



La Evaluación Docente y sus instrumentos: discriminación del desempeño docente y asociación con los resultados de los estudiantes.*

*Macarena Alvarado, Gustavo Cabezas, Denise Falck, María Elena Ortega***

Resumen

El presente trabajo busca explicar la contribución de los diferentes instrumentos de la Evaluación Docente a los resultados de la misma, así como analizar la asociación de éstos con el desempeño de los estudiantes en las pruebas del Sistema de Medición de la Calidad de la Educación (SIMCE).

Respecto a la relación entre Evaluación Docente y rendimiento en las pruebas SIMCE, utilizando diferentes metodologías, se encuentra una asociación positiva entre el desempeño de los docentes y el de sus estudiantes. Es así como, 1 punto adicional promedio en la Evaluación Docente de los profesores entre 1° y 4° básico de un estudiante se relaciona con 15,8 puntos más en la prueba SIMCE de Matemática y con 9,5 puntos más en la de Lectura para ese alumno. Por otro lado, se encuentra que estar vinculado a profesores con resultado en el nivel Básico versus Insatisfactorio no representa ninguna diferencia significativa en los resultados de las pruebas del SIMCE de los estudiantes. Sin embargo, existe una asociación positiva y creciente tanto en los resultados de Matemática como de Lectura si es que los estudiantes estuvieron vinculados con profesores con resultados clasificados en el nivel de Competente o Destacado respecto de uno Insatisfactorio.

A nivel de los instrumentos, el portafolio de enseñanza pedagógica muestra la mayor asociación con los resultados de las pruebas SIMCE. Un punto adicional se asocia con 8 puntos más en Matemática y 6 en Lectura. En el informe de referencia de terceros 1 punto adicional se asocia con un 4,2 y 2,2 puntos más respectivamente, mientras que la entrevista por un evaluador par se asocia con 2,6 puntos más en Matemática y 1 punto más en Lectura. La autoevaluación muestra ser no significativa en ambas pruebas.

Palabras Claves: Docentes, Evaluación Docente, Desempeño Estudiantil.

* Este trabajo se desarrolló en el marco del convenio de colaboración MINEDUC/PNUD. Los autores agradecen a Violeta Arancibia, Sylvia Eyzaguirre, Francisco Lagos, Osvaldo Larrañaga, Fernando Reimers y Carolina Velasco por sus comentarios y recomendaciones, reservándose la responsabilidad por lo expuesto en el presente documento.

** Macarena Alvarado, Magister en Economía, Investigador asociado del Centro de Estudios del Ministerio de Educación. Gustavo Cabezas, Magister en Economía, Consultor PNUD. Denise Falck, Magister en Economía, Consultor PNUD. María Elena Ortega, Candidata a Doctor en Educación, Harvard Graduate School of Education.

1. Introducción

Una de las interrogantes más importantes en educación está relacionada con cómo elevar su calidad, siendo uno de los elementos de interés el rol de los docentes y cómo su desempeño influye en el aprendizaje de los estudiantes.

En este sentido, muchos autores han afirmado que la calidad de los docentes es uno de los factores más importantes al interior de la escuela para predecir el desempeño de los estudiantes. Hanushek (1992), señala que, manteniendo todo el resto constante, los estudiantes que han tenido docentes de alta calidad logran un nivel de aprendizaje tres veces mayor que los estudiantes que tienen docentes de baja calidad. Por otro lado, Rivers y Sanders (1996) encuentran que la influencia individual de los docentes sobre el desempeño académico de los estudiantes es acumulativa y aditiva. Esto es, los estudiantes que han tenido profesores de alto desempeño durante tres años de formación (3° a 5° grado) obtienen 50 puntos porcentuales más en pruebas estandarizadas que los estudiantes que han tenido docentes de bajo desempeño durante esos años.

Sin embargo, medir la calidad de los docentes no tiene una respuesta única. Goe (2007) comenta que la calidad de los docentes ha sido estudiada a partir de distintas dimensiones como: (i) las credenciales, experiencia y desarrollo profesional del maestro, (ii) las actitudes y aptitudes del maestro, (iii) las prácticas del maestro en el aula y (iv) la contribución del maestro a los resultados de aprendizaje de los estudiantes. Un ejemplo de estos estudios lo realizaron Goldhaber y Anthony (2004), quienes utilizando una base de datos de Carolina del Norte observan la relación entre la certificación de los docentes por el National Board for Professional Teaching Standards (NBPTS) y el nivel de logro de los estudiantes. Los autores encuentran que los docentes certificados fueron más efectivos en aumentar el desempeño de sus estudiantes, tanto en Matemática como en Lectura, que los docentes no certificados. Sin embargo, posteriormente Goe y Stickler (2008) argumentan que, fuera de la experiencia en los 5 primeros años de servicio y las competencias matemáticas, no hay evidencia contundente sobre qué aspectos de la calidad docente son los más importantes para el aprendizaje de los estudiantes. Por otro lado, Rivkin, Hanushek y Kain (2005), quienes analizan datos de panel de estudiantes del estado de Texas, encuentran diferencias significativas entre establecimientos en el aprendizaje de sus estudiantes, sugiriendo a partir de la heterogeneidad del desempeño de los estudiantes dentro de cada establecimiento, que las diferencias son explicadas por el efecto de los profesores más que por diferencias organizacionales de los establecimientos. Pero, a su vez, encuentran que poco de la varianza de la calidad de los docentes es explicada por características observables de los mismos como educación o experiencia.

A pesar que la calidad de los docentes ha sido el centro de atención de múltiples investigaciones, políticas y programas educativos, no existe un claro consenso sobre cómo identificarla. En este sentido, establecer la relación entre el desempeño del docente y el de sus estudiantes tampoco resulta trivial en el sistema chileno, considerando que la discusión sobre la calidad de los docentes en Chile, a nivel de investigación, es todavía incipiente y poco concluyente. Una de las aristas de debate entre los actores del sistema escolar tiene relación con cuáles son las metodologías más adecuadas para la medición de la calidad y/o efectividad de los docentes.

Otra limitación se vincula con los posibles sesgos asociados al modelo de selección de estudiantes/establecimientos y docentes/establecimientos (entre e intra establecimientos). Si bien uno de los métodos más utilizados para juzgar la efectividad del docente es observar el desempeño de sus estudiantes, en la mayoría

de los casos no es posible realizar un análisis de causalidad por los problemas de selección no aleatoria de estudiantes y docentes entre e intra establecimientos, así como por la falta de información longitudinal. Un ejemplo de la selección no aleatoria son los establecimientos de mayor excelencia, de cualquier tipo de dependencia administrativa, que eligen a los mejores estudiantes dentro un conjunto de postulantes¹. Aun así existe, puede existir heterogeneidad al interior de un mismo establecimiento como lo señala Ramírez (2007), que encuentra que un 70% de la varianza de los resultados SIMCE se explica por la diferencia entre estudiantes que asisten a una misma escuela.

De tal manera, si los establecimientos tienen altos/bajos resultados en su totalidad o en un curso respectivo surgen las siguientes interrogantes: Los buenos/malos resultados, ¿son producto de buenos/malos establecimientos, buenos/malos docentes y/o buenos/malos estudiantes?; las diferencias de resultados al interior de un establecimiento, ¿pueden deberse a segregación interna, a efectos de diferentes docentes o a diferentes capacidades de los estudiantes?; ¿cuánto del resultado final es explicado por cada uno de los factores mencionados?

Estas son preguntas a responder, sin dejar de considerar que las limitaciones de los datos y de los mecanismos de selección dificultan el ejercicio. Por lo que, aun cuando no es posible establecer causalidad si es posible identificar asociaciones entre fenómenos.

En Chile existen diversos instrumentos de medición de los docentes que se desempeñan en el sector educativo, siendo el de mayor cobertura el de la Evaluación Docente que se aplica de manera obligatoria a los docentes del sector municipal. Algunas investigaciones han intentado responder las interrogantes planteadas usando éste y otros instrumentos. Eisenberg (2008), utilizando información del Sistema de Medición de la Calidad de la Educación (SIMCE)² de 4° básico 2005 y los resultados de la Evaluación Docente hasta el año 2005, encuentra que una alta proporción de docentes de alto desempeño tiende a ejercer en los establecimientos de mejor nivel socioeconómico y mejor desempeño de los estudiantes. A su vez, encuentra que un punto más en el resultado de desempeño docente está asociado con 19 puntos más del estudiante en el SIMCE promedio.

En un estudio similar, Bravo y otros (2009) analizan la Evaluación Docente así como los programas de Asignación Variable por Desempeño Individual (AVDI) y Asignación de Excelencia Pedagógica (AEP) para los años 2002, 2003 y 2006. Con el fin de corregir en parte el problema de distribución no aleatoria de docentes entre establecimientos, los autores controlan por los resultados del SIMCE del establecimiento en años anteriores. Los autores señalan que los docentes bien evaluados tienen una mayor probabilidad de tener estudiantes con un alto desempeño en la prueba SIMCE. En particular indican que los estudiantes de 4° básico que tienen un docente calificado como Destacado obtienen entre 11% y 20% desviaciones estándar más en los resultados de la prueba SIMCE que aquellos que tienen un docente con nivel Insatisfactorio.

En este sentido, un ejercicio para determinar el desempeño de los docentes requiere comprender cómo los mecanismos de evaluación dan cuenta de su calidad y como se relacionan los resultados de esas evaluaciones con los aprendizajes de los estudiantes. A lo largo de este documento se analizará la Evaluación Docente, estudiando como ésta da cuenta del desempeño de los docentes y como se relaciona ese desempeño con rendimiento de sus estudiantes en la prueba SIMCE.³⁴

¹ En Chile, sólo a partir de la entrada en vigencia de la Ley General de Educación en el año 2009 se prohíbe la selección por parte de los establecimientos de estudiantes hasta 6° básico.

² Mayor detalle sobre este instrumento de evaluación en sección 3, apartado 3.1.

³ Esto también puede ser entendido como la efectividad de los docentes.

Por lo tanto, en el marco de este diagnóstico y de los desafíos del Ministerio de Educación (MINEDUC), entre los cuales se considera estudiar mejoras al sistema de Evaluación Docente municipal de manera de avanzar en el vínculo entre el desempeño de los estudiantes y los conocimientos de los docentes, se realiza el presente estudio. Éste intenta responder, de manera principal, dos interrogantes:

- ¿Cuánto discrimina y/o explica cada instrumento de la Evaluación Docente (portafolio, entrevista por un evaluador par, pauta de autoevaluación, informes de referencia de terceros) el resultado final de ésta?
- ¿Cómo se relaciona el desempeño en la Evaluación Docente –a nivel agregado y el de cada instrumento- con el rendimiento de los estudiantes en SIMCE?

En síntesis, este trabajo analiza los resultados de la Evaluación Docente, profundizando en cada instrumento. Para ello, en la siguiente sección se caracteriza brevemente a los docentes y se describe la Evaluación Docente y sus instrumentos. Luego, en la sección 3 se presentan las fuentes de información y las principales variables utilizadas. La sección 4 describe la cobertura y principales resultados de la Evaluación Docente, así como la relación incondicionada entre sus resultados y el desempeño de los estudiantes. En la sección 5 se describen los modelos empíricos utilizados en la investigación y se muestran los resultados de los análisis a los instrumentos, así como la relación existente entre éstos y el desempeño de los estudiantes. Finalmente, en la sección 6 se presentan las conclusiones de los autores. Las secciones 7 y 8 presentan las referencias bibliográficas y los anexos.

2. Los docentes en Chile y la Evaluación Docente en el sistema municipal

2.1 Docentes en Chile y el sector municipal

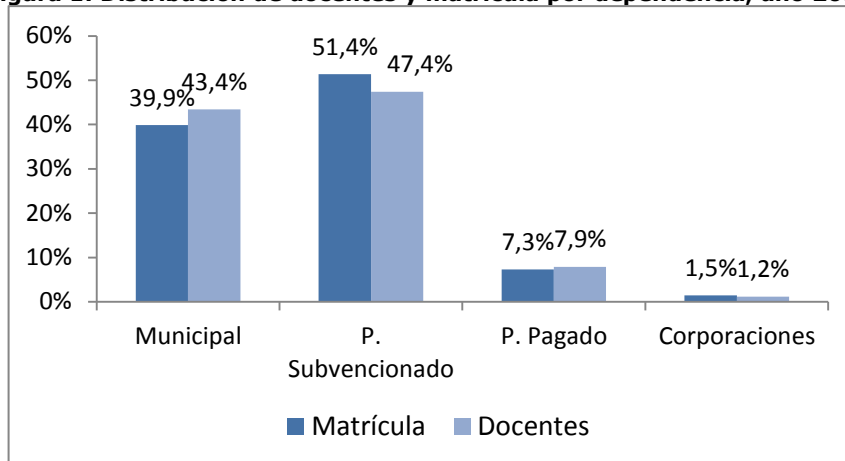
En Chile, en el año 2011 se registra un total de 191.868 docentes en el sistema educacional, que se distribuyen entre las distintas dependencias administrativas según se muestra en la Figura 1. Se observa que los establecimientos subvencionados por el Estado concentran el mayor número de docentes, lo que se relaciona estrechamente con la mayor concentración de matrícula. Específicamente, en los establecimientos municipales ejercen 84.570⁵ docentes, donde 81% corresponde a docentes de aula.

Los docentes en ejercicio el año 2011 tienen en promedio 43 años de edad, 15 años de servicio en el sistema y 8,5 años en el mismo establecimiento (Tabla 1). La edad promedio es mayor en los establecimientos municipales, lo que es consistente con contar en su dotación con 19,5% (16.500) de docentes en edad de jubilar o cercanos a ella, mientras que en los establecimientos particulares estos se encuentran en torno a 8%.

⁴ Si bien sólo se aborda la Evaluación Docente en este estudio, es posible recalcar que es el único instrumento de evaluación obligatorio para los docentes de establecimientos municipales. Todos los otros instrumentos son voluntarios y tienen baja cobertura. Por lo que su estudio es fundamental para entender el desempeño de los docentes.

⁵ De ellos 3% (2.541) también ejerce en otra dependencia administrativa.

Figura 1: Distribución de docentes y matrícula por dependencia, año 2011.



Fuente: Elaboración propia en base a registros administrativos del MINEDUC.

Nota 1: Si el docente ejerce en más de un establecimiento, se considera la dependencia del establecimiento donde tiene más carga horaria. Si tiene una carga horaria superior a 40 horas en más de 1 establecimiento el docente es considerado en todas las dependencias donde iguale o supere dicha carga. En el caso de tener la misma máxima carga horaria en dos o más establecimientos, se considera al docente en todas las dependencias en que esto ocurra.

