

Un Estudio Sobre la Calidad Docente en Chile: El Rol del Contexto en Donde Enseña el Profesor y Medidas de Valor Agregado

A Study About Teacher Quality in Chile: The Role of the Context Where Teachers Work and Value-Added Measures

María Verónica Santelices, Pilar Galleguillos, Jorge González y Sandy Taut
Pontificia Universidad Católica de Chile

Este estudio espera contribuir a la discusión sobre calidad docente en Chile a través de dos enfoques: (a) uno basado en estimaciones del valor agregado del profesor o *efecto profesor* en el aprendizaje, explorando el rol de las características del profesor y del contexto en que trabaja (escuela y municipio); y (b) el otro intenta predecir el desempeño de los alumnos, utilizando variables de contexto y características del profesor. Se utilizaron modelos lineales jerárquicos. Los datos provienen del Sistema de Evaluación del Desempeño Profesional Docente, del Sistema de Medición de la Calidad de la Educación y del Sistema Nacional de Información Municipal. En la muestra existe una mayor proporción de escuelas pobres y de mayores ingresos que en la población total. Los resultados muestran inestabilidad en las estimaciones del valor agregado, ajustando distintos modelos que incorporan covariables progresivamente. De las variables de contexto, la escuela es más importante que el municipio en la predicción del desempeño de los alumnos y, a nivel del profesor, solo resulta estadísticamente significativa su participación en perfeccionamientos profesionales, desestimándose su experiencia y formación inicial.

Palabras clave: modelos de valor agregado, calidad docente, efecto profesor, Chile

This paper aims to contribute to the discussion about teacher quality in Chile through two approaches: (a) one based on estimates of the value added by teachers or *teacher effect* on students' learning, gauged by exploring the role of teachers' characteristics and the context in which they work (school and municipality); and (b) another which seeks to predict student performance using context variables and teacher characteristics. Hierarchical linear models were used in both analyses. The data came from Chile's National Teacher Evaluation System, from the Education Quality Measurement System and the National System of Municipal Information. The sample includes a higher proportion of the poorest and richest schools compared to the total population. The results show instability in the estimation of the teacher effect, adjusting different models that incorporate control variables gradually. The analysis of students' scores shows that, among contextual variables, the school is more important than the municipality for the prediction and that, at the teacher level, only participation in professional development programs is statistically significant, with no evidence supporting the value of experience and initial education.

Keywords: value-added models, teacher quality, teacher effect, Chile

El presente artículo tiene como objetivo contribuir a la discusión sobre la definición de calidad docente en Chile a través de estimaciones del valor agregado del profesor, al que llamaremos *efecto profesor*. Las medidas de valor agregado buscan destacar la contribución que hacen los profesores en el aprendizaje de los estudiantes durante un período de tiempo específico, teniendo en cuenta el nivel inicial de aprendizaje de los estudiantes (McCaffrey, Lockwood, Koretz, Louis & Hamilton, 2004).

El Informe McKinsey (Barber & Mourshed, 2007) y el White Paper de la Academia Nacional de Educación norteamericana (Wilson, 2009) señalan que la calidad de la enseñanza es crucial para el aprendizaje de los alumnos. Esta conclusión, si bien presente en la literatura educacional anteriormente, en años más recientes

María Verónica Santelices, Facultad de Educación, Pontificia Universidad Católica de Chile; Pilar Galleguillos, Instituto de Economía, Pontificia Universidad Católica de Chile; Jorge González, Facultad de Matemáticas y Centro de Medición MIDE UC, Pontificia Universidad Católica de Chile; Sandy Taut, Escuela de Psicología y Centro de Medición MIDE UC, Pontificia Universidad Católica de Chile, Santiago, Chile.

Esta investigación fue financiada por el proyecto FONDECYT INICIACIÓN No 11110262, por el proyecto Anillo SOC1107 y por la Pontificia Universidad Católica de Chile, a través de financiamiento interno para la investigación. Los autores agradecen a Irma Coronado, por su asistencia en el estudio.

La correspondencia relativa a este artículo debe ser dirigida a María Verónica Santelices, Facultad de Educación, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna 4860, Macul, Santiago, Chile. E-mail: vsanteli@uc.cl

proviene principalmente de estudios que han utilizado la metodología de valor agregado. Aunque los investigadores han analizado el papel de las características personales de los profesores en las estimaciones de valor agregado, tales como educación, certificación y experiencia (Hanushek & Rivkin, 2006), sus conclusiones no han sido concluyentes. También se ha estudiado el papel que las características de los alumnos y la escuela juegan en estas estimaciones (Newton, Darling-Hammond, Haertel & Thomas, 2010). No menos importante pero sí menos explorado es el rol del contexto educacional en el que se desempeña el docente, el que considera factores tales como la escuela y el municipio (Everson, Feinauer & Sudweeks, 2013; Graue, Delaney & Karch, 2013).

El estudio de la calidad docente a través de valor agregado, el rol de las características del docente y su contexto se encuentran poco explorados en Chile. Estudios anteriores han investigado el rol de las características de los estudiantes, el de algunas prácticas docentes y de algunas características de la escuela y el municipio en la predicción del desempeño de los alumnos en pruebas estandarizadas, utilizando datos transversales, lo que limita la obtención de medidas de valor agregado (Lara, Mizala & Repetto, 2010; Ortúzar, Flores, Milesi & Cox, 2009). Por otro lado, quienes han realizado estimaciones de valor agregado en Chile utilizando datos longitudinales lo han hecho solo a nivel del establecimiento y no han estudiado el rol de las características del profesor y del contexto en el que realiza su trabajo (González, San Martín, Manzi & Del Pino, 2010, Agosto).

Esta investigación pretende aportar a la discusión sobre calidad docente en Chile, desde la perspectiva de valor agregado, explorando la importancia de las características observables del docente, de la escuela y del municipio. Para ello se utilizaron datos longitudinales y se estimaron modelos lineales jerárquicos, considerando como variable dependiente el puntaje final de los alumnos (puntaje SIMCE de 2° año medio) y controlando por su puntaje inicial (puntaje SIMCE 8° año básico), ambos durante un período en que han trabajado con un mismo profesor.

¿Qué es Calidad Docente?

La discusión sobre calidad docente en la literatura internacional se había desarrollado, hasta recientemente, desde tres líneas principales de investigación: (a) el estudio de características observables del profesor que pueden tener un impacto sobre el desempeño de los alumnos (Hanushek & Rivkin, 2006); (b) el análisis de calidad docente, definida como prácticas profesionales basadas en estándares profesionales y medidas usualmente por medio de pautas de observación de clases (Danielson, 2007; Darling-Hammond, 2012; Papay, 2012) y (c) estimaciones de valor agregado, basadas en el rendimiento de los estudiantes en pruebas estandarizadas (Millman, 1997).

