

LAS HORAS DE INGLÉS

En noviembre 2012, los alumnos de III medio participaron en la Prueba Simce de Inglés que comprendía una parte de comprensión lectora y otra de comprensión auditiva. Si el puntaje promedio alcanzado en ambas habilidades es igual o superior a 70 puntos -de un máximo estandarizado de 100-, el alumno se considera certificado en su nivel de inglés para un nivel A2 o B1 en la escala de niveles del Marco Común Europeo para el Aprendizaje de Idiomas¹.

Este año, 18% de los alumnos que rindieron la prueba se certificó. No es sorpresa que el hecho de que un alumno se certifique o no, esté asociado a características socioeconómicas de su establecimiento o su familia. Sin embargo, también hay otros factores que parecieran estar incidiendo en un mejor nivel de inglés. Entre ellos, la mayor exposición al idioma inglés aparece como un factor asociado positivamente a la probabilidad de certificarse. Hay un mayor porcentaje de alumnos certificados entre los que tienen más horas de clases de inglés, las comenzaron desde más temprano en su vida escolar y/o cuyo profesor actual habla la mayor parte de la clase en el idioma que está enseñando².

Como ocurre en muchos casos, es probable que estas variables estén correlacionadas con otras que no son observadas y que son las que efectivamente están incidiendo en los resultados. Más horas de clases de inglés y/o desde grados menores, pueden ser recursos necesarios, pero en ningún caso son suficientes para mejorar el nivel de inglés de los alumnos. El que exista una asociación entre más horas de inglés y una mayor probabilidad de certificación puede deberse a que más horas adicionales de inglés están correlacionadas con un mayor compromiso de los establecimientos, y también de los padres o apoderados, con el aprendizaje de este idioma. No por instalar en el currículo de un establecimiento más horas de inglés o desde más temprano, los alumnos necesariamente van a aprender más. La relación no es causal y, sin embargo, es difícil que se logre el aprendizaje de un idioma distinto al materno si la exposición a este es muy baja.

Debido a la asociación positiva, y con el resguardo de que no es una relación causal, este Apunte revisa las horas semanales de inglés que tienen los alumnos de III medio desde dos enfoques complementarios. Primero, se aborda la relación entre más horas de inglés³ y mayor probabilidad de certificación con las preguntas ¿Es el efecto positivo de tener más horas de inglés sobre la probabilidad de certificarse según su puntaje en la Prueba Simce de Inglés III medio 2012 un resultado robusto? ¿Esta relación persiste a través de la medición en distintas poblaciones y con diferentes especificaciones? Posteriormente, y dado que las horas de inglés aparecen siempre con una asociación positiva aunque con distintas magnitudes, se investiga la distribución de este recurso con las preguntas ¿Cómo están repartidas estas horas de inglés en la población? ¿Además de los establecimientos particulares pagados bilingües, hay más horas de inglés en otro tipo de establecimientos?

¹ Según los niveles del Marco Común Europeo para el Aprendizaje de Idiomas (CEFR), el nivel A2 (elemental) corresponde a la capacidad de comprender los puntos principales de información clara estándar sobre temas familiares que aparecen habitualmente en el trabajo, el establecimiento, el tiempo libre u otros y, en general, resolver la mayoría de las situaciones que pueden generarse al viajar a un lugar en el cual se habla el idioma. El nivel B1 (intermedio bajo) significa que la persona puede comprender oraciones y expresiones utilizadas frecuentemente, relacionadas con áreas de relevancia inmediata, por ejemplo, información personal y familiar muy básica, compras, geografía local y trabajo.

² La información para la construcción de estas variables es recogida en respuestas autorreportadas a cuestionarios aplicados a los alumnos que participaron en el Simce de Inglés III medio 2012 (Ver Lista de Referencias).

³ Las horas corresponden a horas pedagógicas de 45 minutos.