Nota 2: Los establecimientos particulares pagados declaran voluntariamente su dotación, por lo que los datos podrían no ser exactos.

Nota 3: Corporaciones se refiere a las corporaciones de administración delegada.

Tabla 1: Años de servicio y edad de los profesores, según dependencia administrativa de sus establecimientos educacionales, año 2011.

Dependencia	Años de servicio	Años de servicio en mismo establecimiento	Edad	N° de docente mayores de 59 años
Municipal	18,8	10,7	46,7	16.542 (19,5%)
Particular Subvencionado	11,8	6,7	39,8	7.015 (7,6%)
Particular Pagado	14,4	8,1	42	1.274 (8,1%)
Corporaciones	18,8	13,4	46,8	577 (23,7%)

Fuente: Elaboración propia en base a registros administrativos del MINEDUC.

Nota: Los docentes son contabilizados sólo una vez dentro de cada dependencia administrativa en la que trabajan (ver nota 1 en la Figura 1).

Respecto al género de los docentes, 72% son mujeres. La participación femenina en los establecimientos municipales es de 70%, siendo mayor particularmente en los sectores de educación de párvulos (99%) y educación básica (74%). La composición de género de la planta docente es más balanceada en educación media Científico-Humanista (54,4% son mujeres) y Técnico Profesional (51,4% son mujeres).⁶

En su mayoría, los docentes que ejercen al interior del aula se encuentran titulados en el área de la Educación (94,6%), sólo 3% están titulados en otras áreas y menos de 3% no cuenta con ningún título profesional. Esta distribución es similar entre los establecimientos, con la excepción de las corporaciones de administración delegada (Tabla 2).

⁶ La menor participación es en el área Técnico Profesional Industrial con 35%. Sin embargo, esa área sólo agrupa a 2.400 docentes.

Tabla 2: Caracterización de docencia en aula, según condición de titulación y dependencia administrativa 2011.

Tipo de Título	Dependencia administrativa				
	Nº docencias aula	Municipal	P. Subvencionada	P. Pagada	Corporaciones
Nº de docencias en aula	165.246	70.846	78.665	13.690	2.045
Titulados en educación	156.371	67.275	73.994	13.398	1.704
	94,6%	95,0%	94,1%	97,9%	83,3%
Titulados en otras áreas	5.018	2.204	2.369	187	258
	3,0%	3,1%	3,0%	1,4%	12,6%
No titulados	3.857	1.367	2.302	105	83
	2,3%	1,9%	2,9%	0,8%	4,1%

Fuente: Elaboración propia en base a registros administrativos del MINEDUC.

Nota 1: Si el docente ejerce en más de un establecimiento, se considera la dependencia del establecimiento donde tiene más carga horaria. Si tiene una carga horaria superior a 40 horas en más de 1 establecimiento el docente es considerado en todas las dependencias donde iguale o supere dicha carga. En el caso de tener la misma máxima carga horaria en dos o más establecimientos, se considera al docente en todas las dependencias en que esto ocurra.

Nota 2: Corporaciones se refiere a las corporaciones de administración delegada.

2.2 Evaluación Docente

La Evaluación Docente es un sistema de carácter formativo que evalúa a los profesionales de la educación municipal que se desempeñan en funciones de docencia de aula. Está orientada a fortalecer la profesión docente y contribuir a mejorar la calidad de la educación. Se realiza atendiendo los dominios, criterios y descriptores fijados en el Marco para la Buena Enseñanza aprobado por el Ministerio de Educación, mediante la aplicación de los instrumentos definidos por el Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP).⁷

El Sistema de Evaluación del Desempeño Profesional Docente comienza en junio del año 2003 a partir del Acuerdo Marco tripartito suscrito por el Ministerio de Educación, la Asociación Chilena de Municipalidades y el Colegio de Profesores de Chile. Posteriormente, este acuerdo se plasmó a través de la ley N° 19.933, del 30 de enero de 2004, que reemplaza el antiguo sistema de calificaciones por un sistema de evaluación de carácter formativo basado en estándares de desempeño para los profesionales de la educación que se desempeñan en funciones de docencia de aula.

La ley N° 19.961, del 9 de agosto de 2004, sobre Evaluación Docente, determina la estructura central del sistema de evaluación del desempeño profesional docente y autoriza la dictación de un reglamento para definir los aspectos técnicos del sistema. Con posterioridad, el sistema de evaluación ha sido modificado por la ley N° 20.158, del 27 de diciembre de 2006, que introduce la facultad de eximirse de la Evaluación Docente a quienes se encuentren a menos de tres años de cumplir la edad legal para jubilar. Recientemente, la ley N° 20.501, del 8 de febrero de 2011, modifica las consecuencias de la evaluación para los docentes que obtienen desempeños de nivel Básico o Insatisfactorio en el proceso de evaluación del año 2011 en adelante, estableciendo que:

"Cada vez que un profesional de la educación resulte evaluado con desempeño insatisfactorio, deberá ser sometido al año siguiente a una nueva evaluación, pudiendo el sostenedor exigirle que deje la responsabilidad de curso para trabajar durante el año en su plan de superación profesional, debiendo el empleador asumir el gasto que representa el remplazo del docente en aula. Si el desempeño en el nivel insatisfactorio se mantuviera en la segunda evaluación consecutiva, el profesional de la educación dejará de pertenecer a la

⁷ Ver www.docentemas.cl

dotación docente. Los profesionales de la educación que resulten evaluados con desempeño básico deberán evaluarse al año subsiguiente, pudiendo el sostenedor exigirle que deje la responsabilidad de curso para trabajar durante el año en su plan de superación profesional, debiendo el empleador asumir el gasto que representa el remplazo del docente en aula. En caso de que resulten calificados con desempeño básico en tres evaluaciones consecutivas o en forma alternada con desempeño básico o insatisfactorio durante tres evaluaciones consecutivas, dejará de pertenecer a la dotación docente."

Adicionalmente, la nueva ley postula que "para los efectos de lo dispuesto en el inciso segundo letra a) del artículo 7° bis de esta ley, se entenderá por mal evaluado a quienes resulten evaluados con desempeño insatisfactorio o básico". Por medio de esta ley se le entrega al director la facultad de remover hasta el 5% de los docentes del establecimiento si ellos han sido mal evaluados.

Los estándares de desempeño y las normas son elementos clave en el proceso de Evaluación Docente. En términos generales, un estándar puede ser tanto una meta como una medida de progreso hacia esa meta. En específico, para la Evaluación Docente cada estándar se subdivide en un conjunto de cuatro niveles de desempeño:⁸

- **Destacado:** Indica un desempeño profesional que clara y consistentemente sobresale con respecto a lo que se espera en el indicador evaluado. Suele manifestarse por un amplio repertorio de conductas respecto a lo que se está evaluado o bien por la riqueza pedagógica que se agrega al cumplimiento de los indicadores.
- **Competente:** Indica un desempeño profesional adecuado en el indicador evaluado que cumple con lo requerido para ejercer profesionalmente el rol docente. Aun cuando no es excepcional, se trata de un buen desempeño.
- **Básico:** Indica un desempeño profesional que cumple con cierta irregularidad el conjunto de indicadores evaluados o con regularidad la mayoría de estos.
- **Insatisfactorio:** Indica un desempeño que presenta claras debilidades para el conjunto de indicadores evaluados que afectan significativamente el quehacer docente.

Para los docentes la evaluación significa un proceso de auto-reflexión respecto de su desempeño profesional, que les permite identificar tanto fortalezas como debilidades, y proyectar un camino para potenciar las primeras y superar las segundas. Específicamente:

- Los profesores evaluados como Destacados y Competentes tienen acceso prioritario a oportunidades de desarrollo profesional: ventajas en concursos, pasantías en el extranjero, opción a ser profesores guías de talleres, posibilidad de participación en seminarios académicos, entre otros. Al mismo tiempo tienen la posibilidad de optar a percibir la Asignación Variable por Desempeño Individual (AVDI), previa rendición de una prueba escrita de conocimientos disciplinarios y pedagógicos.
- Los profesores que obtienen en su evaluación global el nivel de Básico o Insatisfactorio disponen de Planes de Superación Profesional⁹ gratuitos destinados a superar sus debilidades. Los planes son diseñados y ejecutados por los respectivos sostenedores municipales y financiados por el Ministerio de Educación.
- Previo a las modificaciones legales de febrero de 2011, el reglamento estipulaba que los docentes que resultaban evaluados con desempeño Insatisfactorio debían someterse a nueva evaluación al año siguiente

⁸ Ver Reglamento Evaluación Docente, con fecha 30 de Agosto de 2004, N° 192.

⁹ Ver Apéndice A.

conforme a los planes de superación profesional que determine el reglamento. Si en la segunda evaluación el resultado era nuevamente insatisfactorio, el docente dejaría la responsabilidad de curso para trabajar durante el año en su plan de superación profesional. No era hasta la tercera evaluación insatisfactoria que el profesional dejaba de pertenecer a la dotación docente. A su vez, los docentes que obtuvieran resultados Básicos debían evaluarse luego de cuatro años en lugar del año subsiguiente como lo estipula la Ley 20.501 que rige desde febrero del año 2011.¹⁰

La evaluación se realiza basada en las evidencias del desempeño profesional de cada docente recogidas a través de cuatro instrumentos: (i) pauta de autoevaluación, (ii) entrevista por un evaluador par, (iii) informe de referencia de terceros y (iv) portafolio de desempeño pedagógico. Todos estos instrumentos deben guardar relación con los dominios, criterios y descriptores fijados en el Marco para la Buena Enseñanza y contar con la aprobación previa del CPEIP del Ministerio de Educación.

- a. *Autoevaluación* tiene como principal función que el docente reflexione sobre su práctica pedagógica y valore su propio desempeño profesional a partir de una pauta previamente determinada por el CPEIP. El docente puede incluir información contextual que influya en su práctica. Su ponderación en la evaluación final es del 10% si es la primera evaluación cada 4 años y de 5% en caso de ser la segunda o tercera evaluación dentro de ese período.
- b. *Entrevista por un evaluador par* se aplica mediante un evaluador par previamente capacitado que utiliza una pauta de entrevista estructurada para recoger información respecto de la práctica del docente evaluado. La entrevista consta de dos partes: preguntas acerca la propia práctica pedagógica y preguntas respecto del contexto de trabajo del docente. Una vez finalizada la entrevista, el evaluador par aplica una rúbrica asignando en base a ella, alguno de los cuatro niveles de desempeño a cada una de las preguntas de la entrevista. Esto permite contar con la apreciación del desempeño de un docente de aula del mismo nivel y modalidad que el docente evaluado. Su ponderación en la evaluación final es de 20% si es la primera evaluación cada 4 años y de 10% en caso de existir una evaluación previa de nivel Insatisfactorio.¹¹
- c. *Informe de referencia de terceros* consiste en una pauta estructurada que debe ser completada por el director y jefe de la Unidad Técnico Profesional (UTP) del establecimiento. Esta pauta contiene distintas preguntas referidas al quehacer del docente evaluado y para cada pregunta, el director y el jefe de UTP deben emitir, por separado, su evaluación del desempeño del docente en uno de los cuatro niveles de desempeño. Esto permite conocer la evaluación de los superiores jerárquicos respecto a la práctica del docente evaluado. Este instrumento también recoge información contextual. Su ponderación en la evaluación final es de 10% si es la primera evaluación cada 4 años y de 5% en caso de existir una evaluación previa de nivel Insatisfactorio. Adicionalmente esta información es utilizada para retroalimentar al docente en el informe individual que recibe al inicio del año escolar siguiente de la evaluación.

¹⁰ Cabe mencionar que existen cláusulas de Eximición – Suspensión de la Evaluación Docente (siguen vigentes luego de las modificaciones de la Ley 20.501). Son susceptibles de eximición (i) los evaluadores pares de ese proceso anual, teniendo obligatoriedad de evaluación al año siguiente, (ii) los docentes durante su primer año de ejercicio profesional. Asimismo, son susceptibles de suspensión los docentes que soliciten hacerlo por (i) fuerza mayor, (ii) traslado de establecimiento posterior al 30 de abril, (iii) permiso sin goce de remuneraciones superior a tres meses y (iv) encontrarse fuera del territorio nacional por perfeccionamiento profesional.

¹¹ El evaluador par debe ser un docente de aula con al menos 5 años de experiencia, pertenecer al mismo nivel escolar, sector del currículo y modalidad que el docente evaluado, pero que no ejerza funciones directivas o técnicas pedagógicas en el mismo establecimiento educacional del docente evaluado. Debe haber sido previamente evaluado como destacado o competente o ser docente con derecho a percibir la AEP o ser docente integrante de la Red Maestros de Maestros. Para mayor detalle ver Reglamento Evaluación Docente, con fecha 30 de Agosto de 2004, N° 192.

- d. *Portafolio de desempeño pedagógico* tiene como función recoger, a través de productos estandarizados, evidencia verificable respecto de las prácticas de desempeño del docente evaluado. Contempla la presentación de productos escritos, en los cuales el docente debe reportar distintos aspectos de su quehacer profesional, así como un registro audiovisual de su desempeño mediante la grabación en video de una clase de cuarenta minutos de duración. La corrección de las evidencias del instrumento Portafolio es realizada por profesores correctores del mismo nivel y subsector de aprendizaje, entrenados para el desempeño de esta función. Su ponderación en la evaluación final es del 60% si es la primera evaluación cada 4 años y de 80% en caso de existir una evaluación previa de nivel Insatisfactorio. Mediante este instrumento se espera conocer la práctica pedagógica del docente evaluado. El portafolio se compone de 8 dimensiones y cada dimensión de entre 2 y 5 sub-dimensiones.¹²

3. Fuentes de datos y variables

Para el análisis de la Evaluación Docente respecto a sus instrumentos y a la relación de sus resultados con el desempeño de los estudiantes se utilizan principalmente tres fuentes de información: SIMCE 2008, resultados Evaluación Docente 2004-2010 y, a fin de identificar a los docentes y sus trayectorias laborales, la base de datos de Idoneidad Docente 2005-2008. A su vez, para las estimaciones posteriores se debió construir otras bases de datos, así como seleccionar el conjunto de variables a ser utilizadas como se detalla posteriormente.

3.1 Fuentes de Información

Sistema de Medición de la Calidad de la Educación (SIMCE)¹³

El SIMCE es el sistema nacional de evaluación de resultados de aprendizaje del Ministerio de Educación de Chile. Su propósito principal es contribuir al mejoramiento de la calidad y equidad de la educación, informando sobre el desempeño de los estudiantes en diferentes áreas de aprendizaje del Currículum Nacional y relacionando estos desempeños con el contexto escolar y social en que aprenden.