La literatura más reciente se ha aproximado a la discusión de calidad docente, proponiendo una integración de las distintas fuentes de información sobre el desempeño docente y su utilización integral en su evaluación (Cantrell & Kane, 2013; Darling-Hammond, 2012; Goe, Bell & Little, 2008; Martínez, 2013). Desde esta aproximación, las tres líneas de investigación mencionadas no constituirían diferentes medidas de calidad docente, sino que serían subdimensiones de una única medida. El proceso de evaluación debería incorporar, entonces, las características observables del profesor, sus prácticas profesionales y el impacto sobre el desempeño de los alumnos.

En este trabajo se profundiza en la línea de valor agregado, poco estudiada aún en Chile, y en aquellas características observables del docente.

Medidas de Valor Agregado

El desempeño de los alumnos en pruebas estandarizadas durante dos o más períodos consecutivos ha sido utilizado como base para el desarrollo de mediciones indirectas de la calidad docente, llamadas *efecto profesor* o *valor agregado del profesor*. Estas mediciones intentan capturar la contribución específica que cada profesor hace al proceso de aprendizaje de sus alumnos, tomando en consideración las variables que están fuera de su control directo, tal como el nivel socioeconómico (NSE) de sus alumnos (McCaffrey et al., 2004). Para ello, se requiere de datos longitudinales con al menos dos puntos de información para un alumno, idealmente al comienzo y al final del año escolar.

Las estimaciones del efecto profesor han resultado atractivas, debido a lo intuitivo de su interpretación, al estar directamente relacionadas con mediciones de aprendizaje, y debido al amplio rango de desempeño identificado, en comparación con mediciones de calidad basadas en criterios de desempeño profesional, que se caracterizan por resultados más homogéneos (Kersting, Chen & Stigler, 2013; Milanowski, 2011, Abril).

Sin embargo, también se critican las medidas de valor agregado, por la sensibilidad de su estimación debido al tamaño pequeño de las muestras (Hanushek & Rivkin, 2006), la dificultad de atribuir las ganancias de aprendizaje a más de un profesor en el aula (Isenberg & Hock, 2011, Abril) o a los profesores que están solo parcialmente presentes en la sala (Pullin, 2013) y porque las escalas utilizadas con frecuencia no están verticalmente escaladas y evalúan solo habilidades básicas, excluyendo áreas importantes del plan de estudios (Hill & Grossman, 2013; Hill, Kapitula & Umland, 2011). Además, varios trabajos empíricos sobre la estabilidad de las medidas de valor agregado, es decir, el grado en que dichas estimaciones cambian como consecuencia de distintas especificaciones del modelo ajustado y los factores que pueden afectar dicha estabilidad, han mostrado que se observan diferencias importantes dependiendo de los modelos estadísticos utilizados y de la presencia o ausencia de control de características demográficas de los alumnos y efectos fijos de la escuela (Darling-Hammond, Amrein-Beardsley, Haertel & Rothstein, 2012; Newton et al., 2010).

Las estimaciones del efecto profesor también son criticadas por el procedimiento de estimación empleado, el cual supone asignación aleatoria de los profesores a las escuelas y en las aulas, así como asignación aleatoria de los alumnos a las escuelas, lo que no se observa en la realidad (Rothstein, 2010). Además, Rothstein (2010) informa que las estimaciones del valor que agrega un docente en un período de un año se correlacionan débilmente con las estimaciones del valor agregado de ese mismo docente si se consideran períodos más largos y la interacción que tendrá el alumno con otros docentes durante su trayectoria escolar. Chetty, Friedman y Rockoff (2011), sin embargo, rebaten el hallazgo de Rothstein respecto de la asignación no aleatoria de alumnos a escuelas y muestran que alumnos asignados a profesores con alto valor agregado presentan mejores resultados de largo plazo, tales como mayor probabilidad de asistir a la universidad y de asistir a universidades de mejor calidad.

A pesar de las críticas, en algunos países, como Estados Unidos, se ha estimulado el uso de estimaciones de valor agregado como base para la evaluación de los planes de remuneración basados en el desempeño (Winters & Cowen, 2013), en parte como resultado de la legislación “*No Child Left Behind*” (Koretz & Hamilton, 2006). La legislación de “*Race to the Top*” (Baker, Oluwole & Green, 2013; McGuinn, 2012), anunciada por el Presidente de Estados Unidos en 2009, ha fomentado el desarrollo de este tipo de medidas y su uso en la evaluación de profesores.

El Rol del Contexto en la Calidad Docente

La calidad docente no solo está condicionada por las características del profesor, sino también por el contexto. Diversos autores (Bryk, Harding & Greenberg, 2012; Darling-Hammond, 2012; Moore Johnson, 2012) señalan que la calidad de la enseñanza —los resultados del proceso educativo— está fuertemente influenciada por el contexto de instrucción y que los profesores deberían ser evaluados considerando la institución, la población estudiantil y los recursos con los que trabajan (Everson et al., 2013).

En la definición de contexto, la investigación ha distinguido entre (a) el rol de las características físicas, tales como seguridad, recursos educativos e infraestructura, y (b) las características de las organización, tales como el liderazgo de la escuela y el distrito, la interacción docente y dinámicas en grupo, la relación con los padres y la relación con los alumnos (Heneman III & Milanowski, 2004; Odden, Borman & Fermanich, 2004). Kennedy (2010) sugiere que la investigación debe examinar el papel de las situaciones organizacionales particulares en la que los docentes tienen que enseñar y que determinan qué y cómo enseñan en sus clases. Graue, Delaney y Rasch (2013) se refieren específicamente a la coherencia entre la forma en que se usan los recursos y las acciones que apoyan el uso efectivo de esos recursos.

El rol del contexto en la estimación del efecto profesor y estimación del desempeño de los alumnos será el foco del análisis de este artículo.

Calidad Docente en Chile

En Chile la discusión respecto a qué es calidad docente ha estado centrada principalmente en dos conceptualizaciones: (a) la calidad docente como las características propias del profesor que afectan el

desempeño de los alumnos en pruebas estandarizadas (Bravo, Falck, González, Manzi & Peirano, 2008, Septiembre; León, Manzi & Paredes, 2008, Septiembre) y (b) la calidad docente basada en la evaluación de competencias profesionales (Taut, Santelices & Stecher, 2012; Valencia & Manzi, 2011).

