

CONCORDANCIA ENTRE LAS VARIACIONES DE PUNTAJES SIMCE Y PUNTAJES DE PISA

RESUMEN

Las tablas de equivalencia entre los puntajes Simce y PISA para Lectura y Matemática¹, se usan para construir funciones de predicción continuas. Se aplican dichas funciones a los resultados de la prueba Simce 2012 para obtener predicciones para la prueba PISA 2012 Matemática. Los resultados obtenidos se comparan con los puntajes efectivos obtenidos por la muestra de grado que rindió la prueba PISA de 2012. Así se determina la calidad de las funciones de predicción de manera global y según distintos grupos de interés. Al comparar predicciones, se puede determinar el efecto en los puntajes Simce de ser cohorte que rinde la prueba nacional (cohortes 2006 y 2012) versus cohorte que no la rinde (2009), lo que se traduciría en una estimación de la instrucción adicional que reciben los estudiantes de II medio, cuando el sistema educacional sabe que se rinden las pruebas Simce.

1. ANTECEDENTES

Uno de los principales propósitos de la participación de Chile en pruebas como PISA y TIMSS es el monitoreo de la evolución de los logros educacionales de estudiantes chilenos en base a referentes internacionales (Matus et al., 2012). Chile ha sumado varios ciclos de participación, tanto en la prueba PISA (2000, 2006, 2009 y 2012), como en la prueba TIMSS (1999, 2003 y 2011), lo que ha permitido determinar los avances en Matemáticas, Lectura y Ciencias para las distintas poblaciones de estudiantes que participan de ellas. Así, los resultados de la prueba PISA reflejan un aumento significativo de puntajes en Lectura y Ciencias, y un aumento no significativo de los puntajes de Matemática para alumnos de quince años. Similarmente, a partir de la comparación de los resultados TIMSS entre 2003 y 2011, se observa un aumento significativo en los resultados de Matemática y Ciencias, para estudiantes de 8° básico².

Del mismo modo, las pruebas Simce también tienen por objeto monitorear el progreso educativo de nuestro sistema escolar. Los resultados de esta prueba muestran un aumento significativo de los puntajes en las pruebas de Lectura y Matemática en 4° básico³. De manera similar a lo observado en TIMSS, en 8° básico, se aprecia un aumento significativo en los resultados de Matemática y Ciencias Naturales entre 2004 y 2013⁴. Para los II medio se ha observado estabilidad en los puntajes Simce en Lectura⁵ y un aumento sostenido en Matemática⁶. Esto pareciera estar en contradicción con lo advertido

¹ Desarrolladas para las muestras de grado de II medio, que rindieron la prueba PISA los años 2006 y 2009.

² En 2011 por primera vez se participó en 4° básico en TIMSS por lo que no se dispone de tendencias.

³ Se aumento 30% de desviación estándar en Matemática y Lectura, entre 2007 y 2010 (pese a que para Lectura después se haya verificado una baja), y en Matemática entre 2007 y 2012.

⁴ El aumento ha sido de 20% de desviación estándar en Matemática y 32% para Ciencias. La prueba TIMS no mide Lectura.

⁵ Aunque entre el año 2003 y 2010 se aumentó en un 11% de desviación estándar, posteriormente el puntaje promedio bajó hasta llegar al mismo nivel de 2003.

⁶ Entre 2003 y 2013 se ha aumentado 34% de desviación estándar.

en la prueba PISA. Sin embargo, hay que considerar que la prueba PISA tiene como población objetivo a los estudiantes de quince años que asisten a establecimientos regulares desde 7° básico en adelante. El año 2000, cuando Chile participó por primera vez en dicha prueba, el porcentaje de estudiantes de quince años que asistía a un nivel inferior a II año medio era de 39%, y en el año 2012 dicho porcentaje bajó a 27%. Lo anterior constituye una mejora del sistema educacional en su conjunto, ya que significa una disminución de la cantidad de estudiantes que presentan algún retraso, con respecto de su edad, en su trayectoria académica.

Con el objeto de estudiar más profundamente los factores que inciden en el rendimiento académico, a partir de la aplicación del año 2006 Chile participa en las pruebas PISA con una muestra de grado, además de la muestra regular de estudiantes de quince años. La muestra de grado no solo entrega información sobre las competencias cognitivas medidas por la prueba PISA, sino que también de los índices actitudinales y contextuales que permiten analizar estudios de factores asociados al rendimiento en estudiantes de II medio en un contexto internacional. Estas muestras agrupan a estudiantes de II medio, seleccionados dentro de los mismos establecimientos educacionales en los cuales se determina la muestra regular que participa en la prueba PISA, y son representativas de la población de estudiantes chilenos de dicho grado (Fernández, 2008).

En Matus et al. (2012) se determinaron tablas de equivalencia PISA-Simce basadas en las muestras de grado que rindieron las pruebas PISA 2006 y 2009, y que simultáneamente rindieron una prueba Simce de II medio. En dicho trabajo se detectó que las tablas de equivalencia para la prueba de Matemática para las muestras de 2006 y de 2009 eran cualitativamente distintas, ya que un mismo puntaje Simce no entregaba similar puntaje PISA. No así en Lectura, donde estas presentaban diferencias menores. Considerando que las pruebas Simce y las pruebas PISA utilizan la Teoría de Respuesta al Ítem (TRI) tanto para su análisis, como para mantener la escala de medición⁷, lo esperable era encontrar coincidencias entre ambas tablas. En efecto, bajo la TRI, existe invarianza⁸ tanto para las preguntas, como para la estimación de la habilidad que ellas miden. En consecuencia, dado que las pruebas PISA se equiparan⁹ entre cada aplicación, y lo mismo acontece con las pruebas Simce, se esperaría que la relación entre ambas pruebas para distintas aplicaciones, se mantuviera. Se observa sin embargo, que los estudiantes de la muestra de grado 2009 obtienen un rendimiento bastante más bajo que sus pares del año 2006 en la prueba Simce¹⁰, de manera que para un mismo puntaje PISA, la equivalencia del puntaje Simce según muestra considerada podía variar más de un 12%¹¹.

A partir de este hallazgo surge la necesidad de plantear hipótesis que permitan explicar estas discordancias. Para elaborar estas hipótesis, es importante tomar en cuenta que tanto PISA como Simce presentan una calidad estadística similar y estable en lo que respecta a su construcción¹², como también a las muestras a las que se aplican en el caso de PISA. Por otro lado, las cohortes estudiadas

⁷ De manera de permitir la comparabilidad en el tiempo y por ende el monitoreo de los logros educacionales.

⁸ La invarianza de las preguntas se refiere a que sus características estadísticas serán estable en el constructo o rasgo latente medido por la prueba, independiente del nivel de habilidad de la muestra de estudiantes a las que se les aplica. Para una clarificación del concepto ver Apunte N.º6 sobre la Calidad de la Educación, disponible en: https://s3-us-west-2.amazonaws.com/documentos-web/Papers/2013_06_Relacion_TIMSS_SIMCE_Calidad_psicometrica_de_puntajes_y_parametros.pdf

⁹ Se pueden usar de manera intercambiada (Kolen y Brennan 2004).

¹⁰ Los estudiantes de la muestra de grado 2009 rindieron la prueba Simce 2008 dentro del marco de la aplicación de PISA 2009.

¹¹ Las variaciones para Lectura eran aproximadamente de un 5%.

¹² Por un lado las pruebas Simce y PISA se construyen y analizan bajo la Teoría de Respuesta al Ítem (TRI) y estrictos controles de calidad, lo que permite asegurar comparabilidad en el tiempo y, por otro, las muestras de grado también satisfacen estrictos protocolos de selección.

(alumnos de II medio en los años 2006 y 2009), estuvieron sometidas a los mismos requerimientos curriculares¹³. La primera hipótesis que se podría plantear es que para alguna de las dos pruebas no se estuvieran verificando los supuestos de TRI. Esta hipótesis era poco plausible, considerando los procesos de construcción de ambas pruebas (OECD [2014], Agencia de Calidad de la Educación [2013]), y quedó descartada a la luz de los resultados de este trabajo.

Una de las posibles diferencias entre dichas cohortes responde al hecho de que, si bien la participación en las pruebas Simce de Lenguaje y Matemática de los alumnos de II medio del año 2006 fue programada y difundida con anterioridad a su aplicación¹⁴, la aplicación del 2009 fue sorpresiva para la muestra de estudiantes¹⁵. Dada la importancia que se asigna a la prueba Simce para efectos comparativos entre establecimientos escolares, las dos cohortes de estudiantes pueden haber recibido un trato diferente de parte de docentes y directivos. En este sentido, el nivel de instrucción que entrega el sistema escolar chileno es sensible a una futura aplicación planeada de la prueba Simce. En esta dirección, el grupo de 2006 habría recibido una instrucción adicional a la que recibió la cohorte 2009, respondiendo a las presiones de la aplicación de la prueba Simce, lo que no sucedió con la cohorte posterior.