METODOLOGÍA

La variable dependiente es el resultado de certificación del estudiante, por lo que es dicotómica: se certifica/no se certifica. Como variables explicativas, se escogieron tres variables que en análisis preliminares aparecían con un efecto significativo y positivo sobre la proporción de alumnos certificados: horas de clases de inglés (en adelante "horas"), que el profesor hablara principalmente en inglés durante la clase (en adelante "idioma") y en qué curso el alumno tuvo su primera clase de inglés (en adelante "inicio"). En este último caso, mientras más temprano es la primera clase, mayor es la probabilidad del alumno de certificarse. La variable que aparece más robusta a todo tipo de especificaciones y para todas las submuestras en que fue medida, fue la variable "horas".

La otra variable robusta fue la de "inicio", pero la construcción de esta variable supone que una vez que un alumno tuvo su primera clase de inglés nunca más dejó de tenerlas. Además, el seguimiento que se le puede hacer a la primera clase de inglés es menos directo, ya que depende del alumno más que del establecimiento. En cambio, el reporte de la variable "horas" es más simple y su construcción se basa en menos supuestos⁴.

Como la variable dependiente es binaria, se empleó un modelo probit para obtener el signo y magnitud de los regresores. Las regresiones probit se llevaron a cabo con la muestra completa de alumnos y por submuestras que se detallarán más adelante. En todas las regresiones se controló por variables socioeconómicas del establecimiento y de las familias, las que obtuvieron coeficientes con los signos esperados. La heteroscedasticidad se corrigió calculando la significancia con errores robustos y haciendo *clusters* (conglomerados) por establecimiento. Los datos descriptivos de todas las variables empleadas se encuentran en las Tablas 1.1 a 1.4 (para Tablas y Glosario, ver Anexos).

RESULTADOS

1. Hay una relación positiva entre certificación y más horas de inglés. El coeficiente de "horas" aparece como positivo y significativo aun controlando por características socioeconómicas del alumno y del establecimiento.

En la primera fila de la Tabla 2 (ver Anexos) se presenta el resultado de la muestra completa. El efecto marginal es calculado con el promedio de "horas" de la muestra con datos completos (0,29). Se incluyeron como dependientes las otras dos variables que también tuvieron una asociación positiva con la probabilidad de certificación, "inicio" e "idioma". A nivel de alumnos, los controles empleados fueron el género y dos características socioeconómicas: los años de educación de la madre y si tiene o no computador en su casa. Además se controló por la modalidad del establecimiento (si es técnico-profesional o científico-humanista), si el establecimiento es de dependencia municipal y además si es o no rural. Adicionalmente se controló por el tamaño del establecimiento, a través de la variable aproximante del tamaño del nivel III medio⁵.

El efecto marginal en toda la población es de 6,4 puntos porcentuales, lo que corresponde a un 17% de

⁴ Para construir la variable "horas", se agrupó la respuesta a la pregunta hecha a los estudiantes sobre cuántas horas de inglés tenían a la semana. Se generó una variable dicotómica: el valor 0 lo reciben los alumnos que declaran tener 3 o menos horas de inglés y el valor 1 los que declaran tener 4 o más horas de inglés. La variable "inicio" consideró la respuesta de los alumnos a la pregunta sobre cuándo tuvieron su primera clase de inglés. Luego se imputaron años de clases con el supuesto que estas fueron ininterrumpidas desde el año en que se iniciaron. La variable "idioma" también agrupó las respuestas de los alumnos para genera una variable dicotómica: 0 si el profesor habla durante toda o la mayor parte de la clase en español y 1 si habla toda o la mayor parte en inglés.

⁵ Por tamaño se entiende el número total de alumnos en III medio.

la desviación estándar de la variable de certificación. Como resultado de un análisis inicial sugiere que sería una variable interesante de investigar en mayor profundidad. En particular, porque no parece ser un efecto homogéneo en toda la población. Esto último se revisa a continuación.

2. La relación positiva y significativa entre certificación y más horas de inglés persiste en todas las submuestras. Sin embargo, y considerando los intervalos de confianza calculados con errores robustos⁶, hay alguna evidencia de que el efecto de horas es mayor en las submuestras que cuentan con mayores recursos⁷.

