analizado es la edad de abandono. En la última ronda del estudio considerada (2011/12) se preguntó si el niño continúa estudiando en ese momento y, en caso que haya dejado los estudios, la edad de abandono. Esta variable es reportada por la madre del niño o niña ex-post a las mediciones de la trayectoria nutricional. Para el JLCMM, el desempeño escolar

¹ De los casos entrevistados en la tercera ola, algunos no cuentan con información antropométrica y otros niños adicionales fueron medidos en 2009 pero no se logró concretar la entrevista en el hogar en 2011/12.

² Los valores de *z-score* menores a -5 y mayores que 5 no se consideran humanamente posibles según la OMS. Por tanto, en las estimaciones del presente trabajo no fueron consideradas las ocasiones donde el *z-score* del IMC pasara los umbrales de valores posibles dados por la OMS

³ Las habilidades cognitivas se recogieron en la última ola del estudio en 2015/16 en tanto que las habilidades no cognitivas se exploraron en las dos últimas entrevistas, 2011/12 y 2015/16

se analiza en el marco de un modelo de sobrevivencia, por tanto, es necesario tener la edad de abandono del niño, que es una variable disponible en esta ola del ELBU.

La revisión de antecedentes indica que en el análisis del desempeño escolar es importante considerar el fracaso. Por tanto, en este trabajo se introduce en el modelo una variable que indica si el niño repitió algún año durante el ciclo educativo previo o no al momento de la última entrevista. Esta información se tiene únicamente para los niños entrevistados en 2011/12 y dado que la repetición es una covariante en el modelo conjunto, este análisis se realiza con la muestra de niños y niñas entrevistados en 2011/12 en sus hogares y cuentan con información sobre la repetición previa.

3.5.2. Estrategia empírica

En este sección se presentan los modelos estimados y algunas consideraciones metodológicas que se realizan para la estimación de estos modelos.

Análisis de la trayectoria nutricional

Este análisis se realiza estimando los modelos LCMM para niñas y niños por separado. La idea es indagar en la estructura de grupos que mejor captura la heterogeneidad de las trayectorias nutricionales de niños y niñas. Para ello se estiman cinco modelos LCMM fijando el número de grupos de uno hasta cinco. Luego de la estimación, mediante criterios de bondad de ajuste y la clasificación *a posteriori*, se elige el número de grupos más adecuado. La estimación se realiza fijando como máximo cinco grupos dado que aumentarlo implica aumentar el número de parámetros y se puede comprometer la convergencia de los modelos¹.

Fijado el número de grupos, cada modelo se estima suponiendo una **función de enlace lineal** y una **I-spline cuadrática** con tres nodos en los cuantiles (Proust-Lima *et al.*, 2016b; Ramsay, 1988). Por tanto, se tienen diez parametrizaciones por sexo para el análisis de la trayectoria nutricional.

La especificación de los LCMM se realiza de la siguiente forma:

- trayectoria cuadrática según edad diferenciando coeficientes para cada grupo
- efectos aleatorios en el intercepto y coeficiente lineal de la trayectoria
- matriz de covarianzas de los efectos aleatorios no diagonal (*unstructured*)
- matriz de covarianzas de los efectos aleatorios proporcional entre grupos

Se especifica una trayectoria cuadrática que resulta adecuada para representar la evolución del IMC en este rango de edad. Además, se cuenta con tres medidas del indicador como máximo por individuo, lo que no permite especificar formas no lineales de mayor orden. Respecto a los efectos aleatorios, se considera finalmente esta especificación luego de considerar otras que no aportaban diferencias significativas para los objetivos del trabajo y que comprometían la convergencia en el proceso de estimación.

Por tanto, las ecuaciones que representan el LCMM finalmente estimado es:

¹ En algunos casos, para un número elevado de grupos no se logró la convergencia de los modelos y esto puede ser porque la heterogeneidad de la población no corresponde con un número tan elevado de grupos.

$$\begin{cases} \pi_{ig} = P(c_i = g | X_{Ci}) = \frac{e^{\lambda_g}}{\sum_{l=1}^G e^{\lambda_l}} \\ \text{IMC}_{ij}|_{c_i=g} = \beta_{0g} + \beta_{1g}\text{Edad}_{ij} + \beta_{2g}\text{Edad}_{ij}^2 + b_{0ig} + b_{1ig}\text{Edad}_{ij} + \varepsilon_{ij} \end{cases} \quad (10)$$

donde la probabilidad de pertenencia a las clases no está condicionada por ninguna característica observable de los individuos, el modelo de medición en la segunda ecuación representa la trayectoria del IMC para el individuo i en las ocasiones $j = 1, \dots, n_i$. En la ecuación de trayectoria del IMC con función de enlace con *splines* se debe incluir la función H con sus correspondientes parámetros. Por último, en cuanto a los efectos aleatorios y la perturbación se supone que $\mathbf{b}_{ig} \sim N(\mathbf{0}, \sigma_g^2 \mathbf{D})$, $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ y son independientes.

Relación entre trayectoria nutricional y desempeño escolar

Este análisis se realiza mediante los JLCMM por separado para niñas y niños. Se destaca que las muestras utilizadas para estimar estos modelos no coinciden con las muestras de los LCMM. Esto es porque niñas y niños deben tener información sobre abandono y repetición escolar¹ para ser incluidos en la estimación de los JLCMM pero no es necesario que la tengan para los LCMM. Dado que las muestras no coinciden, se vuelven a estimar los LCMM utilizando la muestra de los JLCMM. De esta forma, los modelos que consideran únicamente las trayectorias nutricionales (LCMM), permiten analizar la estructura de grupos y comparar sus resultados con los del JLCMM.

Los JLCMM tienen las siguientes características:

- trayectoria cuadrática según edad diferenciando coeficientes para cada grupo
- efectos aleatorios en el intercepto y coeficiente lineal de dicha trayectoria
- matriz de covarianzas de efectos aleatorios no diagonal (*unstructured*)
- función de riesgo de línea de base *Weibull* igual para todos los grupos

La especificación correspondiente a la trayectoria nutricional es la misma que la utilizada para los LCMM. En cuanto al modelo de sobrevivencia se especifica una función de riesgo de línea de base sencilla sin diferencias entre grupos.

El modelo se escribe entonces:

$$\begin{cases} \pi_{ig} = P(c_i = g | X_{Ci}) = \frac{e^{\lambda_g}}{\sum_{l=1}^G e^{\lambda_l}} \\ \text{IMC}_{ij}|_{c_i=g} = \beta_{0g} + \beta_{1g}\text{Edad}_{ij} + \beta_{2g}\text{Edad}_{ij}^2 + b_{0ig} + b_{1ig}\text{Edad}_{ij} + \epsilon_{ij} \\ h_i(t|c_i = g) = h_0(t)e^{\{\gamma_{0g} + \gamma_{1g}\text{Repitio}_i\}} \end{cases} \quad (11)$$

donde la ecuación de pertenencia a las clases y la trayectoria para el IMC son análogas a las del modelo en (10) aunque aquí solamente se considera una función de enlace lineal, por lo tanto no aparece la función H . Por su parte, el modelo de sobrevivencia se estima incluyendo la covariante Repitio_i en un modelo de riesgo proporcional y $h_0(t)$ es la función de riesgo de línea de base.

¹ La pregunta sobre repetición y edad de abandono escolar en el estudio ELBU fue realizada de forma retrospectiva en la entrevista del 2011/12, por tanto, la muestra utilizada corresponde a los niños y niñas con información en *Repitio*

Estimación de los modelos

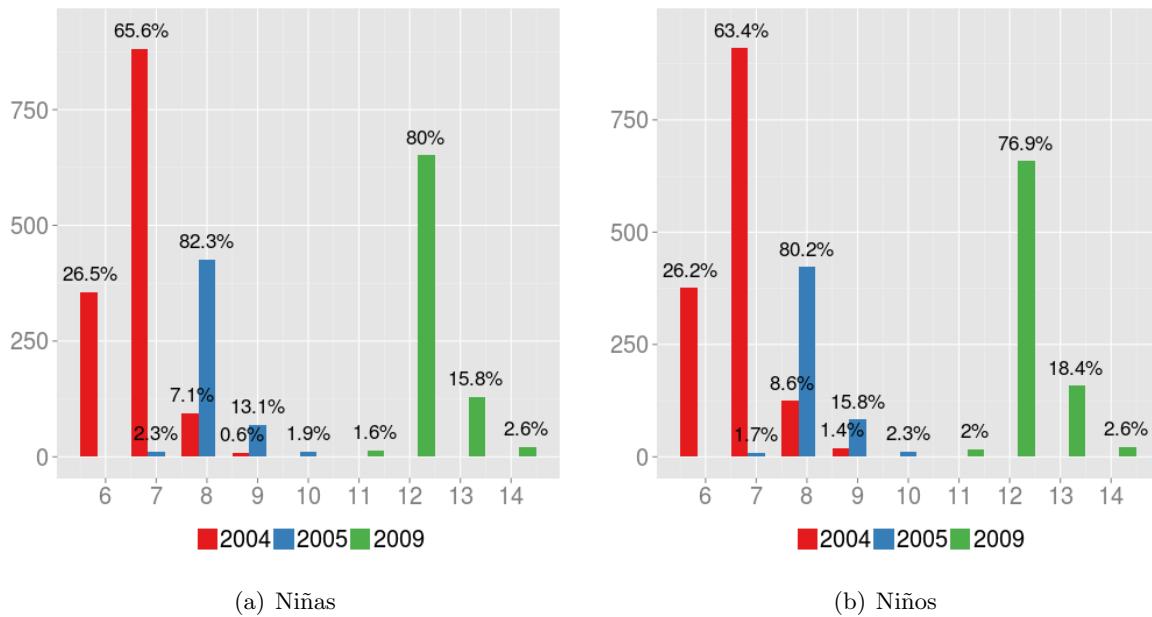
Todas las estimaciones se realizaron utilizando **R**, particularmente las funciones de la librería **lcm** que se describe en Proust-Lima *et al.* (2016a). Los modelos se estiman por el método de máxima verosimilitud, tanto la función **lcm** como **jlcmm**. En el proceso de estimación se toma en cuenta que la función de verosimilitud posiblemente tenga óptimos locales (Muthén, 2004; Proust-Lima *et al.*, 2016b; Rizopoulos, 2012), por tanto, el modelo se estima varias veces imponiendo diferentes puntos iniciales. Se realiza el procedimiento propuesto en Biernacki *et al.* (2003) para disminuir la probabilidad de identificar óptimos locales. Esto es, los resultados finales de la estimación se obtienen estimando 30 modelos partiendo de diferentes puntos iniciales, luego se consideran las estimaciones de los parámetros en el que tiene mayor valor en la función de verosimilitud en el óptimo y estas estimaciones se utilizan como valores iniciales de una nueva iteración finalmente reportada como resultado del trabajo.

4. Resultados

4.1. Estadísticas descriptivas

A continuación, se presenta una descripción de la información relevante para el análisis contenida en el ELBU. En primer lugar, dado que la cohorte de niños y niñas se conforma con los que estaban en primer año de escuela pública en 2004, las edades entre los participantes del estudio pueden diferir en cada momento de la medición. Las medidas antropométricas utilizadas para analizar la situación nutricional de los niños y niñas tienen estándares para la edad al momento de realizar las medidas, por ello se utiliza la edad como variable que indica el paso del tiempo en los modelos analizados.

Figura 2: Distribución según edades simples en los años de medición



Nota: Porcentaje de niños o niñas en esa edad del total de medidas ese año.

Fuente: Elaboración propia

La Figura 2 muestra la distribución según género de las edades simples de los participantes en los años donde se realizaron las medidas antropométricas, es decir en 2004, 2005 y 2009. Además,

se debe tener en cuenta la pérdida de casos en el estudio que llega a un 35 %, distribuida de forma similar según género (ver Cuadro 2).

De la Figura 2 se desprende que en 2004 más del 90 % tenía entre 6 y 7 años tanto para niñas como niños. En la segunda ronda de medición, realizada en 2005 únicamente para las escuelas de Montevideo y el área metropolitana, más del 90 % tenía entre 8 y 9 años. Por último, en 2009 más del 90 % tenía entre 12 y 13 años. En suma, el estudio ELBU nos permite analizar la trayectoria nutricional de los niños entre los 6 y los 12 años aproximadamente.

Cuadro 2: Distribución de casos por género

| | Niñas | Niños |
|---------------------------------------|-------|-------|
| Medidas en 2004 ^a | 1341 | 1432 |
| Medidas en 2005 ^a | 518 | 525 |
| Medidas en 2009 ^a | 805 | 835 |
| Total participantes ELBU ^b | 1570 | 1691 |
| Total muestra LCMM | 1392 | 1492 |
| Total muestra JLCMM | 971 | 1026 |

^a Cantidad de medidas válidas según OMS ^b En 2004

Fuente: Elaboración propia

Por su parte, en el Cuadro 2 se observa que la muestra utilizada para los LCMM y JLCMM es menor al conjunto de niños entrevistados en el ELBU en la primer ronda ya que como fue mencionado, no se cuenta con información antropométrica para todos los niños y niñas entrevistados (Cuadro 2). Además, en las estimaciones fueron excluidas las ocasiones donde el *z-score* del IMC se encontraba fuera de los límites humanamente posibles según la OMS. Por último, la muestra para el JLCMM excluye además los casos con dato faltante en la covariable si repitió algún previamente que fue relevada en la entrevista realizada en 2011/12.

Cuadro 3: Descriptivas del IMC en los años de medición

| | Niñas | | | Niños | | |
|-----------------|------------|------------|------------|------------|------------|------------|
| | 2004 | 2005 | 2009 | 2004 | 2005 | 2009 |
| $z < -2^a$ | 0.8 % | 1.0 % | 2.4 % | 1.5 % | 1.5 % | 3.4 % |
| $-2 < z < -1^b$ | 9.8 % | 6.8 % | 11.0 % | 9.3 % | 7.3 % | 12.1 % |
| $-1 < z < 1^b$ | 65.9 % | 61.5 % | 58.5 % | 66.7 % | 63.0 % | 52.3 % |
| $1 < z < 2^c$ | 16.5 % | 21.2 % | 21.1 % | 14.0 % | 17.2 % | 21.3 % |
| $2 < z^d$ | 7.0 % | 9.5 % | 7.0 % | 8.5 % | 11 % | 10.9 % |
| Media (SD) | 16.1 (2.2) | 17.0 (2.8) | 19.5 (3.5) | 16.1 (1.9) | 16.8 (2.4) | 19.1 (3.4) |
| N | 1341 | 518 | 805 | 1432 | 525 | 835 |

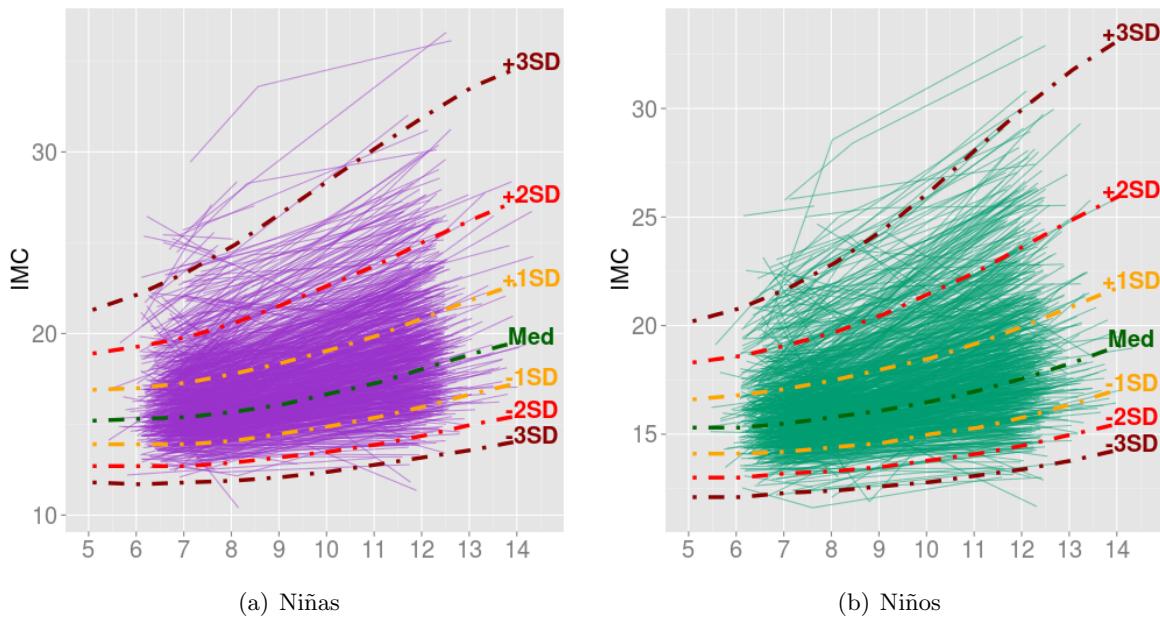
^a Bajo peso ^b Normal ^c Sobre peso ^d Obesidad

Fuente: Elaboración propia

El Cuadro 3 muestra la proporción de niños y niñas en los distintos umbrales definidos por el *z-score* para la edad del IMC. Se observa que en 2004, más del 20 % presenta problemas de sobre peso u obesidad, mientras que aproximadamente un 1 % tiene bajo peso tanto para niñas como niños. A su vez, en las siguientes olas del estudio el porcentaje de niñas y niños con bajo peso se incrementa y, al mismo tiempo lo hace el de sobre peso. En efecto, en 2009 hay una mayor proporción de niños y niñas con problemas nutricionales de bajo peso y de sobre peso, según el *z-score* para la edad.