La participación chilena en PISA 2012 brinda la oportunidad de testear la hipótesis general de que existe una instrucción adicional a los grupos que rinden Simce, como también aquella que apunta a la violación de supuestos de TRI en alguna de las dos evaluaciones. Esta prueba proporciona una muestra que comparte las características de la muestra 2006, en tanto que los alumnos que el año 2012 cursaban II medio rindieron además las pruebas Simce. Es posible comparar, en esta nueva muestra de grado, independiente de las anteriores, el comportamiento de las tablas de equivalencia construidas para las muestras 2006 y 2009 para Matemática, que es aquella que presentó mayores discordancias en el análisis de Matus et al. (2012)¹⁶.

En este documento se presenta el resultado de dicha comparación, como también se pone en discusión y profundiza en la diferencias de resultados obtenidos para las distintas muestras de grado. Este documento se estructura presentando los objetivos y la metodología a continuación, para seguir con los resultados obtenidos. Finalmente, se cierra con las conclusiones y alcances del trabajo.

¹³ El año 2009 se modificaron las bases curriculares para enseñanza media pero solo fueron aplicables a partir del año 2010.

¹⁴ El Mineduc publica con años de antelación el calendario de pruebas Simce.

¹⁵ Ambas cohortes rindieron las pruebas Simce de 8º básico por lo que puede considerarse que han estado expuestos de similar manera a la prueba Simce en educación básica.

¹⁶ Se decidió trabajar solo con la prueba de Matemática, ya que para la prueba de Comprensión de Lectura no presentaban mayores diferencias entre las tablas de alineamiento.

2. Objetivos

Objetivo Principal

Determinar si hay evidencia para considerar comportamiento diferencial en pruebas Simce, según tipo de cohorte de estudiantes.

Objetivos específicos

- Construir funciones continuas de equivalencia entre las pruebas PISA y Simce, basadas en las tablas de equivalencia¹⁷ de las cohortes 2006 y 2009.
- Determinar para la muestra de grado PISA 2012 la mejor función continua de equivalencia que permita predecir los resultados PISA 2012, basados en rendimiento en la prueba Simce 2012.
- Evaluar si existen comportamientos diferenciales de grupos de estudiantes de acuerdo a la función de equivalencia obtenida, esto es si la calidad predictiva de la función de equivalencia escogida en el punto anterior es diferente para distintos grupos de estudiantes.

3. Metodología

En base a las tablas de alineamiento determinadas en Matus et al. (2012) se procedió a determinar funciones continuas que cumplieran con las tablas de equivalencia usando mínimos cuadrados ordinarios. La evaluación de los ajuste de las distintas funciones de equivalencia, como también los comportamiento diferenciales para sub-poblaciones, se realizó calculando los errores cuadráticos medios (ECM), errores absolutos y errores relativos puntuales, analizando los resultados obtenidos reales versus los resultados predichos por el modelo seleccionado. Considerando que el diseño muestral utilizado para la selección de las muestras de grado PISA es complejo¹⁸, el tratamiento de los datos involucró los cálculos de las desviaciones estándares bajo técnicas de re-muestreo, usando estimadores de varianza según el método de Balanced Repeated Replicated (BRR), con ajuste de Fay. Para el caso de la muestra 2012, el tratamiento de los resultados se concretó mediante los “plausible values”, tal como se indica que debe hacerse de acuerdo a los reportes técnicos de las pruebas PISA y a los manuales de usuarios correspondientes (OECD [2012], OECD[2009]).

¹⁷ Las pruebas Simce y Pisa entregan estimaciones de la habilidad de estudiantes que las rinde en base a las distribuciones de dichas habilidades se establecen equivalencias entre los puntajes de ambas pruebas en puntos discretos de las respectivas escalas de medición. Se requiere que las función de equivalencia continuas que se respeten o pasen por dichos puntos.

¹⁸ Estratificado, multi-etápico con distinta probabilidad de conglomerados.

4. Resultados

4.1 Características de las muestras de grado 2006, 2009 y 2012

La Tabla 4.1 detalla los estadísticos descriptivos de las tres muestras de grado analizadas con respecto de variables contextuales¹⁹. A partir de esta tabla, se observa que las muestras para los distintos años exhiben valores similares en las distintas variables consideradas. La edad promedio de estos alumnos de segundo medio es cercana a los 16 años, aunque con tendencia a disminuir. El nivel social, económico y cultural ha ido aumentando paulatina y comparativamente, con respecto del de los países que conforman la OECD, lo mismo que la riqueza de las familias. Las posesiones culturales y recursos educacionales de las familias tienen un comportamiento de parábola invertida: sin embargo, el máximo nivel educacional de los padres presenta aumento a lo largo de estos años.

En la Tabla 4.2 se advierte también un ligero incremento en la proporción de mujeres en las muestras para los 3 años considerados.

Tabla 4.1 *Características de las distintas muestras de grado PISA según índices contextuales*

Muestra		Edad	Nivel social, económico y cultural ^a	Posesiones culturales ^b	Recursos educativos en el hogar ^c	Riqueza familiar ^d	Máximo nivel educativo de padres ^e
2006	N	253.577	252.375	252.167	252.594	252.639	250.243
	Mín.	13,67	-4,13	-1,41	-4,47	-3,89	0
	Máx.	19,83	2,87	1,05	0,93	2,11	6
	Media	16	-0,65	-0,07	-0,61	-0,99	3,8
	Desv. típ.	0,65	1,17	0,86	1,12	0,96	1,64
2009	N	249.472	247.321	246.720	247.036	247.207	243.894
	Mín.	14,67	-4,52	-1,38	-4,01	-4,02	0
	Máx.	20,42	2,62	1,14	1,25	2,5	6
	Media	16,03	-0,57	0,01	-0,14	-0,64	4,01
	Desv. típ.	0,68	1,16	0,87	1,01	0,94	1,53
2012	N	223.724	221.194	222.141	222.604	222.695	218.579
	Mín.	14,75	-3,97	-1,52	-3,98	-4,71	0
	Máx.	16,33	2,62	1,22	1,01	2,8	6
	Media	15,88	-0,51	-0,23	-0,57	-0,44	4,11
	Desv. típ.	0,37	1,03	0,96	1,02	0,96	1,51

Nota: Las índices presentadas (salvo la Edad y Máximo Nivel educativo de los padres), se calculan utilizando el conjunto de los países OECD para cada aplicación y están normalizados.

- Índice ESCS (Economic, Social and Cultural Status index). Corresponde a un índice de nivel económico, social y cultural, construido para cada ciclo de PISA. Toma en cuenta, entre otros, el estatus ocupacional y el nivel educacional más alto de entre ambos padres.
- Índice CULTPOSS (Cultural Possessions index). Se basa en las respuestas de los estudiantes acerca de la presencia de algunos bienes en su casa como literatura clásica, libros de poesía y obras de arte.

¹⁹ La tabla presenta valores ponderados que dan cuenta de la naturaleza estratificada de las muestras consideradas.

- c. Índice HEDRES (Home Educational Resources index). Se basa en los ítems que miden la existencia de recursos educativos en el hogar tales como un escritorio y un lugar tranquilo para estudiar, un computador para hacer tareas, software educacional, libros que lo ayuden en sus trabajo escolar, entre otros.
- d. Índice WEALTH (Family Wealth index). Se basa en la respuesta de los estudiantes respecto de ciertos bienes que se encuentran en su hogar, como un cuarto propio, conexión a internet y otros.
- e. Índice HISEI (Highest Occupational level of parents). Corresponde al nivel ocupacional más alto de los padres según un puntaje asignado por ocupación.

Fuente: Elaboración propia en base a los resultados de las pruebas PISA 2006, 2009 y 2012.

Tabla 4.2 *Características de las distintas muestras de grado PISA según la distribución por sexo*

Distribución por sexo		Mujer	Hombre	Total
2006	N	118.449	136.099	254.548
	%	46,53	53,47	100,00
2009	N	122.704	126.768	249.472
	%	49,19	50,81	100,00
2012	N	114.407	109.317	223.724
	%	51,14	48,86	100,00

Nota: Fuente: Elaboración propia en base a los resultados de las pruebas PISA 2006, 2009 y 2012.