El efecto de horas sobre la probabilidad de certificarse, aparece como positivo y significativo en las diferentes submuestras consideradas. En general, se aprecia una mayor magnitud donde hay más recursos, insinuando complementariedad de los insumos. En cambio, no hay mayor diferencia en las muestras por género.

Si se considera una no superposición de los intervalos de confianza como un criterio inicial al comparar la magnitud de los coeficientes, se aprecia que el efecto "horas" sobre la probabilidad de certificarse es menor al promedio en los establecimientos de dependencia municipal, similar al promedio en los particulares subvencionados, y mayor en los particulares pagados. Podemos suponer una gradiente positiva del efecto "horas", en función de otros recursos disponibles.

Una relación positiva entre recursos y el efecto que las horas tienen sobre la probabilidad de certificarse, también se observa al separar la muestra por establecimientos de distinto grupo socioeconómico. En este caso son los establecimientos del grupo D (el segundo con más recursos) y grupo A (el de menores recursos), los que presentan la mayor y menor magnitud, respectivamente, del efecto "horas"⁸.

Asimismo, se observa mayor (resp. menor) magnitud en los establecimientos grandes (resp. pequeños), urbanos (resp. rurales) y de modalidad científico-humanista (resp. técnico-profesional). Esta evidencia inicial sugiere que habría un cierto grado de complementariedad entre una mayor cantidad de horas de inglés y otros recursos con que cuenta un establecimiento.

A nivel individual, la muestra se separó por género y por el nivel educacional de la madre de los alumnos.⁹ No hay diferencias en los coeficientes de las submuestras por género, pero sí las hay en aquellas por nivel educacional de la madre. El efecto de "horas" es de mayor magnitud en los alumnos cuyas madres alcanzan más altos niveles de educación que enseñanza media completa.

Por lo tanto, hay evidencia inicial de que el recurso de horas de inglés se complementa con otros recursos disponibles tanto en el hogar como en el establecimiento.

⁶ La tabla muestra los intervalos de confianza contruados con errores estándar calculados por *clusters* (conglomerados) de establecimientos.

⁷ A nivel individual sería un recurso sociocultural como el nivel de educación de la madre. A nivel de establecimiento el recurso es socioeconómico (grupo socioeconómico al que pertenece el establecimiento, por ejemplo). También se puede hacer el supuesto de que los establecimientos urbanos cuentan con más recursos socioeconómicos enfocados al aprendizaje del idioma que los rurales, o que los de modalidad científico humanista cuentan con más que los de modalidad técnico profesional.

⁸ Los establecimientos se clasifican en cinco grupos socioeconómicos: Bajo (A), Medio-Bajo (B), Medio (C), Medio-Alto (D) y Alto (E). Para detalles sobre la construcción de estos grupos ver el siguiente documento: Departamento de Pruebas Nacionales, División de Evaluación de Logros de Aprendizaje, Agencia de Calidad de la Educación (2012). *Metodología de construcción de Grupos Socioeconómicos Pruebas SIMCE 2012*. Santiago. Autor. Recuperado de <http://www.agenciaeducacion.cl/wp-content/uploads/2013/02/Metodologia-de-Construccion-de-Grupos-Socioeconomicos-SIMCE-2012.pdf>.

⁹ La agrupación del nivel de escolaridad de las madres se hizo entre las que tienen más de 12 años de educación (más de educación media completa), y las que tienen igual cantidad o menos de 12 años de educación (educación media completa o menos).

De comprobarse esta complementariedad, se plantea un desafío adicional a las políticas públicas en educación. Si se desea mejorar los resultados en la adquisición del idioma inglés de manera equitativa, las políticas deberán más que compensar los menores recursos disponibles en un porcentaje importante de establecimientos¹⁰. Esto también puede incluir propuestas asociadas a recursos culturales complementarios en el hogar o en otras instancias extracurriculares.

3. La distribución de las horas de inglés no se concentra exclusivamente en los establecimientos de mayores recursos. Al clasificar a los establecimientos se aprecia que en todas las categorías hay al menos un 15% donde, de acuerdo a sus alumnos, se imparten 4 o más horas de inglés.