La Figura 3 muestra las trayectorias del IMC para la edad según género para los participantes en el ELBU que fueron medidos al menos en alguna ocasión y que la medida correspondiente

Figura 3: Trayectorias del IMC para la edad según género



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Fuente: Elaboración propia

resultó dentro de los umbrales válidos según la OMS.

Por otro lado, se presenta la tasa de abandono de niños y niñas del ELBU según si repitieron o no algún año del ciclo educativo previo a la entrevista realizada en 2011/12. El Cuadro 4 muestra la tasa de abandono reportada en 2011/12 donde para los participantes del ELBU asciende a 9 %, cuando más del 90 % tenía entre 13 y 14 años. Este valor resulta superior al que surge de la ECH para 2011 que es 6 % aproximadamente en el mismo tramo etario. Además, resulta notable la diferencia en esta tasa según si experimentaron repeticiones previas ya que abandona un 3 % de los que no repitieron mientras que el abandono representa el 18.4 % de los que repitieron. Por su parte, las tasas de abandono también presentan diferencias según género y éstas se acentúan al considerar los episodios de repetición previa de los individuos.

Cuadro 4: Tasa de abandono según repetición por género

| | No repitió | | Repetió | | Total | |
|-------|------------|------|------------|-----|------------|------|
| | % Abandono | N | % Abandono | N | % Abandono | N |
| Niñas | 2.4 % | 681 | 16.6 % | 359 | 7.3 % | 1040 |
| Niños | 3.8 % | 605 | 19.6 % | 502 | 10.9 % | 1107 |
| Total | 3.1 % | 1286 | 18.4 % | 861 | 9.1 % | 2147 |

Fuente: Elaboración propia

Cuadro 5: Tasa de abandono según edad por género

| Edad | 11 años | 12 años | 13 años | 14 años | N Abandono | N |
|-------|---------|---------|---------|---------|------------|------|
| Niñas | 5.5 % | 17.3 % | 37.3 % | 34.6 % | 75 | 1040 |
| Niños | 1.7 % | 19.5 % | 35.6 % | 40.7 % | 118 | 1107 |
| Total | 3.1 % | 18.7 % | 36.3 % | 38.3 % | 193 | 2147 |

Fuente: Elaboración propia

Por último, el Cuadro 5 muestra distribución de las niñas y niños que declaran haber abandonado los estudios en 2011/12 según edad. Del los que abandonan, más del 70 % lo hace a los 13 o 14 años de edad tanto para niñas como niños. Además, esta cohorte de niños y niñas no es representativa de niños mayores a 14 años por tanto se omitieron los abandonos presentes en este tramo¹.

4.2. Análisis de la trayectoria nutricional

En primer lugar, se detallan los resultados relativos a la elección del número de grupos según la trayectoria nutricional. Se estiman los modelos, por separado para niñas y niños, con función de enlace lineal y con una *splines* cuadrática en el modelo de medición. En todos los casos, se fija el número de grupos de uno a cinco y se comparan indicadores de ajuste y las probabilidades *a posteriori*. Para la elección del número de grupos que mejor representa la heterogeneidad poblacional, se considera el porcentaje de casos en cada grupo y se sigue un criterio de parsimonia para la elección del número de grupos. Este criterio implica que ante modelos que arrojan resultados similares en términos de ajuste y distribución de los individuos entre grupos, se elige el de menor número de parámetros a estimar. Es decir, el de menor número de grupos.

Por último, una vez elegido el número de grupos se analizan diferencias entre los niños y niñas pertenecientes a diferentes grupos, en algunas características asociadas al estado nutricional según se ha reportado en otros trabajos.

Como criterio general para la presentación de los resultados, la etiqueta de los grupos es de forma ascendente con el tamaño de los mismos. De esta forma, “G 1” es el grupo con mayor porcentaje de niñas o niños del total de la muestra que fueron clasificados *a posteriori* en él, “G 2” es el segundo grupo en términos de este porcentaje y así sucesivamente con los restantes grupos.

4.2.1. Elección del número de grupos

Los resultados con función de enlace lineal para niñas (Cuadro 6) muestran que al aumentar el número de grupos hasta tres disminuye el BIC, luego aumenta para cuatro y baja nuevamente para cinco. Esto indica que, en términos de ajuste, resulta adecuado clasificar a las niñas en tres grupos. Por su parte, la entropía para esta clasificación es 0.513, indicando que la separación de niñas en tres grupos es medianamente nítida (Muthén, 2004). Las trayectorias nutricionales de las niñas según tres grupos pueden observarse en la Figura 4 panel *b* donde se aprecia cierta superposición de las trayectorias de niñas en distintos grupos.

El Cuadro 6 contiene además el porcentaje clasificado *a posteriori* en cada grupo. Con una estructura de dos grupos se observa que uno contiene más del 80 % de las niñas, con tres grupos los dos más grandes contienen un porcentaje de niñas similar y el más pequeño no alcanza el 10 %. Luego, para cuatro y cinco grupos la distribución de niñas entre ellos es muy similar, ya que el quinto grupo contiene menos del 2 % de las niñas.

También debe observarse la tabla de clasificación ya que aporta información acerca de la capacidad de discriminación del número de grupos. El Cuadro 7 contiene información respecto a la clasificación para tres grupos: probabilidad media *a posteriori* (panel *a*) y porcentaje de niñas que superan cierto umbral en su probabilidad máxima (panel *b*). En el panel *a*, se observa que, dado el grupo al que pertenece la niña, la probabilidad de ser clasificada en él supera para

¹ En el ELBU la tasa de abandono a los 15 años es entre 3 y 5 % dependiendo del género y en la ECH para es entorno a 10 %.

Cuadro 6: Estadísticos y clasificación por grupos de niñas con LCMM y enlace lineal

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niñas en cada grupo | | | | |
|--------|---------------------|---------|---------|-----------------|-----------------------|------|-----|-----|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -5775.9 | 11602.5 | | 7 | 100 % | | | | |
| 2 | -5524.2 | 11135.4 | 0.644 | 12 | 86 % | 14 % | | | |
| 3 | -5471.4 | 11065.9 | 0.513 | 17 | 49 % | 42 % | 9 % | | |
| 4 | -5457.3 | 11073.9 | 0.564 | 22 | 57 % | 32 % | 8 % | 3 % | |
| 5 | -5409.2 | 11013.8 | 0.538 | 27 | 51 % | 34 % | 8 % | 5 % | 2 % |

N=1392

^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

Fuente: Elaboración propia

Cuadro 7: Probabilidades *a posteriori* para tres grupos de niñas con LCMM y enlace lineal

| | (a) Prob. media | | | (b) % >umbral ^a | | |
|-----|-----------------|-------|-------|----------------------------|-------|-------|
| | G 1 | G 2 | G 3 | > 0.7 | > 0.8 | > 0.9 |
| G 1 | 0.745 | 0.239 | 0.016 | 69 % | 37 % | 2 % |
| G 2 | 0.134 | 0.785 | 0.081 | 70 % | 55 % | 27 % |
| G 3 | 0.005 | 0.138 | 0.857 | 78 % | 69 % | 56 % |

N=1392

^a % de niñas con probabilidad *a posteriori* que supera el umbral

Fuente: Elaboración propia

todos los grupos 0.74 en promedio. Esto es, la diagonal de el Cuadro de clasificación muestra sustento en la discriminación que se logra con tres grupos de niñas.

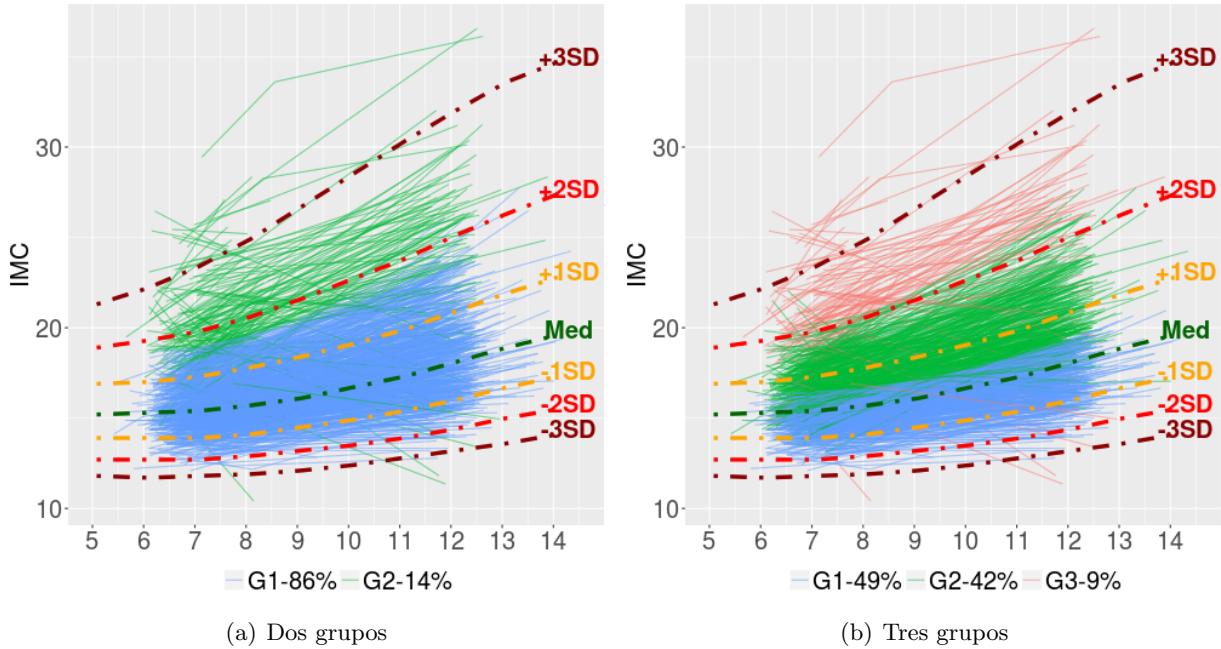
Por otro lado, en el panel *b* se observa que el “G 3” tiene más de la mitad de las niñas con probabilidad máxima superior a 0.9, lo que indica clara pertenencia a este grupo. Para “G 2” la mitad de las niñas superan 0.8 mientras que en el “G 1” aproximadamente el 70 % supera 0.7 de probabilidad, si bien hay solo un 2 % que supera el 0.9 en este grupo. Por tanto, si bien la entropía es intermedia para la clasificación en tres grupos de niñas el análisis de las probabilidades *a posteriori* muestra que tres grupos de niñas logra una buena discriminación.

En las Figuras 4 y 5 pueden verse las trayectorias nutricionales de cada niña según el grupo. Estas trayectorias se representan junto con los umbrales del IMC para la edad que proporciona la OMS, estos son, la mediana del IMC para la edad (*Med*) y los desvíos estándar ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$)¹.

Como ya fue mencionado, existe cierta superposición de las trayectorias de las niñas pertenecientes a los diferentes grupos. Sin embargo, en todas las clasificaciones, el grupo con mayor cantidad de niñas contiene las trayectorias en los niveles más bajos del IMC para la edad, respecto a los restantes grupos. A su vez, para la clasificación en tres, cuatro y cinco grupos se observa que los tres primeros grupos (“G 1”, “G 2” y “G 3”) son similares en términos del porcentaje de niñas que lo integran y los niveles de IMC que representan. En este sentido, el “G 1” tiene los niveles más bajos del IMC respecto al resto de grupos. El “G 2” tiene niveles superiores de IMC que el “G 1”. Por último, el “G 3” con niveles altos del IMC o bien trayectorias que decrecen

¹ Los umbrales se obtienen del Programa *Growth reference 5-19 years* disponibles en la Web de la OMS. La interpretación de estos valores es que un IMC $>1SD$ (equivalente a un IMC de 25 kg/m^2 a los 19 años) corresponde a problemas de sobrepeso y si es $>2SD$ (equivalente a un IMC de 30 kg/m^2 a los 19 años) el problema es obesidad. Por su parte, un IMC $<-2SD$ conlleva problemas de desnutrición (*thinnes*) en tanto si el índice es $<-3SD$ el problema de desnutrición es severo.

Figura 4: Trayectoria nutricional según dos y tres grupos de niñas con LCMM y enlace lineal



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

Fuente: Elaboración propia

en el rango de edad considerado y con poca proporción de niñas en él.

La clasificación en cuatro o cinco grupos no parece aportar en la separación de las trayectorias nutricionales de las niñas ya que los grupos “G 4” y “G 5” representan una proporción muy baja de niñas (apéndice 6.1)¹.

Por su parte, los resultados utilizando como función de enlace *spline* cuadrática no arrojan diferencias sustantivas con los obtenidos utilizando la función lineal (apéndice 6.3). Sin embargo, con la *spline* el ajuste mejora ya que el BIC es menor al obtenido con la función lineal dado el mismo número de grupos.

En la especificación con *splines* se debe tener en cuenta que el modelo con cuatro grupos no converge², por tanto según el criterio de BIC tres grupos sería adecuado. Si bien luego con cinco vuelve a disminuir la diferencia entre los respectivos valores del BIC es muy pequeña. La entropía para la clasificación de niñas en tres grupos con *splines* es 0.582, levemente mayor a la obtenida con función lineal (0.513) pero continúa siendo un valor medio respecto a las recomendaciones prácticas.

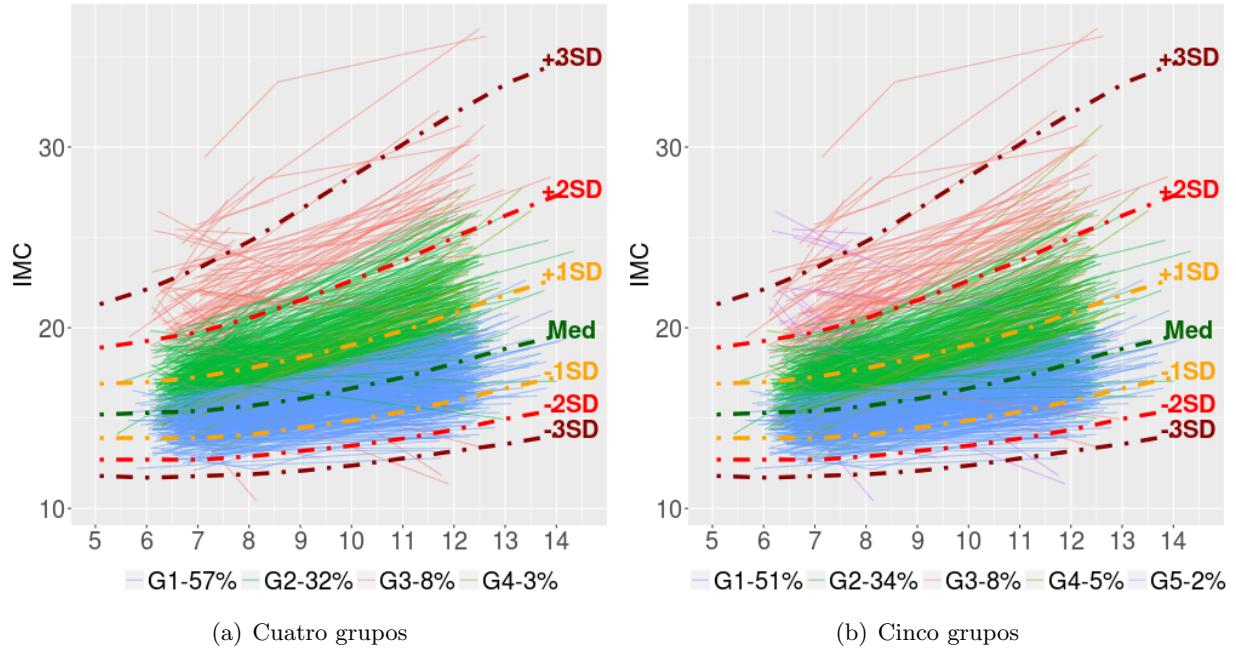
Incorporar esta función *splines* hace más flexible el modelo, ya que permite separar el estructural del correspondiente a la medida longitudinal observada, pero el aumento de los parámetros compromete la estimación, su interpretación y los resultados no aportan diferencias sustanciales respecto a la especificación con una función lineal.

En conclusión, del análisis previo se desprende que la clasificación de las niñas según sus trayectorias nutricionales en tres grupos resulta adecuada y el modelo especificado con una función de enlace lineal logra un ajuste adecuado.

¹ En el apéndice 6.1 se presentan los gráficos que separan las trayectorias según grupos para observar más claramente las diferencias.