La prueba SIMCE es el único test estandarizado y censal que se realiza en el país. Esta prueba es aplicada cada año en diferentes niveles: 4° básico todos los años, así como 8° básico y 2° medio, intercalando un año por medio. Se evalúan diferentes subsectores dependiendo del curso a analizar, pero en todos ellos se evalúa Matemática y Lectura.

Las pruebas están diseñadas para que sus resultados sean comparables en el tiempo en un mismo nivel y sector. Son aplicadas de manera tal que los resultados sean representativos para el establecimiento, respondiendo los alumnos de un mismo nivel preguntas sobre diferentes ámbitos de los contenidos abordados en cada sector del aprendizaje. Esto podría traer limitaciones a la hora de estudiar los efectos individuales en las pruebas dado que estas no necesariamente son comparables entre estudiantes, pero aun así es la mejor aproximación a nivel nacional del desempeño de cada estudiante.

Para los efectos de este trabajo, se opta por estudiar el SIMCE de 4° básico, porque comúnmente en este nivel un solo profesor es el responsable de enseñar todas las

¹² Para mayor detalle ver Apéndice A2.

¹³ Para mayor información ver www.simce.cl

materias (profesor generalista), por lo tanto la vinculación entre profesor y alumno es más directa y completa.¹⁴

Las bases de datos incluyen información sobre los alumnos, sus padres de familia o apoderados y su establecimiento:

- Características individuales de los estudiantes, como género y edad y los resultados que obtuvieron en cada una de las pruebas.
- Características de los establecimientos como comuna, dependencia, nivel socioeconómico (grupo) y número de alumnos en el curso.
- Características de las familias como escolaridad de los padres, ingresos del hogar, número de libros, bienes, prácticas académicas y conocimiento del establecimiento.

Evaluación Docente

La base de datos de la Evaluación Docente agrupa información de los resultados en cada uno de los instrumentos de esta evaluación entre el año 2004 y el 2010. El resultado final de la evaluación se reporta en términos de puntaje total y de la categoría a la que el docente es asignado. También se señala el ciclo de enseñanza y el subsector de aprendizaje donde el docente fue evaluado.

Los instrumentos de evaluación de los pares, informe de referencia de terceros y autoevaluación tienen una escala de calificaciones discretas que va 1 a 4 según el nivel de desempeño del docente, el instrumento de portafolio tiene puntajes continuos ente 1 y 4 que resultan de la combinación de calificaciones de sus ocho dimensiones.

Aun cuando los instrumentos con calificaciones discretas guardan relación 1 a 1 entre puntaje y categoría de desempeño, a nivel agregado los resultados reflejan 2 o menos puntos como Insatisfactorio, entre 2 y 2,5 nivel de Básico, entre 2,5 y 3,5 Competente y mayor a 3,5 Destacado.

Base de Datos de Idoneidad Docente

Almacena información de los docentes en ejercicio en todo el sistema educacional desde 2003 a 2011 que permite corroborar el carácter de idóneo del docente según lo especificado en la Ley General de Educación.

En esta fuente de información administrada por el Ministerio de Educación es posible identificar el establecimiento en el cual ejerce el docente, su función principal, el área de estudio que enseña, si se encuentra o no titulado y los años de experiencia, entre otras características.

Otras fuentes de información

Las bases de datos de los Registros de Estudiantes de Chile (RECH) y el Sistema de Información General de Estudiantes (SIGE) que contienen información de los alumnos y profesores como resultados académicos de los estudiantes por asignatura y promedio final de notas, si fue o no promovido, asistencia a clases, así como otras variables desde 2003 a la fecha. Del mismo modo se informa de todos los cursos y asignaturas en que el profesor ha impartido clases.¹⁵

¹⁴ En 8° básico y enseñanza media hay un profesor por asignatura, lo que hace difícil establecer el aporte de cada docente a un resultado específico por alumno. Por ejemplo, en el caso de los resultados SIMCE de Matemática podría estar influyendo no sólo el desempeño del profesor de Matemática, sino que también el de Lectura y Comunicación, Ciencias o de otras disciplinas afines. Por otro lado, el SIMCE de 4° básico igualmente puede tener limitaciones porque el estudiante puede cambiar de profesor entre 1° y 4°, pero esto se salvaguarda con los modelos de estimación presentados más adelante.

¹⁵ Respecto de la validez de los datos, sólo la identificación de los alumnos y profesores es validada por el Registro Civil e Identificación de Chile. El resto de la información pasa por algunos filtros de validación

3.2 Construcción de las bases de datos

Para llevar a cabo el análisis de la relación entre el desempeño de los estudiantes y el de los docentes se construyen dos bases de datos adicionales. La primera vincula a cada estudiante de cuarto básico del año 2008 con sus respectivos docentes entre 1° y 4° básico. La segunda, es una base de datos a nivel de establecimiento que vincula los resultados promedio de SIMCE 2008 en cada establecimiento con el desempeño de los docentes de primer ciclo en la Evaluación Docente.

Para el primer enfoque, el vínculo individual de estudiantes con sus docentes entre 1° básico (año 2005) y 4° básico (año 2008), se utilizan los resultados de las pruebas SIMCE tanto de Matemática como de Lectura a nivel individual. Para el segundo enfoque, se requieren sólo los resultados a nivel de establecimiento.¹⁶

3.3 Principales variables

Docentes

- *Desempeño promedio en la Evaluación Docente:* A partir del puntaje en cada instrumento, del puntaje final de la evaluación y de la categoría de evaluación correspondiente a su desempeño se construyen variables que promedian el resultado de los distintos docentes del primer ciclo básico que imparten clases a cada estudiante entre los años 2005 y 2008 en el establecimiento educacional o el de todos los docentes de primer ciclo año 2008. Esta variable será representada mediante diferentes indicadores, continuos y discretos, según como se explica en el apartado 5.
- *Años de experiencia promedio en el sistema educacional:* Se calcula el promedio de los años de experiencia total en el sistema educacional de los distintos docentes del primer ciclo básico que han dado clases a cada estudiante entre los años 2005 y 2008.
- *Edad promedio.* Se calcula el promedio de edad de los distintos docentes del primer ciclo básico que han dado clases a cada estudiante entre los años 2005 y 2008.
- *Concentración de docentes insatisfactorios, básicos, competentes y destacados.* Se calcula el porcentaje de docentes evaluados y clasificados en cada categoría que han ejercido en el primero ciclo básico de cada establecimiento educacional entre los años 2005 y 2008. Para cada categoría de desempeño se construye una variable que refleja la proporción.

Estudiantes

- *SIMCE individual.* Dos variables continuas para cada estudiante. Una para el resultado SIMCE de Matemática y otra para el de Lectura.
- *Ranking promedio.* Se calcula como el promedio de los rankings del estudiante en cada uno de los años en su curso ordenados por el promedio final de notas.
- *Dummy de asistencia a pre escolar.* Variable dicotómica que toma valor 1 cuando el estudiante asistió a pre básica y 0 en caso contrario.

Establecimientos

- *SIMCE promedio 2005-2007.* Se calcula el resultado promedio de SIMCE 4° básico de cada establecimiento donde el alumno estudió entre los años 2005 y 2007.¹⁷
- *Tamaño promedio de la cohorte de 4° Básico del establecimiento.* Indica el número promedio de estudiantes de 4° básico en cada establecimiento.

en el Ministerio de Educación. Con esta base es posible vincular a cada alumno con sus profesores en cada año.

¹⁶ Para más detalle sobre las bases de datos ver Apéndice C.

¹⁷ Para evitar duplicidad de niños por haber estado en más de un establecimiento, se utiliza la base de datos de actas, filtrada por asistencia y promedio general de notas, luego se revisa por consistencia con la base del SIMCE a nivel individual, quedando sólo aquellos donde hay certeza del establecimiento en que finalizaron el año escolar.

- *Grupo socioeconómico del establecimiento.* Esta definición es realizada por el Ministerio de Educación. Se define como establecimiento tipo A, B, C, D o E; donde A corresponde a establecimientos de bajo nivel socioeconómico y E de alto nivel socioeconómico.

4. Evaluación Docente: Evolución y principales resultados

4.1 Caracterización de la Evaluación Docente

Desde el año 2004 a 2010 se han realizado 79.763 evaluaciones, siendo la educación básica el primer nivel educacional en incorporarse al sistema, y la educación de párvulos y la educación especial los más recientes (Tabla 3).

Tabla 3: Número de evaluaciones de docentes por ciclo y período.

Ciclo	Período							Total
	2004	2005	2006	2007	2008	2009	2010	
1er ciclo e. básica	920	8.027	4.937	3.137	2.061	5.164	3.269	27.515
2do ciclo e. básica	799	2.291	8.544	3.650	5.557	4.421	5.194	30.456
E. media	0	347	709	3.626	4.336	3.631	1.697	14.346
E. especial	0	0	0	0	0	1.584	612	2.196
E. parvularia	0	0	0	0	4.061	900	289	5.250
Total	1.719	10.665	14.190	10.413	16.015	15.700	11.061	79.763

Fuente: Elaboración propia con registros del CPEIP.

Las 79.763 evaluaciones corresponden a un total de 65.950 docentes. A su vez, éstos corresponden a 53.164 docentes que se encontraban en ejercicio en el sector municipal el año 2011, lo que equivale a un 63% de los docentes registrados en el sistema municipal ese año. Cabe mencionar que la menor cantidad de docentes que evaluaciones se debe a que algunos docentes han sido evaluados más de una vez, como se observa en la Tabla 4.

Tabla 4: Número de evaluaciones a docente en ejercicio el año 2011, según ciclo. 2004-2010.

Nivel	N° de evaluaciones				Total
	1	2	3	4	
E. parvularia	4.619	35	1	0	4.655
1er ciclo e. básica	9.382	6.038	82	0	15.502
2do ciclo e. básica	14.103	5.797	75	3	19.978
E. media	10.496	695	12	0	11.203
E. especial	1.769	57	0	0	1.826
Total	40.369	12.622	170	3	53.164

Fuente: Elaboración propia con registros CPEIP.

Nota: Se registran 79.763 evaluaciones y 53.164 docentes evaluados, lo que se debe a que hay docentes que tienen más de una evaluación como se observa en esta tabla.

Del total de docentes evaluados y que se encuentran en ejercicio en los distintos sectores al año 2011¹⁸ se observa que la educación de párvulos, si bien ha sido una de las últimas en incorporarse a la Evaluación Docente, es la que tiene mayor tasa de participación (Tabla 5).

¹⁸ Hay docentes que pueden ejercer en más de un sector, habiendo sido evaluados en alguno de los sectores donde ejerce pero no en todos.

Tabla 5: Número y porcentaje de docentes del sector municipal evaluados según nivel en el que ejercen, año 2011.

	Docentes no evaluados	Docentes evaluados	Total
E. parvularia	1.045	4.580	5.625
	19 %	81%	100%
Educación básica	16.905	36.806	53.711
	31%	69%	100%
Educación especial	2.112	213	2.325
	91%	9%	100%
Educación media HC	4.753	8.577	13.330
	36%	64%	100%
Educación media TP	5.078	2.994	7.072
	63%	37%	100%
Educación básica adultos	423	120	543
	78%	22%	100%
Educación media HC adultos	1.163	437	1.600
	73%	27%	100%
Educación media TP adultos	264	64	328
	81%	19%	100%
Total	31.743	53.791	85.534
	37%	63%	100%

Fuente: Elaboración propia con registros de la base de datos de Idoneidad Docente 2011 (censal).

Nota 1: Se considera a los docentes tantas veces según sectores ejerzan, pero sólo una vez dentro de cada sector. Si se encuentra evaluado en uno de esos sectores se le clasifica en la categoría de evaluado en todos los sectores que ejerce. Aproximadamente 482 (0,6%) docentes ejercen en más de un sector.

Nota 2: La sigla HC: Científico Humanista. La sigla TP: Técnico Profesional.

Respecto a los docentes que aún no han sido evaluados, 68% (21.215) no ejerce funciones en el aula, 2,6% tiene menos de 2 años en el sistema y 6% se encuentra en sectores diferentes a educación básica o educación media científico humanista.¹⁹

Surge entonces la necesidad de identificar si existen diferencias observables entre los evaluados y no evaluados que se encuentran en condiciones similares. Se identifica que los docentes evaluados difieren en sus características observables respecto de aquellos que no se han sometido a ninguna evaluación (Tabla 6). Si se considera sólo a los docentes del ciclo básico que cumplen funciones de docente en aula y que tienen al menos 2 años de participación en el sistema²⁰, se encuentran diferencias importantes en la proporción de docentes sin título a favor de los no evaluados, así como para los evaluados es mayor la cantidad de años de servicio en el establecimiento aunque no necesariamente en el sistema. Esto puede ser explicado por docentes que ejercían anteriormente en un establecimiento de otra dependencia administrativa. A su vez, se observa que, en términos relativos, en la Región Metropolitana los docentes se han evaluado más respecto a los docentes las otras regiones. Cabe mencionar que no se observan diferencias relevantes en los resultados de la Prueba de Selección Universitaria (PSU) o de la Prueba de Aptitud Académica (PAA) de los docentes evaluados y de aquellos que no lo han sido.

¹⁹ Estos sectores llevan más tiempo sometidos a la evaluación, por lo que la ausencia de docentes de otros sectores puede responder a la temporalidad de la misma.

²⁰ Incluir a todos los ciclos de enseñanza y todas las posibles funciones que el docente desarrolla dentro de un establecimiento podría esconder distintas razones de no evaluó, como una situación intrínseca a la periodicidad de ingreso a la evaluación, o no haber participado por no cumplir la función de docentes en aula.

Tabla 6: Diferencias entre evaluados y no evaluados, considerando a docentes de aula del 1er ciclo básico que se encuentran en ejercicio en el año 2011.

	Evaluado (1) N=28.449	No Evaluado (2) N=4.680	Diferencia (1) - (2)
Mujer (d=1) %	75,7	78,5	-2,8
Años de servicio total	21,7	20,7	1
Años de servicio mismo establecimiento	13,5	7,2	6,3
Sin título Profesional (d=1) %	0,5	2,6	-2,1
Con título en Educación (d=1) %	99,7	98,8	0,9
PAA 1989-2003*	481	484	-3
PSU 2003-2010**	482	479	3
Impartir clases también en otra dependencia. (d=1) %	1,8	2,2	-0,4
Impartir clases en RM. (d=1) %	23,6	18,2	5,4

Fuente: Elaboración propia con registros MINEDUC (censal).

Nota: Se considera una sola vez a cada docente e incluye sólo a docentes que tienen al menos 2 años de ejercicio.