4.2 Presentación de las funciones de alineamiento

Las tablas de equivalencia (Matus et al., 2012), relacionan puntajes Simce con puntajes PISA de manera discreta, es decir punto a punto. Para poder aplicar las tablas de equivalencia para puntajes Simce en una nueva muestra independiente, se debe obtener una función que relacione ambos puntajes y que respete las tablas de equivalencia ya obtenidas. Se escogió ajustar las funciones por mínimos cuadrados, ya que la tabla de equivalencia se obtuvo utilizando una metodología equipercantil (Kolen y Brennan, 2004), por lo que se puede considerar que cada entrada en la tabla corresponde a un valor esperado, y en consecuencia no se puede considerar que sean valores medidos sin incertidumbre, como cuando se usan métodos de ajuste funcionales que interpolan los datos como ajuste por polinomios, funciones splines, etc.

Las figuras 4.1 y 4.2 muestran los gráficos de dispersión de la tabla de equivalencia Simce-PISA 2006 y 2009, respectivamente. Cada gráfico incluye la misma función lineal para resaltar la diferencia entre las relaciones entre cada prueba según cohorte. En los gráficos se puede apreciar que los datos para 2006 se acercan más a la función lineal que los datos para 2009. Asimismo, para puntajes bajos de Simce (inferiores a 200) existe un alejamiento con respecto a dicha función que es más notorio en 2009 que en 2006. Adicionalmente, en ambos gráficos observamos que un ajuste de una función lineal no sería satisfactorio, ya que no daría cuenta de los cambios en las inclinaciones de ambas curvas.

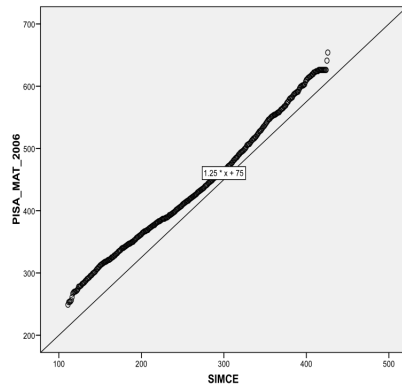


Figura 4.1 Gráfico de dispersión de la tabla de alineamiento Simce-PISA 2006, conjuntamente con una función lineal.
Fuente: Elaboración propia.

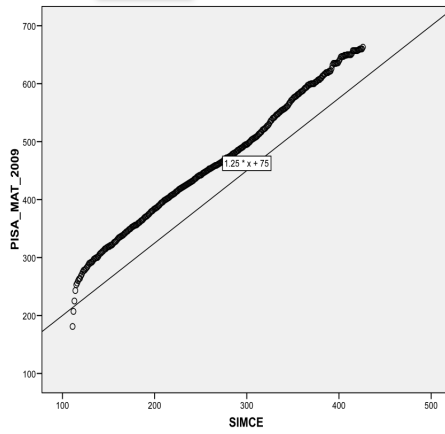


Figura 4.2 Gráfico de dispersión de la tabla de alineamiento Simce-PISA 2009, conjuntamente con una función lineal. Fuente: Elaboración propia.

4.2.1 Funciones de ajuste

Para determinar la función continua de alineamiento se probaron diversos ajustes, funciones lineales, cuadráticas y cúbicas por tramos²⁰ y en rango completo (ver Anexo). A continuación se presenta el mejor ajuste logrado, considerando la minimización del error cuadrático medio y el ajuste a los datos de las tablas de equivalencia (Matus et al., 2012).

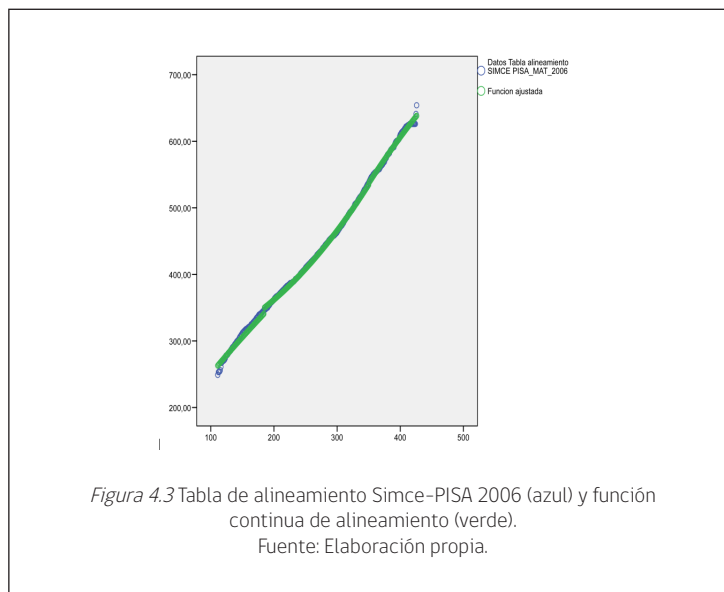
Para el alineamiento Simce–Pisa 2006 Matemática se obtiene:

$$PISA_{2006}(\text{Simce}) = \begin{cases} 15,63 + 1,5657 \text{ Simce} - 0,0026 \text{ Simce}^2 + 4 \cdot 10^{-6} \text{ Simce}^3 & \text{para Simce} \leq 185 \\ 281,5 + 0,00193 \text{ Simce}^2 + 3,88 \cdot 10^{-7} \text{ Simce}^3 & \text{para } 185 < \text{Simce} \leq 351 \\ -112,91 + 2,05 \text{ Simce} - 1,58 \cdot 10^{-6} \text{ Simce}^3 & \text{para } 351 < \text{Simce} \end{cases}$$

Para el alineamiento Simce–Pisa 2009 Matemática, en cambio:

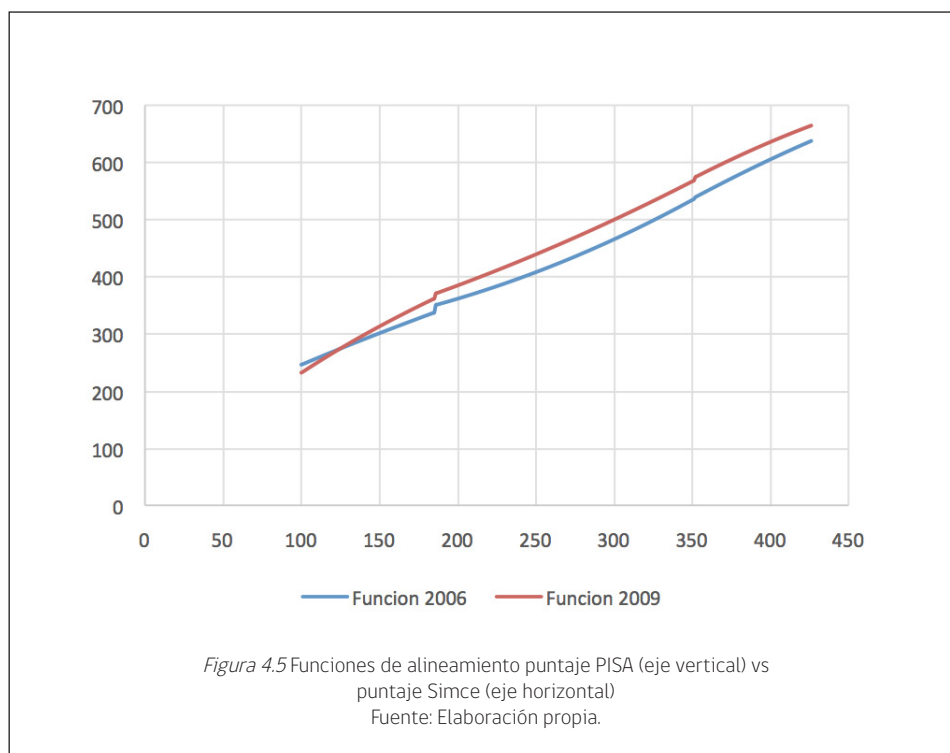
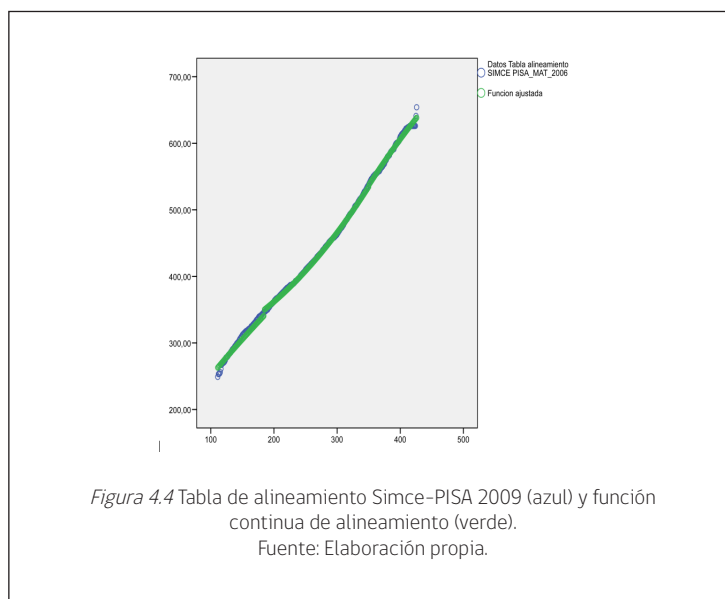
$$PISA_{2009}(\text{Simce}) = \begin{cases} 9,011 + 2,7416 \text{ Simce} - 0,0058 \text{ Simce}^2 - 7 \cdot 10^{-6} \text{ Simce}^3 & \text{para Simce} \leq 185 \\ 231,2 + 0,515 \text{ Simce} + 0,00127 \text{ Simce}^2 & \text{para } 185 < \text{Simce} \leq 351 \\ -97,85 + 2,17 \text{ Simce} - 2,096 \cdot 10^{-6} \text{ Simce}^3 & \text{para } 351 < \text{Simce} \end{cases}$$