² La no convergencia puede deberse a que la función *splines* si bien logra mayor ajuste del modelo aumenta su complejidad ya que se incrementan los parámetros a estimar o a la no identificación del modelo, en el sentido de que no habrían cuatro grupos claramente definidos para niñas

Figura 5: Trayectoria nutricional según cuatro y cinco grupos de niñas con LCMM y enlace lineal



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

Fuente: Elaboración propia

Los resultados con función de enlace lineal para niños (Cuadro 8) muestran que, en términos de ajuste resulta adecuado clasificarlos en tres grupos, ya que aumentar su número disminuye el BIC hasta tres, para cuatro permanece constante y baja nuevamente para cinco, aunque esta disminución es muy pequeña. Por su parte, la entropía para tres grupos es 0.556 mostrando que la separación es medianamente nítida. La forma en que quedan clasificadas las trayectorias de los niños en tres grupos se observa en la Figura 6 panel *b*.

Por otro lado, el Cuadro 8 muestra el porcentaje que queda clasificado *a posteriori* en cada grupo: para dos uno contiene casi el 80 % de los niños, para tres grupos uno de ellos tiene el 70 % mientras que los otros 21 y 8 %. Luego, con cuatro se tienen dos grandes grupos (40 y 34 %) y otros dos más pequeños (17 y 8 %). Por último, la estructura con cinco grupos es similar a la de cuatro ya que el quinto contiene menos del 1 % de los niños.

Cuadro 8: Estadísticos y clasificación por grupos de niños con LCMM y enlace lineal

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niños en cada grupo | | | | |
|--------|---------------------|---------|---------|-----------------|-----------------------|------|------|-----|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -5915.3 | 11881.7 | | 7 | 100 % | | | | |
| 2 | -5622.6 | 11332.8 | 0.535 | 12 | 79 % | 21 % | | | |
| 3 | -5530.2 | 11184.7 | 0.556 | 17 | 71 % | 21 % | 8 % | | |
| 4 | -5511.9 | 11184.7 | 0.456 | 22 | 41 % | 34 % | 17 % | 8 % | |
| 5 | -5446.4 | 11090.1 | 0.516 | 27 | 43 % | 28 % | 20 % | 8 % | 1 % |
| N=1492 | | | | | | | | | |

^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

Fuente: Elaboración propia

Para indagar en la calidad de clasificación que se tiene con tres grupos, se estudian las probabilidades *a posteriori*. El Cuadro 9 muestra la probabilidad media de pertenencia a cada grupo (panel *a*) y el porcentaje de niños cuya probabilidad máxima supera cierto umbral (panel

Cuadro 9: Probabilidades *a posteriori* para tres grupos de niños con LCMM y enlace lineal

| | (a) Prob. media | | | (b) % >umbral ^a | | |
|-----|-----------------|-------|-------|----------------------------|-------|-------|
| | G 1 | G 2 | G 3 | > 0.7 | > 0.8 | > 0.9 |
| G 1 | 0.823 | 0.022 | 0.154 | 85 % | 66 % | 33 % |
| G 2 | 0.018 | 0.826 | 0.155 | 69 % | 60 % | 47 % |
| G 3 | 0.146 | 0.112 | 0.742 | 56 % | 40 % | 24 % |

N=1492

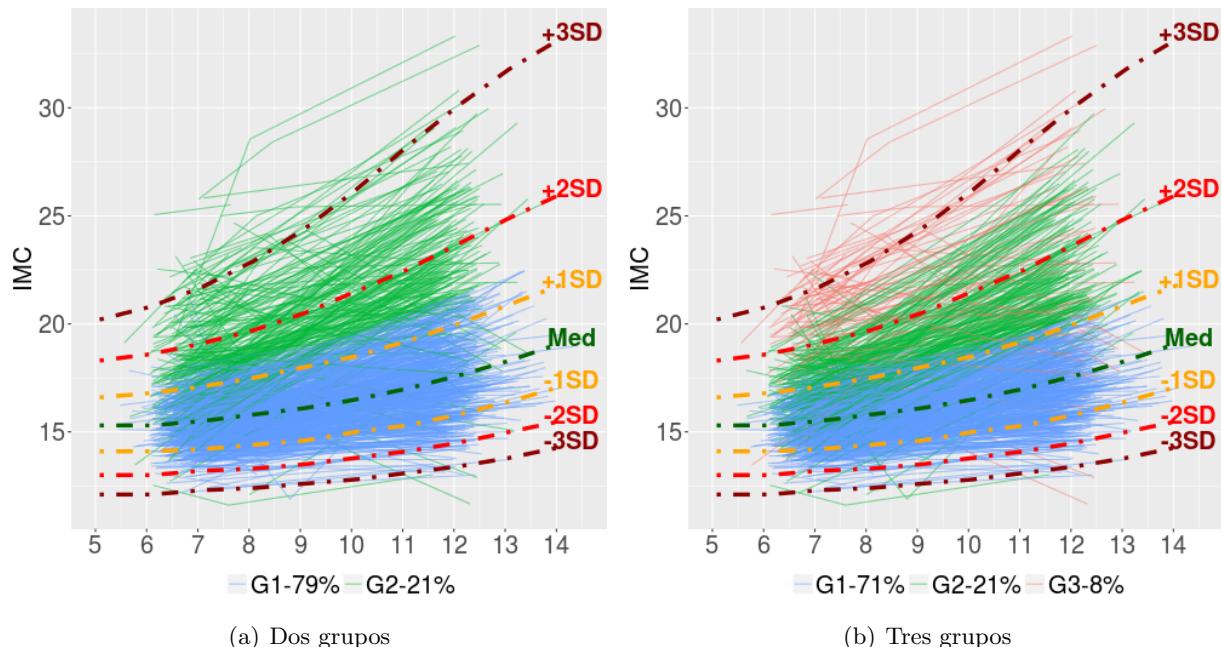
^a % de niños con probabilidad *a posteriori* que supera el umbral
Fuente: Elaboración propia

b). La probabilidad de ser clasificados en su propio grupo supera en promedio 0.74 para todos los grupos, por tanto, la diagonal de el Cuadro de clasificación contiene valores cercanos a uno y sustenta una buena discriminación de los niños en esos grupos latentes. Por su parte, se observa que el “G 1” y “G 2” tienen aproximadamente un 60 % de los niños con probabilidad máxima mayor a 0.8 en tanto que el “G 3” alcanza un 56 % de los niños con probabilidad máxima superior a 0.7. Por lo tanto, al considerar las probabilidades *a posteriori*, también es adecuada la clasificación en tres grupos de niños, con un índice de entropía en el entorno de 0.6.

A su vez, en las Figuras 6 y 7 se puede observar la trayectoria nutricional para cada niño según el grupo en que quedó clasificado, como ya fue mencionado, se aprecia cierta superposición de las trayectorias entre distintos grupos. A diferencia de lo que ocurría para las niñas, ahora los grupos mayoritarios de niños no siempre tienen menores niveles de IMC respecto a los restantes grupos. Esto último se cumple para la clasificación de niños en dos y tres grupos pero no para la clasificación en cuatro y cinco.

Cuando se clasifica en cuatro grupos las trayectorias con nivel de IMC más bajo son las del “G 2” (34 %), le supera en nivel el “G 1” (41 %) y luego el “G 3” (17 %) y “G 4” (8 %).

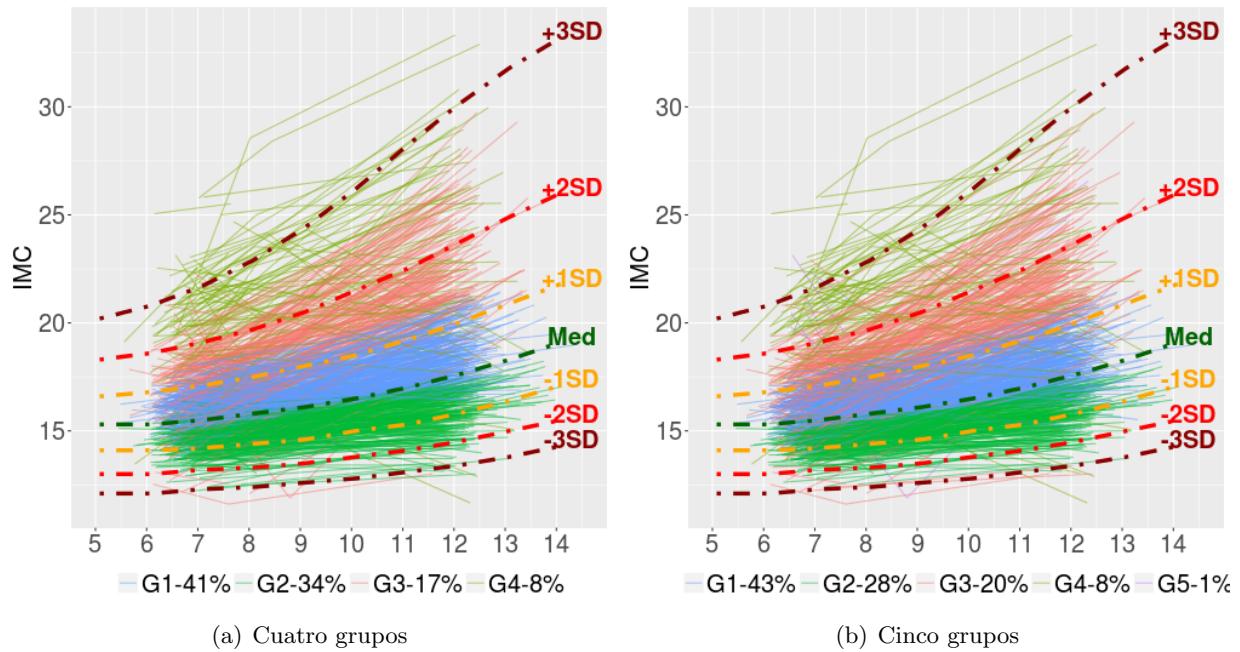
Figura 6: Trayectoria nutricional según dos y tres grupos de niños con LCMM y enlace lineal



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

Fuente: Elaboración propia

Figura 7: Trayectoria nutricional según cuatro y cinco grupos de niños con LCMM y enlace lineal



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

Fuente: Elaboración propia

Por su parte, en la clasificación en cinco grupos los cuatro primeros tienen trayectorias similares a los de la clasificación en cuatro grupos, ya que el quinto contiene menos del 1% de los niños con trayectorias “raras” (apéndice 6.2).

Por otro lado, los resultados utilizando una función de enlace *splines* cuadrática (apéndice 6.4) son similares respecto a los obtenidos utilizando la función lineal y, para el mismo número de grupos, el BIC es menor con *splines*.

Cuando se introduce una función *splines*, el modelo con cinco grupos no converge¹ y el BIC se minimiza en la especificación con cuatro grupos. La distribución de niños *a posteriori* con esta clasificación es similar a la obtenida con tres grupos, ya que el cuarto grupo contiene solamente un 1% de los niños. El valor de la entropía para cuatro grupos es 0.504, situándose en valores medios respecto a las recomendaciones prácticas, mayor al valor para cuatro grupos con función lineal (0.456), pero menor que la obtenida con tres grupos (0.541 con *splines* y 0.556 con lineal).

Al comparar las trayectorias de los niños en función de los grupos obtenidos con enlace lineal o *splines* (apéndice 6.2), la clasificación en cinco grupos separa en el grupo adicional, un conjunto de trayectorias “raras” (1%). La clasificación en cuatro grupos con *splines*, muestra uno con trayectorias en forma de “V”, al igual que las clasificaciones en cinco grupos con función lineal o *splines*. Al comparar las trayectorias de la clasificación en cuatro grupos con función lineal y con *splines* se observa que en esta última se separa un cuarto grupo con trayectorias en forma de “V” y los grupos “G 1” y “G 2” del caso lineal parecen fusionarse en el “G 1” del caso con *splines* (apéndice 6.2).

Como ya fue mencionado, incorporar la función *splines* hace más flexible el modelo pero compromete la estimación y su interpretación. Los resultados, presentan leves diferencias en el agrupamiento que resulta de aplicar la función lineal o la *splines*, pero en las clasificaciones con cuatro y cinco grupos. En particular, el grupo de niños que presenta trayectorias nutricionales

¹ Las razones de la no convergencia son similares a las expuestas para las niñas. Una razón es el aumento del número de parámetros y además un problema de identificación del modelo, ya que puede exigirse identificar un número de grupos que podría ser no adecuado para clasificar a los niños.

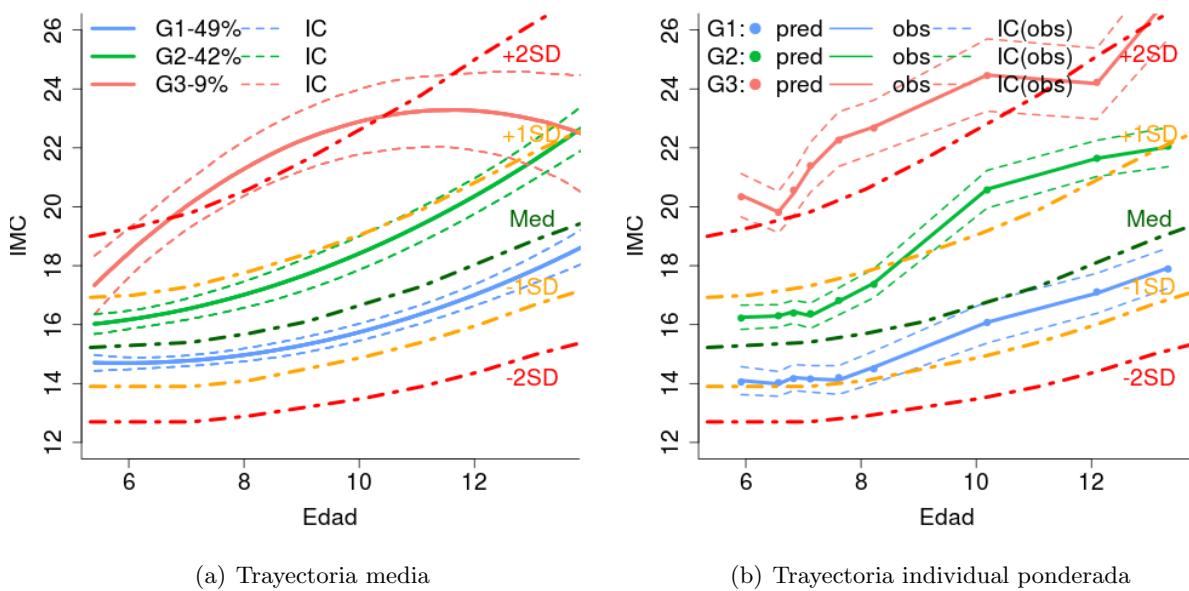
en forma de “V” no se identifica de igual forma con ambas funciones. Sin embargo, se considera un ajuste adecuado con tres grupos de niños según medidas de bondad de ajuste y, además, las probabilidades *a posteriori* arrojan una estructura adecuada para tres grupos, tanto para el caso lineal como *splines* (apéndice 6.3).

En definitiva, para el análisis de la trayectoria nutricional de los niños se opta por una clasificación en tres grupos que resulta adecuada, en términos de ajuste y probabilidades *a posteriori*, y se presentan los resultados del modelo especificado con función de enlace lineal.

4.2.2. Trayectoria nutricional según grupos

En esta sección se muestran características de las niñas y niños según el grupo en el que fueron clasificados. El objetivo es caracterizar los grupos anteriormente identificados, según factores que se asocian con el estado nutricional. Para ello, se presentan las trayectorias predichas para cada grupo de niñas y niños. A su vez, se muestran descriptivas sobre el peso y la talla al nacer y educación materna de las niñas y niños según el grupo al que pertenecen.

Figura 8: Trayectoria nutricional media e individual según grupos de niñas con LCMM y enlace lineal



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Fuente: Elaboración propia

En la Figura 8 se presentan las trayectorias medias (panel *a*) para cada grupo de niñas y específicas de los individuos ponderadas con sus intervalos de confianza (panel *b*) y los umbrales de la OMS. Se observa que el “G 1” tiene una trayectoria creciente por debajo de la mediana. El “G 2” presenta una trayectoria en niveles superiores, que sobrepasa en algún tramo el umbral de normalidad ($+1\text{SD}$) y que crece con tendencia similar a la del “G 1”. El “G 3” tiene trayectoria en forma de “U” invertida y por encima del límite de sobrepeso ($+1\text{SD}$) e incluso obesidad ($+2\text{SD}$).