Si bien en la primera aproximación algunos estudios han incluido en sus análisis variables del establecimiento, el foco de la discusión ha estado puesto en la capacidad predictiva de la evaluación docente (Lara et al., 2010; Ortúzar et al., 2009) y en el ordenamiento no aleatorio de los profesores entre escuelas (Toledo, 2010). Los estudios han encontrado evidencia apoyando la importancia de la experiencia del profesor, su grado académico, género y el porcentaje de contenido que declara haber cubierto en clase (Lara et al., 2010; León et al., 2008; Toledo, 2010). Ortúzar et al. (2009) encuentran evidencia para las mismas variables, pero limitadas solo a profesores jóvenes.

En términos comparativos, se ha puesto menos atención al papel que las variables propias del establecimiento educativo y de la municipalidad pudiesen tener sobre el desempeño de los alumnos. Hasta donde se conoce, al predecir desempeño de los alumnos, solo Ortúzar et al. (2009) consideran el NSE de la escuela, el tipo de administración escolar y el porcentaje de matrícula municipal sobre la matrícula particular subvencionada, además de variables del alumno y del profesor. El análisis, sin embargo, se realiza utilizando un modelo de solo dos niveles, que no considera apropiadamente la estructura anidada de la información.

Los estudios presentados, además, no han abordado el rol de distintos factores en el progreso del alumno durante el año que trabaja con un determinado profesor y han utilizado solo datos transversales. Para abordar este aspecto se requiere de mediciones repetidas del mismo grupo de alumnos en una misma prueba estandarizada, idealmente al principio y a final del mismo año.

Las visiones más recientes en la literatura internacional sobre calidad docente, lejos de rechazar las estimaciones de valor agregado, las incorporan en sus mediciones multidimensionales de calidad docente. En Chile, hasta la fecha, no se dispone de estudios publicados sobre medidas de valor agregado de calidad docente basadas en mediciones longitudinales del aprendizaje de los alumnos.

Preguntas de Investigación

El principal interés de este trabajo fue explorar el rol del contexto y de las características propias del profesor en las estimaciones del valor agregado del profesor. Específicamente, se pretendió responder las siguientes preguntas de investigación:

1. ¿Es la estimación del efecto profesor estable frente a la incorporación de diversos controles que consideran las características observables del docente y del contexto escolar en que se desempeña, es decir, de la escuela y del municipio?
2. ¿Cuál es la importancia de las características del profesor y de su contexto en la predicción del desempeño del alumno?

Método

Fuentes de Información

Los datos analizados provienen de tres fuentes de información: (a) Sistema de Evaluación del Desempeño Profesional Docente, la cual es una evaluación obligatoria para los docentes que se desempeñan en el sector de educación municipal, (b) SIMCE, cuyos resultados entregan información de los estándares de aprendizaje logrados por los estudiantes en los diferentes niveles de enseñanza, además de entregar el nivel educacional y el nivel de ingresos de los padres, y (c) Sistema Nacional de Información Municipal. Tras realizar el cruce de información, se obtuvo una muestra de 19.225 alumnos, 484 profesores del sector municipalizado o particular subvencionado, 292 escuelas y 204 municipios.

Análisis de Datos

El análisis de los datos se dividió en dos partes. La primera aborda la primera pregunta, utilizando un modelo lineal jerárquico de dos niveles. El modelo incluye un coeficiente aleatorio que representa el efecto profesor y el desempeño de los estudiantes en la prueba SIMCE de Matemáticas de 8° año básico y 2° año medio como variables independiente y dependiente, respectivamente (Snijders & Bosker, 1999). Las variables

incorporadas en el nivel del alumno corresponden a $POST_{ij}$, que es el puntaje del alumno en la prueba SIMCE del año 2006, y PRE_{ij} , que es su puntaje en la prueba SIMCE del año 2004. En esta notación, los subíndices i y j representan a cierto estudiante y profesor, respectivamente. El resto de las características del alumno, denotadas como $VAR_ESTUDIANTE$, corresponde a educación de la madre y educación del padre (categorías 0 a 19), repitencia (1/0) y cambio de colegio (1/0). El segundo nivel incorpora las variables del profesor, resumidas como $VAR_PROFESOR$, incluyendo género, años de experiencia, estudio en un régimen de tiempo completo (sí/no), asistencia a programas de formación inicial de cinco años (sí/no); asistencia a desarrollo profesional, genéricamente llamados PPF, (sí/no), asistencia a otros desarrollos profesionales (sí/no) y grados académicos (sin título profesional/con título profesional o escuela normal/magíster o doctorado). El modelo también considera las características de la escuela y el municipio en que esta se encuentra. Las características de la escuela, resumidas como $VAR_ESCUELA$, corresponden a ruralidad (urbano = 1; rural = 0), NSE de la escuela y el puntaje promedio obtenido en las pruebas SIMCE del año 2004. Esta última variable se incluyó como una forma de controlar la asignación no aleatoria de alumnos y docentes a escuelas. El supuesto es que mejores alumnos y mejores docentes optan por escuelas de mejor puntaje SIMCE. Las características del municipio, resumidas como $VAR_MUNICIPALES$, son porcentaje de personas viviendo en pobreza, la contribución porcentual de la municipalidad al gasto en educación, el número de alumnos matriculados al inicio del año escolar y el tipo de administración escolar. Se hipotetizó que estas últimas están relacionadas con la calidad y foco de la gestión municipal en el área educacional, las que inciden a su vez en el desempeño de los docentes y estudiantes.

Para abordar la segunda pregunta se estimó un modelo lineal jerárquico de tres niveles que incluyó, además, un efecto aleatorio a nivel de escuela (Raudenbush & Bryk, 2002). Las estimaciones consideraron solo a los estudiantes que tenían el mismo profesor durante el 1° y 2° año medio.

Todos los modelos fueron estimados utilizando máxima verosimilitud en el programa SAS. A continuación se presentan detalles de cada una de estas estrategias.

Estimación del Efecto Profesor en un Modelo de Dos Niveles y Estabilidad de las Estimaciones. El modelo jerárquico de dos niveles entrega un orden relativo de la contribución de todos los profesores incluidos en la muestra. El modelo con el que se inició el análisis fue el siguiente:

$$(1.1) POST_{ij} = \alpha_{0j} + \alpha_1 PRE_{ij} + e_{ij}$$

$$(1.2) \alpha_{0j} = \beta_{00} + u_{0j}$$

donde el primer nivel corresponde al nivel de alumno (i) y el segundo, al de profesor (j). El error a nivel del alumno (e_j) se supone que se distribuye en forma normal con media 0 y varianza σ^2 . El error a nivel del profesor (μ_{0j}) también se supone que se distribuye en forma normal con media 0 y varianza τ_α . En este modelo el efecto del profesor corresponde a la predicción del efecto aleatorio μ_{0j} , término que captura la diferencia entre el desempeño estimado de los alumnos de un profesor y el desempeño observado para ese grupo de alumnos. Si bien esta conceptualización se ha utilizado en otros estudios (Kersting et al., 2013; Newton et al. 2010), es importante explicitar que podría contar con un problema de endogeneidad en la estimación (para más detalles ver Manzi, San Martín & Van Belleghem, 2014). El análisis de estabilidad de las estimaciones se realiza mediante la comparación de modelos de dos niveles que incorporan progresivamente diversas variables, primero en el nivel del alumno y luego en el nivel del profesor/sala de clases. Los modelos ajustados se muestran en la Tabla 1.