En las figuras 4.3 y 4.4 se presentan los resultados en representaciones gráficas²¹, donde se observa que la función continua concuerda con los valores de la tabla de alineamiento casi en todo punto para ambas cohortes.



²⁰ Se escogieron tres tramos de puntaje Simce: los inferiores a 185, los entre 185 y 351 y los superiores a 351.

²¹ Los gráficos por tramos se encuentran en el anexo.



La Figura 4.5 entrega las funciones de alineamiento continuas para los datos 2006 y los datos 2009. Observamos que dichas funciones son bastante similares hasta los 150 puntos Simce, para luego comenzar a diferenciarse. Esto significa que sobre este umbral, el mismo puntaje Simce se asocia con un puntaje PISA diferente según el año. Así, un puntaje Simce de 250 el año 2006 se asocia con un

puntaje PISA de alrededor de 408 puntos, en cambio el mismo puntaje Simce de 250 para la cohorte 2009 se asocia con un puntaje PISA de sobre 430 puntos.

4.3 Ajuste de las funciones a la muestra de grado 2012 y resultados para grupos de estudiantes

Para los estudiantes de la muestra de grado de PISA 2012, a quienes también les correspondió rendir la prueba Simce, se estimó el puntaje PISA en Matemática usando las funciones de alineamiento obtenidas para 2006 y 2009. A estos valores se los llamó "Estimación PISA al 2006"²² y "Estimación PISA al 2009", según la función proveniente de la muestra empleada para el cálculo. Lo obtenido se contrastó con los resultados reales observados en PISA Matemática²³ de la muestra de grado 2012, tal como se describe en la Tabla 4.3 para la población general:

Tabla 4.3 *Resultados observados en PISA 2012 de la muestra de grado 2012 y resultados estimados usando las funciones de alineamiento 2006 y 2009*

	N	Mínimo	Máximo	Media	Error Est.	Desv. Est.
Estimación PISA al 2006	194.803	268,91	633,51	427,35	3,04	73,07
Estimación PISA al 2009	194.803	260,47	660,43	455,86	3,28	79,18
Promedio PISA 2012 Matemática	223.724	211,8	734,47	426,97	2,87	77,68

Nota: Fuente: elaboración propia.

Se observa claramente que el resultado promedio de la muestra de grado efectivamente obtenido ("Promedio PISA 2012 Matemática") es muy cercano al promedio de lo que pronostica el alineamiento 2006. Sin embargo, el de la predicción que utiliza la función de alineamiento 2009 sobreestima en casi 29 puntos el rendimiento PISA observado en 2012. En la práctica, los intervalos de confianza al 95% para el valor promedio de los resultados PISA 2012, y el valor promedio de las estimaciones PISA basado en el alineamiento 2006 son los mismos. En cambio existe diferencia significativa entre el resultado promedio los resultados PISA 2012 y el valor promedio de las estimaciones PISA basado en el alineamiento 2009.

La Tabla 4.4 describe los resultados PISA 2012 observados y predichos para estudiantes hombres y estudiantes mujeres. La tabla sugiere que la estimación basada en el alineamiento 2006 pronostica mejor el resultado PISA para estudiantes de ambos sexos, en comparación a la predicción, que basada en la función de alineamiento 2009. De hecho, esta última predicción sobrestima la media observada de PISA 2012 en 21 puntos para los hombres y 34 puntos para las mujeres.

²² "Estimación PISA al 2006" es por alineamiento 2006.

²³ Los entregados por el Consorcio de Instituciones a cargo de PISA 2012.

Tabla 4.4 *Resultados observados en PISA 2012 de la muestra de grado 2012 y resultados estimados usando las funciones de alineamiento 2006 y 2009 según sexo*

Sexo	Estimación/Puntaje	N	Mínimo	Máximo	Media	Desv. típ.
Hombre	Estimación PISA al 2006	95.165	268,91	633,51	432,52	74,99
	Estimación PISA al 2009	95.165	260,47	660,43	461,38	80,49
	Promedio PISA Matemática	104.677	262,59	734,47	440,07	76,05
Mujer	Estimación PISA al 2006	99.637	270,84	633,51	422,42	70,83
	Estimación PISA al 2009	99.637	263,35	660,43	450,59	77,54
	Promedio PISA Matemática	110.530	211,80	713,98	416,60	71,49

Nota: Fuente: Elaboración propia.

La Tabla 4.5 describe la distribución de los valores observados y las predicciones para PISA 2012 basadas en las funciones de alineamiento 2006 y 2009, para estudiantes de acuerdo con el máximo nivel educacional de ambos padres (educación básica incompleta, básica completa, educación media, formación terciaria no universitaria o formación universitaria y de postgrado)²⁴. A partir de esta tabla, observamos que, para todos los grupos de educación de los padres, la estimación de PISA 2012, usando el alineamiento 2006, pronostica mejor el resultado observado que aquella que se basa en la función de alineamiento 2009. Esta última, sobrestima el valor observado para estudiantes de todos los grupos en más de 26 puntos en promedio. En cambio, la estimación de PISA 2012 que se basa en la función de alineamiento 2006, predice el resultado observado con menos de 4 puntos de diferencia para cada nivel educativo, salvo para la categoría de "Básica incompleta" (grupo que corresponde al 3% de la población), donde la diferencia es de 13 puntos.

24 Reportado por el estudiante de II medio.

Tabla 4.5 Resultados observados en PISA 2012 de la muestra de grado 2012 y resultados estimados usando las funciones de alineamiento 2006 y 2009 según máximo nivel educacional de los padres

Máximo Nivel Educacional Padres	Estimación/Puntaje	N	Mínimo	Máximo	Media	Error Est.	Desv. típ.
Básica incompleta	Estimación PISA al 2006	5.461	273,24	580,22	374,51	5,18	52,42
	Estimación PISA al 2009	5.461	266,91	612,51	397,81	6,03	60,31
	Promedio PISA Matemática	6.439	224,96	546,51	361,67	6,36	60,76
	N válido (según lista)	5.461					
Básica Completa	Estimación PISA al 2006	31.484	272,37	633,51	394,88	4,31	59,06
	Estimación PISA al 2009	31.484	265,62	660,43	420,83	4,84	66,83
	Promedio PISA Matemática	37.052	217,57	634,53	390,99	4,03	65,11
	N válido (según lista)	31.484					
Media (H/C o TP)	Estimación PISA al 2006	80.067	268,91	628,44	418,74	3,32	65,36
	Estimación PISA al 2009	80.067	260,47	655,99	447,10	3,61	71,88
	Promedio PISA Matemática	92.222	211,80	674,80	421,40	3,20	68,60
	N válido (según lista)	80.067					
Formación Terciaria no Universitaria	Estimación PISA al 2006	23.271	274,61	613,30	432,57	4,37	67,99
	Estimación PISA al 2009	23.271	268,93	642,58	462,04	4,70	73,58
	Promedio PISA Matemática	26.659	218,58	673,32	431,30	3,52	67,40
	N válido (según lista)	23.271					
Universitaria o Post-grado	Estimación PISA al 2006	47.077	282,01	633,51	474,30	4,02	75,72
	Estimación PISA al 2009	47.077	279,81	660,43	505,54	4,17	79,05
	Promedio PISA Matemática	52.518	231,35	734,47	476,35	3,71	81,77
	N válido (según lista)	47.077					

En conclusión, los resultados promedios observados en PISA 2012 no son significativamente distintos a los predichos con la función de alineamiento 2006 y sí son significativamente distintos de los valores predichos con la función de alineamiento 2009. Asimismo, observamos que las diferencias entre ambas predicciones son consistentes para hombres y mujeres, y para estudiantes cuyos padres tienen distintos niveles educativos. Esto indica que se puede concluir, por un lado, que la función de alineamiento que mejor pronostica el resultado PISA para los estudiantes de 2012 es la basada en la muestra 2006, y que por otro lado, por los bajos errores de predicción observados la naturaleza del alineamiento entre PISA y Simce es la misma para las cohortes 2006 y 2012.