Por tanto, se identifican dos grupos que representan el 90% de las niñas que en general transitan dentro de los umbrales de normalidad: el grupo normal-bajo, con casi 50% de las niñas, tiene una trayectoria por debajo de la mediana; el normal-alto, con 40%, transita por encima de la mediana y, en cierto rango de edad, supera el umbral de sobrepeso ($+1\text{SD}$). El último grupo con sobrepeso, representa menos del 10% de niñas, sigue una trayectoria en forma

Cuadro 10: Descriptivas para tres grupos de niñas con LCMM y enlace lineal

| | Normal-bajo | Normal-alto | Sobrepeso | Total |
|---------------------|-------------|-------------|------------|------------|
| % niñas | 49 % | 42 % | 9 % | |
| N | 679 | 588 | 125 | 1392 |
| Educación materna | | | | |
| Primaria | 48 % | 46 % | 44 % | 47 % |
| Secundaria | 45 % | 47 % | 49 % | 46 % |
| Terciaria | 6 % | 7 % | 8 % | 7 % |
| N | 569 | 558 | 84 | 1211 |
| p=0.99 | | | | |
| Peso al nacer (g) | | | | |
| Media (SD) | 3111 (570) | 3246 (557) | 3329 (542) | 3188 (567) |
| N | 635 | 556 | 115 | 1306 |
| p=0.00*** a | | | | |
| Talla al nacer (cm) | | | | |
| Media (SD) | 48.4 (4.3) | 48.9 (3.4) | 48.9 (2.7) | 48.7 (3.8) |
| N | 452 | 420 | 94 | 966 |
| p=0.08* b | | | | |

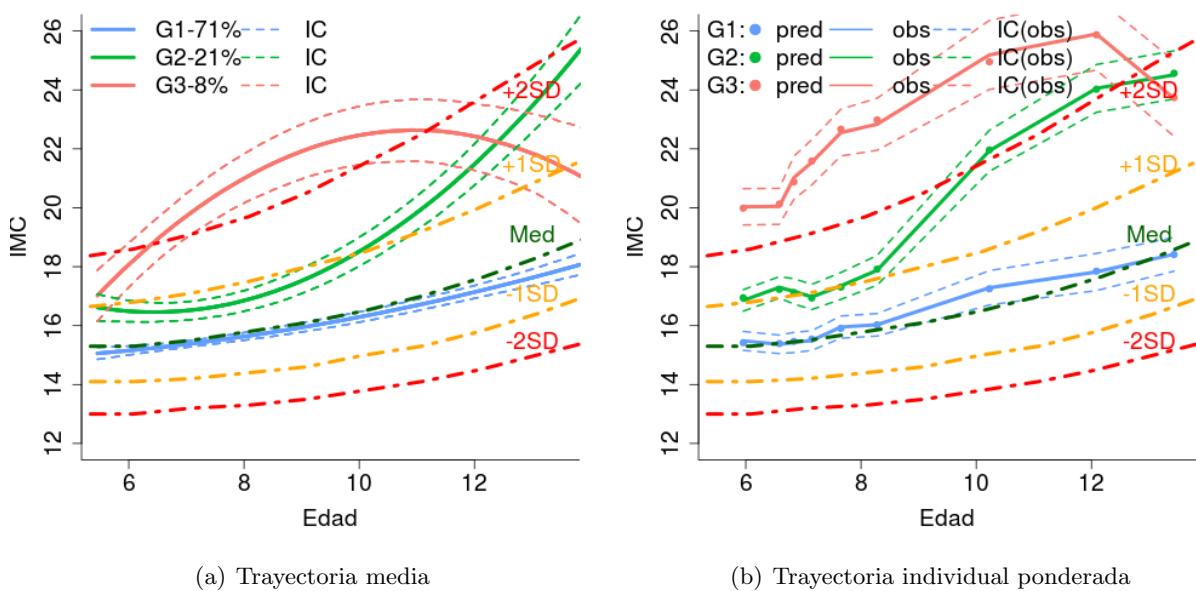
^a No se rechaza igual varianza entre grupos ^b Se rechaza igual varianza entre grupos

* p<0.10 ** p<0.05 *** p<0.01

Fuente: Elaboración propia

de “U” invertida que se sitúa por encima del umbral de sobrepeso (+1SD) e incluso obesidad (+2SD).

Figura 9: Trayectoria nutricional media e individual según grupos de niños con LCMM y enlace lineal



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Fuente: Elaboración propia

Además, se indaga en las posibles diferencias entre los grupos según nivel de educación

materna, el peso y la talla al nacer de las niñas. Respecto a la educación materna, no se observan diferencias significativas entre los tres grupos de niñas (Cuadro 10). En cuanto al peso y la talla al nacer, sí se aprecian diferencias significativas entre los grupos. Por tanto, el peso al nacer de las niñas del grupo normal-bajo en promedio menor a las del grupo normal-alto y éste, a su vez, menor en promedio a las del grupo de sobrepeso. En tanto que, la talla al nacer es en promedio más alta para el grupo de sobrepeso, le sigue el grupo de normal-alto y el normal-bajo (Cuadro 10).

En cuanto al análisis de las características para los niños la Figura 9 muestra las trayectorias medias (panel *a*) para cada grupo de niños y específicas de los individuos ponderadas con sus intervalos de confianza (panel *b*) y los umbrales de la OMS. El “G 1” tiene una trayectoria creciente que transita entorno a la mediana, el “G 2” con trayectoria creciente pero pendiente más inclinada que el “G 1” transita en los niveles normales y por encima del umbral de sobrepeso (+1SD) y, el “G 3” tiene trayectoria en forma de “U” invertida por encima del nivel de obesidad (+2SD).

Cuadro 11: Descriptivas para tres grupos de niños con LCMM y enlace lineal

| | Normal | Normal-Sobrepeso | Obesidad | Total |
|---|------------|------------------|------------|------------|
| % niños | 71 % | 21 % | 8 % | |
| N | 1064 | 310 | 118 | 1492 |
| Educación materna | | | | |
| Primaria | 52 % | 44 % | 39 % | 49 % |
| Secundaria | 44 % | 47 % | 51 % | 45 % |
| Terciaria | 4 % | 9 % | 10 % | 6 % |
| N | 938 | 270 | 104 | 1312 |
| $\chi^2=16.1$ df=9 p=0.06* | | | | |
| Peso al nacer (g) | | | | |
| Media (SD) | 3306 (574) | 3372 (645) | 3427 (578) | 3330 (590) |
| N | 999 | 295 | 113 | 1407 |
| ANOVA ^a F=3.1 df=(2,1404) p=0.04** | | | | |
| Talla al nacer (cm) | | | | |
| Media (SD) | 49.4 (4.0) | 49.9 (3.3) | 49.9 (3.5) | 49.6 (3.9) |
| N | 713 | 241 | 90 | 1044 |
| ANOVA ^b F=2.1 df=(2,1041) p=0.12 | | | | |

* p<0.10 ** p<0.05 *** p<0.01

^a Contraste igualdad de varianzas (Bartlett's), rechaza que las varianzas son iguales ($K^2=6.5$ df=2 p=0.04**)

^b Contraste igualdad de varianzas (Bartlett's), rechaza que las varianzas son iguales ($K^2=14.8$ df=2 p=0.00***)

Fuente: Elaboración propia

Por tanto, se identifica un grupo normal, que representa más del 70 % de los niños que transita dentro de los umbrales de normalidad con trayectoria creciente entorno a la mediana. Un grupo normal-sobrepeso con el 20 % de los niños, con crecimiento más pronunciado y por encima del umbral de sobrepeso para algún tramo de edad. Un pequeño grupo con obesidad que contiene el 8 % de los niños.

En cuanto a las características de estos grupos de niños, existen diferencias en los niveles de educación materna (significativas al 10%). En este sentido, los niños del grupo normal tienen más proporción de madres con nivel educativo hasta primaria que los niños del grupo normal-sobrepeso y con obesidad. Mientras tanto, el grupo con obesidad tiene mayor proporción de niños con madres de nivel terciario que los otros dos grupos (Cuadro 11).

Por otro lado, en cuanto al peso y la talla al nacer de los niños existe una diferencia significativa en el peso medio al nacer entre los grupos aunque la talla no presenta diferencias significativas. Por tanto, el peso al nacer de los niños del grupo normal es en promedio menor al de los grupos normal-sobrepeso y obesidad, mientras que la talla al nacer promedio entre los grupos no difiere de forma significativa (Cuadro 11).

4.3. Vínculo entre trayectoria nutricional y desempeño escolar

En esta sección se presentan los resultados del estudio de la vinculación entre los grupos según trayectoria nutricional y la edad de abandono escolar. Para ello se estiman modelos conjuntos de clases latentes especificando de uno a cinco el número de grupos, igual que en la sección 4.2.

Se realiza un análisis análogo para elegir el número adecuado de grupos. Es decir, se estiman los modelos fijando el número de grupos de uno a cinco y se comparan indicadores de bondad de ajuste, las probabilidades *a posteriori* y el porcentaje de casos en cada grupo. También se sigue un criterio de parsimonia para la elección del número de grupos una vez estimados los modelos conjuntos para niños y niñas. Se presentan resultados de los JLCMM que permiten analizar la relación entre desempeño escolar y trayectoria nutricional.

Por último, se destaca que los resultados de las estimaciones para los modelos JLCMM de esta sección, se obtienen luego de realizar 30 réplicas con distintos puntos iniciales, para disminuir la probabilidad de convergencia a óptimos locales, según se expuso en la sección 3.5.2.

4.3.1. Trayectoria nutricional y desempeño escolar según grupos

A continuación, se presentan los resultados de los JLCMM estimados para niñas. En primer lugar, se presentan los estadísticos y clasificación *a posteriori* en grupos según trayectorias nutricionales que resulta de los modelos conjuntos. En la clasificación de grupos, los resultados no difieren sustancialmente con los obtenidos en los LCMM¹. Como se observa en el Cuadro 12 el BIC se minimiza con cinco grupos, sin embargo para cuatro este estadístico es muy similar. Además, la entropía es mayor para cuatro y tres grupos que para cinco.

Cuadro 12: Estadísticos y clasificación según grupos de niñas con JLCMM

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niñas en cada grupo | | | | |
|--------|---------------------|---------|---------|-----------------|-----------------------|------|------|------|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -4971.1 | 10011.0 | | 10 | 100 % | | | | |
| 2 | -4791.6 | 9686.5 | 0.691 | 15 | 88 % | 12 % | | | |
| 3 | -4747.3 | 9632.2 | 0.542 | 20 | 46 % | 45 % | 9 % | | |
| 4 | -4715.2 | 9602.4 | 0.610 | 25 | 49 % | 40 % | 10 % | 1 % | |
| 5 | -4697.5 | 9601.3 | 0.537 | 30 | 46 % | 31 % | 12 % | 10 % | 1 % |
| N=971 | | | | | | | | | |

^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

Fuente: Elaboración propia

Por último, respecto a la proporción de niñas en cada grupo, las clasificaciones en mayor número de grupos parecen separar trayectorias “raras”, al igual que en el caso de los LCMM, pero no cambian sustancialmente las proporciones contenidas en los tres grupos más grandes. Por tanto, en el modelo conjunto, para el análisis del desempeño escolar y su relación con la

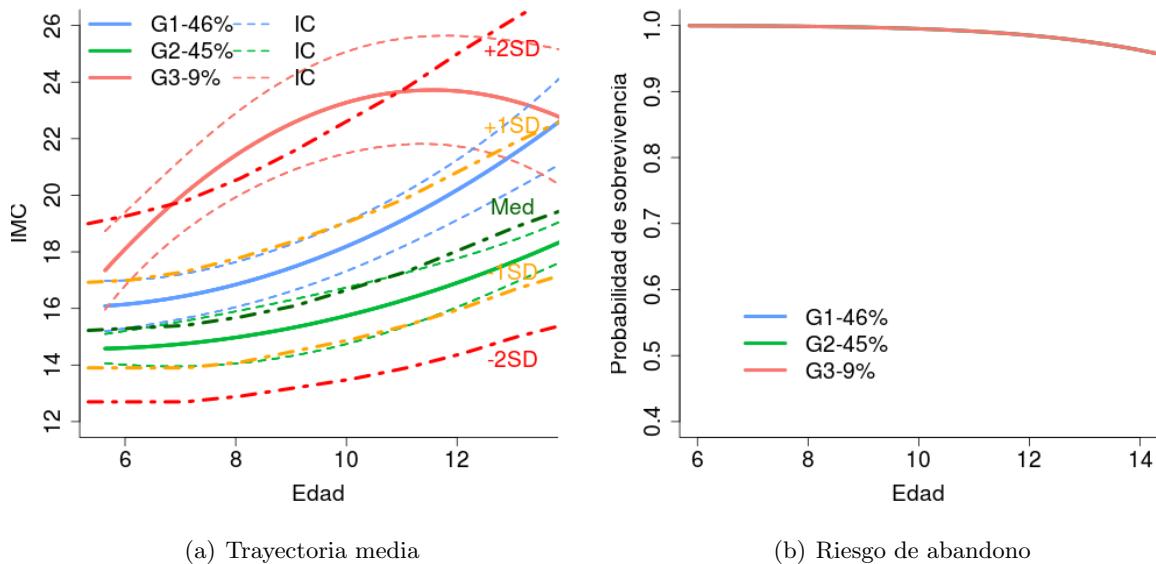
¹ Los resultados de los LCMM para la muestra utilizada en los modelos conjuntos se explican en el trabajo de tesis que dio origen a este documento.

trayectoria nutricional, es adecuado agrupar a las niñas en tres grupos según sus trayectorias nutricionales.

Las trayectorias predichas en cada grupo (Figura 10), según el JLCMM, muestran que el “G 1” (46 % de las niñas) está por encima de la trayectoria mediana y umbral de sobrepeso para cierto rango (Normal-alto) y el “G 2” sigue una trayectoria por debajo de la media con 45 % de las niñas (Normal-bajo). El grupo más pequeño, 9 % de las niñas presenta una trayectoria de sobrepeso e incluso obesidad.

Por último, se aprecia que la probabilidad de continuar en el sistema educativo, según la edad, para los distintos grupos se reduce con la edad y este comportamiento no difiere entre grupos (Figura 10). Por tanto, las diferentes trayectorias nutricionales de las niñas, capturada por la pertenencia a los grupos latentes, no resultan asociadas con el riesgo de abandono escolar para la edad de las niñas.

Figura 10: Trayectoria nutricional media y riesgo de abandono según grupos de niñas



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Fuente: Elaboración propia

A continuación se presentan los resultados de los modelos conjuntos JLCMM estimados para niños. Se presentan los estadísticos y clasificación *a posteriori* en los grupos según trayectorias nutricionales que resulta de los modelos conjuntos.

Cuadro 13: Estadísticos y clasificación según grupos de niños - JLCMM

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niños en cada grupo | | | | |
|----------------|---------------------|---------|---------|-----------------|-----------------------|------|------|-----|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -5371.0 | 10811.3 | | 10 | 100 % | | | | |
| 2 | -5168.0 | 10439.9 | 0.514 | 15 | 72 % | 28 % | | | |
| 3 | -5085.3 | 10309.2 | 0.593 | 20 | 67 % | 26 % | 7 % | | |
| 4 | -5019.8 | 10213.0 | 0.671 | 25 | 66 % | 26 % | 7 % | 1 % | |
| 5 ^c | -5003.9 | 10215.9 | 0.656 | 30 | 40 % | 26 % | 25 % | 8 % | 1 % |
| N=1026 | | | | | | | | | |

^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

^c No converge

Fuente: Elaboración propia

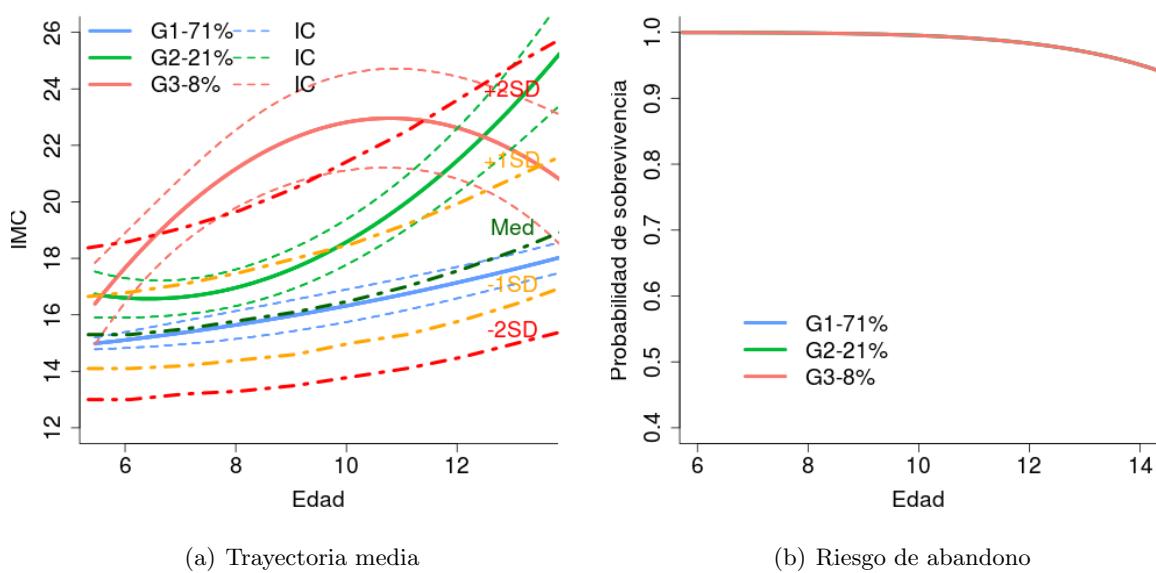
En la clasificación de grupos por JLCMM, los resultados no muestran diferencias significativas con los LCMM¹, al igual que ocurría para las niñas. El Cuadro 13 muestra que el BIC se minimiza con cinco grupos, pero este modelo no converge, por tanto no es considerado. Entre las demás estructuras el mínimo BIC y el mayor índice de entropía se alcanza con cuatro grupos.