Tabla 1
Análisis de Estabilidad de las Medidas de Valor Agregado: Variables Consideradas Según Modelo

Modelo	VARIABLES INCORPORADAS EN EL MODELO
M0	Modelo incondicional
M1	M0 + puntaje en primera medición del alumno (SIMCE 8° básico)
M2	M1 + características del alumno
M3	M2 + características del profesor
M4	M3 + características de la escuela
M5	M4 + características del municipio

Con el objetivo de estudiar la estabilidad en las estimaciones del efecto profesor entre modelos que incorporan progresivamente más controles por características del docente y de su contexto, se presentan los coeficientes de correlación de Spearman para la relación entre el ranking obtenido por un mismo docente en distintos modelos. Interesados en explorar la relación entre las estimaciones de valor agregado de distintas salas de clases enseñadas por un mismo profesor, se presentan, además, análisis que comparan: (a) la variabilidad del ranking de un profesor en un modelo en que se estima el valor agregado para cada profesor y (b) la variabilidad de los rankings de las salas de clases en las que enseñan esos profesores. Para el caso (b), el modelo estadístico utilizado es idéntico al indicado en la ecuación 1.1 y 1.2, en donde μ_{oj} representa ahora el efecto sala de clases. Además, con el objetivo de estudiar la capacidad del modelo de valor agregado de efectivamente controlar el contexto en el que se desempeña el docente, se analizó la evolución de la relación entre estimaciones de valor agregado obtenidas de distintos modelos con las características de sus salas de clases. Dicha relación se estudió utilizando la correlación de Pearson.

Modelos de Tres Niveles Para Predecir el Desempeño del Alumno. En la segunda parte del análisis se exploran las variables relevantes para predecir el desempeño de los alumnos en la prueba SIMCE de 2° medio utilizando un modelo lineal de tres niveles (alumno, profesor y escuela). La incorporación de un efecto aleatorio a nivel de escuela permite tomar en consideración la correlación existente entre profesores que enseñan en un mismo establecimiento educacional (González & San Martín, 2009). Si bien se utilizaron datos longitudinales, lo que en principio permitiría la obtención de medidas de valor agregado, con la incorporación al modelo de un efecto aleatorio adicional (efecto escuela), es incierto si la predicción del efecto aleatorio asociado al profesor puede aún ser considerada como una medida del efecto profesor. Así, las conclusiones que pueden obtenerse de este modelo estarán basadas solo en el contexto de predicción del desempeño del alumno.

Es importante aclarar que en esta sección el objetivo no es medir el efecto profesor, sino que explorar una modelación del desempeño del alumno que considere al docente y su contexto, y que permita comparar con los resultados de otros trabajos en la misma línea realizados previamente en Chile. Metodológicamente hablando, la omisión de niveles relevantes puede llevar a sesgo en las estimaciones de varianza del nivel de profesor y nivel de alumno, tal como mostró Martínez (2012) respecto del efecto escuela (*schooling effect*). Además, Cervini (2009, 2012) señala que ignorar los niveles relevantes tendría efectos sobre la estimación del coeficiente de efectos fijos y aleatorios, además de afectar los errores estándar. Sus resultados muestran la importancia de incorporar el nivel inmediatamente anterior y posterior al nivel de interés en la estimación. El modelo de tres niveles con dos interceptos aleatorios que se utiliza en este trabajo se presenta en las ecuaciones (2.1) a (2.3).

$$(2.1) POST_{ijs} = \alpha_{0js} + \alpha_1 PRE_{ijs} + \sum_{p=2}^n \alpha_p (VAR_ESTUDIANTE_{ijs}) + e_{ijs}$$

$$(2.2) \alpha_{0js} = \beta_{0s} + \sum_{k=1}^K \beta_k (VAR_PROFESOR/CLASE_{js}) + u_{0js}$$

$$(2.3) \beta_{0s} = \phi_{00} + \sum_{i=1}^T \phi_i (VAR_ESCUELA_s) + \sum_{z=1}^Z \theta_z (VAR_MUNICIPALES_m) + \sigma_{00s}$$

donde el primer nivel corresponde al nivel de alumno (i), el segundo a nivel profesor o sala de clases (j) y el tercero a nivel escuela (s). Las variables explicativas son idénticas a las consideradas en la sección anterior. El error a nivel del alumno (e_{ijs}) y el error a nivel del profesor/sala de clases (u_{0js}) se distribuyen igual que en el modelo de dos niveles. El error a nivel escuela (σ_{00s}) se distribuye en forma normal con media 0 y varianza τ_β .

Para evaluar el ajuste de los modelos, se reporta el valor de la función de máxima verosimilitud, el Criterio de Información de Akaike (AIC y AICC), el Criterio de Información Bayesiano (BIC) y la proporción de varianza observada a la que corresponde cada nivel de interés. Esto permite comparar las diferentes especificaciones e inferir qué modelo se ajusta mejor a los datos. Además, se reporta la significación estadística de las variables incluidas en cada uno de los niveles y sus coeficientes estandarizados.

Resultados

Medidas de Valor Agregado con Modelo de Dos Niveles y Análisis de Sensibilidad

La Tabla 2 muestra un análisis de correlaciones del efecto profesor estimado por medio de los distintos modelos propuestos en la Tabla 1. En ella es posible observar que la correlación disminuye a medida que se incorporan controles en la estimación. Ello ilustra la sensibilidad de la medida de valor agregado a la especificación del modelo a estimar. El cambio en magnitud es mayor cuando se incluyen las características de la escuela. Al incluir dichas variables de control, la correlación entre las estimaciones de valor agregado pasa de mostrar una magnitud en torno a 0,99 a ser no mayor a 0,84.