* N° evaluados=6.652, N° No evaluados=2.046.

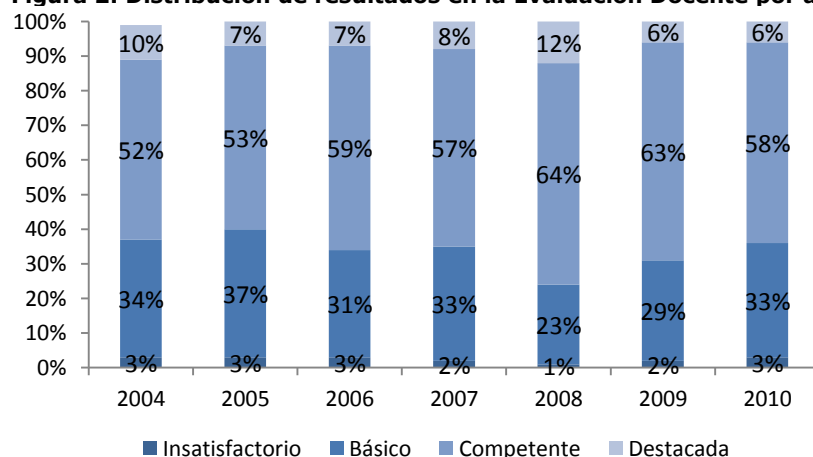
** N° evaluados=400, N° No evaluados=598.

4.2 Análisis descriptivo de los resultados en la Evaluación Docente

En esta sección se presentan los resultados de la Evaluación Docente en los distintos años y desagregando por características de los docentes.

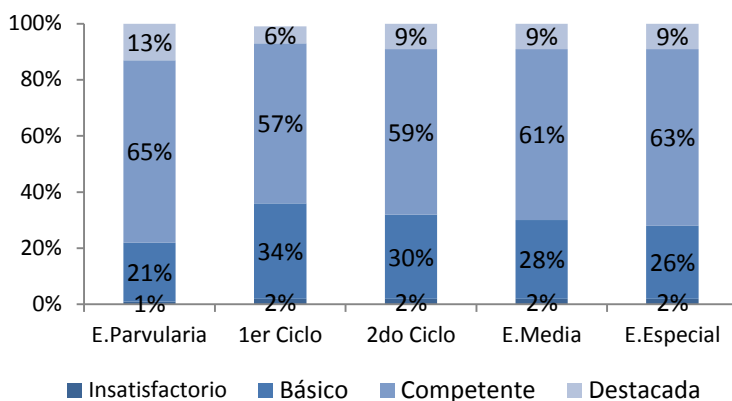
Según la distribución de resultados de la Evaluación Docente en el período 2004-2010, el porcentaje de docentes calificados como Insatisfactorio no sobrepasa el 3% anual, mientras que el porcentaje de docentes Destacados fluctúa entre 6 y 12%. En el año 2008 los docentes presentan mejor desempeño que el resto de los períodos, con un 76% calificado como Competente o Destacado (Figura 2). Cabe notar que el año 2008, cuando la educación parvularia comenzó a ser evaluada, la mayor parte de los docentes de dicho nivel participaron en la evaluación, lo que podría haber afectado a los resultados, puesto que, al comparar los resultados de los docentes entre los distintos ciclos se observa que educación de párvulos es la que presenta una mayor proporción de profesores con resultados en los niveles Competente y Destacado (78%), mientras que primer ciclo básico presenta la menor proporción (63%) (Figura 3).

Figura 2: Distribución de resultados en la Evaluación Docente por año.



Fuente: Elaboración propia con registros del CPEIP.

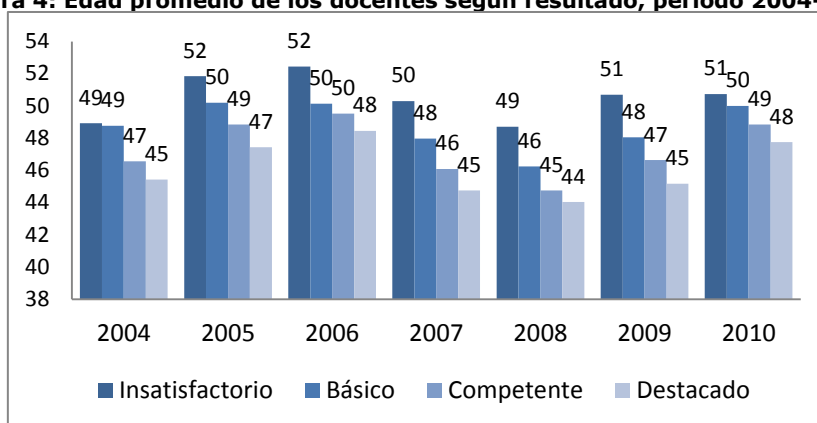
Figura 3: Distribución de resultados en la Evaluación Docente por nivel educativo, 2004-2010.



Fuente: Elaboración propia con registros del CPEIP.

Adicionalmente, entre los años 2004 y 2010 ha existido una relación inversa entre el resultado en la Evaluación Docente y la edad del docente evaluado (Figura 4). Es decir, a mayor edad del docente, menor desempeño en la evaluación. En particular, los docentes evaluados como Destacado son, en promedio, 5 años más jóvenes que los evaluados como Insatisfactorio. Esto sugiere que los nuevos docentes tienen una mejor preparación en términos de las prácticas pedagógicas medidas en la Evaluación Docente²¹ o bien que a más años de servicio disminuye el rendimiento de los docentes en este tipo de mediciones. Esto es consistente con la comparación entre el grupo de docentes en edad de jubilar (mayores a 59 años) con el resto de los docentes, indicando mientras 73% de los docentes más jóvenes logran un resultado de Competente y Destacado, los docentes en edad de jubilar que obtienen esas categorías de resultados corresponden a 63%. La tasa de docentes en nivel Insatisfactorio aunque pequeña, es casi el doble para aquellos en edad de jubilar (1,1%) respecto a los otros docentes (0,6%).

Figura 4: Edad promedio de los docentes según resultado, período 2004-2010.



Fuente: Elaboración propia con registros del CPEIP.

Como se señala en la sección 2, para el período analizado en este documento los docentes son evaluados cada 4 años a menos que su resultado haya sido calificado como Insatisfactorio, en cuyo caso deben ser medidos al año siguiente. Considerando esto y que el inicio de las mediciones registradas fue el año 2004, es posible deducir que aquellos docentes que nunca han tenido un resultado Insatisfactorio podrían haber sido evaluados a lo sumo 2 veces dentro del período 2004-2010 mientras que quienes hayan obtenido 1 ó 2 resultados Insatisfactorio podrían tener 3 ó 4 evaluaciones dentro del mismo período.

²¹ Ver Apéndice B, Figura B.1.

Se observa en la Tabla 7 que del total de docentes que deberían haber sido re-evaluados entre los años 2005 y 2010, 51% lo ha hecho. Del total de docentes evaluados como Insatisfactorio, 69% ha sido re-evaluado.

Tabla 7: Número de docentes evaluados y re-evaluados. Primera evaluación entre años 2004 – 2009 y re-evaluación entre años 2005-2010*.

	2004	2005	2006	2007**	2008**	2009**	Total***
1° Evaluación	1.719	10.631	13.931	10.178	14.890	10.131	61.480
N° Insatisfactorio	52	366	340	174	154	167	1.253
N° Re-evaluados	1.168	6.366	5.765	128	99	86	13.612
N° Insatisfactorio Revaluados	45	290	222	128	99	86	870
% Re-evaluación General	68%	60%	41%	74%	64%	52%	51%
% Re-evaluación Insatisfactorio	87%	79%	65%	74%	64%	52%	69%

Fuente: Elaboración propia con registros del CPEIP.

*: El docente re-evaluado el año 2010 puede haber sido evaluado cualquier año anterior, por lo que se asigna al año en el que fue su primera evaluación anterior.

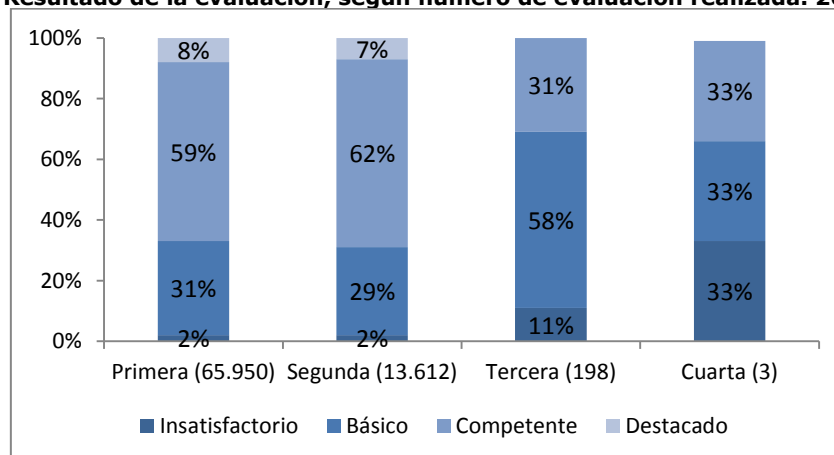
** : En los años 2007, 2008, 2009 sólo deben ser revaluados en el rango de años de estudio los docentes que obtienen un resultado insatisfactorio en la 1ª evaluación, por lo que sobre esa base es calculado el porcentaje de re-evaluación.

***: No es considerado en el análisis el año 2010 debido a que no es posible obtener información de sus re-evaluados. Por lo tanto, el total no considera las 4.470 evaluaciones realizadas por primera vez el año 2010.

De aquellos profesores que no se han re-evaluado, 38% no se encuentra registrado en el sistema y 10% se encuentra trabajando en un establecimiento particular subvencionado y particular pagado. El 52% restante ejerce en el sector municipal, lo que se descompone en 14% (2.026) que no se encuentra ejerciendo labores de docencia en aula y 38% (5.500) que sí lo hace. En otras palabras, del total de docentes evaluados que se encuentran en el sistema municipal el año 2011 (53.164) al menos al 10% le correspondería ser re-evaluado pero no lo ha hecho.

Respecto a los resultados por evaluación, se observa que éstos son similares entre la primera y la segunda, pero varían en la tercera y cuarta evaluación, donde los resultados son en promedio más bajos (Figura 5). Lo anterior es esperable considerando que los docentes con más de dos evaluaciones tienen al menos un resultado Insatisfactorio, por lo que se podría suponer que tienen desempeño promedio menor que los docentes que nunca han obtenido un resultado Insatisfactorio.

Figura 5: Resultado de la evaluación, según número de evaluación realizada. 2004 – 2010.



Fuente: Elaboración propia con registros del CPEIP.

Como se observa en la matriz de transición presentada en la Tabla 8, del total de evaluados por segunda vez, 53% mantiene su resultado inicial, 28% mejora y 14%

empeora. Si se considera exitoso el aumento de desempeño entre evaluaciones y el mantenimiento en caso de los docentes Destacados y Competentes, el nivel de éxito sería de 68%. Por otro lado, si se considera como fracaso la disminución de calificación llegando a niveles de Básico o Insatisfactorio, así como la persistencia en dicho niveles, se tiene un nivel de fracaso de 27%. Específicamente, dentro de la categoría de Insatisfactorio en la 1ª evaluación, 18% vuelve a obtener la misma calificación en su 2ª evaluación mientras que un 66% mejora logrando pasar al nivel de Básico y 16% al nivel de Competente. Estas cifras son relevantes puesto que con la nueva legislación, aquellos maestros que por segunda vez consecutiva obtengan una calificación en el nivel Insatisfactorio deberán salir del sistema (18% en este ejercicio).

Tabla 8: Relación entre 1ª y 2ª evaluación, 2004 – 2010.

		2ª evaluación				Total
		Insatisfactorio	Básico	Competente	Destacado	
1ª evaluación	Insatisfactorio	160	570	138	2	870
		18,4%	65,5%	15,9%	0,2%	100%
	Básico	85	1.688	2.391	119	4.283
		2,0%	39,4%	55,8%	2,8%	100%
	Competente	46	1.615	5.205	638	7.504
		0,6%	21,5%	69,4%	8,5%	100%
	Destacado	1	73	685	196	955
		0,1%	7,6%	71,7%	20,5%	100%
Total		292	3.946	8.419	955	13.612
		100%	100%	100%	100%	100%

Fuente: Elaboración propia con registros del CPEIP.

La Tabla 9 presenta los resultados en la 3ª y 4ª evaluación, considerando todas las combinaciones de resultados previos que se originaron. En este sentido, y tal como se señala anteriormente, los docentes que se encuentran evaluados por tercera y cuarta vez son aquellos que han sido evaluados como Insatisfactorio al menos 1 vez entre la primera y segunda evaluación. De ellos, ninguno logra llegar a un nivel de Destacado en evaluaciones posteriores, mientras que en torno al 60% transita entre un desempeño Insatisfactorio o Básico en sus evaluaciones posteriores. Para el caso de docentes evaluados dos veces consecutivas como Insatisfactorios, sólo 16,5% logra un nivel de Competente en su tercera evaluación. Esto es de interés si se consideran las nuevas consecuencias de los resultados Básico e Insatisfactorio de la Evaluación Docente, bajo las cuales un docente evaluados con éste desempeño puede ser dado de baja del establecimiento educacional donde ejerce a partir de la primera vez que obtenga ese resultado.

Tabla 9: Resultados en 3ª y 4ª evaluación, considerando todas las combinaciones de resultados anteriores de los docentes que se someten a estas. 2004 – 2010.

Nº Evaluación	1ª - 2ª - 3ª	Resultado en evaluación siguiente			Total
		Insatisfactorio	Básico	Competente	
1ª - 2ª	I-I	13	53	13	79
		16.5%	67.1%	16.5%	100%
	I-B / B-I	6	49	38	93
		6.5%	52.7%	40.9%	100%
I-C/C-I	3	12	11	26	
	11.5%	46.2%	42.3%	100%	
1ª - 2ª - 3ª	I-I-B / B-I-I	1	1	0	2
		50%	50%	0%	100%
	I-I-C	0	0	1	1
		0%	0%	100%	100%

Fuente: Elaboración propia con registros del CPEIP.

Nota: Insatisfactorio (I), Básico (B), Competente (C).

4.3 Relación simple entre Evaluación Docente y SIMCE 4° básico 2008

En esta sección se presentan los resultados de la Evaluación Docente relacionados con los resultados de SIMCE 4° básico del año 2008, pero sin controlar por otras variables.