Por último, respecto a la proporción de niños en cada grupo, las clasificaciones en mayor número de grupos parecen separar trayectorias “raras”, al igual que en el caso de los LCMM, pero no cambian sustancialmente las proporciones contenidas en los grupos más grandes.

Por tanto, en el modelo conjunto y para el análisis del desempeño escolar y su relación con la trayectoria nutricional, es adecuado agrupar a los niños en tres grupos según sus trayectorias nutricionales.

Las trayectorias predichas en cada grupo (Figura 11), según el JLCMM, muestran que el 67 % de los niños sigue una trayectoria entorno a la mediana (grupo normal) y un 26 % está en el entorno de +1SD (grupo normal-sobrepeso). El grupo más pequeño con 7% de los niños presenta una trayectoria por encima de +2SD (grupo con obesidad), si bien para edades finales el IMC disminuye².

Figura 11: Trayectoria nutricional media y riesgo de abandono según grupos de niños



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

Fuente: Elaboración propia

Por último, la probabilidad de continuar en el sistema educativo, según la edad, para los distintos grupos (Figura 11) se reduce con la edad, al igual que para las niñas, y este comportamiento no difiere significativamente entre grupos. Sin embargo, a diferencia de las niñas, la probabilidad de continuar en el sistema educativo decrece a edades más tempranas.

¹ Los resultados de los LCMM para la muestra utilizada en los modelos conjuntos fue omitida en este documento si bien se encuentra explicada en el trabajo de tesis que dio origen a este documento.

² Esto puede ser por el crecimiento en estatura muy alto que tienen los niños a esta edad y que puede mostrar disminución del IMC

4.3.2. Análisis de sensibilidad

Para el análisis de sensibilidad de los resultados se optó por estimar el modelo conjunto sin incluir la variable *Repetición*. Si bien los antecedentes y el análisis descriptivo de los datos muestran que esta variable es relevante para discriminar entre las tasas de abandono la forma en que fue relevada en la base de datos utilizada hace que la muestra se vea restringida para la estimación de estos modelos. Por tanto, se estiman modelos conjuntos sin incorporar esta covariable en el modelo de sobrevivencia.

Trayectoria nutricional y desempeño escolar sin repetición

Los resultados de la estimación de los JLCMM para niñas no muestran diferencias importantes con las obtenidas incluyendo la repetición previa como covariable en el modelo de sobrevida. Los estadísticos para los modelos y la clasificación *a posteriori* en grupos se presentan en el Cuadro 14. Se observa que el BIC se minimiza con cinco grupos, sin embargo para cuatro este estadístico es muy similar. Además, la entropía es mayor para cuatro y tres grupos que para cinco.

Cuadro 14: Estadísticos y clasificación según grupos de niñas - JLCMM - Sin Repetición

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niñas en cada grupo | | | | |
|--------|---------------------|---------|---------|-----------------|-----------------------|-----|-----|-----|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -6250.2 | 12565.5 | | 9 | 100 | | | | |
| 2 | -5999.5 | 12100.3 | 0.644 | 14 | 86 | 14 | | | |
| 3 | -5947.4 | 12032.4 | 0.506 | 19 | 48 | 43 | 9 | | |
| 4 | -5902.5 | 11978.7 | 0.552 | 24 | 50 | 38 | 10 | 2 | |
| 5 | -5882.6 | 11975.1 | 0.514 | 29 | 40 | 37 | 18 | 3 | 2 |
| N=1392 | | | | | | | | | |

Fuente: Elaboración propia

^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

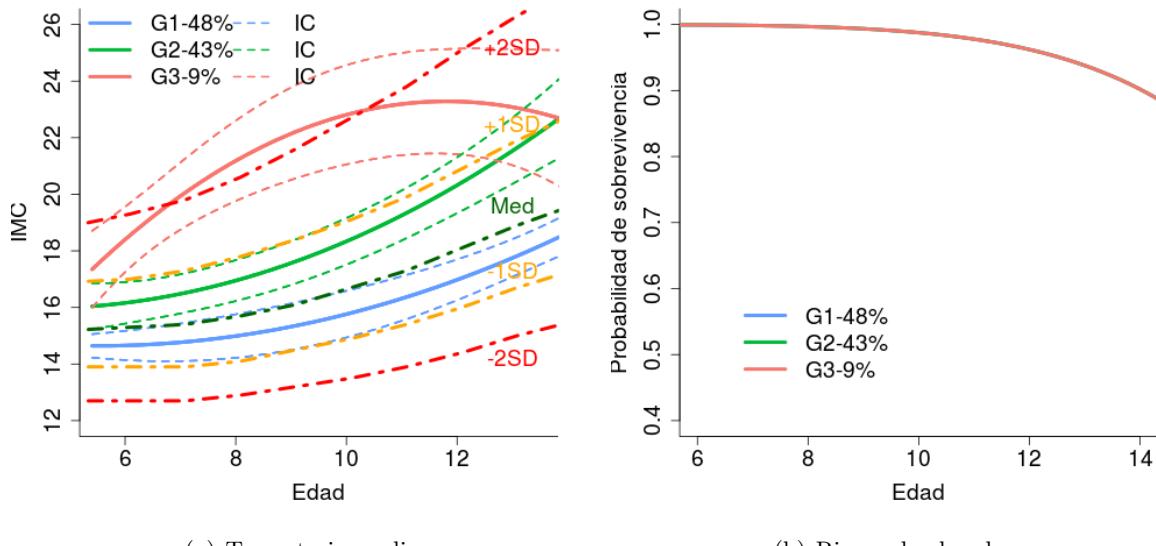
Por último, respecto a la proporción de niñas en cada grupo, las clasificaciones en mayor número de grupos parecen separar trayectorias “raras” pero no cambian sustancialmente las proporciones contenidas en los tres grupos más grandes. Por tanto, estos resultados son consistentes con una clasificación de las niñas en tres grupos para analizar el desempeño escolar y su relación con la trayectoria nutricional.

Las trayectorias predichas en cada grupo (Figura 12), según el JLCMM, muestran que el “G 1” (48 % de las niñas) está por encima de la trayectoria mediana y umbral de sobrepeso para cierto rango (Normal-alto) y el “G 2” sigue una trayectoria por debajo de la media con 43 % de las niñas (Normal-bajo). El grupo más pequeño, 9 % de las niñas presenta una trayectoria de sobrepeso e incluso obesidad.

Por último, la probabilidad de continuar en el sistema educativo para distintas edades y según grupos (Figura 12), se reduce con la edad y el comportamiento no difiere entre grupos, como ocurría en el modelo que incluye la repetición previa como covariable para el modelo de sobrevida.

Los resultados de los modelos conjuntos JLCMM estimados para niños tampoco difieren respecto a la especificación que incluye la covariable *Repetición*. Los estadísticos y clasificación *a posteriori* en grupos se presentan en el Cuadro 15. Se observa que el BIC se minimiza con cuatro aunque la diferencia con el BIC de tres grupos es pequeña. A su vez, la entropía para cuatro grupos es 0.634.

Figura 12: Trayectoria nutricional media y riesgo de abandono según grupos de niñas



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Fuente: Elaboración propia

En este caso, respecto a la proporción de niños en cada grupo, también se observa que las clasificaciones con mayor número de grupos separan trayectorias “raras” pero no cambian sustancialmente las proporciones contenidas en los grupos más grandes. Además, la distribución de niños entre los grupos es similar a la que obtenida sin incluir *Repetición*.

Cuadro 15: Estadísticos y clasificación según grupos de niños - JLCMM - Sin Repetición

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niños en cada grupo | | | | |
|--------|---------------------|---------|---------|-----------------|-----------------------|-----|-----|-----|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -6601.3 | 13268.3 | | 9 | 100 | | | | |
| 2 | -6309.2 | 12720.7 | 0.535 | 14 | 79 | 21 | | | |
| 3 | -6213.8 | 12566.6 | 0.553 | 19 | 71 | 21 | 8 | | |
| 4 | -6144.8 | 12465.0 | 0.634 | 24 | 70 | 22 | 8 | 1 | |
| 5 | -6130.2 | 12472.2 | 0.673 | 29 | 43 | 27 | 21 | 8 | 1 |
| N=1492 | | | | | | | | | |

Fuente: Elaboración propia

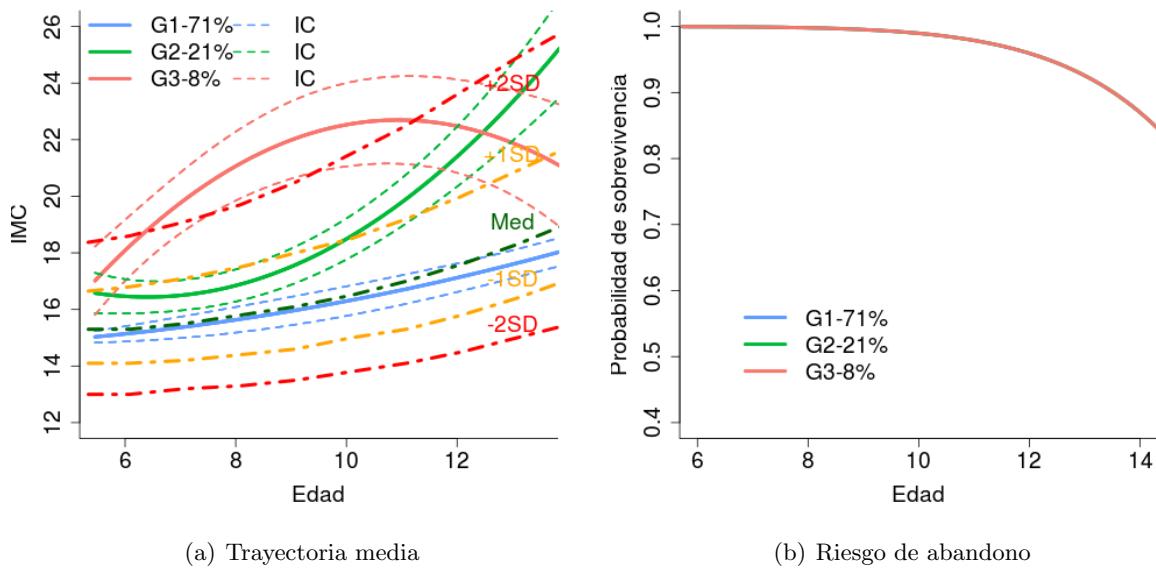
^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

Las trayectorias predichas (Figura 13) según la clasificación en tres grupos, muestran que el 71 % de los niños sigue una trayectoria entorno a la mediana (grupo normal) y un 21 % está por encima de $+1\text{SD}$ (grupo con sobrepeso). El grupo más pequeño con 8 % de los niños presenta una trayectoria por encima de $+2\text{SD}$ (grupo con obesidad), si bien para edades finales el IMC cae.

Para finalizar, la Figura 13 muestra la probabilidad de no abandonar el sistema educativo para distintas edades y según los tres grupos de trayectorias nutricionales de niños. En este caso, también se observa que dicha probabilidad disminuye con la edad pero no se ve modificada por la pertenencia a los grupos nutricionales.

Figura 13: Trayectoria nutricional media y riesgo de abandono según grupos de niños



Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Fuente: Elaboración propia

5. Conclusiones

La realización de este trabajo permitió aplicar una técnica estadística moderna a un problema de relevancia en Uruguay y ha dado lugar a nuevas interrogantes y desafíos respecto a los temas vinculados con la trayectoria nutricional y el desempeño escolar de los niños y niñas. La aplicación de esta metodología constituye un avance ya que los trabajos existentes en Uruguay utilizan técnicas para datos de corte transversal para estudiar este tema. Los LCMM y JLCMM utilizados en el presente trabajo han permitido analizar la trayectoria nutricional con una perspectiva temporal, es decir, estudiar el cambio con la edad del estado nutricional y no únicamente el estado en una edad determinada.

Del análisis de la trayectoria nutricional, se desprende la existencia de grupos de niños y niñas en edad escolar, entre 6 y 12 años aproximadamente, con una evolución diferente del IMC. La diferencia en la evolución del índice a lo largo del tiempo según grupos puede tener consecuencias sobre otros aspectos, como por ejemplo enfermedades que se han encontrado vinculadas a problemas nutricionales durante la infancia. Los resultados de los modelos estimados muestran que tanto niños como niñas pueden caracterizarse adecuadamente en términos de su trayectoria nutricional en tres grupos.

Estos tres grupos de niñas se caracterizan por:

- **Normal-bajo:** representa un 50 % de las niñas, con una trayectoria por debajo de la mediana
- **Normal-alto:** representa un 40 % de las niñas, transita por encima de la mediana y en algún rango de edad por encima del umbral de sobrepeso
- **Sobrepeso:** representa 10 %, con trayectoria en forma de “U” invertida en niveles de sobrepeso e incluso obesidad

Mientras que los tres grupos de niños se caracterizan por:

- **Normal:** representa un 70 % de los niños, con una trayectoria en el entorno de la mediana
- **Normal-Sobrepeso:** representa un 20 % de los niños, transita por encima de la mediana con tendencia creciente más inclinada y en algún rango de edad por encima del umbral de sobrepeso
- **Obesidad:** representa 10 %, con trayectoria en forma de “U” invertida en niveles de obesidad

Los resultados muestran diferencias de los grupos en otras características relevantes. Para las niñas las diferencias son significativas en el peso y la talla al nacer. En este sentido, las niñas con trayectoria nutricional normal-baja presentan peso y talla al nacer más bajos que los grupos normal-alto y con sobrepeso. Sin embargo, no son significativas las diferencias en los niveles de educación materno entre grupos. Por su parte, para los niños se encuentran diferencias significativas en el peso, pero no en la talla al nacer, y también en los niveles educativos maternos. El grupo normal se compone de niños con menor peso al nacer y cuyas madres tienen menor nivel educativo. Asimismo, para los niños con obesidad se encuentra que hay una mayor presencia de madres con nivel educativo más alto.

Los resultados de este trabajo para Uruguay no difieren sustancialmente de los reportados por Stuart & Panico (2016) para Reino Unido, si bien se trata de otro rango de edad y para un país desarrollado. Stuart & Panico (2016) identifican grupos de niños y niñas con sobrepeso y obesidad pero no con problemas de bajo peso o trayectorias del IMC decrecientes con la edad. Surge entonces como desafío el análisis de la trayectoria nutricional mediante la metodología propuesta en esta investigación para otros países en desarrollo. De esta forma, se pueden comparar los resultados obtenidos en el presente trabajo con estudios para países con similar nivel de desarrollo.

En cuanto al análisis de la relación entre trayectoria nutricional y desempeño escolar la evidencia muestra que los niños en Uruguay tienden a abondar el sistema educativo a edades más tempranas que las niñas. Además, la pertenencia a los grupos según trayectoria nutricional no modifica la probabilidad de continuar en la escuela de los niños y las niñas en el rango de edad analizado. Por tanto, la evidencia no arroja una vinculación entre la trayectoria nutricional y la edad de abandono escolar de los niños y niñas del ELBU.

Este resultado puede estar afectado por el rango de edad que queda representado en el ELBU. Este estudio longitudinal de niños y niñas es una cohorte que abarca el tramo de edades entre seis y trece o catorce años en 2011/12. Como las tasas de abandono son crecientes con la edad y es a partir de los trece o catorce años que tienen un aumento pronunciado, la sub representación de las edades donde el problema de abandono se agrava puede afectar los resultados.

Esta debilidad, da lugar a futuros trabajos que incorporen la cuarta ola del ELBU que contendrá el rango de edad completo en el sistema educativo formal, de seis a diecisiete años aproximadamente. Además, se pueden estudiar otros desempeños educativos mediante otras técnicas estadísticas, que permitan entender la vinculación entre la trayectoria nutricional y el desempeño escolar medido por la repetición, por ejemplo.

A su vez, ninguno de los grupos identificados tiene una trayectoria nutricional caracterizada por el déficit nutricional o que tenga un comportamiento decreciente con la edad¹, si bien existe variabilidad en las trayectorias de los niños al interior de los grupos. En este sentido, otros trabajos reportaron que los problemas nutricionales de déficit se asocian con un peor desempeño escolar, particularmente con problemas de aprendizaje que pueden finalmente resultar en el abandono.

¹ Salvo los grupos con trayectorias “raras”, ninguno de los grupos identificados con una cantidad relevante de niños o niñas tiene este tipo de trayectoria nutricional.