Lo anterior permite afirmar que la hipótesis de no cumplimiento de supuestos de TRI de las pruebas PISA o Simce para explicar la discordancia entre las tablas de alineamiento no se sustenta, ya que se puede predecir con 95% de confianza los resultados PISA de una muestra independiente conociendo solamente sus resultados Simce. Asimismo se comprueba que la predicción funciona con la misma calidad estadística para distintos grupos de estudiantes. Observamos también que hay evidencia para afirmar que los resultados Simce de la muestra de grado 2009 no permiten predecir adecuadamente

los resultados PISA. En la siguiente sección se ahondará en el estudio de este comportamiento diferente de las muestras de grados 2006, 2009 y 2012.

4.4 Análisis de Comportamiento diferenciales de las muestras

Hemos observado que para la población y distintos grupos de estudiantes la función de predicción de resultados PISA basada en el alineamiento 2006 entrega resultados más cercanos y no significativamente diferentes a los efectivamente obtenidos en 2012, en tanto que los resultados obtenidos por la predicción basada en el alineamiento 2009 entrega resultados significativamente distintos a los realmente obtenidos. Se puede concluir que la muestra de grado 2012 se comporta de manera más similar a la muestra 2006 que a la 2009.

La Tabla 4.6 describe los resultados Simce y PISA Matemática para 2006, 2009 y 2012 para las muestras de grado de PISA. Estos resultados se obtienen usando los ponderadores o pesos finales de muestreo de cada muestra de grado.

Tabla 4.6 *Resultados PISA y Simce para las muestras de grado 2006, 2009 y 2012*

	N	Mínimo	Máximo	Media	Error Est.	Desv Est.
Muestra 2006						
Simce Matemática	231.152	97,35	426,58	253,61	3,52	64,65
PISA Matemática	254.548	170,74	690,99	416,65	3,67	80,53
Muestra 2009						
Simce Matemática	224.928	111,45	425,64	236,65	2,55	58,98
PISA Matemática	249.472	180,64	663,04	423,53	3,16	77,26
Muestra 2012						
Simce Matemática	194.803	116,33	422,05	261,78	2,62	63,22
PISA Matemática	223.724	211,8	734,47	426,97	2,87	77,68

Nota: Fuente: Elaboración propia.

Los resultados de la muestra 2006 para Simce son iguales a los obtenidos en la aplicación censal del mismo año (promedio nacional de 252 puntos). Sin embargo, los resultados de la muestra 2009²⁵ son muy inferiores y significativamente distintos a los resultados censales de segundo medio para los años 2008 y 2010²⁶ (promedios nacionales 250 y 256 respectivamente). La muestra 2012 vuelve a recuperar los resultados censales (promedio nacional de 265). Por otro lado, las muestras de grado 2006, 2009 y 2012 no presentan mayores variaciones con los resultados de PISA de las muestras de quince años²⁷ (resultados promedios 411, 421 y 423 respectivamente²⁸). Observamos entonces que mientras los

²⁵ La que, recordemos, tiene las mismas características de calidad para representar a la población de segundo medio que la muestra 2006.

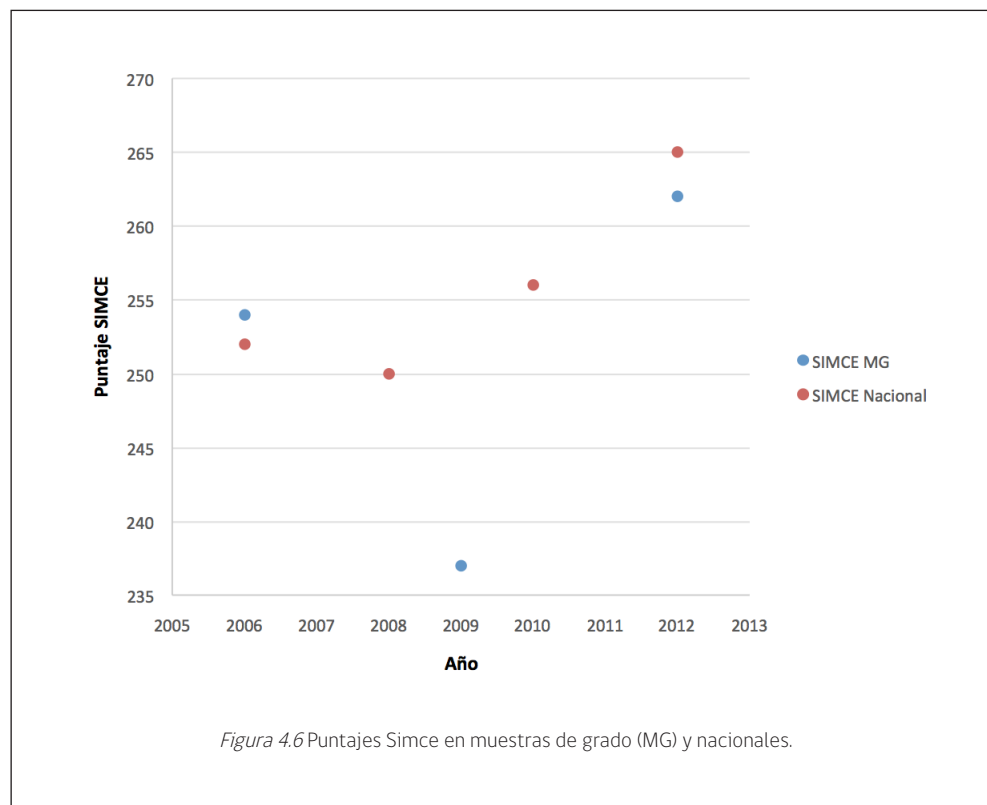
²⁶ Mencionamos los resultados censales anteriores ya que a la cohorte de la muestra de grado 2009 no le tocaba rendir Simce. La muestra de grado rindió la prueba Simce 2008 dentro de la aplicación de la prueba PISA 2009.

²⁷ Los estudiantes de quince años de la muestra de grado componen la muestra regular de PISA. Se espera que la muestra de grado obtenga resultados superiores a la muestra regular ya que en esta última participan estudiantes de niveles educativos inferiores que tienen quince años, y que por ende presentan retraso y no han sido expuestos a los contenidos y enseñanza de segundo medio.

²⁸ Los promedios de las muestras de grado no son significativamente distintos de los promedios de la muestra de quince años.

resultados PISA muestran un aumento consistente y sistemático entre 2006 y 2012²⁹, no acontece lo mismo para la prueba Simce, donde la muestra 2009³⁰ presenta un retroceso importante de 15 puntos con respecto de la cohorte 2006, de 13 puntos con respecto de la de 2008 y de 19 puntos con respecto de la de 2010³¹.

Todas las diferencias de resultados para la muestra de grado 2009 con los resultados censales 2008 y 2010 son significativas al 95%, y se reflejan en el gráfico de la Figura 4.6.



Nota: Fuente: Elaboración propia.

La muestra de grado 2009 respondió la prueba Simce 2008 en las mismas condiciones y bajos los mismos protocolos de control de calidad que la prueba PISA 2009. Las diferencias en resultado, por lo tanto, no son fruto de diferencias en las condiciones de aplicación ni en las características constitutivas de la muestra. Se observa entonces que estos estudiantes demuestran menos dominio de las competencias que mide la prueba Simce de Matemática que sus pares de los años 2006, 2008, 2010 y 2012. La diferencia de la cohorte 2009 con las anteriores radica principalmente en que en ese año el sistema escolar estaba al tanto que los alumnos que cursaban II medio no rendirían la prueba Simce de dicho grado, a diferencia de las otras cohortes.