Tabla 2
Predicción del Efecto Aleatorio en Modelo de Dos Niveles: Correlación de Spearman Entre Percentil de Cada Profesor en M1 a M5

Modelo	M1	M2	M3	M4	M5
M1	1				
M2	0,9921	1			
M3	0,9839	0,9923	1		
M4	0,8171	0,8434	0,8526	1	
M5	0,8133	0,8389	0,8477	0,9937	1

Nota. Todas las correlaciones son significativas con $p < 0,01$; $n = 457$

La Tabla 3 muestra una comparación entre las estimaciones de valor agregado que consideran el efecto profesor y el efecto sala de clase, respectivamente. Los resultados muestran que existen diferencias entre la estimación del efecto profesor a nivel de sala y las estimaciones a nivel de profesor. Si se toma a los profesores del primer decil de las estimaciones de valor agregado anidadas a nivel de profesor y se sigue a dichos profesores en una estimación anidada a nivel de sala de clases, se observa que sus salas se ubican entre el percentil 1 y 39. Si un profesor se encuentra en el decil 10 en la estimación anidada en profesor, su estimación anidada en sala de clases puede tomar valores entre 48 y 100. Esto implica que la estimación del profesor muestra variabilidad dependiendo del grupo de alumnos con los que está trabajando, aun después de haber incorporado controles a la estimación.

Tabla 3
Relación Entre el Percentil del Profesor de una Estimación de Valor Agregado Anidada en Profesor y el Decil del Efecto Profesor de una Estimación Anidada en Sala de Clase

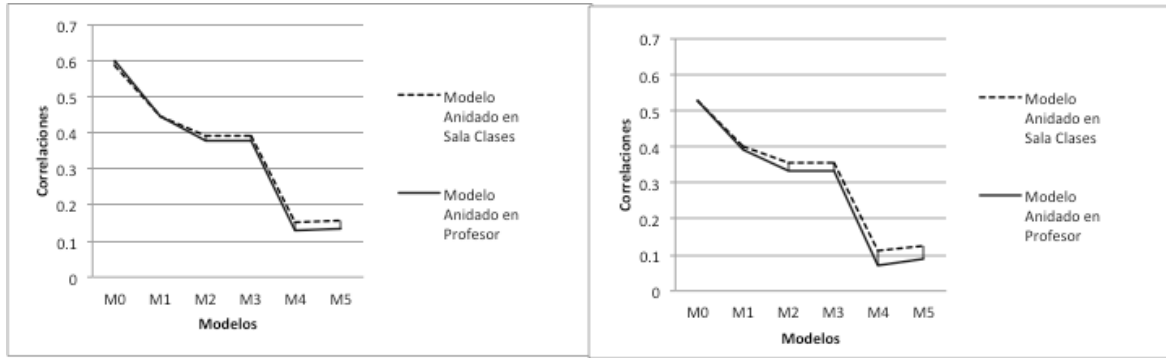
Modelo pre test	<i>n</i>	Media	Desviación estándar	Mediana	Mínimo	Máximo
Decil 1	67	8,31	7,22	6	1	39
Decil 2	68	18,62	8,45	17	2	47
Decil 3	69	27,61	11,60	27	2	66
Decil 4	73	36,85	13,87	37	4	89
Decil 5	69	47,42	14,73	48	13	94
Decil 6	68	54,78	14,98	57	8	79
Decil 7	72	62,13	15,92	65	6	95
Decil 8	80	71,95	15,83	75	1	100
Decil 9	66	83,65	7,56	85	61	97
Decil 10	65	92,06	9,98	95	48	100

La Tabla 4 presenta un análisis de correlaciones entre las estimaciones del efecto profesor y características de la sala de clases en que enseñan. Las correlaciones directas de la Tabla 4 y de las Figuras 1 y 2 que la complementan indican que aquellos profesores que enseñan a alumnos que tienen padres más educados y familias con mayores ingresos obtienen posiciones relativas más altas frente a profesores con escenarios más adversos. Sin embargo, la correlación disminuye en la medida que el modelo incorpora más antecedentes del profesor y mayor número de características del contexto en que el profesor enseña. Esto se cumple hasta llegar al modelo que incluye el nivel escuela y las variables de las escuelas en la que enseña el docente (desde M3 a M4). Destaca la disminución observada entre las variables al incluir las características de la escuela, ya que, al incluir dichas variables de control, la correlación pasa de mostrar una magnitud media a una magnitud pequeña (Cohen, 1988).

Tabla 4
Modelo Lineal de Dos Niveles: Correlaciones Entre Características de las Salas de Clases y la Predicción de Efectos Aleatorios

Variable de anidación	Modelo	Promedio educación madre	Promedio ingreso familiar
Modelo anidado en sala de clases <i>n</i> = 695	M0	0,5867*	0,5268*
	M1	0,4475*	0,4023*
	M2	0,3923*	0,3558*
	M3	0,3909*	0,3570*
	M4	0,1524*	0,1118*
	M5	0,1589*	0,1271*
Modelo anidado en profesor <i>n</i> = 456	M0	0,6019*	0,5269*
	M1	0,4476*	0,3905*
	M2	0,3797*	0,3318*
	M3	0,3791*	0,3314*
	M4	0,1284*	0,0700
	M5	0,1358*	0,0893

* $p < 0,01$



Figuras 1 y 2. Modelo lineal de dos niveles: correlaciones entre características de las salas de clases (izquierda-promedio de educación de la madre, derecha-promedio ingreso familiar) y la predicción de efectos aleatorios.

Estimaciones Considerando el Contexto: Modelo de Tres Niveles

Al examinar los resultados de un modelo de tres niveles en que se controla solo el test previo del alumno, se observa que la escuela representa el 5% de la varianza explicada, mientras el profesor representa el 21%. El nivel que más aporta a la varianza explicada, sin embargo, es el nivel del estudiante, con 74% de la varianza.

En la Tabla 5 se puede apreciar que, al introducir las variables del estudiante, del profesor y de la escuela en cada uno de los niveles correspondientes, además del desempeño en el test anterior, hay una reducción en el tamaño de la estimación de varianzas a nivel escuela y a nivel profesor. Esta reducción no se observó en el modelo que presenta las características a nivel municipal, por lo que el modelo más apropiado es el (4). Esto es ratificado por los estadísticos AIC y BIC. Todas las estimaciones de varianzas son significativas a valores de $p < 0,001$.