Una de las principales limitaciones de la investigación se encuentra en la elección del número de grupos mediante el análisis de medidas de bondad de ajuste y probabilidades *a posteriori* de los modelos estimados. Este procedimiento no cuenta con sustento formal que permita detectar de forma unívoca el número de grupos que mejor se ajusta a los datos. Sin embargo, en este trabajo se han estimado los modelos de varias formas¹ y no se han encontrado diferencias importantes en los resultados obtenidos. Además, la evidencia internacional en otros trabajos muestra resultados similares a los encontrados en esta investigación.

Un desafío que resultó de la aplicación de esta metodología fue el proceso de estimación de los modelos, que resultó muy demandante computacionalmente. El método de estimación MLE que aplica la librería de **R** utilizada optimiza mediante un algoritmo de Newton-Raphson extendido, conocido como Algoritmo de Marquardt. Este último, resulta exigente en los criterios de convergencia utilizados. En este sentido, algunos de los modelos presentados no lograron converger aunque se realizaron varios ajustes en las estimaciones para hacer más sencilla la parametrización. De todas maneras, no se logró en algunos casos la convergencia ya que los modelos tienen muchos parámetros, particularmente cuando el número de grupos es alto y se especifican variaciones entre grupos².

Otra discusión que surge de los resultados del análisis de las trayectorias nutricionales proviene de observar las gráficas del IMC para la edad según las distintas clasificaciones entre grupos y los umbrales de la OMS. En los gráficos, parece que podrían diferenciarse las trayectorias de niñas y niños por los cuantiles de la distribución del IMC ya que no se evidencian claras diferencias en las tasas de cambio del IMC para la edad según los grupos.

Este resultado puede deberse en parte por la propia estructura de datos. El ELBU contiene solamente tres momentos de medición longitudinal, por lo que se cuenta con una baja cantidad de medidas para el IMC por individuo. Esto da poco margen para la identificación de trayectorias con distintos comportamientos en la pendiente lineal o cuadrática, lo que puede resultar en la separación de los grupos por el nivel del índice más que por la tasa de cambio según la edad de los niños y niñas. En general, se observa que las diferencias en los grupos provienen por estar encima o no de la mediana y los umbrales de obesidad y sobrepeso de la OMS. Por su parte, en futuros trabajos se puede realizar un análisis más detallado de las trayectorias identificadas como raras, para ver de qué forma pueden estar condicionando los resultados.

Se destaca que la implementación de estos modelos ha generado experiencia práctica que puede ser de utilidad en futuras aplicaciones. La estimación de los LCMM y JLCMM resulta exigente computacionalmente y se ha logrado implementar de forma tal que la experiencia puede ser útil en la aplicación de otros trabajos. En esta investigación se ha indagado en distintas formas de parametrización de los modelos. En particular, cuando se tenga disponible la nueva ronda del ELBU, se contará con un tramo de edad donde el abandono escolar es más severo y, por tanto, resulta de interés analizar su vinculación con la trayectoria nutricional aplicando la misma metodología que en el presente trabajo.

La identificación del grupo con trayectoria nutricional de sobrepeso o incluso obesidad, según los estándares de la OMS, podría dar lugar a futuras investigaciones en términos de su vínculo con otros aspectos de la salud de los niños, características socioeconómicas del entorno u otras, que den cuenta de las características particulares de los niños y niñas con estos problemas nutricionales durante el período escolar. Identificar las características de los grupos con sobrepeso u obesidad resulta relevante para entender la población de niños y niñas que enfrentan estos problemas nutricionales que pueden asociarse también con otros desempeños de los niños y niñas.

¹ Con distintas funciones de enlace para distintas muestras de niños y niñas.

² Además, los criterios de convergencia del algoritmo son rigurosos (Proust-Lima *et al.*, 2016b).

Referencias

- ACC/SCN (2000). *Fourth Report on the World Nutrition Situation*. Geneva: Administrative Committee on Coordination - Sub-Committee on Nutrition in collaboration with IFPRI, WHO.
- Alderman, H., Behrman, J., Lavy, V., & Menon, R. (1997). *Child Nutrition, Child Health, and School Enrollment*. Policy Research Working Paper, The World Bank.
- Amarante, V., Arim, R., Severi, C., Vigorito, A., & Aldabe, I. (2007). *El estado nutricional de los niños/as y las políticas alimentarias*. PNUD, Uruguay.
- Arim, R., Failache, E., Salas, G., Severi, C., & Vigorito, A. (2014). *Trayectorias del bienestar multidimensional en la infancia. Un estudio en base a datos de panel*. Proyectos de Investigación e Innovación Orientados a la Inclusión Social. Llamado 2010. Modalidad 1. CSIC. IECON, UdelaR, Uruguay.
- Aristimuño, A. & deArmas, G. (2012). *La transformación de la educación media en perspectiva comparada: Tendencias y experiencias innovadoras para el debate en Uruguay*. UNICEF, Uruguay.
- Arredondo, J., Carranza, N., Vázquez, M., & Rodríguez, M. (2003). *Transición epidemiológica*. Acta Pediátrica de México, 24(1):46-56.
- Barker, D., Winter, P., Osmond, C., Margetts, B., & Simmonds, S. (1989). *Weight in infancy and death from ischaemic heart disease*. Lancet, 9;2(8663):577-80.
- Behrman, J. (1990). *The Action of Human Resources and Poverty on One Another. What We Have Yet to Learn*. Working Paper no. 74. The World Bank.
- Behrman, J. & Wolfe, B. (1987). *Investments in Schooling in Two Generations in Pre-Revolutionary Nicaragua: The Roles of Family Background and School Supply*. Journal of Development Economics, 27(1-2): 395-420.
- Bérgolo, M., Leites, M., & Salas, G. (2006). *Privaciones nutricionales: su vínculo con la pobreza y el ingreso monetario*. Serie Documentos de Trabajo, DT 03/2006, Instituto de Economía, FCEA, UdelaR.
- Biernacki, C., Celeux, G., & Govaert, G. (2003). *Choosing Starting Values for the EM Algorithm for Getting the Highest Likelihood in Multivariate Gaussian Mixture models*. Computational Statistics and Data Analysis, 41(3-4), 561-575.
- Boado, M. & Fernández, T. (2010). *Trayectorias académicas y laborales de los jóvenes uruguayos. El panel PISA 2003-2007*. Facultad de Ciencias Sociales, UdelaR.
- Bucheli, M. & Casacuberta, C. (2000). *Asistencia escolar y participación en el mercado de trabajo de los adolescentes en Uruguay*. El Trimestre Económico, Vol. LXVII (3), México.
- Butte, N., Garza, C., & deOnis, M. (2007). *Evaluation of the Feasibility of International Growth Standards for School-Aged Children and Adolescents*. Journal of Nutrition, 137:153-7.
- Cameron, A. & Trivedi, P. (2010). *Microeconometrics using STATA*. STATA Press, Revised edition.
- Cueto, S. & Chinen, M. (2000). *Impacto Educativo de un Programa de Desayunos Escolares en Escuelas Rurales del Perú*. Documento de Trabajo 34., Grupo de Análisis para el Desarrollo, GRADE. Perú.

- deMelo, G., Failache, E., & 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.
- deOnis, M., Blössner, M., & Borghi, E. (2004). *Methodology for estimating regional and global trends of child malnutrition*. International Journal of Epidemiology, 33:1260-1270.
- deOnis, M. & Habicht, J.-P. (1996). *Anthropometric reference data for international use: recommendations from a World Health Organization Expert Committee*. The American Journal of Clinical Nutrition, Vol. 64:650-8.
- deOnis, M. & Lobstein, T. (2010). *Defining obesity risk status in general childhood population: Which cut-offs should we use?* International Journal of Pediatric Obesity, 5:458-460.
- deOnis, M., Onyango, A., Borghi, E., Siyam, A., Nishida, C., & Siekmann, J. (2007). *Development of a WHO growth reference for school-aged children and adolescents*. Bulletin of the World Health Organization, 85:660-7.
- Engle, P., Menon, P., & Haddad, L. (1997). *Care and Nutrition: Concepts and Measurement*. International Food Policy Research Institute, Washington, D.C.
- Failache, E., Salas, G., & Vigorito, A. (2015). *Trayectorias educativas de los adolescentes en Uruguay. Un estudio en base a datos de panel*. Serie Documentos de Trabajo, DT 06/2015, Instituto de Economía, FCEA, Udelar.
- Faucett, C. & Thomas, D. (1996). *Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach*. Statistics in Medicine 15, 1663-1685, doi:10.1002/(SICI)1097-0258(19960815)15:15<1663::AID-SIM294>3.0.CO;2-1.
- Frees, E. (2004). *Longitudinal and Panel Data: Analysis and Applications for the Social Sciences*. Cambridge University Press.
- Georgiadis, A., Benny, L., Crookston, B., Duc, L. T., Hermida, P., Mani, S., Woldehanna, T., Stein, A., & Behrman, J. (2016). *Growth trajectories from conception through middle childhood and cognitive achievement at age 8 years: Evidence from four low- and middle-income countries*. SSM - Population Health, Elsevier, Vol. 2 pp. 43-54.
- Glewwe, P. (2005). *The impact of child health and nutrition on education in developing countries: Theory, econometric issues, and recent empirical evidence*. Food and Nutrition Bulletin, Vol. 26, no.2 (supplement 2). The United Nations University.
- Gluckman, P., Hanson, M., & Pinal, C. (2005). *The developmental origins of adult disease*. Matern Child Nutr. 1(3):130-41.
- Johnson, W. (2015). *Analytical Strategies in Human Growth Research*. American Journal of Human Biology. 27:69-83, Wiley.
- Laird, N. & Ware, J. (1982). *Random-effects models for longitudinal data*. Biometrics 38, 963-974. The International Biometric Society.
- Levinger, B. (1992). *Nutrition, Health and Learning: Current Issues and Trends*. Education Development Center and UNDP.
- Lin, H., Turnbull, B., McCulloch, C., & Slate, E. (2002). *Latent Class Models for Joint Analysis of Longitudinal Biomarker and Event Process Data: Application to Longitudinal Prostate-Specific Antigen Readings and Prostate Cancer*. Journal of the American Statistical Association. Vol. 97, No. 457, pp.53-65.

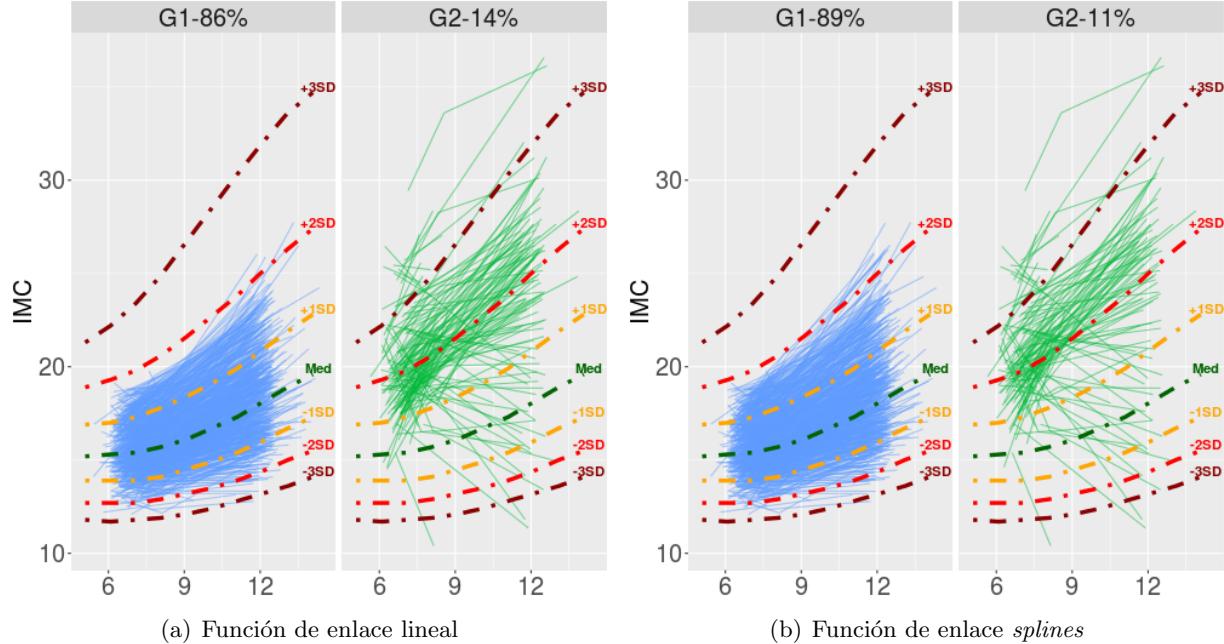
- Massa, F. (2015). *Efecto de valores faltantes en estudios longitudinales en adultos mayores*. Trabajo de Tesis para la Maestría en Ingeniería Matemática, Facultad de Ingeniería, UdelaR.
- Muthén, B. (2004). *Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data*. In D. Kaplan (ed.), *Handbook of quantitative methodology for the social sciences*, Newbury Park, CA: Sage Publications, in press.
- Proust-Lima, C., Dartigues, J.-F., & Jacqmin-Gadda, H. (2015). *Joint modelling of repeated multivariate cognitive measures and competing risks of dementia and death: a latent process and latent class approach*. Statistics in Medicine, 35(3). arXiv:1409.7598.
- Proust-Lima, C. & Jacqmin-Gadda, H. (2005). *Estimation of linear mixed models with a mixture of distribution for the random-effects*. Computer Methods and Programs in Biomedicine, Elsevier, Vol. 78 (2), pp.165-73.
- Proust-Lima, C., Joly, P., Dartigues, J.-F., & Jacqmin-Gadda, H. (2009). *Joint modelling of multivariate longitudinal outcomes and a time-to-event: A nonlinear latent class approach*. Computational Statistics and Data Analysis, 53 1142-1154.
- Proust-Lima, C., Phillipps, V., Diakite, A., & Liquet, B. (2016a). *Package 'lcmm'*. <https://cran.r-project.org/web/packages/lcmm/lcmm.pdf>.
- Proust-Lima, C., Phillipps, V., & Liquet, B. (2016b). *Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package lcmm*. arXiv:1503.00890.
- Proust-Lima, C., Séne, M., Taylor, J., & Jacqmin-Gadda, H. (2014). *Joint latent class models for longitudinal and time-to-event data: A review*. Statistical Methods in Medical Research, Vol. 23(1) 74-90.
- Ramsay, J. (1988). *Monotone Regression Splines in Action*. Statistical Science, Vol. 3, No. 4, pp. 425-441.
- Regnault, N. & Gillman, M. (2014). *Importance of Characterizing Growth Trajectories*. Annals of Nutrition and Metabolism. 65:110-113.
- Rizopoulos, D. (2012). *Joint Models for Longitudinal and Time-to-event Data*. Chapman and Hall. CRC Biostatistics Series.
- Sen, A. (1999). *Desarrollo y libertad*. Planeta (1ra edición en español 2000).
- Stuart, B. & Panico, L. (2016). *Early-childhood BMI trajectories: evidence from a prospective, nationally representative British cohort study*. Nutrition and Diabetes, doi:10.1038/nutd.2016.6.
- Therneau, T., Grambsch, P., & Fleming, T. (1990). *Martingale-Based Residuals for Survival Models*. Biometrika. Vol. 77 (1), pp. 147-160.
- Tsiatis, A. & Davidian, M. (2004). *Joint modeling of longitudinal and time-to-event data: An overview*. Statistica Sinica, 14:809–834.
- Verbeke, G. & Lesaffre, E. (1996). *A Linear Mixed-Effects Model with Heterogeneity in the Random-Effects Population*. Journal of the American Statistical Association, 91(433), 217–221.
- Verbeke, G. & Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, New York.

- Victora, C., Adair, L., Fall, C., Hallal, P., Martorell, R., Richter, L., & Sachdev, H. (2008). *Maternal and child undernutrition: consequences for adult health and human capital*. Lancet. 371:340–57, Series 2. doi:10.1016/S0140-6736(07)61692-4.
- Victora, C., DeOnis, M., Hallal, P., Blössner, M., & Shrimpton, R. (2010). *Worldwide timing of growth faltering: Revisiting implications for interventions*. Pediatrics. Vol. 125 (3), pp. e473-e480.
- Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press: Cambridge, MA.
- Wooldridge, J. (2010). *Correlated Random Effects Models with Unbalanced Panels*. Department of Economics. Michigan State University.
- Wu, L., Liu, W., Yi, G., & Huang, Y. (2012). *Analysis of Longitudinal and Survival Data: Joint Modeling, Inference Methods, and Issues*. Journal of Probability and Statistics, Vol. 2012, Article ID 640153, 17 pages. Doi:10.1155/2012/640153.
- Wulfsohn, M. & Tsiatis, A. (1997). *A joint model for survival and longitudinal data measured with error*. Biometrics 53, 330-339.
- Zajacova, A. & Ailshire, J. (2012). *Body Mass Trajectories and Mortality Among Older Adults: A Joint Growth Mixture-Discrete-Time Survival Analysis*. The Gerontologist Vol. 54, No. 2, 221-231.
- Zhao, M. & Glewwe, P. (2010). *What determines basic school attainment in developing countries? Evidence from rural China*. Economics of Education Review, 29(3): 451-460.