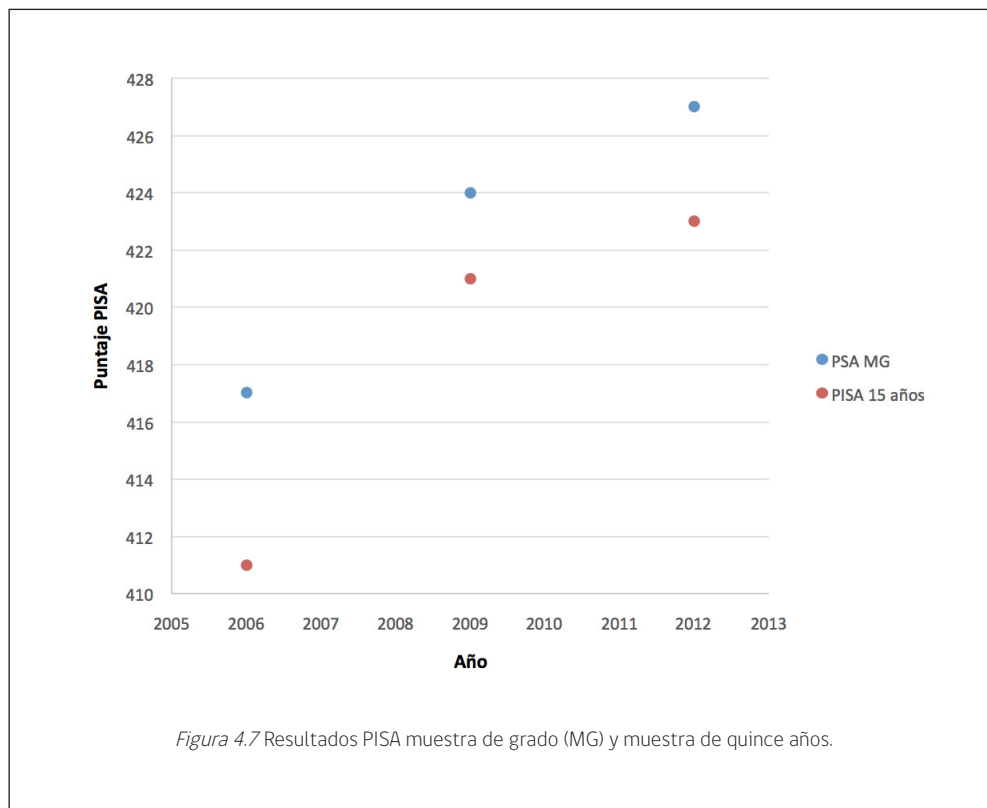
Las cohortes 2006, 2008, 2010 y 2012 habrían estado mejor preparadas para rendir la prueba Simce que la de 2009. Dicha preparación, sin embargo, no influenciaría o afectaría de igual manera las competencias que mide

²⁹ En similar magnitud a lo que indica la muestra de quince años.

³⁰ Todos estos valores corresponden a diferencias significativas.

³¹ Todos estos valores corresponden a diferencias significativas.

la prueba PISA ya que no se observa un patrón de comportamiento diferente entre las muestras PISA de grado 2006, 2009 y 2012 con respecto de las muestras PISA regulares de quince años. Esto se observa en el gráfico de la Figura 4.7, donde las diferencias entre los puntajes promedios de ambas muestras no son significativas, y donde la muestra de grado 2009 no se aparta del comportamiento de las otras dos muestras, 2006 y 2012.



Nota: Fuente: Elaboración propia.

La muestra de grado 2009 obtuvo 424 puntos PISA, aplicando la función de alineamiento 2006 este puntaje correspondería a 249 puntos Simce, lo que está muy por encima de los 237 que realmente obtuvieron. Este brecha de rendimiento se podría explicar por el tratamiento distinto al que fueron expuestos estos estudiantes; no se los apoyó de la misma manera que a sus compañeros de las cohortes 2006 y 2012.

5. Conclusiones

Este trabajo ha permitido dilucidar que la función de alineamiento más predictiva de los puntajes PISA, para las pruebas Simce y PISA de Matemática, corresponde a la función de alineamiento obtenida para el año 2006. También se ha podido verificar la invarianza de esta función de alineamiento entre cohortes que se sabe que deben rendir la prueba Simce de II medio, tal como lo predice la TRI, con lo que se reafirma que ambas pruebas se enmarcan correctamente en la TRI cada una según su propio modelo psicométrico. Se ha comprobado también que la muestra de grado 2009 obtuvo un resultado Simce significativamente diferente a lo que hubiera predicho su resultado PISA.

La diferencia de puntajes Simce entre las muestras 2006, 2012 y 2009 podría corresponder a un indicador del grado de instrucción adicional en competencias matemáticas que reciben los estudiantes. Sin embargo, debe descontarse a estas diferencias la tendencia de los puntajes. Las pruebas Simce de Matemática variaron aproximadamente 13 puntos³² entre 2006 y 2012, considerando cohortes preparadas, lo que corresponde al 20% de desviación estándar de la prueba Simce. En ese mismo lapso la prueba PISA de Matemática aumentó 10 puntos³³, lo que corresponde al 13% de su desviación estándar. De lo anterior, se podría concluir que el efecto de la instrucción adicional o preocupación del sistema escolar por la instrucción matemática, debido a que rinde prueba Simce, correspondería a aproximadamente 7% de desviación estándar, lo que equivale aproximadamente a 5 puntos en dicha prueba.

Además, se ha constatado que los resultados de pruebas PISA no se ven mayormente afectados por esta instrucción adicional, por lo que esta prueba internacional permite ser utilizada como referente de comparación para las pruebas Simce.

También ha surgido de este estudio la necesidad de analizar en más detalle la construcción de las pruebas Simce de Matemática, de manera de determinar aquellas áreas o contenidos que han permitido el alza de los puntajes. Se pueden realizar estudios psicométricos comparativos entre las pruebas PISA y Simce para determinar aquellas áreas o competencias que psicométricamente se comporten de manera diferenciada en ambas pruebas. Este análisis se puede complementar realizando un estudio en el marco de la teoría clásica de análisis de las pruebas, observando los patrones y porcentajes de respuestas correctas por contenidos o procesos cognitivos, que puedan considerarse similares en ambas pruebas.

³² Tomando los resultados nacionales.

³³ Para las muestras de grado.

Referencias

- Agencia de Calidad de la Educación. (2013) *Informe Técnico Simce 2012*. Santiago de Chile: autor.
- Kolen, M. y Brennan, R. (2004). *Test Equating, scaling, and linking: methods and practices*, segunda edición. Nueva York: Springer. pp. 548.
- Matus, C., Guzmán V., Stevenson M., y Valencia M. (2012). Alineamiento de las puntuaciones Simce 2008 y PISA 2009 en muestras de estudiantes de 2º Medio. Lectura y Matemática. *Evidencias para Políticas Públicas en Educación: Selección de Investigaciones. Concurso Extraordinario FONIDE-PISA*. Santiago de Chile: Centro de Estudios Mineduc.
- Fernández C. (2009). Distribución de puntajes Simce en PISA y Simce 2006, en perspectiva comparada. *Nuevos análisis y perspectivas sobre los resultados en PISA 2006*. Santiago de Chile: Unidad de Currículum y Evaluación.
- OECD. (2014). *PISA 2012, Technical Report*. Autor.
- OECD. (2012). *PISA 2009 Technical Report*. Autor. Rescatado de: <http://dx.doi.org/10.1787/9789264167872-en>
- OECD. (2009). *PISA Data Analysis Manual SPSS*, segunda edición. Autor. Rescatado de: <http://browse.oecdbookshop.org/oecd/pdfs/free/9809031e.pdf>

Anexo

1. Resultado de ajuste por mínimos cuadrados ordinarios

1.1 Simce-PISA Matemática 2006

Se consideraron los siguientes tramos para el ajuste:

Tramo = 1 para valores Simce < 185.

Tramo = 2 para valores Simce \geq 185 and Simce < 351.

Tramo =3 para valores Simce \geq 351.

Estimación curvilínea: resultados de los ajustes.

Tabla A.1 *Tabla de descripción. Resumen del modelo*

Nombre del Modelo		MOD_4
Variable dependiente	1	PISA_MAT_2006
Ecuación	1	Lineal
	2	Cuadrático
	3	Cúbico
Variable independiente		Simce
Constante		Incluidos
Variable cuyos valores etiquetan las observaciones en los gráficos		Sin especificar
Tolerancia para la entrada de términos en ecuaciones		,0001

Tabla A.2 *Resumen del procesamiento de los casos*

Tramo	N
1,00 Total de casos	74
Casos excluidos ^a	0
Casos pronosticados	0
Casos creados nuevos	0
2,00 Total de casos	166
Casos excluidos ^a	0
Casos pronosticados	0
Casos creados nuevos	0
3,00 Total de casos	76
Casos excluidos ^a	0
Casos pronosticados	0
Casos creados nuevos	0

Nota: ^a Los casos con un valor perdido en cualquier variable se excluyen del análisis.