Tabla 5
Modelos de 3 Niveles: Estimaciones de Varianzas

	(0) Porcentaje de la varianza con test previo	(1) Controlando por el test previo	(2) Covariables del estudiante	(3) Covariables del estudiante y profesor	(4) Covariables del estudiante, profesor y escuela	(5) Covariables del estudiante, profesor, escuela y municipio
	Coeficientes aleatorios					
Varianza nivel escuela	5	89,9	76,4	73,5	29,0	31,5
Varianza nivel profesor	21	342,5	309,4	306,0	245,6	239,9
Varianza nivel estudiante	74	1237,2	1218,9	1218,9	1218,6	1218,6
-2 Log likelihood		180768,5	180458,8	180451,7	180316,1	180311,7
AIC		180778,5	180476,8	180487,7	180362,1	180365,7
AICC		180778,5	180476,8	180487,7	180362,2	180365,8
BIC		180799,1	180513,9	180561,9	180457,0	180477,1

En la Tabla 5 se puede observar específicamente la reducción de la varianza en los distintos niveles al introducir en el modelo características observables de los docentes y de su contexto. Esa reducción de la varianza muestra que la predicción del desempeño de los alumnos es sensible a las características propias de cada estudiante, a los rasgos de los profesores y a las particularidades de cada escuela.

Para profundizar en el análisis de las variables observables propias del docente y aquellas de su contexto que afectan la estimación del desempeño de los alumnos, se analizaron los coeficientes de los efectos fijos incluidos en el modelo. Al hacerlo, se observa (ver Tabla 6) que todos los coeficientes obtenidos para las características del estudiante son estadísticamente significativos y se asocian de la forma esperada al desempeño en el post-test. En el ámbito docente, el desarrollo profesional del profesor es la única variable estadísticamente significativa. Entre las características de la escuela, se observa como única variable estadísticamente significativa el promedio de la escuela en las pruebas SIMCE del año 2004. Una consecuencia de la incorporación de esta variable es la obtención de un intercepto negativo en la ecuación de regresión. Se realizaron análisis complementarios en que el puntaje SIMCE de la escuela no fue considerado. En ellos se obtuvieron coeficientes significativos para el NSE de la escuela. Por último, ninguna de las características del municipio resultó significativa.

Discusión y Conclusión

Las correlaciones observadas en la Tabla 2 son consistentes con los hallazgos de Newton et al. (2010), quienes también encontraron correlaciones distintas de cero entre el ranking de profesores y las características de los alumnos, aun controlando por dichas características. Los resultados fueron interpretados por los autores como evidencia de la importancia del contexto en estimaciones de valor agregado y la insuficiente capacidad del modelo de controlar las características del lugar y alumnos con los que trabaja un docente, aun cuando se incluyan controles en la regresión. Es importante recalcar que la magnitud de las correlaciones observadas en el presente estudio es mayor a las reportadas por Newton et al. (2010).

Ahora bien, la menor correlación observada entre las estimaciones del efecto profesor y las características de los alumnos, al incluir características de la escuela en el modelo, muestran la importancia de considerar dicho nivel en la modelación de la efectividad docente. Este hallazgo se ve reforzado por la significancia estadística del componente de varianza incluido a nivel escuela en el modelo de tres niveles que predice desempeño del alumno y la proporción de varianza explicada atribuible a ese nivel (5%). La importancia de este nivel ha sido subrayada por otros estudios que, comparativamente, encuentran una mayor proporción de varianza del aprendizaje de los alumnos asociada a nivel escuela: Martínez (2012) reporta 16,2% y Cervini (2009) reporta 25,5%. Estos resultados van en la línea con lo descrito por Moore Johnson (2012), al describir la efectividad de un docente como una efectividad *situada en un contexto organizacional*. Entre los efectos fijos incluidos disponibles en las bases de datos, la variable de mayor importancia a nivel escuela fue el desempeño previo en el SIMCE; el NSE de la escuela no resulta significativo una vez que se ha controlado el nivel educacional de los padres y el SIMCE.

Los resultados relativos a la no significancia estadística de las variables municipales incluidas en el modelo de tres niveles sugieren la menor importancia relativa de la gestión municipal sobre el desempeño de los alumnos y efectividad docente, en comparación con la gestión de la escuela. Estos resultados también podrían indicar, sin embargo, que hay variables relevantes a nivel municipal que no fueron consideradas en el análisis. Si bien los resultados reportados en este estudio en relación a variables del contexto (NSE de la escuela, tipo de administración escolar y porcentaje de matrícula municipal sobre la matrícula subvencionada) difieren de los hallazgos de Ortúzar et al. (2009), es importante considerar que estos autores utilizan un modelo de solo dos niveles para el análisis y en él no se intenta corregir por asignación no aleatoria de alumnos y profesores a escuelas y cursos.

Tabla 6
Modelo de Tres Niveles: Estimación de Efectos Fijos

Efectos fijos	Coefficiente	Error estándar
Intercepto	-83,53*	16,61
Características del estudiante		
Puntaje previo	0,81*	0,01
Educación de la madre	0,28*	0,10
Educación del padre	0,41*	0,10
Repitencia (categoría base: no)	-12,29*	0,89
Cambio de colegio (categoría base: no)	-6,21*	1,15
Características del profesor		
Género (categoría base: hombre)	1,32	1,71
Experiencia	-0,19	0,10
Formación inicial en programas de 5 años	4,61	2,77
Formación inicial en régimen tiempo completo	-5,52	4,17
Formación inicial con especialización en educación media	-3,93	3,52
Desarrollo profesional (PPF) (categoría base: no)	-0,30	3,08
Otros desarrollos profesionales (categoría base: no)	4,12*	1,86
Grado		
Sin título profesional	-11,48	10,94
Magíster o doctorado	-2,27	3,25
Título profesional o escuela normal (categoría base)		
Características de la escuela		
Urbana	-4,36	5,44
NSE B	-2,27	2,22
NSE C	-2,65	4,23
NSE D	-0,65	8,98
Promedio de la escuela en el test previo	0,55*	0,06
Características municipales		
Porcentaje de personas viviendo en pobreza	10,53	13,09
Contribución municipal al ingreso total de educación de la comuna	-12,71	9,63
Número de alumnos matriculados al inicio del año escolar	0,00	0,00
Tipo de administración escolar (categoría base: municipal)	-1,01	2,32

* $p < 0,001$

La proporción de varianza asociada al nivel profesor reportada en este estudio (21%) es algo mayor que la proporción de varianza observada a nivel sala de clases observada por Martínez (2012; 15,1%) y Cervini (2009; 12,7%). A nivel del profesor resulta interesante que ninguna de las variables asociadas a formación inicial sea significativa. Estos resultados se relacionan con los hallazgos de Ortúzar et al. (2009), quienes encontraron apoyo para la relación entre desempeño de los alumnos y algunas variables limitadas de la formación inicial, pero solo entre profesores jóvenes. Tampoco se encontró evidencia en el presente estudio a favor de los postgrados o de la experiencia, lo que difiere de los resultados reportados por Hanushek y Rivkin (2006), León et al. (2008) y Ortúzar et al. (2009). Se debe mencionar, sin embargo, que los últimos dos estudios no usan medidas longitudinales de desempeño del alumno, sino que medidas transversales (es decir, no controlan el puntaje inicial). La única variable observable significativa propia del profesor fue experiencias de desarrollo profesional distintas a los PPF. Este hallazgo subraya la importancia de la formación continua. Los resultados encontrados, si bien no apoyan la importancia de la formación inicial de los docentes ni la certificación educacional posterior, pueden deberse a la calidad promedio de la formación recibida por el grupo de docentes incluidos en la muestra y no deben interpretarse, necesariamente, como evidencia de que la formación en el futuro no sería relevante. Su relevancia podría cambiar en la medida en que la formación que se imparte hoy en Chile y se siga impartiendo en el futuro en las facultades y escuelas de pedagogía sea diferente —ojalá mejor— que aquella recibida por el grupo de docentes estudiado. Sin embargo, este estudio no está en condiciones de hacer inferencias causales respecto de dicha futura relación.