6. Apéndice de Resultados

6.1. Clasificación de niñas con enlace lineal y *splines*

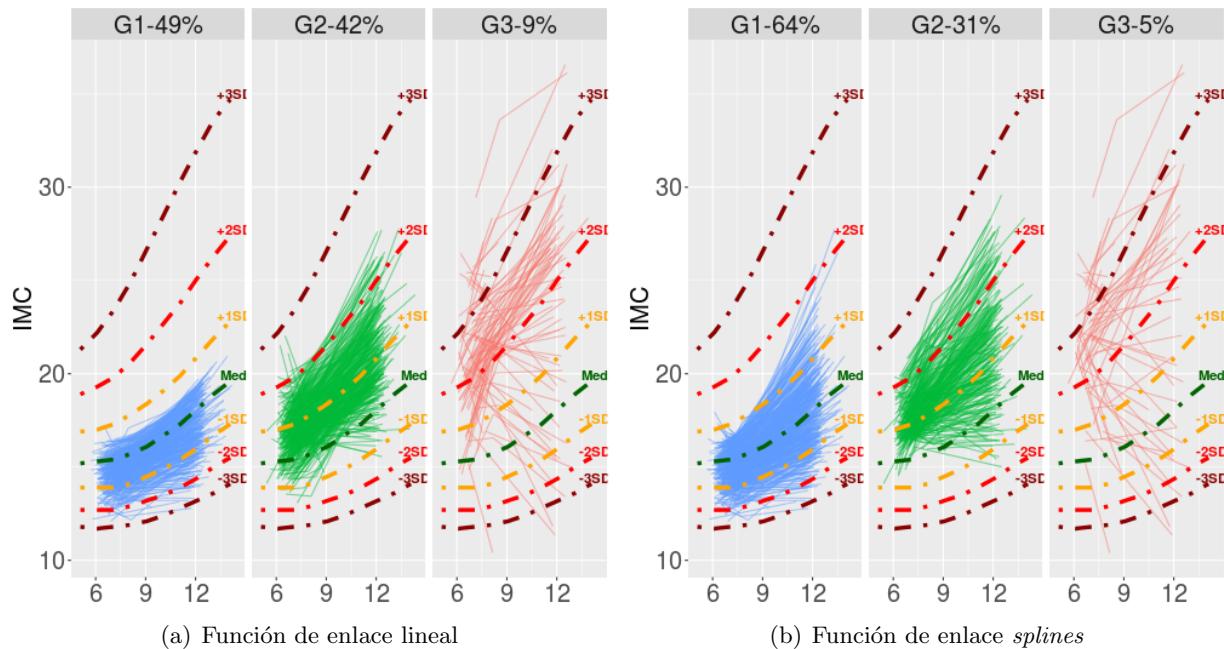
Figura 14: Trayectoria nutricional según dos grupos de niñas - LCMM



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

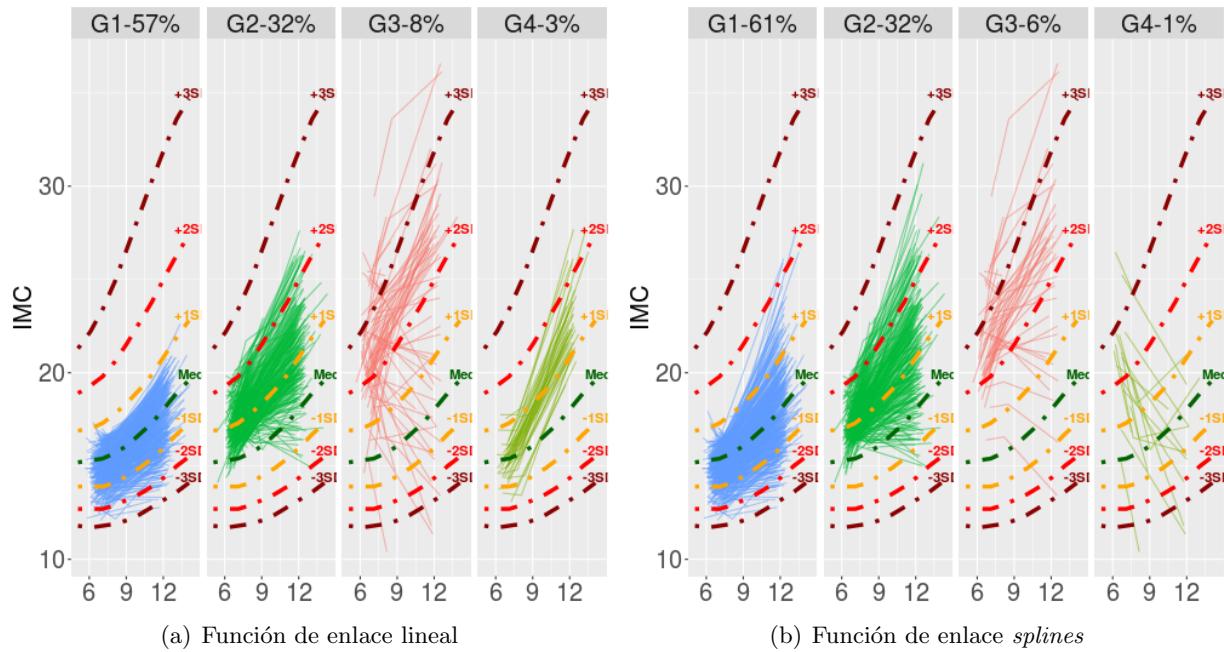
Figura 15: Trayectoria nutricional según tres grupos de niñas - LCMM



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

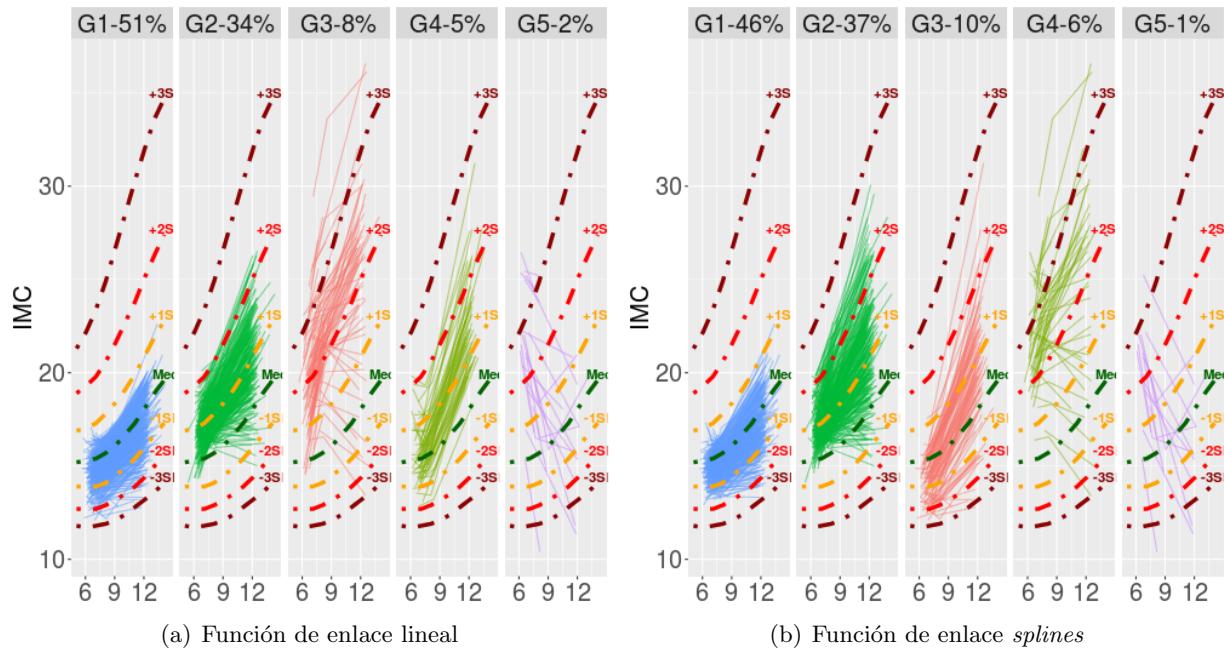
Figura 16: Trayectoria nutricional según cuatro grupos de niñas - LCMM



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

Figura 17: Trayectoria nutricional según cinco grupos de niñas - LCMM

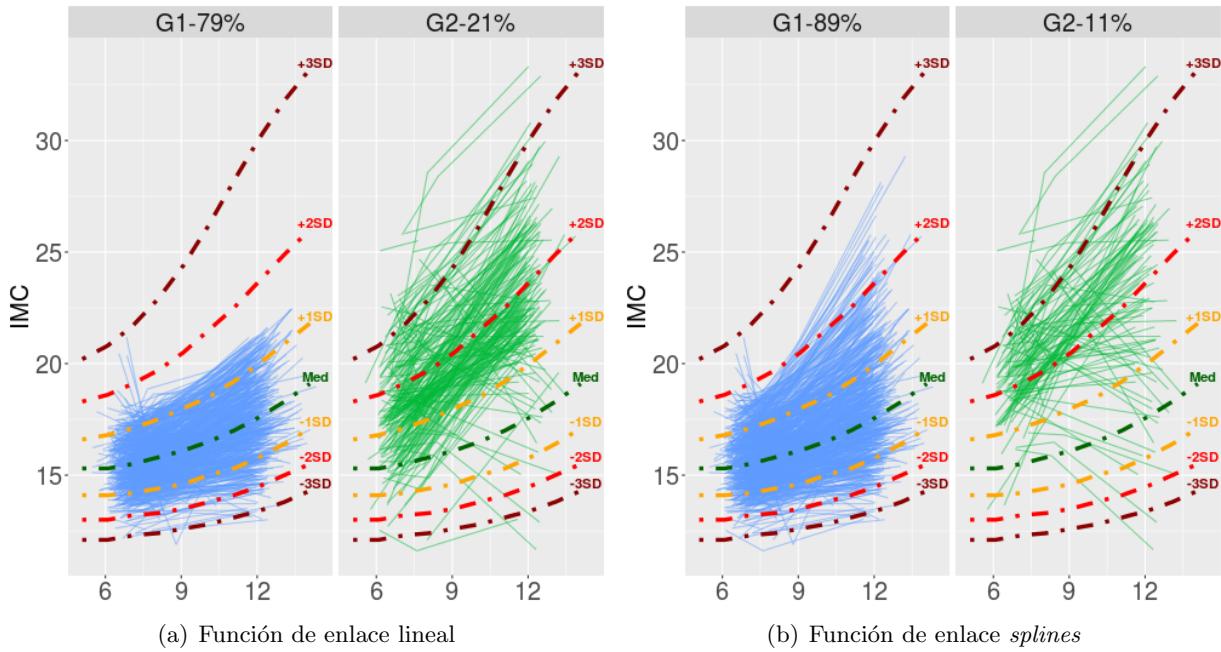


Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

6.2. Clasificación niños con enlace lineal y *splines*

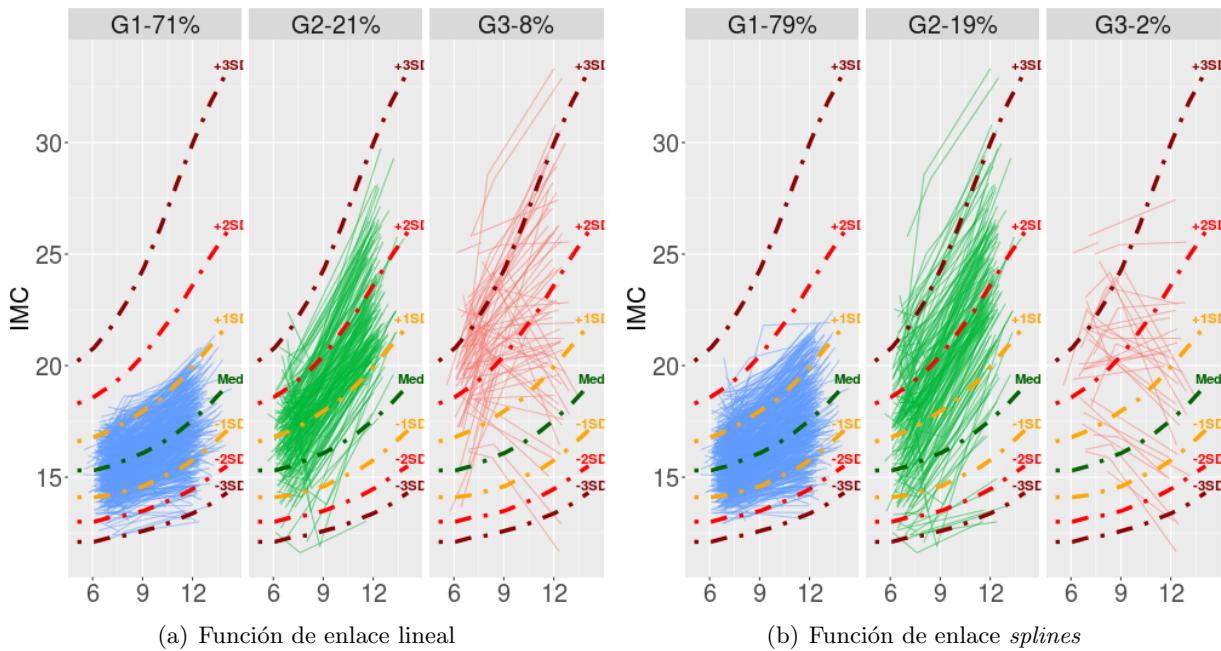
Figura 18: Trayectoria nutricional según dos grupos de niños - LCMM



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

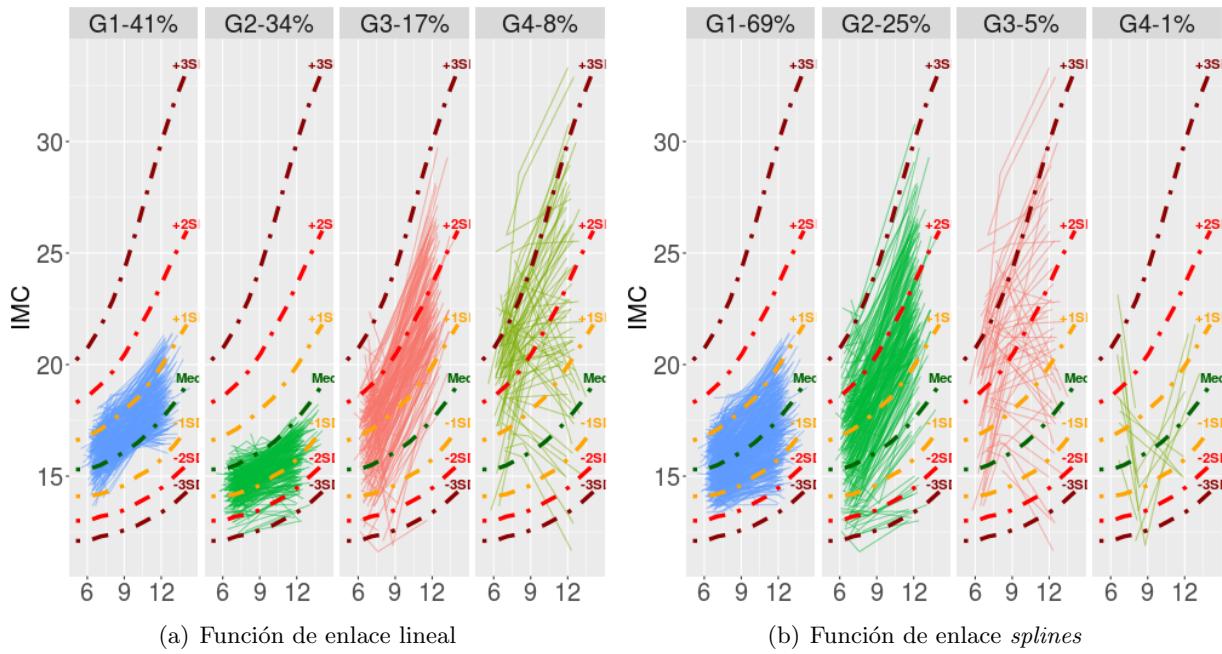
Figura 19: Trayectoria nutricional según tres grupos de niños - LCMM



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

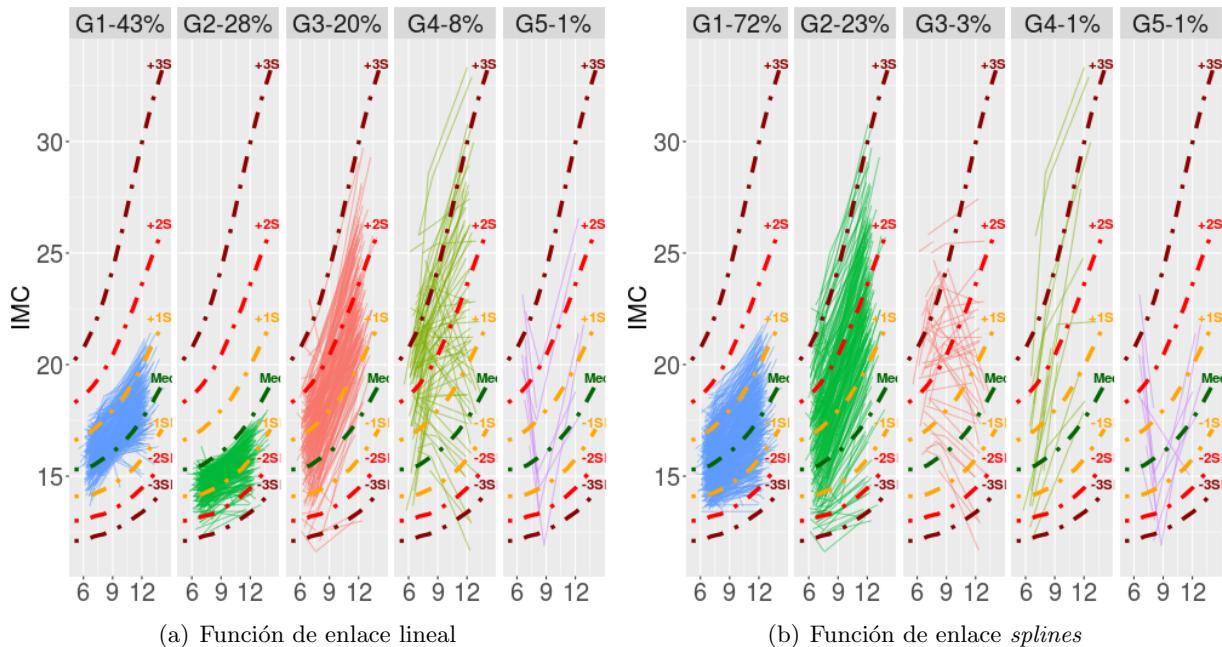
Figura 20: Trayectoria nutricional según cuatro grupos de niños - LCMM



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

Figura 21: Trayectoria nutricional según cinco grupos de niños - LCMM



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1SD$, $\pm 2SD$ y $\pm 3SD$).