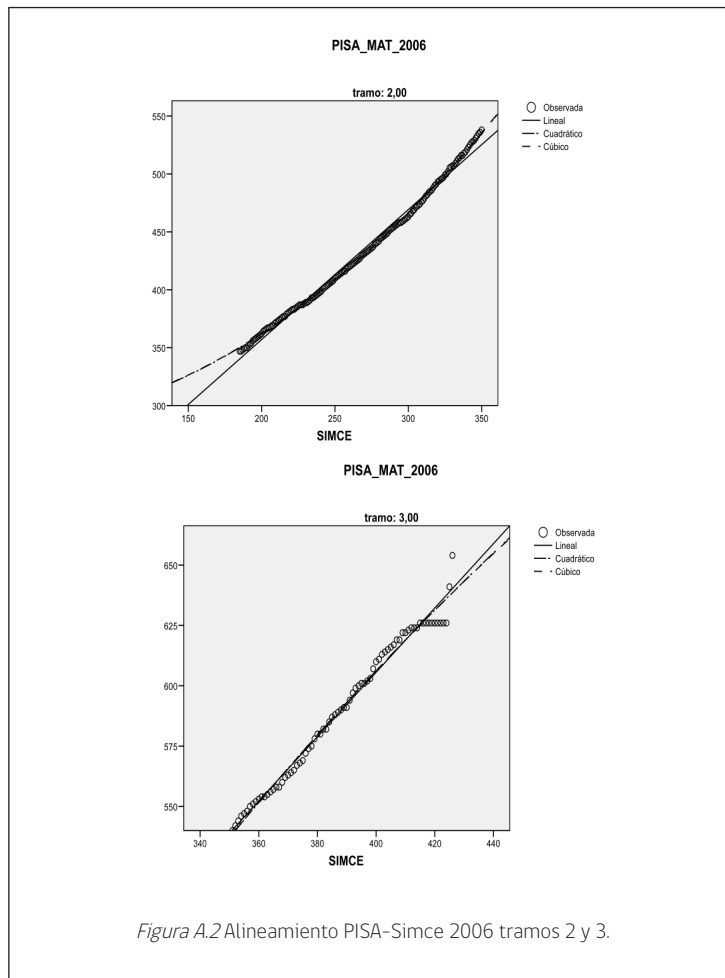
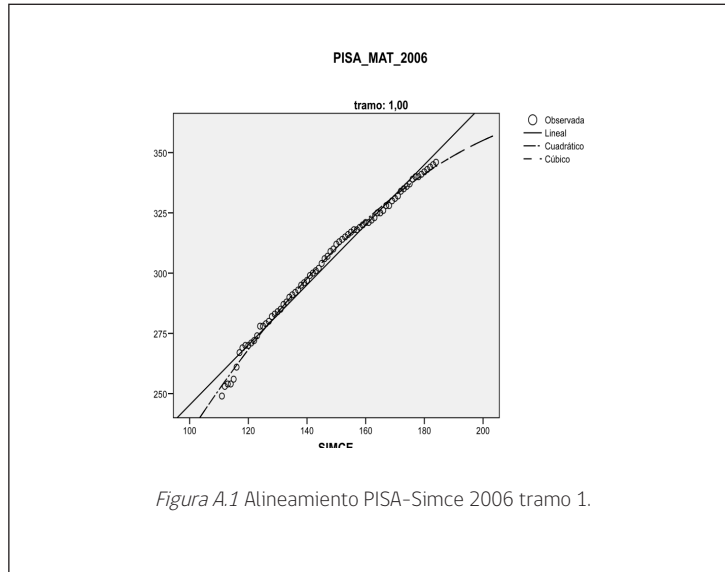
Tabla A.3 *Resumen del procesamiento de las variables*

Tramo			Variables	
			Dependiente	Independiente
			PISA_MAT_2006	SIMCE
1,00	Número de valores positivos		74	74
	Número de cero		0	0
	Número de valores negativos	Perdidos definidos por el usuario	0	0
	Número de valores perdidos	Perdidos del sistema	0	0
2,00	Número de valores positivos		166	166
	Número de cero		0	0
	Número de valores negativos	Perdidos definidos por el usuario	0	0
	Número de valores perdidos	Perdidos del sistema	0	0
3,00	Número de valores positivos		76	76
	Número de cero		0	0
	Número de valores negativos	Perdidos definidos por el usuario	0	0
	Número de valores perdidos	Perdidos del sistema	0	0

Tabla A.4 *Resumen del modelo y estimaciones de los parámetros*

Tramo	Ecuación	Resumen del modelo					Estimaciones de los parámetros			
		R cuadrado	F	gl1	gl2	Sig.	Constante	b1	b2	b3
1,00	Lineal	,987	5.319,007	1	72	,000	120,732	1,246		
	Cuadrático	,996	9.495,106	2	71	,000	-16,528	3,147	-,006	
	Cúbico	,996	9.495,106	2	71	,000	-16,528	3,147	-,006	,000
2,00	Lineal	,992	21.067,221	1	164	,000	132,700	1,122		
	Cuadrático	,999	105.266,000	2	163	,000	284,406	-,050	,002	
	Cúbico	,999	108.243,000	2	163	,000	281,508	,000	,002	3,88E-007
3,00	Lineal	,985	4.857,300	1	74	,000	72,284	1,333		
	Cuadrático	,986	2.510,895	2	73	,000	-195,959	2,719	-,002	
	Cúbico	,986	2.519,360	2	73	,000	-112,910	2,051	,000	-1,6E-006

Nota: Variable dependiente: PISA_MAT_2006. La variable independiente es Simce.



2. Simce- PISA Matemática 2009

Tabla A.5 *Descripción del modelo*

Nombre del modelo		MOD_5
Variable dependiente	1	PISA_MAT_2009
Ecuación	1	Lienal
	2	Cuadrático
	3	Cúbico
Variable independiente		SIMCE
Constante		Incluidos
Variable cuyos valores etiquetan las observaciones en los gráficos		
Tolerancia para la entrada de términos en ecuaciones		Sin especificar
		,001

Tabla A.6 *Resumen del procesamiento de los casos*

Tramo	N
1,00 Total de casos	74
Casos excluidos ^a	0
Casos pronosticados	0
Casos creados nuevos	0
2,00 Total de casos	166
Casos excluidos ^a	0
Casos pronosticados	0
Casos creados nuevos	0
3,00 Total de casos	76
Casos excluidos ^a	0
Casos pronosticados	0
Casos creados nuevos	0

Tabla A.7 Resumen del procesamiento de las variables

Tramo					Variables		
					Dependiente	Independiente	
					PISA_ MAT_2009	SIMCE	
1,00	Número de valores positivos			74	74		
	Número de cero			0	0		
	Número de valores negativos	Perdidos definidos por el usuario		0	0		
	Número de valores perdidos	Perdidos del sistema		0	0		
2,00	Número de valores positivos			166	166		
	Número de cero			0	0		
	Número de valores negativos	Perdidos definidos por el usuario		0	0		
	Número de valores perdidos	Perdidos del sistema		0	0		
3,00	Número de valores positivos			76	76		
	Número de cero			0	0		
	Número de valores negativos	Perdidos definidos por el usuario		0	0		
	Número de valores perdidos	Perdidos del sistema		0	0		

Tabla A.8 Resumen del modelo y estimaciones de los parámetros

Tramo	Ecuación	Resumen del modelo					Estimaciones de los parámetros			
		R cuadrado	F	gl1	gl2	Sig.	Constante	b1	b2	b3
1,00	Lineal	,903	673,602	1	72	,000	65,793	1,656		
	Cuadrático	,944	600,709	2	71	,000	-326,510	7,089	-,018	
	Cúbico	,944	600,709	2	71	,000	-326,510	7,089	-,018	,000
2,00	Lineal	,995	30.076,500	1	164	,000	143,102	1,196		
	Cuadrático	,997	24.210,969	2	163	,000	231,233	,515	,001	
	Cúbico	,997	24.210,969	2	163	,000	231,233	,515	,001	,000
3,00	Lineal	,988	5,968,786	1	74	,000	147,500	1,333		
	Cuadrático	,989	3.340,025	2	73	,000	-209,768	2,719	-,002	
	Cúbico	,989	3.366,582	2	73	,000	-97,854	2,051	,000	-2,1E-006

Nota: Variable dependiente: PISA_MAT_2009. La variable independiente es Simce.