Al clasificar a los establecimientos en cinco grupos socioeconómicos, siendo A el más bajo y E el más alto, se aprecia que el 15% de los establecimientos del grupo A, ofrecen 4 o más horas de inglés a sus alumnos de III medio. Este porcentaje va creciendo, hasta alcanzar el 60% de los establecimientos del grupo E (ver Tabla 4.1 en Anexos).

El 20% de los establecimientos de dependencia municipal ofrecen 4 o más horas de inglés a sus alumnos de III medio. En los establecimientos particulares subvencionados este porcentaje sube levemente a 27%, pero luego sube más hasta alcanzar el 50% de los establecimientos particulares pagados (ver Tabla 4.2 en Anexos).

Entre los establecimientos con modalidad técnico-profesional, el 15% ofrece 4 o más horas de inglés, en cambio, de los con modalidad científico-humanista, el 39% lo hace. La diferencia entre establecimientos urbanos y rurales es menor: un 22% de los rurales y un 32% de los urbanos ofrecen más horas de inglés. También es menor la diferencia entre establecimientos grandes y pequeños, aunque la hay. El 24% de los establecimientos grandes y el 38% de los pequeños ofrecen más horas de inglés (ver Tablas 4.3 a 4.5 en Anexos).

Al clasificar a los alumnos según dos niveles de educación de la madre (hasta 12 años de educación y sobre 12 años), nuevamente se aprecia que el 41% de los hijos de madres con mayor nivel de educación asisten a establecimientos con 4 o más horas de inglés, en cambio solo el 23% de los hijos de madres con menor nivel de educación lo hace (ver Tabla 4.6 en Anexos).

En general, se aprecia que la distribución de las horas de inglés es más equitativa que los recursos con los que se correlaciona¹¹. A pesar de que hay una relación directa entre más horas y otros recursos disponibles en el hogar o el establecimiento, siempre hay un subconjunto de establecimientos o alumnos, no menor al 15%, que accede a tener una inmersión más completa en otro idioma.

Si las horas verdaderamente representan mayores esfuerzos de los establecimientos para que sus alumnos puedan comunicarse en inglés y ampliar el horizonte de sus experiencias de vida, entonces sería recomendable investigar la magnitud del efecto con mayor profundidad.

¹⁰ La Tabla 3 del Anexo muestra la distribución de los alumnos en distintos tipos de establecimientos clasificados por dependencia administrativa y grupo socioeconómico.

¹¹ Las correlaciones entre horas y los otros recursos muestran una asociación positiva entre un 10% y 20%.

LISTA DE REFERENCIAS

Departamento de Estudios de la Calidad de la Educación, División de Estudios, Agencia de Calidad de la Educación. (2012). [Respuestas a Cuestionarios de Contexto Simce de Inglés III medio 2012].

Departamento de Pruebas Nacionales, División de Evaluación de Logros de Aprendizaje, Agencia de Calidad de la Educación (2012). *Metodología de construcción de Grupos Socioeconómicos Pruebas SIMCE 2012*. Santiago: Autor. Recuperado de <http://www.agenciaeducacion.cl/wp-content/uploads/2013/02/Metodologia-de-Construccion-de-Grupos-Socioeconomicos-SIMCE-2012.pdf>

Instituto Cervantes. (2002). *Marco común europeo de referencia para las lenguas: enseñanza, aprendizaje, evaluación*. Madrid: MEC-D-Anaya. [Versión original en inglés: Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Strasbourg: Council of Europe, 2001]. Recuperado de http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Massachusetts: The MIT Press.

Anexo

Glosario aplicado

Modelo probit: Modelo de respuesta binaria en el que la variable dependiente es dicotómica y la función de probabilidad acumulada es normal estándar¹. Dado que el objetivo de este estudio es estimar de manera inicial cómo el pasar de tener 3 o menos horas² a 4 o más horas de inglés afecta la probabilidad de certificarse, se emplea este modelo que permite tomar en cuenta propiedades específicas de los modelos con variables dependientes binarias. La variable dependiente “certificación” toma valores 1 si el alumno se certifica y 0 si no lo hace.