6.3. Trayectoria nutricional niñas con enlace *splines*

Cuadro 16: Estadísticos y clasificación por grupos de niñas con LCMM y enlace *splines*

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niñas en cada grupo | | | | |
|----------------|---------------------|---------|---------|-----------------|-----------------------|------|------|-----|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -5532.1 | 11136.7 | | 10 | 100 % | | | | |
| 2 | -5454.0 | 11016.6 | 0.523 | 15 | 89 % | 11 % | | | |
| 3 | -5420.8 | 10986.4 | 0.582 | 20 | 64 % | 30 % | 6 % | | |
| 4 ^c | -5397.3 | 10975.5 | 0.660 | 25 | 61 % | 32 % | 6 % | 1 % | |
| 5 | -5384.0 | 10985.1 | 0.539 | 30 | 46 % | 37 % | 10 % | 6 % | 1 % |

N=1392

Fuente: Elaboración propia

^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

^c No Converge

Cuadro 17: Probabilidades *a posteriori* para tres grupos de niñas con LCMM y enlace *splines*

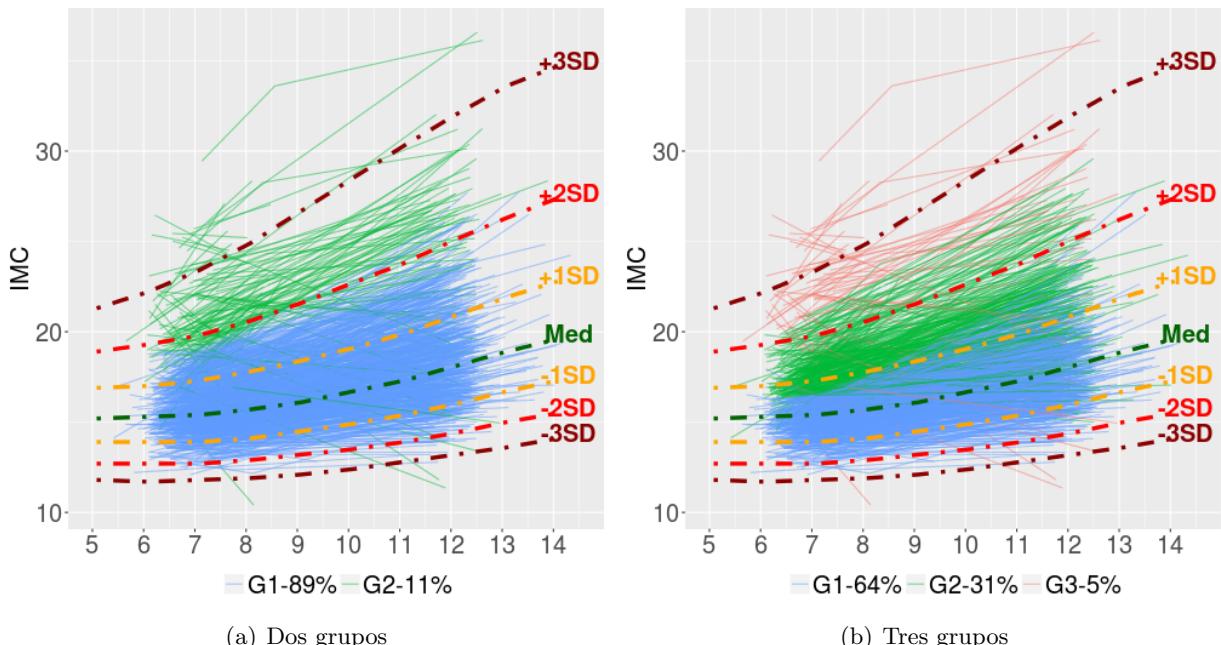
| | (a) Prob. media | | | (b) % >umbral ^a | | |
|-----|-----------------|-------|-------|----------------------------|-------|-------|
| | G 1 | G 2 | G 3 | > 0.7 | > 0.8 | > 0.9 |
| G 1 | 0.750 | 0.150 | 0.099 | 68 % | 42 % | 10 % |
| G 2 | 0.129 | 0.851 | 0.020 | 83 % | 71 % | 53 % |
| G 3 | 0.138 | 0.004 | 0.859 | 77 % | 71 % | 58 % |

N=1392

Fuente: Elaboración propia

^a % de niñas con probabilidad *a posteriori* que supera el umbral

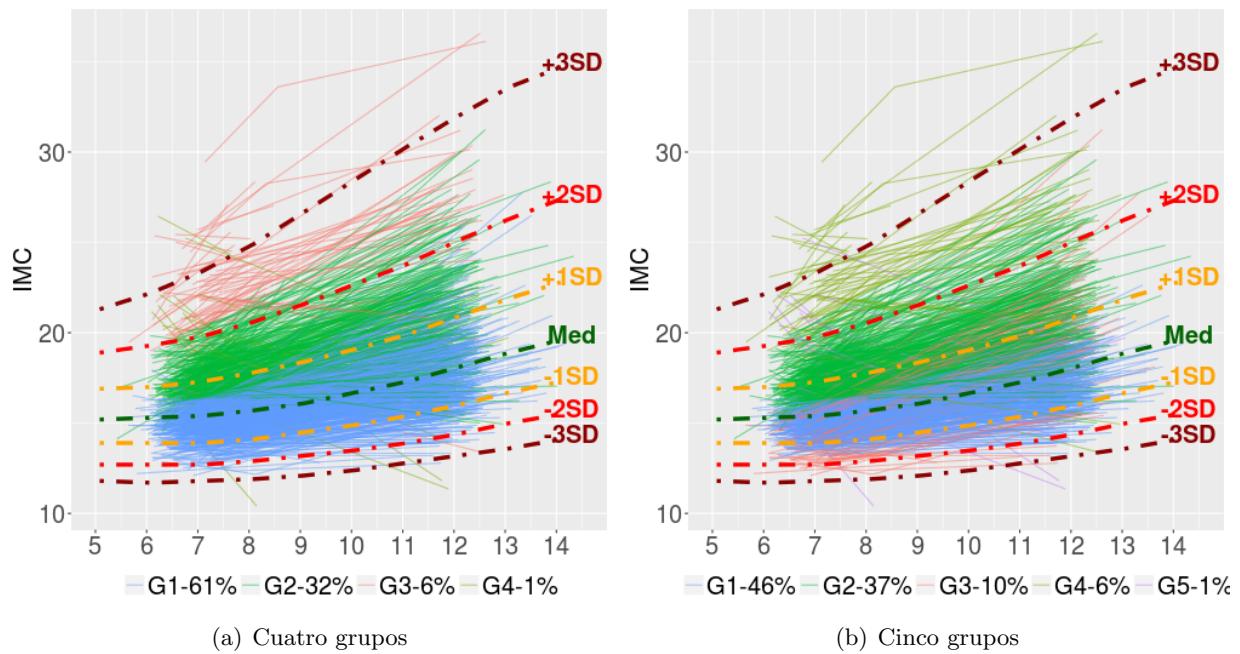
Figura 22: Trayectoria nutricional según dos y tres grupos de niñas con LCMM y enlace *splines*



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Figura 23: Trayectoria nutricional según cuatro y cinco grupos de niñas con LCMM y enlace *splines*



6.4. Trayectoria nutricional niños con enlace *splines*

Cuadro 18: Estadísticos y clasificación por grupos de niños con LCMM y enlace *splines*

| Grupos | loglik ^a | BIC | Entropy | NP ^b | % Niñas en cada grupo | | | | |
|----------------|---------------------|---------|---------|-----------------|-----------------------|------|-----|-----|-----|
| | | | | | G 1 | G 2 | G 3 | G 4 | G 5 |
| 1 | -5592.0 | 11257.1 | | 10 | 100 % | | | | |
| 2 | -5509.7 | 11129.0 | 0.628 | 15 | 89 % | 11 % | | | |
| 3 | -5483.0 | 11112.1 | 0.541 | 20 | 79 % | 19 % | 2 % | | |
| 4 | -5418.8 | 11020.3 | 0.504 | 25 | 69 % | 25 % | 5 % | 1 % | |
| 5 ^c | -5408.9 | 11036.9 | 0.586 | 30 | 72 % | 23 % | 3 % | 1 % | 1 % |

N=1492

Fuente: Elaboración propia

^a Logaritmo de la verosimilitud

^b Número de parámetros estimados

^c No Converge

Cuadro 19: Probabilidades *a posteriori* para tres grupos de niños con LCMM y enlace *splines*

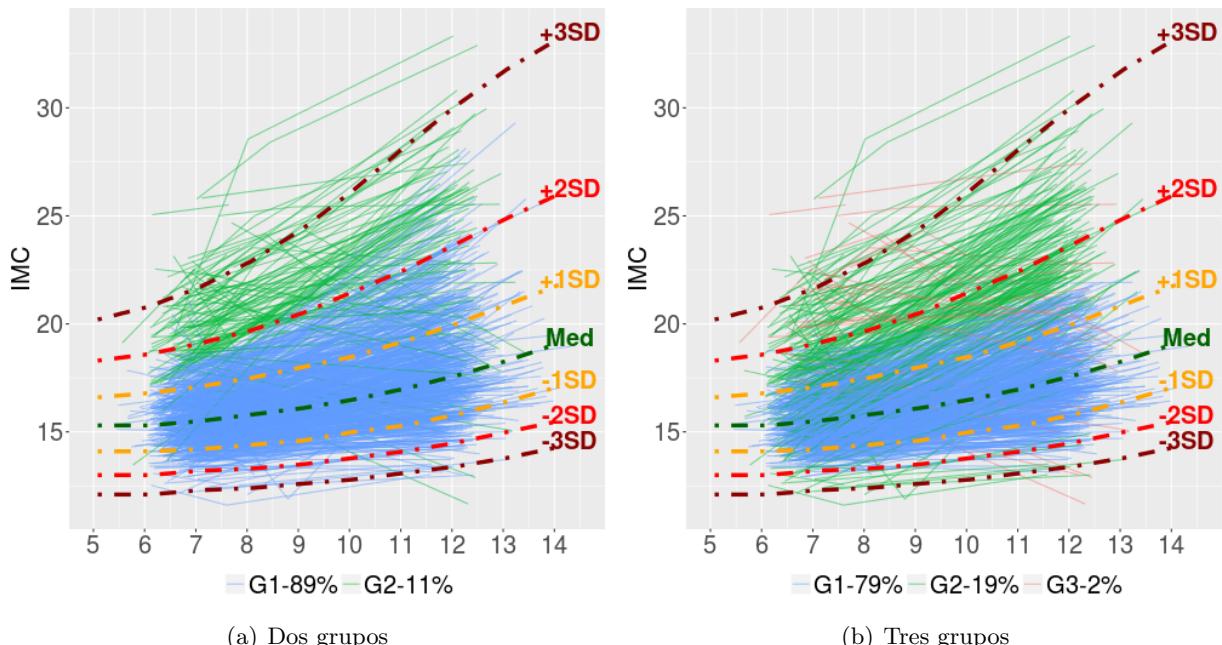
| | (a) Prob. media | | | (b) % >umbral ^a | | |
|-----|-----------------|-------|-------|----------------------------|-------|-------|
| | G 1 | G 2 | G 3 | > 0.7 | > 0.8 | > 0.9 |
| G 1 | 0.806 | 0.014 | 0.179 | 84 % | 59 % | 51 % |
| G 2 | 0.061 | 0.774 | 0.166 | 63 % | 51 % | 40 % |
| G 3 | 0.177 | 0.078 | 0.746 | 20 % | 41 % | 29 % |

N=1492

Fuente: Elaboración propia

^a % de niños con probabilidad *a posteriori* que supera el umbral

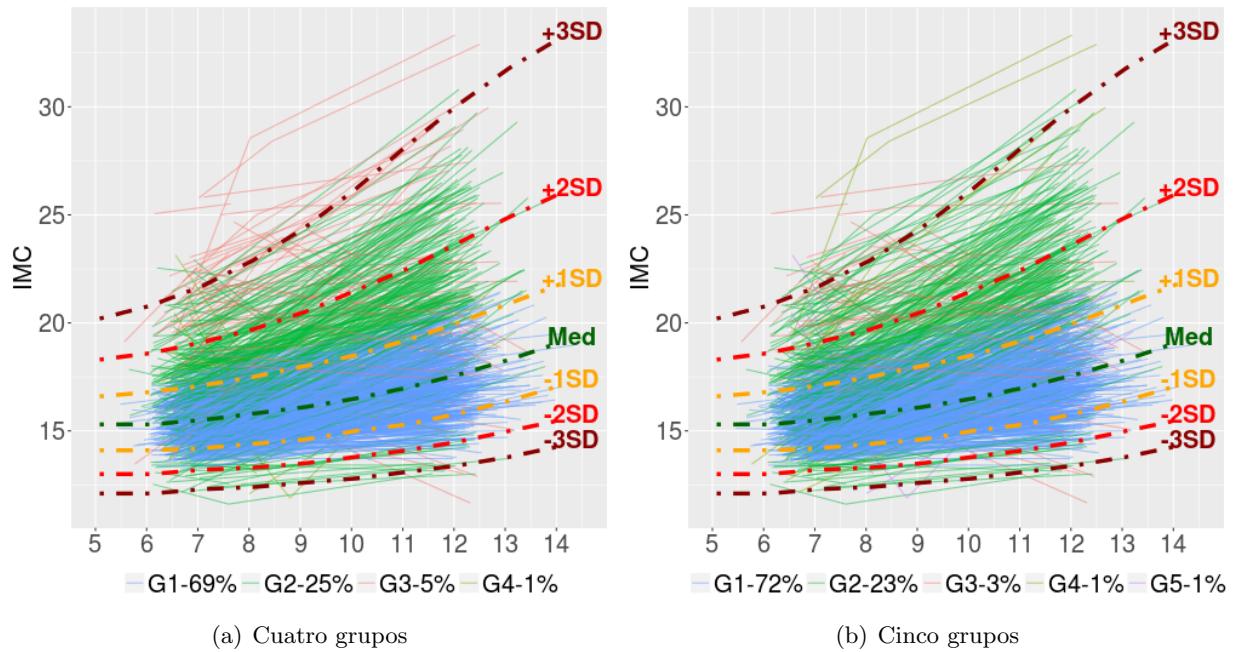
Figura 24: Trayectoria nutricional según dos y tres grupos de niños con LCMM y enlace *splines*



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

Figura 25: Trayectoria nutricional según cuatro y cinco grupos de niños con LCMM y enlace *splines*



Fuente: Elaboración propia

Nota: Valores de referencia del IMC: mediana (Med) y desvíos ($\pm 1\text{SD}$, $\pm 2\text{SD}$ y $\pm 3\text{SD}$).

INSTITUTO DE ECONOMÍA

Serie Documentos de Trabajo

Febrero, 2018
DT 02/2018



Instituto de Economía

Facultad de Ciencias Económicas y de Administración
Universidad de la República - Uruguay