La importancia del nivel estudiantes observado en este estudio (74%) es coherente con los hallazgos reportados por Martínez (2012) y Cervini (2009), de 68,6% y 61,8% respectivamente. El hecho de que todas las características de los alumnos consideradas hayan resultado estadísticamente significativas para la estimación del desempeño estudiantil y calidad docente revela que son múltiples los factores que pueden relacionarse con la labor del docente, no solo factores socioeconómicos, sino que también el haber repetido alguna vez durante la trayectoria académica, por ejemplo.

Este trabajo se encuentra sujeto a limitaciones propias de los datos con que se trabajó. En primer lugar, la muestra consideró solo a los alumnos de escuela municipales o particulares subvencionadas que tuvieron el mismo profesor durante dos años, tanto en lenguaje como en matemáticas, por lo que podría ser poco representativa de la población estudiantil chilena evaluada por el SIMCE. La diferencia más importante en esta muestra se observa en la mayor participación de escuelas pobres y de mayores ingresos respecto de la participación de la población general (grupo A, 40% versus 28% y grupo E, 40% versus 18%), pero, al ser sobrepresentados ambos extremos, en desmedro de los grupos medios, no se invalidarían los resultados. En segundo lugar, por no contar con la información administrativa necesaria, no fue posible incorporar en el modelo situaciones en que ha habido más de un profesor enseñando en una misma sala de clases o situaciones en que el profesor solo ha enseñado a ese grupo de alumnos durante una fracción del período analizado, lo que podría afectar las estimaciones.

Los resultados de esta investigación indican que la estimación de la calidad docente se relaciona no solo con las características de los alumnos que el profesor atiende, sino también con las capacitaciones a las que ha asistido dicho profesor, las características de la escuela en la que enseña y el desempeño promedio de la escuela en el SIMCE anterior. Los hallazgos pueden interpretarse como una demostración de la dificultad de medir el desempeño de un docente separado del contexto en el que se desempeña. Si bien no se pretende postular que el desempeño docente no debiera medirse, es importante explorar formas en que el rol del contexto sea apropiadamente incluido en dicha medición. Para ello podrían considerarse propuestas recientes (“Symposium: Teacher Effectiveness”, 2013), tales como la implementación de metodologías estadísticas complejas para la evaluación de profesores (Everson et al., 2013) o el desarrollo de pautas de observación más holísticas que permitan capturar dimensiones de la práctica docente situadas en el contexto de cada profesor y centradas en contenidos relevantes (Hill & Grossman, 2013).

Referencias

- Baker, B. D., Oluwole, J. O. & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the race-to-the-top era. *Education Policy Analysis Archives*, 21(5), 1-65. doi:10.14507/epaa.21n5.2013 Extraído de <http://epaa.asu.edu/ojs/article/view/1298/1043>
- Barber, M. & Mourshed, M. (2007). *How the world's best-performing school systems come out on top*. London, Reino Unido: McKinsey.