La distribución de la variable latente que subyace a la decisión se supone normal estándar³. La estimación de los parámetros se hace a través del método de máxima verosimilitud.

Heteroscedasticidad: Propiedad de una variable aleatoria que tiene varianza finita no constante. Es muy posible que la varianza del error de la regresión no sea constante para todas las observaciones. Los alumnos de diferentes establecimientos pueden presentar desviaciones de la media distintas, lo que significa que la probabilidad de certificarse de un alumno no es independiente del establecimiento al cual asiste. Esto genera la heteroscedasticidad de la varianza del error de estimación.

Clusters: Análisis por conglomerados o grupos que presentan características homogéneas. Para corregir la heteroscedasticidad, se emplea un método que calcula errores estándar por *clusters* (conglomerados). Se asume que las observaciones están correlacionadas al interior de los colegios, que sería el conglomerado considerado en este caso.

Errores estándar robustos: Método por el cual la matriz de varianza-covarianza de los estimadores es consistente en presencia de distintas formas de heteroscedasticidad. En este trabajo se calcula con errores estándar robustos y no se observa mayor diferencia con las significancias obtenidas calculándolos por *clusters*⁴.

¹ Los modelos de respuesta binaria son de la forma $P(y = 1|x) = G(x\beta)$ donde $x\beta$ corresponde al índice construido con las variables explicativas x y los coeficientes β . El modelo de variable latente subyacente es $y^* = x\beta + e$

De modo que cuando la variable no observada e , que distribuye normal estándar ($e \sim N(0,1)$) toma valores mayores que un umbral, la función indicador toma el valor 1. De lo contrario, toma el valor 0. En el caso de los modelos probit tenemos que:

$$G(x\beta) \equiv \Phi(x\beta) \equiv \int_{-\infty}^{x\beta} \phi(v) dv$$

donde $\phi(v)$ es la función de densidad normal estándar:

$$\phi(v) = (2\pi)^{-1/2} \exp(-v^2/2)$$

² Las horas pedagógicas son de 45 minutos.

³ Esto diferencia los modelos probit de los modelos logit, también de respuesta binaria, pero que suponen una distribución logística de la variable latente.

⁴ Se emplea Stata 12 (2012) como paquete de software estadístico.

Glosario aplicado

Efectos marginales: Efecto que expresa el cambio de la variable dependiente provocado por un cambio unitario en una de las variables independientes, manteniendo el resto constante. A diferencia de los modelos lineales⁵, en los modelos de respuesta binaria el efecto marginal no se obtiene directamente de los coeficientes β y no es constante para todos los valores de x .

En el modelo probit, los efectos marginales se obtienen mediante el siguiente cálculo:

$$\frac{\partial \Phi(x\beta)}{\partial x_j} = \phi(x\beta) \cdot \beta_j$$

Solo el *signo* del efecto se obtiene directamente de los coeficientes β , pero la *magnitud* deber ser calculada posteriormente para valores específicos de x . En general se calcula empleando la media muestral de la variable.

Pseudo R-cuadrado: Medida de bondad de ajuste propuesta por McFadden (1974) que puedes ser aplicada a modelos no lineales como en el caso del modelo probit. Su fórmula es⁶

$$\text{pseudo R} - \text{cuadrado} \equiv R^2 = 1 - \frac{\ln \hat{L}(M_{Full})}{\ln \hat{L}(M_{Intercepts})}$$

Donde \hat{L} es la verosimilitud estimada y $\ln \hat{L}(M_{Full})$ es el logaritmo de la verosimilitud, denominado función soporte o log-verosimilitud (*log likelihood*).

$\ln \hat{L}(M_{Full})$ es la log-verosimilitud del modelo completo, con variables predictivas, y $\ln \hat{L}(M_{Intercepts})$ es la log-verosimilitud del modelo sin variables predictivas, por lo que solo se estima el intercepto. La razón entre las dos log-verosimilitudes sugiere el aporte del modelo completo, por sobre el modelo con intercepto.