La Tabla 10 presenta el resultado promedio en las pruebas del SIMCE de los alumnos según el nivel de desempeño en la Evaluación Docente de sus profesores en 4° básico en el año 2008. Como se puede observar, existe una relación positiva entre los resultados de los docentes y el de los alumnos. En la prueba SIMCE de Matemática, un estudiante cuyo profesor obtiene un resultado que lo asigna a la categoría de Básico en la Evaluación Docente obtiene en promedio 10 puntos menos que un estudiante cuyo profesor está en la categoría de Competente, siendo la diferencia aún mayor si se compara con un Docente destacado (cerca a 20 puntos). En la prueba SIMCE de Lectura persisten las diferencias pero en menor magnitud. Un estudiante con un docente de la categoría de Básico obtiene en promedio 5 puntos menos que un estudiante con un profesor Competente y 12 puntos menos respecto a otro estudiante con un docente evaluado como Destacado.

Tabla 10: SIMCE promedio según resultado en Evaluación Docente de los profesores de 4° básico 2008.

Resultado Evaluación Docente	Matemática			Lectura		
	N° profesores	N° alumnos	Media puntaje SIMCE	N° profesores	N° alumnos	Media puntaje SIMCE
Insatisfactorio	22	196	214,49	26	322	225,90
Básico	1.213	15.867	228,93	1.240	16.146	228,23
Competente	2.305	35.669	239,27	2.319	35.032	233,04
Destacado	236	4.079	248,42	252	4.110	240,20
Total	3.776	55.811	236,92	3.837	55.610	232,16

Fuente: Elaboración propia con registros del MINEDUC.

Al realizar la misma comparación anterior, pero para las categorías promedio de profesores (Tabla 11),²² se observa que estudiantes vinculados a docentes con resultados clasificados como Insatisfactorio obtienen un puntaje promedio en Matemática y Lectura menor que aquellos estudiantes vinculados con docentes que en promedio clasifican como Competente o Destacado. De este modo, un estudiante vinculado durante sus cuatro primeros años con docentes que en promedio obtienen resultados clasificados como Destacado obtiene 27 puntos adicionales en Matemática y 22 en Lectura.

Tabla 11: SIMCE promedio según resultado en Evaluación Docente de los profesores entre 1° y 4° básico.

Resultado Evaluación Docente	Matemática			Lectura		
	N° profesores	N° alumnos	Media puntaje SIMCE	N° profesores	N° alumnos	Media puntaje SIMCE
Insatisfactorio	20	224	222,85	21	221	240,80
Básico	1.270	16.990	229,23	1.285	17.022	246,63
Competente	2.350	36.023	240,19	2.349	35.781	255,43
Destacado	98	1.676	249,19	100	1.717	262,88
Total	3.738	54.913	237,01	3.755	54.741	252,87

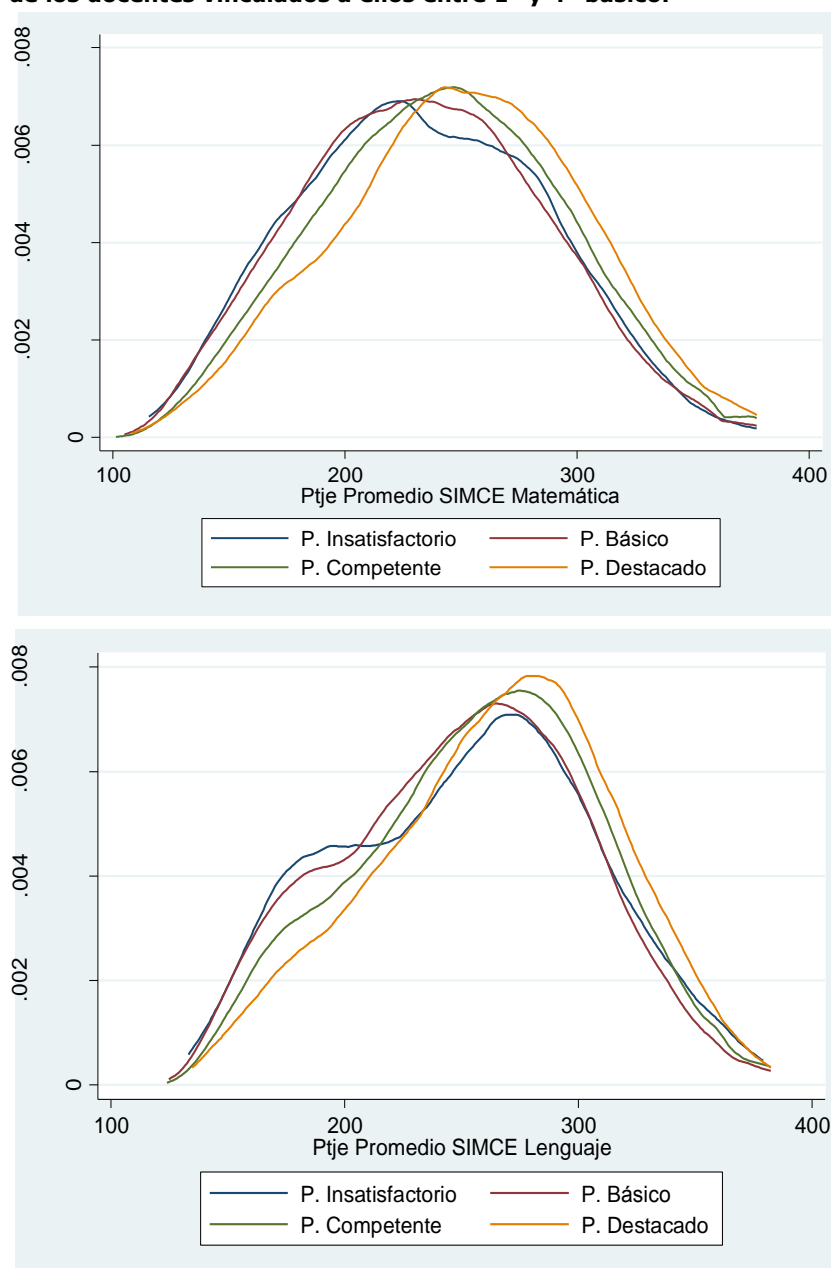
Fuente: Elaboración propia con registros del MINEDUC.

Adicionalmente, es posible observar en la Figura 6 que estudiantes vinculados durante los cuatro años con docentes que en promedio fueron mejor evaluados obtienen mejores resultados en SIMCE Matemática y Lectura. La distribución de los

²² Para calcular la categoría promedio de profesores se considera a todos los docentes del estudiante entre 1° y 4° básico y se promedian sus puntajes, asignando luego el resultado a la categoría que corresponda.

puntajes de los estudiantes se desplaza hacia la derecha a medida que sus docentes tienen mejor desempeño, mostrando que, a mayor desempeño del docente, mayor concentración de estudiantes con mejores resultados.

Figura 6: Distribución de los puntajes SIMCE Matemática y Lectura de los alumnos, según desempeño de los docentes vinculados a ellos entre 1° y 4° básico.



Fuente: elaboración propia en base a registros SIMCE y MINEDUC.

Nota: Para cada alumnos se calculó el promedio del puntaje final en la Evaluación Docente de todos los profesores vinculados a ellos entre 1° y 4° básico y fueron clasificados en cuatro categorías según los puntos de corte que define la evaluación.

Por otro lado, en la Tabla 12 se puede observar que existe una correlación no condicional entre los instrumentos de la Evaluación Docente y los resultados en la prueba SIMCE de Matemática y Lectura, siendo ésta positiva y significativa, pero con coeficientes en gran parte de los casos menores a 0,1. La correlación más alta corresponde al puntaje promedio en la Evaluación Docente, pero a nivel de instrumentos las correlaciones más altas corresponden al informe de referencia de terceros y al portafolio, mientras que la más baja a la autoevaluación.

Cabe mencionar que en el caso del informe de referencia de terceros la correlación no condicionada, tanto en el resultado de SIMCE en Matemática como en Lectura, es mayor a la correlación del portafolio con estos dos instrumentos, pero menor a la correlación con el resultado de la Evaluación Docente.

Tabla 12: Correlaciones incondicionales entre SIMCE promedio y resultado en Evaluación Docente de los profesores entre 1° y 4° básico, desagregado por instrumento. Cohorte 4° básico 2008.

	Promedio Matemática	Promedio Lectura	Promedio ED	Promedio PF	Promedio IRT	Promedio AE	Promedio EP
Promedio Matemática	1						
Promedio Lectura	0.7401*	1					
Promedio ED	0.1227*	0.1007*	1				
Promedio PF	0.0797*	0.0570*	0.7513*	1			
Promedio IRT	0.1134*	0.0918*	0.4614*	0.1837*	1		
Promedio AE	0.0306*	0.0277*	0.3042*	0.0924*	0.1024*	1	
Promedio EP	0.0723*	0.0663*	0.6272*	0.0877*	0.1414*	0.1288*	1

Fuente: Elaboración propia con registros del MINEDUC.

Nota 1: * nivel de significancia 5%.

Nota 2: ED: Evaluación Docente. PF: portafolio de enseñanza pedagógica. IRT: informe de referencia de terceros. AE: prueba de autoevaluación. EP: entrevista por un evaluador par.

Si bien estos resultados son informativos respecto a la relación incondicionada entre el desempeño del docente y el de sus estudiantes, es posible que existan otras variables que afecten los resultados de los estudiantes como características socioeconómicas, tipo de establecimiento, calidad de los docentes anteriores, entre otras. Por esto, es necesario realizar un análisis más riguroso de la relación condicionada entre el desempeño de los estudiantes y los docentes. Dicho análisis se presenta en la siguiente sección.

5. Capacidad discriminatoria de los instrumentos y vinculación entre el desempeño docente y rendimiento de sus estudiantes

Como se ha planteado en los apartados anteriores, este documento presenta dos grandes desafíos. Por un lado investigar si los instrumentos de la Evaluación Docente, tanto por separado como en conjunto, tienen la capacidad de discriminar qué docente es mejor que otro y luego ver si estos resultados se relacionan con el desempeño de los estudiantes. Dicho de otra manera, la Evaluación Docente o alguno de sus instrumentos podrían carecer de sentido si: (i) no tienen la capacidad de discriminar entre profesores, y/o (ii) si los resultados no se relacionan con los aprendizajes de los estudiantes.

Pero analizar la capacidad discriminatoria de la Evaluación Docente puede tener un alcance limitado, dado que al no existir otro instrumento que sirva como medida de comparación, ésta es analizada sobre sí misma. Aun así, se empleará la varianza de los instrumentos como señal de capacidad de discriminar entre los profesores, bajo el supuesto que a mayor varianza mejor discrimina el instrumento y que la evaluación docente es capaz de capturar el desempeño de los maestros (o al menos es un indicador de éste). Esto nos permitirá saber si el instrumento es sensible a las diferencias existentes entre los docentes, permitiendo recogerlas. Este es un primer paso para determinar la efectividad de los instrumentos.

Posteriormente, y aun cuando medir la relación entre el desempeño de los docentes y el de sus estudiantes tiene variadas limitaciones previamente mencionadas, se realizará este análisis utilizando diferentes especificaciones que permitan validar la robustez de los resultados. Este ejercicio no sólo nos responderá la interrogante de si existe o no una asociación entre el desempeño de ambos actores, sino que

también permitirá complementar el análisis a nivel de instrumentos señalando cuál(es) de ellos se relaciona(n) en mayor magnitud con los resultados de los estudiantes, siendo este un segundo paso para determinar la efectividad de los instrumentos respecto a reflejar la relación entre el desempeño de los docentes y desempeño de los estudiantes, que en este documento es entendida como una aproximación de la calidad de los docentes.

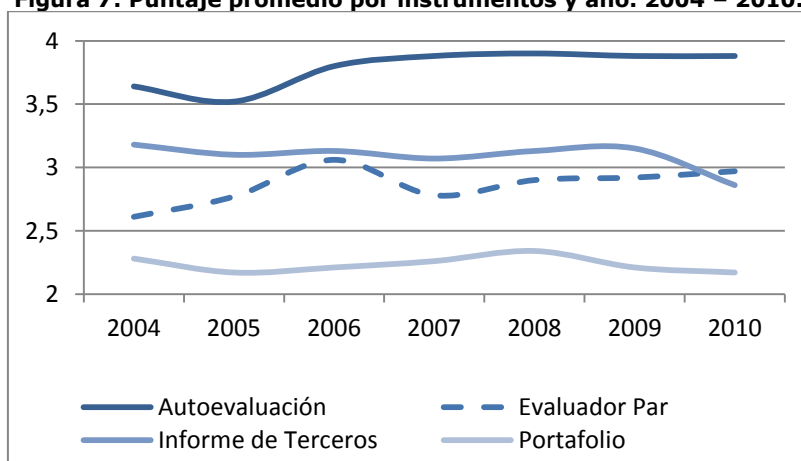
A modo de síntesis, las siguientes secciones abordan los objetivos señalados. Para el enfoque de capacidad discriminativa de instrumentos se realizará un primer apartado de análisis de cada instrumento y sus varianzas. Luego, complementando lo anterior y abordando la relación de la Evaluación Docente y sus instrumentos con el desempeño de los estudiantes, se realizará un análisis de asociación entre el resultado medido por cada instrumento (por separado y en su conjunto) y el resultado de los estudiantes en la pruebas del SIMCE de Matemática y Lectura.

5.1 Análisis de capacidad discriminativa de los instrumentos de evaluación

Para analizar la capacidad discriminativa de los instrumentos de evaluación se lleva a cabo un análisis de varianza de los mismos. El supuesto detrás de este análisis es que aquellos instrumentos que presentan mayor varianza recogen mejor las diferencias existentes entre los docentes. Asimismo, se asume que los resultados de la evaluación docente –y de cada instrumento- son un buen indicador del desempeño de los maestros (de su calidad) y, a su vez, que la evaluación docente y sus instrumentos capturan adecuadamente lo que dicen medir. Esto permitiría ordenar a los docentes respecto a su calidad..

La Figura 7 muestra los puntajes promedio por instrumento de la Evaluación Docente para el período 2004 – 2010. Es importante recordar que todos los instrumentos son evaluados en una escala de 1 a 4, donde 1 es Insatisfactorio y 4 Destacado. Si se observan los resultados promedio en cada año, se identifica que mientras los resultados del portafolio indican que en promedio los docentes no superan el punto de corte general de un nivel Básico (2,5 puntos), el instrumento de autoevaluación en promedio no baja del punto de corte general del nivel de Destacado (3,5 puntos). Es decir, según el portafolio los resultados de los docentes en promedio pertenecen a la categoría de Básico, pero según la autoevaluación a la de Destacado.

Figura 7: Puntaje promedio por instrumentos y año. 2004 – 2010.



Fuente: Elaboración propia con registros del MINEDUC.