- Bravo, D., Falck, D., González, R., Manzi, J. & Peirano, C. (2008, Septiembre). *La relación entre la evaluación docente y el rendimiento de los alumnos: evidencia para el caso de Chile*. Ponencia presentada en el Encuentro Anual de la Sociedad de Economía de Chile, Viña del Mar, Chile.
- Bryk, A., Harding, H. & Greenberg, S. (2012). Contextual influences on inquiries into effective teaching and their implications for improving student learning. *Harvard Educational Review*, 82, 83-106.
- Cantrell, S. & Kane, T. J. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study* (MET Project Research Paper). Seattle, WA: Bill & Melinda Gates Foundation. Extraído de http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf
- Cervini, R. A. (2009). Class, school, municipal, and state effects on mathematic achievement in Argentina: A multilevel analysis. *School Effectiveness and School Improvement*, 20, 319-340. doi:10.1080/09243450802664404
- Cervini, R. A. (2012). El "efecto escuela" en países de América Latina: reanalizando los datos del SERCE. *Education Policy Analysis Archives*, 20(39), 1-26. doi:10.14507/epaa.v20n39.2012 Extraído de <http://epaa.asu.edu/ojs/article/view/1086/1012>
- Chetty, R., Friedman, J. N. & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (NBER Working Paper 17699). Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w17699 Extraído de http://www.nber.org/papers/w17699.pdf?new_window=1
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2ª ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2ª ed.). Alexandria: VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, 93(6), 8-15.
- Everson, K. C., Feinauer, E. & Sudweeks, R. R. (2013). Rethinking teacher evaluation: A conversation about statistical inferences and value-added models. *Harvard Educational Review*, 83, 349-370.
- Goe, L., Bell, C. & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Extraído de <http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>
- González, J. & San Martín, E. (2009). Rendimiento en la prueba PISA: ¿es posible entender los alcances y límites de las comparaciones entre países? En L. Cariola, G. Cares & E. Lagos (Coords.), *¿Qué nos dice PISA sobre la educación de los jóvenes en Chile? Nuevos análisis y perspectivas sobre los resultados en PISA 2006* (pp. 327-351). Santiago, Chile: Ministerio de Educación, Unidad de Curriculum y Evaluación.
- González, J., San Martín, E., Manzi, J. & Del Pino, G. (2010, Agosto). *Mixed modeling approach to value-added analysis: Features, problems, and recent developments in the Chilean case*. Ponencia presentada en la Second Biennial Meeting of the EARLI Special Interest Group 18 "Educational Effectiveness: Models, Methods and Applications", Leuven, Bélgica.
- Graue, M. E., Delaney, K. K. & Karch, A. S. (2013). Ecologies of education quality. *Education Policy Analysis Archives*, 21(8), 1-30. doi:10.14507/epaa.v21n8.2013 Extraído de <http://epaa.asu.edu/ojs/article/view/1163/1044>
- Hanushek, E. A. & Rivkin, S. G. (2006). Teacher quality. En E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (Vol. 2, pp. 1051-1078). Amsterdam, Holanda: Elsevier.
- Heneman III, H. G. & Milanowski, A. T. (2004). Alignment of human resource practices and teacher performance competency. *Peabody Journal of Education*, 79(4), 108-125. doi:10.1207/s15327930pje7904_6
- Hill, H. C. & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*, 83, 371-384.
- Hill, H. C., Kapitula, L. & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, 48, 794-831. doi:10.3102/0002831210387916
- Isenberg, E. & Hock, H. (2011, Abril). *Value-added methods in a high-stakes assessment system*. Ponencia presentada en la Annual Meeting of the American Educational Research Association, New Orleans, LA, Estados Unidos.
- Kennedy, M. M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher*, 39, 591-598. doi:10.3102/0013189X10390804
- Kersting, N. B., Chen, M. -K. & Stigler, J. W. (2013). Value-added teacher estimates as part of teacher evaluations: Exploring the effects of data and model specifications on the stability of teacher value-added scores. *Education Policy Analysis Archives*, 21(7), 1-33. doi:10.14507/epaa.v21n7.2013 Extraído de <http://epaa.asu.edu/ojs/article/view/1167/1049>
- Koretz, D. & Hamilton, L. S. (2006). Testing for accountability in K-12. En R. L. Brennan (Ed.), *Educational Measurement* (4ª ed., pp. 531-578). Westport, CT: American Council on Education/Praeger.
- Lara, B., Mizala, A. & Repetto, A. (2010). Una mirada a la efectividad de los profesores en Chile. *Estudios Públicos*, 120, 147-182. Extraído de http://www.cepchile.cl/dms/archivo_4744_2876/rev120_BLara_AMizala_ARepetto.pdf
- León, M. G. Manzi, J. & Paredes, R. D. (2008, Septiembre). *Calidad docente y rendimiento escolar en Chile: evaluando la evaluación*. Ponencia presentada en el Encuentro Anual de la Sociedad de Economía de Chile, Viña del Mar, Chile.
- Manzi, J., San Martín, E. & Van Belleghem, S. (2014). School system evaluation by value added analysis under endogeneity. *Psychometrika*, 79, 130-153. doi:10.1007/S11336-013-9338-0
- Martínez, J. (2013). Combinación de mediciones de la práctica y el desempeño docente: consideraciones técnicas y conceptuales para la evaluación docente. *Pensamiento Educativo. Revista de Investigación Educativa Latinoamericana*, 50 (1), 4-20. doi:10.7764/PEL.50.1.2013.2
- Martínez, J. F. (2012). Consequences of omitting the classroom in multilevel models of schooling: An illustration using opportunity to learn and reading achievement. *School Effectiveness and School Improvement*, 23, 305-326. doi:10.1080/09243453.2012.678864
- McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A. & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29, 67-101. doi:10.3102/10769986029001067 Extraído de <http://www.jstor.org/discover/10.2307/3701307?uid=3737784&uid=2&uid=4&sid=21104090372037>
- McGuinn, P. (2012). *The state of teacher evaluation reform: State education agency capacity and the implementation of new teacher-evaluation systems*. Washington, DC: Center for American Progress.
- Milanowski, A. T. (2011, Abril). *Validity research on teacher evaluation systems based on the framework for teaching*. Ponencia presentada en la Annual Meeting of the American Educational Research Association, New Orleans, LA, Estados Unidos. Extraído de <http://www.eric.ed.gov/PDFS/ED520519.pdf>

- Millman, J. (1997). *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* Thousand Oaks, CA: Corwin Press.
- Moore Johnson, S. (2012). Having it both ways: Building the capacity of individual teachers and their schools. *Harvard Educational Review*, 82, 107-122.
- Newton, X. A., Darling-Hammond, L., Haertel, E. & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Education Policy Analysis Archives*, 18(23), 1-22. doi:10.14507/epaa.v18n23.2010 Extraído de <http://epaa.asu.edu/ojs/article/view/810/858>
- Odden, A., Borman, G. & Fermanich, M. (2004). Assessing teacher, classroom and school effects, including fiscal effects. *Peabody Journal of Education*, 79(4), 4-32. doi:10.1207/s15327930pje7904_2
- Ortúzar, M. S., Flores, C., Milesi, C. & Cox, C. (2009). Aspectos de la formación inicial docente y su influencia en el rendimiento académico de los alumnos. En Pontificia Universidad Católica de Chile, *Camino al Bicentenario: propuestas para Chile. Concurso Políticas Públicas / 2009* (pp. 155-186). Santiago, Chile: Editor, Vicerrectoría de Comunicaciones y Asuntos Públicos, Dirección de Asuntos Públicos.
- Papay, J. (2012). Refocusing the debate: Assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, 82, 123-141.
- Pullin, D. (2013). Legal issues in the use of students test scores and value-added models (VAM) to determine educational quality. *Education Policy Analysis Archives*, 21(6), 1-27.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2ª ed.). Thousand Oaks, CA: Sage.
- Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, 125, 175-214. doi:10.1162/qjec.2010.125.1.175
- Snijders, T. A. B. & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, Reino Unido: Sage.
- Symposium: Teacher Effectiveness. The role of context [Special issue]. (2013). *Harvard Educational Review*, 83(2).
- Taut, S., Santelices, M. V. & Stecher, B. (2012). Validation of a national teacher assessment and improvement system. *Educational Assessment*, 17, 163-199. doi:10.1080/10627197.2012.735913
- Toledo, G. A. (2010). *Calidad docente y logro escolar: enfrentando el problema de ordenamiento no aleatorio entre características de profesores y alumnos* (Tesis de Magíster, Universidad de Chile, Santiago, Chile). Extraído de http://www.ciee2010.cl/docs/doc/sesiones/61_GToledo_Calidad_docente_y_logro_escolar.pdf
- Valencia, E. & Manzi, J. (2011). Desempeño docente: relaciones con antecedentes de los profesores y su contexto. En J. Manzi, R. González & Y. Sun (Eds.), *La evaluación docente en Chile* (pp. 177-194). Santiago, Chile: Pontificia Universidad Católica de Chile, Escuela de Psicología, Centro de Medición MIDE UC.
- Wilson, S. (Ed.) (2009). *Teacher quality: Education policy white paper*. Washington, DC: National Academy of Education.
- Winters, M. A. & Cowen, J. M. (2013). Who would stay, who would be dismissed? An empirical consideration of value-added teacher retention policies. *Educational Researcher*, 42, 330-337. doi:10.3102/0013189X13496145

Fecha de recepción: Septiembre de 2013.

Fecha de aceptación: Julio de 2014.