Porcentaje predicho: Otra medida de bondad de ajuste que comúnmente se emplea en los modelos probit. En este caso se emplea el cálculo sugerido por Wooldridge (2002). Se computa la probabilidad predicha de que la variable "certificación" de un alumno sea igual a 1 ($y_i = 1$), dado el vector de variables explicativas (x_i). Si $G(x_i, \hat{\beta}) > 0,5$, se predice que $y_i = 1$; si $G(x_i, \hat{\beta}) \leq 0,5$, se predice $y_i = 0$. Se calcula por separado el porcentaje correcto de ceros y unos. El porcentaje predicho corresponde al promedio ponderado de los dos porcentajes calculados, donde los ponderadores son, respectivamente, las fracciones de ceros y unos del total de los resultados.

⁵ En el modelo lineal $G(x\beta) = x\beta$ y, por lo tanto, el efecto marginal de la variable x_j es $\partial y / \partial x_j = \beta_j$, constante para todo valor de x .

⁶ Corresponde a la fórmula original empleada por Stata 12.

Tabla 1.1
Información descriptiva de las variables empleadas

Variable	Tipo	Descripción
Certificado	Dicotómica	1: Nivel B1o A2
		0: Nivel A1 o inferior
Horas de inglés	Dicotómica	1: 4 o más horas a la semana
		0: 3 o menos horas a la semana
Profesor habla inglés en clases	Dicotómica	1: Toda o casi toda la clase en inglés
		0: Toda o casi toda la clase en español
Años con clases de inglés ^a	Dicotómica	Según información de cuándo tuvo su primera clase
Modalidad	Dicotómica	1: Modalidad técnico profesional
		0: Modalidad científico humanista
Ruralidad	Dicotómica	1: Escuela rural
		0: Escuela urbana
Tamaño del nivel III medio	Continua	Número de alumnos en el nivel que tienen puntaje en la prueba
Tamaño del nivel III medio ^b	Dicotómica	1: 61 o más alumnos en nivel
		0: 60 o menos alumnos en nivel
Género	Dicotómica	1: Mujer
		0: Hombre
Años de educación de la madre ^c	Continua	Según transformación estándar de datos SIMCE
Educación de la madre	Dicotómica	1: Más de 12 años de educación
		0: 12 o menos años de educación
Alumno tiene computador en casa	Dicotómica	1: Tiene
		0: No tiene

a Los años con clases se basan en el supuesto de que el alumno tuvo clases continuas de inglés desde el año en que declara haber tenido su primera clase.

b 61 alumnos es la mediana de tamaño a nivel de establecimiento.

c La tabla de conversión se basa en la utilizada por la UNESCO. Al nivel máximo que declara haber alcanzado la madre se le imputa un número específico de años. Al grado de Doctor se le imputan 22 años, al de magíster 19; a un titulado de una universidad se le asignan 17 años y a los que tienen educación incompleta en una universidad, 15 años de estudio. Si es titulado en un centro de formación técnica o un instituto profesional, se le asignan 16 años; y 14 años si tiene educación incompleta en alguno de ellos. Los años de educación básica se imputan de manera directa (por ejemplo, a 2.º básico se le imputan 2 años).

Fuente: Elaboración propia.

Tabla 1.2
Información descriptiva de las variables empleadas (a nivel de alumno)

Variable a nivel de alumno	Obs.	Promedio/ Porcentaje	Desv. Est.	Mín.	Máx.
Certificado ^a	182.449	0,18	0,38	0	1
Horas de inglés	183.154	0,28	0,45	0	1
Profesor habla inglés en clases	182.449	0,57	0,47	0	1
Años con clases de inglés ^a	168.563	7,76	2,66	1	12
Modalidad técnico profesional	182.449	0,37	0,48	0	1
Escuela rural	182.449	0,04	0,19	0	1
Tamaño del nivel (continua)	182.449	154,37	121,29	2	843
Tamaño del nivel (dicotómica)	182.449	0,80	0,40	0	1
Género (femenino)	182.449	0,51	0,50	0	1
Años de educación de la madre	148.120	11,56	3,70	0	22
Educación de la madre (dicotómica)	148.120	0,28	0,45	0	1
Alumno tiene computador en casa	179.170	1,68	0,66	1	3

^a La muestra original se reduce a este número al considerar a los alumnos de los cuales se tiene información de otras variables de interés.