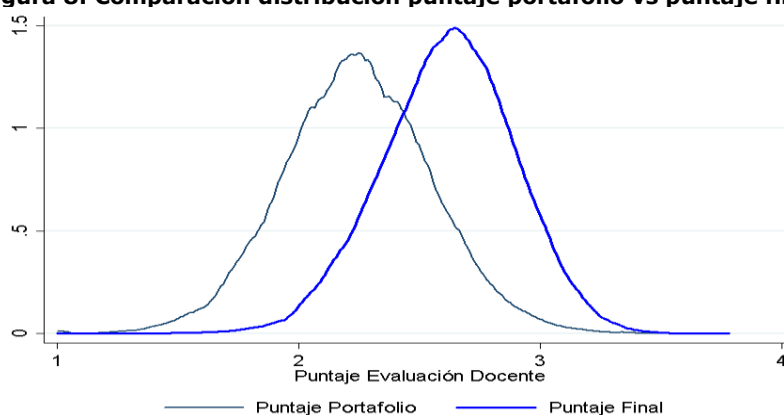
Esto deja la interrogante respecto de si estos puntajes son contradictorios entre sí. La respuesta podría ser negativa si los instrumentos abordaran diferentes dominios, criterios o descriptores fijados por el Marco para la Buena Enseñanza. Pero podría

ser positiva si se considera que los cuatro instrumentos buscan medir la práctica pedagógica del docente con similares escalas de puntaje, utilizando sólo diferentes miradas. Con todo, aun cuando se cumpliera que existen diferencias entre los instrumentos, es cuestionable que éstas sean tan profundas como para que un docente tenga una práctica Destacada en algunos aspectos y Básica o Insatisfactoria en otros.

Adicionalmente, al analizar la concentración de docentes por tramo de puntaje en cada instrumento, se observa que en la autoevaluación al menos 96% de los evaluados tiene un puntaje de 3 o más puntos (Competente o más). Pero, por otro lado, en los resultados del portafolio, 21,6% de las evaluaciones no alcanza los 2 puntos (es decir, son calificados en el nivel Insuficiente) y sólo 1% tiene resultado igual o superior a 3 puntos (Competente o más). Una posible hipótesis detrás de los altos puntajes en la autoevaluación es que al tener el resultado de la evaluación consecuencias laborales para el docente, éste podría intentar asegurar un alto puntaje en éste ítem. Alternativamente, podría estar ocurriendo que los educadores no estén conscientes de sus déficits.

Por otro lado, como se muestra en la Figura 8, las distribuciones del puntaje total de la evaluación y del portafolio tienen una forma similar. Sin embargo, el puntaje total se encuentra desplazado a la derecha respecto al puntaje del portafolio. Esto es, a nivel agregado, los instrumentos de autoevaluación, entrevista del evaluador par y el informe de referencia de terceros desplazan la distribución de resultados hacia la derecha, mejorando así los resultados de todos.

Figura 8: Comparación distribución puntaje portafolio vs puntaje final



Fuente: Elaboración propia con registros del MINEDUC.

Considerando esto, es necesario realizar un análisis más detallado de los instrumentos para determinar si efectivamente contribuyen a la explicación del desempeño de los docentes y cuál de ellos lo hace en mayor medida.

Portafolio de enseñanza pedagógica (portafolio)

Como se observa en la Figura 8, el portafolio tiene una distribución más bien normal y similar a la distribución de los puntajes de la Evaluación Docente dado que éste es el instrumento de mayor ponderación.

Aun cuando su distribución no despierta mayores cuestionamientos se realiza un análisis de descomposición de varianza según ciclo²³ para comprender mejor el comportamiento del portafolio.

²³ 1er ciclo de enseñanza básica, 2do ciclo de enseñanza básica, educación de párvulos, educación media y educación especial.

En la Tabla 13 se observa una gran varianza entre los ciclos, sobre todo entre los años 2006 y 2008, lo que indica que no existe homogeneidad en los resultados. Mientras que al interior de los ciclos el instrumento tiene un comportamiento más estable, teniendo una varianza muy baja. Adicionalmente, como es de esperar, la varianza total del instrumento es similar entre los años²⁴. A partir de esto es posible intuir que al interior de cada ciclo los resultados son bastante parejos, pero varían entre los ciclos, lo que puede tener relación con diferencias en el desempeño de los docentes de cada ciclo.

Tabla 13: Análisis de descomposición de varianza del portafolio por ciclo.

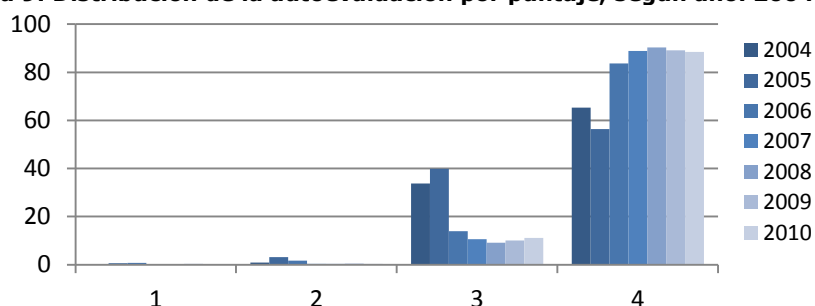
Varianza	2005	2006	2007	2008	2009	2010
Entre Ciclos	3,59	22,76	54,43	10,24	5,14	9,01
Intra Ciclos	0,12	0,08	0,09	0,09	0,08	0,08
Total	0,13	0,09	0,10	0,09	0,08	0,08

Fuente: Elaboración propia con registros del MINEDUC.

Pauta de autoevaluación (autoevaluación)

En la Figura 9 es posible distinguir la concentración de los resultados del instrumento de autoevaluación en los dos puntajes más altos en el tiempo. Una hipótesis detrás de esto responde a un posible condicionamiento o acostumbramiento de los evaluados, donde luego de las dos primeras evaluaciones sistemáticamente más de 80% se autoevalúa como Destacado.

Figura 9: Distribución de la autoevaluación por puntaje, según año. 2004 – 2010.



Fuente: Elaboración propia con registros del CPEIP.

Nota: Eje X contiene los 4 tramos de puntajes posibles en la autoevaluación y eje Y muestra el porcentaje de docentes.

Otra forma de expresar lo anterior es observar cómo se han comportado la media y la varianza en cada período. Es posible identificar que la media se ha ido acercando al valor máximo de la evaluación con el transcurso de los años y la variabilidad entre los resultados ha ido disminuyendo, siendo menor a 0,15 puntos desde el año 2007 en adelante (Tabla 14). No obstante esta tendencia general, cabe mencionar que tanto para la educación media como para la educación de párvulos la tendencia de la varianza ha ido al alza en los últimos dos años.

²⁴ Ver detalle en Apéndice B, Tabla B.1.

Tabla 14: Análisis de varianza anual de la autoevaluación.

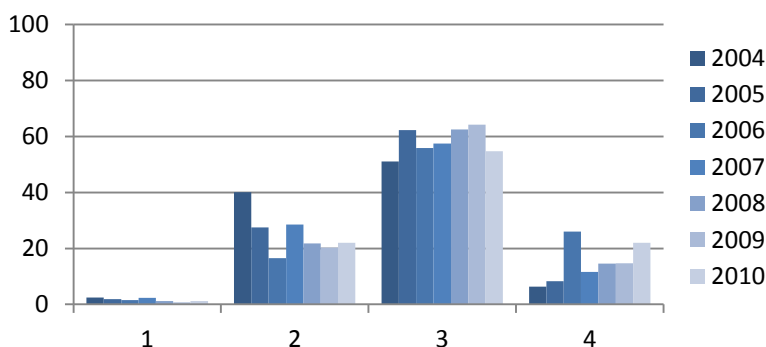
		2004	2005	2006	2007	2008	2009	2010
	Promedio	3,6	3,5	3,8	3,9	3,9	3,9	3,9
Varianza	Todos	0,25	0,35	0,24	0,12	0,11	0,14	0,12
	1er ciclo e. básica	0,26	0,36	0,24	0,13	0,12	0,11	0,08
	2do ciclo e. básica	0,22	0,31	0,23	0,12	0,12	0,14	0,11
	E. especial						0,11	0,12
	E. media		0,35	0,24	0,12	0,11	0,17	0,19
	E. parvularia					0,08	0,12	0,18

Fuente: Elaboración propia con registros del CPEIP.

Entrevista por un evaluador par (evaluación de los pares)

En la Figura 10 se presentan los resultados de la evaluación de los pares. En esta figura se observa que aun cuando la concentración del puntaje es menor a la observada en la autoevaluación, en promedio alrededor de 60% de los evaluados se encuentra en la categoría de Competente, con una tendencia decreciente de los puntajes inferiores y un aumento sistemático de los puntajes superiores en el tiempo.

Figura 10: Distribución de los resultados de la evaluación de los pares por puntaje, según año. 2004 - 2010.



Fuente: Elaboración propia con registros del CPEIP.

Nota: Eje X contiene los 4 tramos de puntajes posibles en la evaluación del par y el eje Y el porcentaje de docentes.

Considerando que, entre otros requisitos, los evaluadores pares deben ser docentes de aula del mismo ciclo²⁵ que el docente evaluado, es posible que algún ciclo educativo sea más estricto para evaluar que otro. Se requiere ver cuánto de la varianza total del instrumento se explica por la varianza entre ciclos y cuánto por la varianza al interior de cada ciclo.

La Tabla 15 presenta un análisis de descomposición de la varianza para la evaluación de los pares entre los años 2005 y 2010. En ésta es posible observar que, para todos los años, la varianza entre ciclos de enseñanza es muy superior a la varianza al interior de cada ciclo. Esto podría validar la hipótesis de diferentes estándares de exigencia entre ciclos. Adicionalmente, es importante señalar que desde que se encuentran los 5 ciclos participando en la Evaluación Docente (año 2009 - 2010), la varianza entre ciclos ha experimentado un aumento. Cabe mencionar que en el 1er y 2do ciclo de e. básica se observan las mayores varianzas del instrumento.²⁶

²⁵ 1er ciclo de enseñanza básica, 2do ciclo de enseñanza básica, educación de párvulos, educación media y educación especial.

²⁶ Ver detalle en Apéndice B, Tabla B.2.

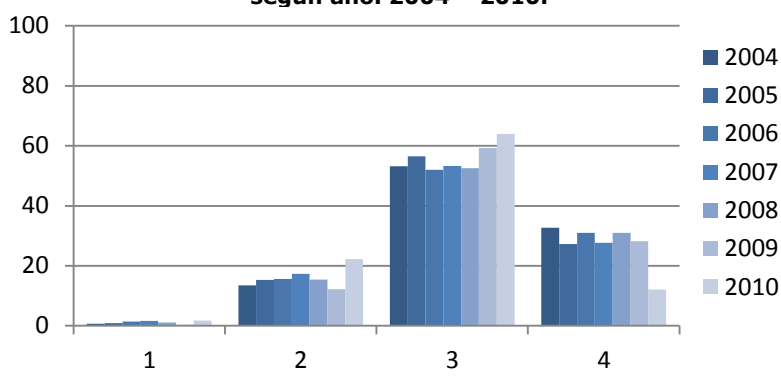
Tabla 15: Análisis de descomposición de varianza de la entrevista del evaluador por ciclo.

Varianza	2005	2006	2007	2008	2009	2010
Entre Ciclos	2,27	8,18	5,85	15,88	18,70	40,89
Intra Ciclos	0,38	0,48	0,45	0,39	0,37	0,47
Total	0,38	0,48	0,45	0,40	0,37	0,49

Fuente: Elaboración propia con registros del MINEDUC.

Informe de referencia de terceros (IRT)

En la Figura 11 se presentan los resultados del informe de referencia de terceros. En ésta es posible ver que se mantiene una distribución altamente concentrada en los dos niveles superiores.

Figura 11: Distribución de los resultados del informe de referencia de terceros por puntaje, según año. 2004 – 2010.

Fuente: Elaboración propia con registros del CPEIP

Nota: Eje X contiene los 4 tramos de puntajes posibles en el informe de referencia de terceros y el eje Y el porcentaje de docentes.

Al ser el informe de referencia de terceros realizado por los directores y jefes de UTP de los establecimientos educacionales en los que trabajan los docentes, resulta interesante analizar si la variabilidad de los resultados responde a la variabilidad que puede existir dentro de los establecimientos o responde más bien a la varianza entre los establecimientos. Es decir, analizar si al interior de cada establecimiento existe menor, igual o mayor variabilidad en los resultados que la existente entre los establecimientos.

La Tabla 16 presenta un análisis de descomposición de varianza para el instrumento informe de referencia de terceros en todos sus años de aplicación. En ésta es posible observar que, para todos los años, la varianza entre establecimientos es superior a la varianza al interior de los establecimientos, lo que puede ser interpretado como: (i) parte de un comportamiento de evaluación donde los directores o jefes de UTP evalúan de manera homogénea a sus docentes, o (ii) puede ser señal de que los docentes se encuentran segregados entre los establecimientos según su calidad.

Tabla 16: Análisis de descomposición de varianza del informe de referencia de terceros.

Varianza	2005	2006	2007	2008	2009	2010
Entre Escuelas	1.16	1.37	1.28	1.28	1.08	0.90
Intra Escuelas	0.28	0.32	0.33	0.34	0.25	0.26
Total	0.45	0.50	0.50	0.49	0.39	0.38

Fuente: Elaboración propia con registros del MINEDUC.

Agregando los instrumentos

Como se observa en la Figura 8, las distribuciones del puntaje del portafolio y del puntaje total (incorporando todos los instrumentos) son similares, pero está última se encuentra desplazada hacia la derecha por la incorporación de los restantes instrumentos. Sin embargo, el ordenamiento de los docentes no es el mismo en la distribución de resultados del portafolio que en el resultado final de la evaluación. Esto se observa con claridad en la Tabla 17 que muestra la matriz de transición entre el cuartil de resultado de cada docente en el portafolio y su puntaje final. Si el ordenamiento de los docentes en ambas distribuciones fuera el mismo, todas las observaciones estarían en la diagonal de la matriz de transición, sin embargo, se observa que sólo 56% de los docentes se mantiene en el mismo cuartil de resultados una vez que se incorporan el resto de los instrumentos; afectando el ordenamiento de los docentes en términos de resultados.