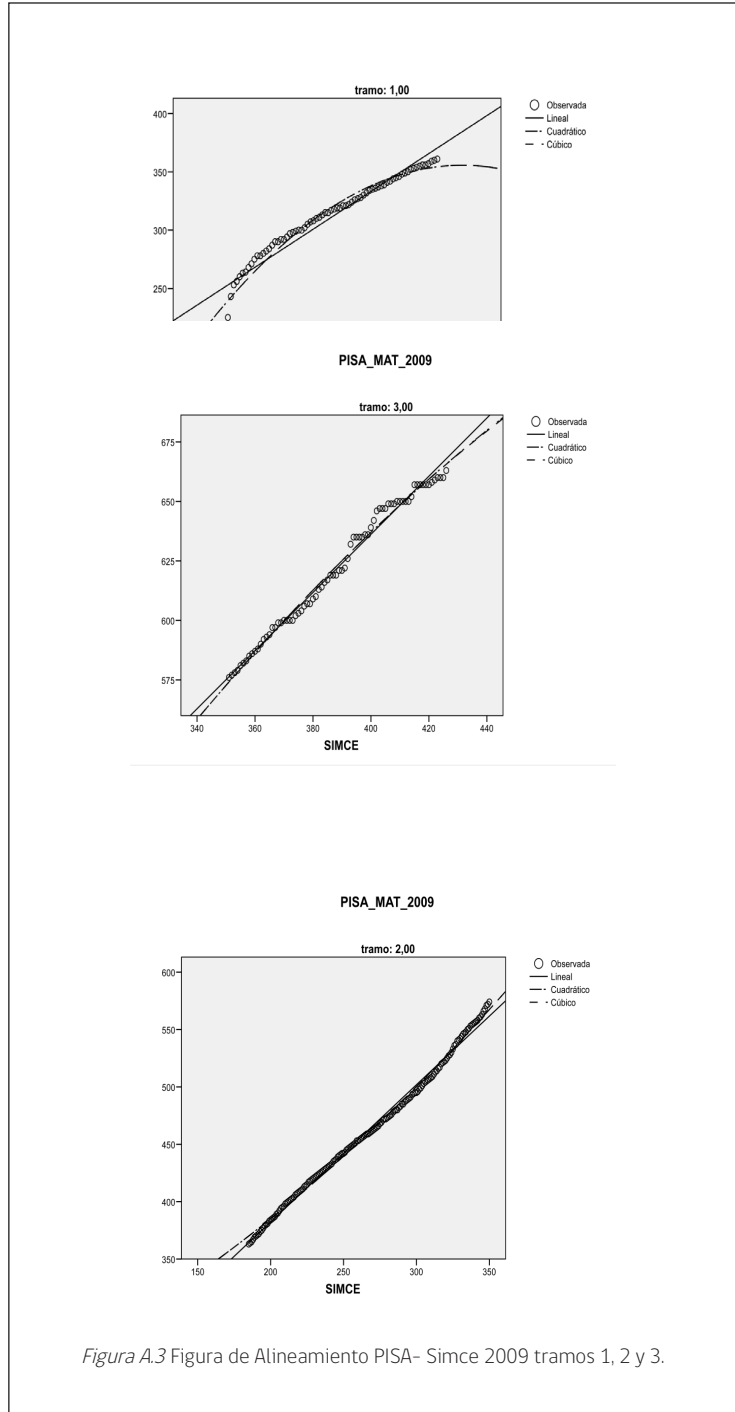


Figura A.3 Figura de Alineamiento PISA- Simce 2009 tramos 1, 2 y 3.

3. Ajustes en rango completo

Tabla A.9 *Resultados alineamiento PISA-Simce 2009*

Ecuación	Resumen del modelo					Estimaciones de los parámetros			
	R cuadrado	F	gl1	gl2	Sig.	Constante	b1	b2	b3
Lineal	0,99	61.759,53	1,00	314,00	0,00	120,62	1,28		
Cuadrático	0,99	30.797,98	2,00	313,00	0,00	118,97	1,30	-2,59 10 ⁻⁵	
Cúbico	1,00	27.772,38	3,00	312,00	0,00	9,01	2,74	-0,01	7,17 10 ⁻⁷

Nota: La variable independiente es Simce.

Tabla A.10 *Resultados alineamiento PISA-Simce 2006*

Ecuación	Resumen del modelo					Estimaciones de los parámetros			
	R cuadrado	F	gl1	gl2	Sig.	Constante	b1	b2	b3
Lineal	0,99	49.209,32	1,00	314,00	0,00	123,71	1,18		
Cuadrático	1,00	70.631,20	2,00	313,00	0,00	183,18	0,68	9,3 10 ⁻⁴	
Cúbico	1,00	64.098,24	3,00	312,00	0,00	115,66	1,57	-2,61 10 ⁻³	4,4 10 ⁻⁶

Nota: La variable independiente es Simce.

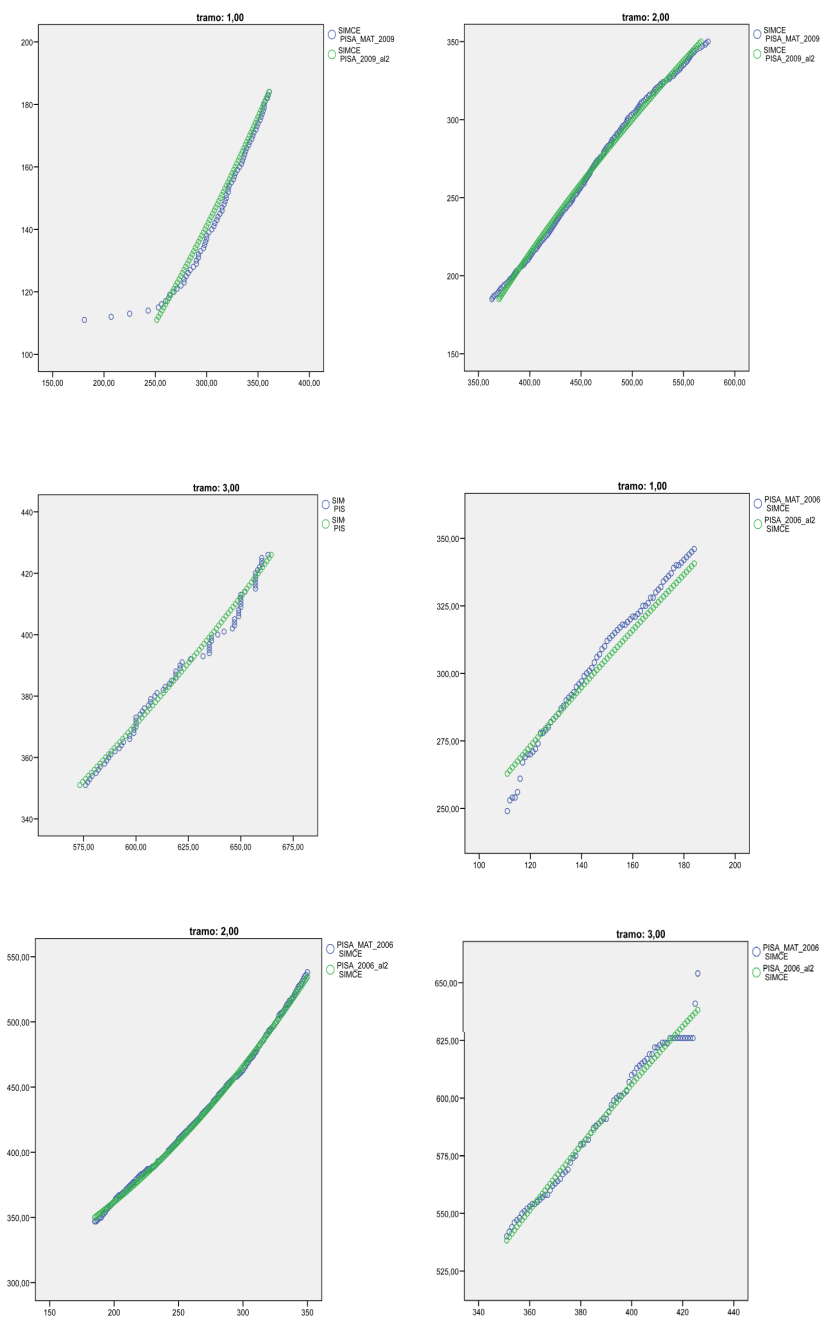


Figura A.4 Gráficos de alineamiento por tramos.