Fuente: Elaboración propia.

Tabla 1.3
Información descriptiva de las variables empleadas (a nivel de curso)

Variable a nivel de alumno	Obs.	Promedio/ Porcentaje	Desv. Est.	Mín.	Máx.
Certificado	7.160	0,17	0,27	0	1
Horas de inglés	7.138	0,28	0,45	0	1
Profesor habla inglés en clases	7.138	0,71	0,45	0	1
Establecimiento rural	6.770	0,41	0,49	0	1
Tamaño del nivel (continua)	7.160	0,04	0,20	0	1

Fuente: Elaboración propia.

Tabla 1.4
Información descriptiva de las variables empleadas (a nivel de establecimiento)

Variable a nivel de alumno	Obs.	Promedio/ Porcentaje	Desv. Est.	Mín.	Máx.
Certificado	2.688	0,21	0,28	0	1
Horas de inglés	2.683	0,31	0,46	0	1
Profesor habla inglés en clases	2.683	0,75	0,43	0	1
Modalidad técnico profesional	2.611	0,28	0,45	1	2
Establecimiento rural	2.688	0,06	0,24	0	1

Fuente: Elaboración propia.

Tabla 2
Efecto marginal de pasar de 3 o menos horas de inglés a la semana a 4 o más horas

Efecto marginal de pasar de 3 o menos de inglés a 4 horas o más a la semana							
Submuestras	Efecto Marginal	95% [C.I.]		Promedio horas ^a	Pseudo R-Cuadrado ^b	Porcentaje Predicho ^c	
Muestra Completa	0,064	0,0506	0,0764	0,2860	0,3004	64,68	
Por dependencia ^d	MUN	0,024	0,0056	0,0431	0,2197	0,2631	67,15
	PSUB	0,052	0,0383	0,0655	0,2775	0,1657	62,94
	PPAG	0,049	0,0165	0,0805	0,6092	0,0753	81,17
Por grupo socioeconómico establecimiento ^e	A	0,005	0,0015	0,0089	0,1359	0,1106	67,63
	B	0,014	0,0075	0,0201	0,2201	0,0588	67,38
	C	0,061	0,0366	0,0850	0,3187	0,0503	63,60
	D	0,123	0,0881	0,1584	0,3906	0,0553	56,42
	E	0,045	0,0133	0,0766	0,6023	0,0683	81,43
Por grupo socioeconómico establecimiento ^f	A	0,007	0,0025	0,0107	0,1420	0,0718	86,39
	B	0,015	0,0083	0,0211	0,2187	0,0430	84,27
	C	0,059	0,0345	0,0837	0,3142	0,0374	75,56
	D	0,114	0,0788	0,1488	0,3915	0,0441	59,43
	E	0,041	0,0099	0,0726	0,5888	0,0507	81,53
Por modalidad	Cient. Hum	0,106	0,0845	0,1270	0,3706	0,2081	62,41
	Téc. Prof.	0,017	0,0072	0,0259	0,1367	0,0846	68,51
Por ruralidad	Urbano	0,065	0,0518	0,0787	0,2895	0,2942	64,69
	Rural	0,023	0,0012	0,0454	0,1871	0,4844	65,11
Por tamaño del nivel	Pequeño	0,093	0,0717	0,1152	0,3975	0,2782	63,03
	Grande	0,053	0,0370	0,0700	0,2576	0,2940	65,14
Por educación de la madre	12 años o menos	0,035	0,0253	0,0451	0,2363	0,1555	81,66
	más de 12 años	0,154	0,1230	0,1841	0,4125	0,1559	65,25
Por género	Hombre	0,063	0,0491	0,0773	0,2702	0,2980	62,93
	Mujer	0,065	0,0490	0,0806	0,3005	0,3055	66,34

Nota: Los efectos marginales son calculados con la media de certificación del grupo correspondiente, lo que corresponde al porcentaje de alumnos certificados. Todos los coeficientes son significativos a un nivel de confianza del 99%, calculando la significancia con errores robustos y con *clusters* (conglomerados) por establecimiento, para corregir por heteroscedasticidad.