Tabla 17: Cuartil de resultado utilizando el puntaje de portafolio versus el puntaje final

Cuartil de puntaje del portafolio	Cuartil de puntaje total				
	1	2	3	4	Total
1	11.338	4.715	977	111	17.141
	66,15%	27,51%	5,7%	0,65%	100%
2	3.975	7.173	3.743	1.284	16.175
	24,57%	44,35%	23,14%	7,94%	100%
3	1.079	4.222	7.211	3.720	16.232
	6,65%	26,01%	44,42%	22,92%	100%
4	96	1.372	3.710	11.224	16.402
	0,59%	8,36%	22,62%	68,43%	100%
Total	16.488	17.482	15.641	16.339	65.950
	25%	26,51%	23,72%	24,77%	100

Fuente: Elaboración propia con registros del MINEDUC.

Lo anterior permite entrever que el resultado final no sólo se explica por el resultado en el portafolio sino que uno o más del resto de los instrumentos tiene efectos en las diferencias entre los docentes.

Por ello, se hace un análisis adicional para ver cuáles son los instrumentos de la evaluación que tienen mayor poder discriminativo en términos de desempeño docente. Un instrumento efectivo no sólo debería ser insesgado, sino también capaz de capturar la variabilidad existente entre los docentes. En decir, un instrumento que tiene poca varianza provee menor información sobre las diferencias entre docentes que un instrumento con mayor varianza. Basados en esta idea y considerando que existe una correlación entre algunos instrumentos de la evaluación, se realiza un análisis de componentes principales (ACP).

Este método no paramétrico ayuda a re-expresar bases de datos complejas en una combinación lineal de variables con diferentes ponderaciones, según la variabilidad de cada variable y la correlación entre ellas.

Cada instrumento de la evaluación mide características/capacidades distintas de los docentes, pero éstas pueden estar correlacionadas generando problemas de precisión en el indicador final. La combinación lineal de componentes principales busca maximizar la varianza y reducir el ruido de los datos, evitando así perder información valiosa. De esta manera, los instrumentos con mayor poder discriminativo reciben un mayor ponderador en el índice, mientras que los instrumentos que casi no permiten diferenciar entre docentes reciben un menor ponderador.

En general, el análisis indica que la combinación ponderada de los instrumentos de la Evaluación Docente explica alrededor de 50%²⁷ de la varianza total observada entre el resultado de los docentes. Adicionalmente, del ejercicio de componentes principales se desprende que el portafolio es el instrumento que debería tener la mayor ponderación en la combinación lineal, mientras que la autoevaluación debería tener la menor ponderación (Tabla 18). Dicho orden es consistente con la forma en que se calcula actualmente el puntaje final de la Evaluación Docente. Cabe mencionar que en los casos de educación media y educación de párvulos el ponderador de la autoevaluación toma el mayor valor respecto a los otros ciclos de enseñanza, lo que es consistente con el análisis de la varianza realizado al instrumento.

Tabla 18: Ponderador de cada instrumento según Análisis de Componentes Principales.

		Ponderador de componentes principales						
		2004	2005	2006	2007	2008	2009	2010
Todos los ciclos	AE	0,35	0,44	0,41	0,43	0,40	0,42	0,39
	EP	0,48	0,47	0,49	0,50	0,51	0,50	0,51
	IRT	0,56	0,54	0,54	0,53	0,53	0,53	0,53
	PF	0,57	0,55	0,56	0,54	0,55	0,53	0,55
Primer ciclo	AE	0,32	0,43	0,42	0,41	0,40	0,39	0,32
	EP	0,51	0,47	0,48	0,44	0,51	0,47	0,50
	IRT	0,57	0,54	0,53	0,55	0,56	0,56	0,58
	PF	0,56	0,55	0,55	0,57	0,52	0,56	0,56
Segundo ciclo	AE	0,38	0,44	0,40	0,42	0,41	0,40	0,39
	EP	0,45	0,42	0,49	0,51	0,50	0,48	0,51
	IRT	0,55	0,55	0,53	0,52	0,54	0,54	0,54
	PF	0,59	0,57	0,57	0,55	0,55	0,56	0,55
E. media	AE	0,42	0,27	0,38	0,40	0,40	0,44	0,36
	EP	0,55	0,49	0,50	0,48	0,50	0,53	0,53
	IRT	0,57	0,56	0,57	0,56	0,51	0,47	0,52
	PF	0,40	0,61	0,53	0,54	0,57	0,56	0,56
E. parvularia	AE					0,35	0,44	0,39
	EP					0,56	0,49	0,48
	IRT					0,51	0,53	0,53
	PF					0,56	0,55	0,59

Fuente: Elaboración propia con registros del MINEDUC.

Nota: AE: pauta de autoevaluación. EP: entrevista por un evaluador par. IRT: informe de referencia de terceros. PF: portafolio de enseñanza pedagógica.

Los resultados presentados en la Tabla 18 también sugieren que los instrumentos de entrevista por un evaluador par y el informe de referencia de terceros añaden información acerca del desempeño docente, probablemente porque miden algún aspecto de la práctica del docente que no es capturado por los otros instrumentos. A pesar de tener un peso relativamente bajo en el puntaje final de la Evaluación Docente, el análisis indica que estos dos instrumentos tienen, para algunos años y ciclos, ponderadores tan importantes como el portafolio. Los resultados de este ejercicio sugieren la posibilidad de darle mayor peso en el puntaje final al informe de referencia de terceros, actualmente en 10%, para incrementar el poder discriminativo de la Evaluación Docente. La evaluación de pares, aun cuando es en menor magnitud, también aporta información que permite distinguir mejor a los docentes de los distintos ciclos. En particular, las prácticas capturadas por la evaluación de pares juegan un papel importante en la educación media.

²⁷ El otro 50% es la varianza no explicada, es decir de aspectos que no estamos siendo medidos con los instrumentos. Esto puede ser desde el género o la edad de los docentes hasta características de la personalidad que no son medidas por los instrumentos.

Finalmente, la autoevaluación es el instrumento que menos aporta a las discriminación entre docentes, esto se debe en gran medida a que, como se explicó antes, la mayor parte de los docentes se autoevalúan con puntajes altos con poca variabilidad entre ellos.

Al ser el portafolio el instrumento que más ayuda a discriminar entre los docentes, resulta relevante analizar cuáles son las dimensiones de éste que más aportan a dicha diferenciación. El puntaje final del portafolio depende de 8 indicadores (A-H) que buscan medir distintas prácticas pedagógicas de los docentes de aula. Al realizar nuevamente el ACP a los indicadores del portafolio, se observa en la Tabla 19 que (B) el análisis de las actividades de las clases, (E) la reflexión pedagógica, (A) la organización de los elementos de la unidad, (D) la reflexión a partir de los resultados de la evaluación y (C) la calidad de la evaluación de la unidad son las dimensiones que tienen los mayores ponderadores. Estos ponderadores tienen magnitudes similares y sugieren que dichas dimensiones tienen poder discriminativo semejante. Por el otro lado, las dimensiones de (G) estructura de la clase, (H) interacción pedagógica y (F) ambiente de la clase para el aprendizaje parecen tener menor poder para discriminar entre docentes, especialmente a aquellos del 1er y 2do ciclo de e. básica. Estos patrones se han mantenido relativamente constantes a través del tiempo con pequeñas variaciones debidas probablemente a cambios en los constructos que han sido evaluados dentro de cada dimensión.²⁸

Tabla 19: Ponderador de cada indicador del portafolio según análisis de Componentes Principales. 2006 – 2010.

	Ponderador de componentes principales					
	Todos	E. parvularia*	1er ciclo e. básica	2do ciclo e. básica	E. media	E. especial**
A: Organización de los elementos de la unidad	0,398	0,358	0,410	0,416	0,382	0,348
B: Análisis de las actividades de las clases	0,427	0,365	0,430	0,434	0,420	0,374
C: Calidad de la evaluación de la unidad	0,366	0,378	0,386	0,371	0,324	0,332
D: Reflexión a partir de los resultados de la evaluación	0,381	0,341	0,407	0,380	0,369	0,352
E: Reflexión pedagógica	0,412	0,395	0,449	0,403	0,380	0,390
F: Ambiente de la clase para el aprendizaje	0,204	0,275	0,111	0,214	0,275	0,329
G: Estructura de la clase	0,291	0,348	0,214	0,294	0,324	0,376
H: Interacción pedagógica	0,290	0,356	0,270	0,250	0,335	0,320

Fuente: Elaboración propia con registros del MINEDUC.

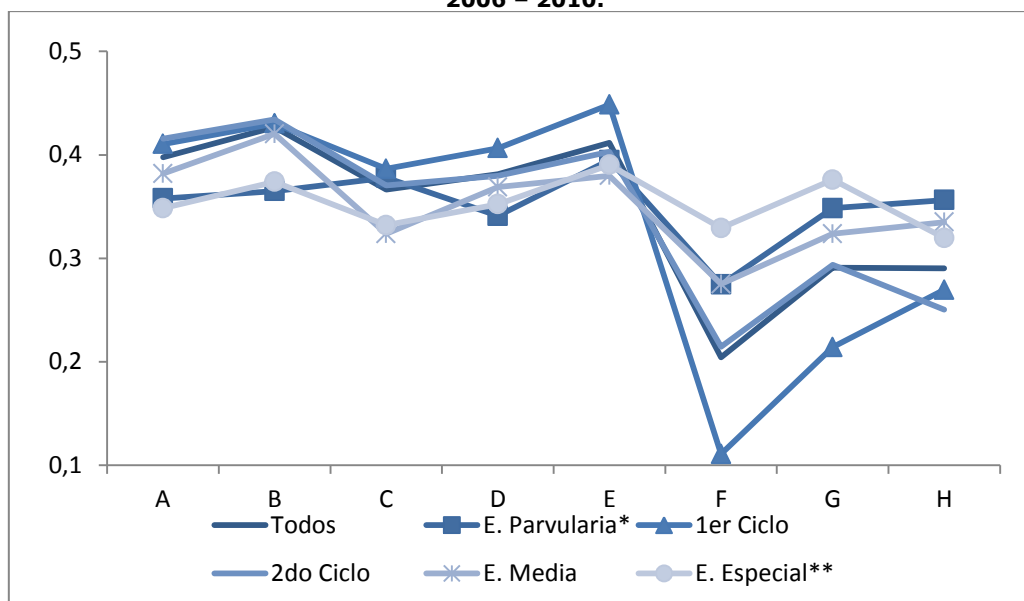
* Disponible para 2008.

** Disponible para 2009 y 2010.

Otra forma de ver lo anterior es gráficamente. La Figura 12 muestra claramente los indicadores con mayor poder de discriminación entre los docentes.

²⁸ Ver Apéndice B, Tablas B.3 a B.8.

Figura 12: Indicadores del portafolio según análisis de Componentes Principales. 2006 – 2010.



Fuente: Elaboración propia con registros del MINEDUC.

Nota 1: Eje Y presenta las dimensiones del portafolio.

* Disponible para año 2008.

** Disponible para años 2009 y 2010.

5.3 Desempeño docente y de los estudiantes. Modelos estimados.

El método comúnmente utilizado para estudiar la efectividad del docente es vincular su desempeño con el de sus estudiantes. Como se ha comentado al comienzo de este documento existen limitaciones con este mecanismo por lo que a continuación se desarrollan distintas estrategias complementarias que permiten elaborar conclusiones más robustas. Con todo, las estimaciones y resultados que se muestran a continuación dan cuenta de la asociación entre el desempeño docente en la Evaluación Docente y el de los estudiantes en las pruebas del SIMCE, sin pretender establecer un efecto causal del desempeño de los profesores medido en la Evaluación Docente sobre el de los alumnos en las pruebas del SIMCE.

En primer lugar, se opta por estudiar los resultados del SIMCE en 4° básico, dado que, si bien hay muchos factores que intervienen en el aprendizaje de los estudiantes, comúnmente en este nivel un solo profesor es el responsable de enseñar todas las materias (profesor generalista), por lo tanto, la vinculación entre profesor y alumno es directa y completa. Por el contrario, en gran medida, en 8° básico y casi siempre en enseñanza media hay un profesor destinado a cada asignatura, por lo que se hace aún más difícil establecer el aporte de cada docente al desempeño de cada alumno. Por ejemplo, en el caso de los resultados del SIMCE de Matemática podría estar influyendo no sólo el desempeño del profesor de Matemática, sino que también el de Lectura y Comunicación, Ciencias o el de otra disciplina.

Por otro lado, si sólo se vinculara el resultado del estudiante en el SIMCE de 4° básico con el desempeño del docente que tuvo en ese nivel, se podría estar obteniendo resultados sesgados al asumir que únicamente el docente de 4° básico es fundamental en el aprendizaje del estudiante en una prueba que no sólo mide conocimientos de 4° básico, sino que todos los asociados a cursos anteriores. Así, no se estaría considerando el efecto acumulativo que pueden tener los distintos docentes en el aprendizaje de los estudiantes, como se señala en el estudio de Rivers y Sanders (1996). Por lo anterior, se decide realizar también las estimaciones considerando el desempeño de todos los docentes que tuvo el estudiante entre 1° y 4° básico.

Otra forma de mirar el desempeño de los docentes está relacionada con el desempeño promedio de los establecimientos. En este sentido, los propios establecimientos pueden tener determinadas funciones de producción que permiten que su desempeño sea mejor que otro. Por esto se considera vincular el desempeño promedio de los estudiantes de un mismo establecimiento con la concentración de docentes bien o mal evaluados al interior de cada establecimiento. Esto permitirá identificar si la composición del staff de docentes en cada establecimiento marca la diferencia en el desempeño promedio de sus estudiantes.

Finalmente, se definen una serie de mecanismos que permiten controlar los problemas de selección de los docentes entre establecimientos y al interior de cada establecimiento identificados a continuación: (i) si se asume que los mejores docentes son atraídos por los mejores establecimientos, entonces un mecanismo de control de la autoselección es considerar el resultado SIMCE de los años anteriores a que el estudiante se encontrara en el establecimiento educacional. (ii) Si en cada establecimiento los docentes se distribuyen según el desempeño de cada curso, esto introduciría endogeneidad en los resultados, lo que sesgaría la estimación, por lo que para chequear la robustez en los resultados se consideran modelos con un solo curso y con más de un curso.

Modelo 1: Estimaciones a nivel individual

Este modelo busca estudiar la relación entre los resultados en las pruebas del SIMCE de los estudiantes de 4° básico del año 2008 con los de sus profesores en la Evaluación Docente de 1° a 4° básico. Para esto se realiza la siguiente especificación:

$$SIMCE_{is} = \beta_1 + \beta_2 X_{is} + \beta_3 Y_s + \beta_4 Z_i + \mu_{is}$$

Donde X_{is} , corresponde a un vector de características individuales del estudiante i que asiste a la escuela S ; Y_s corresponde a un vector de características del establecimiento S y Z_i corresponde a un vector de características de los docentes que han impartido clases al alumno i entre 1° y 4° básico, entre ellas, el resultado en la Evaluación Docente. Finalmente μ_{is} , corresponde al error de la estimación, el cual cumple con los supuestos clásicos de la estimación por Mínimos Cuadrados Ordinarios.²⁹

Entre las características individuales de los estudiantes se utiliza el género, el *ranking* promedio en su curso, una variable dicotómica que se refiere a la asistencia a educación pre escolar y el nivel de escolaridad de la madre. El primero, controla por las diferencias de género observadas en las pruebas del SIMCE. El *ranking* promedio controla por el desempeño del estudiante y la asistencia a educación pre escolar por las diferencias en el capital humano adquirido previo al seguimiento realizado. La escolaridad de la madre se emplea como un *proxy* del nivel socioeconómico de la familia del menor.