a El promedio corresponde a la muestra o submuestra con datos completos.

b El pseudo R-cuadrado se define en el Glosario aplicado.

c El cálculo de los valores predichos se describe en el Glosario aplicado.

d La dependencia administrativa corresponde a municipalizada (MUN), particular subvencionada (PSUB) y particular pagada (PPAG).

e La referencia de la clasificación por grupo socioeconómico (GSE) se mencionó en la nota al pie 8. En este caso, en la regresión de cada submuestra se controló por educación de la madre del alumno, considerando que el GSE es asignado a todo el establecimiento.

f En este caso no se controló por educación de la madre del alumno.

Fuente: Elaboración propia.

Tabla 3.1
Distribución de alumnos según grupo socioeconómico (GSE) y dependencia administrativa del establecimiento donde estudian

Alumnos	MUN	PSUB	PPAG	TOTAL
A	30.950	13.125	0	44.075
B	39.106	43.491	0	82.597
C	10.677	39.785	0	50.462
D	1.819	25.208	676	27.703
E	0	905	17.765	18.670
TOTAL	82.552	122.514	18.441	223.507

Nota: Las letras A, B, C, D, E corresponden a los 5 grupos socioeconómicos, de menor a mayor, que se emplean para clasificar a las escuelas cuya caracterización se encuentra en el documento *Metodología de construcción de Grupos Socioeconómicos Pruebas SIMCE 2012* (ver Lista de referencias).

Fuente: Elaboración propia.

Tabla 3.2
Distribución de cursos según grupo socioeconómico (GSE) y dependencia administrativa del establecimiento al que pertenecen

Cursos	MUN	PSUB	PPAG	TOTAL
A	1.155	467	0	1.622
B	1.208	1.338	0	2.546
C	294	1.1154	0	1.448
D	48	746	31	825
E	0	905	17.765	18.670
	2.705	3.737	725	7.167

Fuente: Elaboración propia.

Tabla 3.3
Distribución de establecimientos según grupo socioeconómico (GSE) y dependencia administrativa

GSE	MUN	PSUB	PPAG	TOTAL
A	362	201	0	563
B	293	426	0	719
C	65	558	0	623
D	8	386	20	414
E	0	18	351	369
	728	1589	371	2.688

Fuente: Elaboración propia.

Tabla 4.1
Porcentaje de establecimientos en cada grupo socioeconómico (GSE) según horas de inglés impartidas

GSE	3 horas o menos	4 horas o más
A	0,854	0,146
B	0,781	0,219
C	0,675	0,325
D	0,572	0,428
E	0,413	0,587

Fuente: Elaboración propia.

Tabla 4.2
Porcentaje de establecimientos en cada dependencia administrativa según horas de inglés impartidas

Dependencia	3 horas o menos	4 horas o más
MUN	0,787	0,213
PSUB	0,729	0,271
PPAG	0,403	0,597

Fuente: Elaboración propia.

Tabla 4.3
Porcentaje de establecimientos en cada modalidad según horas de inglés impartidas

Modalidad	3 horas o menos	4 horas o más
TP	0,847	0,153
CH	0,616	0,384

Fuente: Elaboración propia.

Tabla 4.4
Porcentaje de establecimientos urbanos y rurales según horas de inglés impartidas

Ruralidad	3 horas o menos	4 horas o más
RUR	0,784	0,216
URB	0,683	0,317

Fuente: Elaboración propia.

Tabla 4.5
Porcentaje de establecimientos pequeños y grandes según horas de inglés impartidas

Tamaño	3 horas o menos	4 horas o más
< = 60 alum.	0,616	0,384
> 60 alum.	0,762	0,238

Fuente: Elaboración propia.

Tabla 4.6
Porcentaje de alumnos por educación de la madre según horas de inglés impartidas

Educación de la madre	3 horas o menos	4 horas o más
< = 12 años educ.	0,616	0,384
> 12 años educ.	0,591	0,409

Fuente: Elaboración propia.