Respecto a las características del establecimiento se incluye el tamaño del curso y el resultado promedio del SIMCE en 4° básico a nivel de establecimiento en los años anteriores (2005-2007). El efecto del tamaño de la clase ha sido estudiado en diversas oportunidades, teniendo resultados diferentes en los modelos de función de producción educacional.³⁰ Del mismo modo, controlar por el desempeño anterior del establecimiento es de utilidad para evitar que los efectos encontrados en los resultados sean afectados por algún *shock* en los resultados individuales durante el

²⁹ En los resultados que se presentan en esta sección se asume que los errores son homocedásticos. En el Apéndice D, Tabla D6 se presentan los resultados para estimaciones con errores corregidos por cluster a nivel de establecimiento. Las diferencias no son importantes y no varían los resultados presentados.

³⁰ Para mayor detalle ver Krueger (1997), Summers y Wolfe (1977).

año en análisis y permite controlar, en la medida que los datos lo permiten, por sesgo de selección. Siendo el puntaje en el SIMCE a nivel de establecimiento el resultado de un conjunto de elementos, al incluir los resultados anteriores se está controlando implícitamente por la selección de los profesores que podrían estar realizando algunos establecimientos y por la autoselección de los mismos docentes.³¹

Finalmente, se incluyen las variables de género del profesor, la edad y experiencia promedio (medida como años de servicio en el sistema educacional) y el resultado en la Evaluación Docente de todos los educadores vinculados a los alumnos durante los 4 años.³² El desempeño de los docentes es analizado de tres maneras diferentes:

- A1: agrupa a los docentes³³ en cuatro categorías en función de su resultado final en la Evaluación Docente. Esta variable fue construida de la siguiente forma:
 - Se obtiene el puntaje promedio en la Evaluación Docente de todos aquellos docentes vinculados a un alumno durante los cuatro primeros años de básica.
 - Este puntaje promedio es categorizado según la clasificación oficial: (i) Insatisfactorio si el puntaje promedio es menor a 2; (ii) Básico si el puntaje promedio está entre 2 y 2,5; (iii) Competente si el puntaje promedio está entre 2,51 y 3 y (iv) Destacado si el puntaje promedio es mayor a 3.
- A2: considera el promedio del puntaje final en la Evaluación Docente de los profesores vinculados a los alumnos de 4° básico.
- A3: considera los promedios para cada uno de los cuatro instrumentos de la Evaluación Docente de todos los docentes vinculados a los alumnos de 4° básico.

Adicionalmente, como un mecanismo para evitar sesgos por la selección no aleatoria de los docentes al interior de los establecimientos, se realizan estimaciones considerando: (i) los estudiantes de todos los establecimientos (ii) alumnos de establecimientos con un solo curso y (iii) estudiantes de establecimientos con al menos dos cursos por nivel.³⁴

En la Tabla 20³⁵ se presentan los resultados de las estimaciones para las pruebas del SIMCE de Matemática y Lectura utilizando las alternativas de medición de desempeño docente antes descritas.

Se observa que existe una relación positiva entre el desempeño de los docentes en la Evaluación Docente y el resultado de sus estudiantes en SIMCE 4° básico. En general, las tres especificaciones realizadas, tanto para Matemática como Lectura, son consistentes en mostrar que estudiantes vinculados con profesores de buen desempeño tienen en promedio mejores resultados en la prueba SIMCE de 4° básico que los estudiantes vinculados a docentes de bajo desempeño, aun controlando por características de los estudiantes, docentes y establecimientos. Cabe mencionar que las asociaciones son de mayor magnitud en Matemática que en Lectura.

³¹ Respecto a la selección que pueden hacer las familias se controla por escolaridad de los padres como proxy de acceso a información, asumiendo que padres más informados matriculan a sus hijos en establecimientos de mejor calidad. Por otro lado, cabe mencionar que el segmento de menores ingresos o aquellos alejados de centros urbanos matriculan a sus hijos en establecimientos disponibles cerca de su domicilio.

³² Los docentes pueden haber realizado la Evaluación Docente en diferentes ciclos escolares. Esto porque pueden ejercer en más de un ciclo en el mismo año o haber cambiado de ciclo en los diferentes años.

³³ Todos los educadores vinculados a los alumnos de 4° básico (desde primero a cuarto).

³⁴ Ver Tablas D1 y D2, Apéndice D.

³⁵ Tabla D1 en Apéndice D presenta los resultados para todos los regresores.

Tabla 20: Estimación SIMCE 4° Básico 2008 a nivel de alumno, controlando por características de los estudiantes, docentes y establecimientos.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
P. Básico	-1.245			-0.221		
	(-0.94)			(-0.16)		
P. Competente	3.801**			2.947*		
	(2.90)			(2.15)		
P. Destacado	9.305***			6.601***		
	(5.61)			(3.84)		
Promedio puntajes ED		15.85***			9.531***	
		(17.06)			(10.04)	
Puntaje promedio PF			8.051***			6.193***
			(9.38)			(7.03)
Puntaje promedio IRT			4.205***			2.210***
			(11.94)			(6.07)
Puntaje promedio AE			-0.763			-0.459
			(-1.32)			(-0.78)
Puntaje promedio EP			2.674***			1.021**
			(7.42)			(2.76)
N	51809	51809	51809	51757	51757	51757
R² ajustado	0.423	0.424	0.424	0.362	0.362	0.362

Fuente: Elaboración propia, en base a registros del MINEDUC.

Nota 1: PF: portafolio de enseñanza pedagógica. IRT: informe de referencia de terceros. AE: pauta de autoevaluación. EP: entrevista por un evaluador par.

Nota 2: Error estándar en paréntesis *** Nivel de significancia 0,1%, ** 1% y * 5%.

La descomposición del análisis A1 (agrupación de docentes en cada categoría) al categorizar el puntaje promedio de la Evaluación Docente permite observar que no hay un efecto estadísticamente significativo en el resultado en el SIMCE de los estudiantes de 4° básico si estos estuvieron en promedio vinculados a profesores con un desempeño Básico, respecto de haber estado vinculados a profesores con resultados promedio en el nivel Insatisfactorio. Sin embargo, se observa un efecto positivo y creciente, en Matemática y Lectura, para los estudiantes que estuvieron vinculados con profesores en promedio en los niveles Competente o Destacado en comparación con sus pares que tuvieron profesores en promedio en nivel Insatisfactorio.

La especificación A2 (promedio puntaje final) muestra que, de cuatro puntos posibles de obtener, un punto adicional promedio en la Evaluación Docente de los profesores de 1° a 4° básico está asociado con un incremento promedio de 15,8 puntos en la prueba de Matemática y 9,5 puntos en la de Lectura (del SIMCE de 4° básico). Es decir, cerca de un cuarto y un quinto de sus respectivas desviaciones estándar.

A nivel de instrumentos (A3: puntaje promedio en cada instrumento), es posible observar la contribución relativa de cada uno de ellos al logro de los estudiantes. En orden de magnitud, el portafolio muestra la mayor contribución, en donde un punto adicional de los docentes entre 1° y 4° básico se asocia con un aumento promedio de 8 puntos en Matemática para estudiantes de 4° básico, mientras que la contribución promedio en Lectura es de 6 puntos. En segundo lugar se encuentra el informe referencia de terceros que se asocia con un incremento promedio de 4,2 y

2,2 puntos respectivamente, y, en tercer lugar, la entrevista por un evaluador par con 2,6 puntos en Matemática y 1 punto en Lectura. El instrumento de autoevaluación muestra ser no significativo en ambas pruebas.

Según los resultados obtenidos, la Evaluación Docente estaría asociada con cerca de 30% de la varianza en la prueba SIMCE de Lectura, y aproximadamente con 19% de la varianza en la prueba de Matemática.

Modelo 2: Estimaciones a nivel de establecimientos

Este modelo estudia la relación entre los resultados del SIMCE de 4° básico de los establecimientos (promedio de todos los alumnos de cuarto básico) y el desempeño promedio del profesorado de primer ciclo de cada establecimiento (todos los maestros del primer ciclo básico, hayan o no sido profesores de los alumnos de cuarto básico que rindieron SIMCE). Para esto se utiliza la siguiente especificación:

$$SIMCE_S = \beta_1 + \beta_2 X_S + \beta_3 Y_S + \beta_4 Z_S + \beta_5 W_S + \mu_S$$

Donde X_S corresponde al vector de características promedio de los estudiantes que asisten a la escuela S ; Y_S corresponde a un vector de características del establecimiento S ; Z_S corresponde a un vector de características promedio de los docentes de primer ciclo del establecimiento S , entre ellas el resultado en la Evaluación Docente y W_S corresponde a un vector de características de la comuna. Finalmente μ_S , corresponde al error de la estimación.

Las variables representativas del nivel socioeconómico de los estudiantes de cada establecimiento están representadas por la clasificación de grupos socioeconómicos (GSE) realizada por el SIMCE. A nivel de establecimiento se utiliza la variable del tamaño del curso. En términos de las características comunales se emplean controles por tipo de administración municipal (Corporación Municipal o DAEM), por tamaño de la comuna y por el grado de competencia. El tipo de administración puede afectar la asignación de los recursos disponibles. El tamaño de la comuna tiene relación con el nivel de recursos disponibles y necesidades de gestión. El grado de competencia, entendida como el porcentaje de estudiantes en establecimientos particulares pagados y particulares subvencionados, tiene relación con el grado de presión por calidad a la que se ven expuestos los establecimientos municipales.

A nivel de los docentes se emplea el porcentaje de profesores hombres, el porcentaje de educadores evaluados con la Evaluación Docente³⁶, las horas de contrato, los años de experiencia promedio, medido como el total de años de servicio en el sistema educacional, y el desempeño promedio en la evaluación.

En línea con las estimaciones a nivel individual se emplean tres modelos de análisis alternativos:

- A1: Concentración (%) de docentes evaluados como Competentes y/o Destacados en el primer ciclo de cada establecimiento.
- A2: Promedio del puntaje final en la Evaluación Docente (considerando todos los instrumentos) de los docentes de primer ciclo del establecimiento.
- A3: Promedio del puntaje en cada instrumento de la Evaluación Docente de los docentes evaluados del primer ciclo del establecimiento.

La Tabla 21³⁷ muestra los resultados de estas estimaciones tanto para la prueba del SIMCE de Matemática como de Lectura, utilizando todos los establecimientos que contaban con al menos un docente del primer ciclo evaluado (93%).³⁸

³⁶ Los docentes pueden haber realizado la Evaluación Docente en diferentes ciclos escolares. Esto porque pueden ejercer en más de un ciclo en el mismo año o haber cambiado de ciclo en los diferentes años.

³⁷ Tabla D2 en Apéndice D presenta los resultados para todos los regresores.

Tabla N° 21: Estimación SIMCE 4° básico 2008 a nivel de establecimiento, controlando por características de los estudiantes, docentes, establecimientos y comuna.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
% Básico	0.252*			0.283*		
	(2.05)			(2.49)		
% Competente	0.385**			0.360**		
	(3.14)			(3.19)		
% Destacado	0.502***			0.439***		
	(3.95)			(3.74)		
Puntaje promedio final ED		26.85***			18.72***	
		(7.59)			(5.77)	
Puntaje prom. PF			17.89***			11.91***
			(5.44)			(3.95)
Puntaje prom. AE			4.712*			0.946
			(1.98)			(0.43)
Puntaje prom. IRT			4.641***			4.703***
			(3.90)			(4.32)
Puntaje prom. EP			3.333*			2.031
			(2.37)			(1.57)
N	2378	2378	2377	2379	2379	2378
R² ajustado	0.266	0.269	0.268	0.272	0.272	0.275

Fuente: Elaboración propia, en base a registros del MINEDUC.

Nota 1: ED: Evaluación Docente. PF: portafolio de enseñanza pedagógica. IRT: informe de referencia de terceros. AE: pauta de autoevaluación. EP: entrevista por un evaluador par.

Nota 2: Error estándar en paréntesis *** Nivel de significancia 0,1%, ** 1% y * 5%.

Al descomponer la integración del profesorado en cada establecimiento por el porcentaje de presencia en las cuatro categorías de desempeño en la Evaluación Docente (A1), se observan resultados positivos vinculados a una mayor presencia de docentes mejor evaluados. Esto es, se observa que un aumento de 1 punto porcentual en el porcentaje de docentes evaluados en el nivel Básico (en reemplazo de la misma proporción de educadores en nivel Insatisfactorio) se asocia con un incremento de 0,2 puntos en el resultado promedio del establecimiento en la prueba de Matemática del SIMCE. En otras palabras, un establecimiento con 9 docentes en la categoría de Insatisfactorio y 1 en la de Básico está vinculado a un puntaje SIMCE 2 puntos mayor que un establecimiento con 10 docentes en nivel Insatisfactorio. Del mismo modo, un aumento en un 1 punto porcentual en el porcentaje de docentes evaluados en nivel Competente (en relación a 1 nivel Insatisfactorio) se asocia con un incremento de cerca de 0,4 puntos en el resultado promedio de Matemática y Lectura del establecimiento en el SIMCE. Es decir, un establecimiento con 9 docentes en categoría de Insatisfactorio y 1 docente en el de Competente o Destacado está vinculado a un puntaje SIMCE 4 ó 5 puntos más alto, respectivamente, que un establecimiento con 10 docentes en nivel Insatisfactorio.

Existe una relación positiva y significativa entre el puntaje promedio de los docentes en la evaluación y el resultado promedio de los estudiantes de cuarto básico del establecimiento en la prueba SIMCE, siendo mayor la magnitud de esta

³⁸ Un 34% de los establecimientos en la muestra tiene sólo un profesor del primer ciclo evaluado.

relación para Matemática que para Lectura. La especificación A2 muestra que un profesorado con un punto adicional promedio en la Evaluación Docente (de un máximo de 4) está asociado con un aumento estadísticamente significativo en el puntaje promedio del establecimiento de 27 puntos en la prueba de Matemática y de cerca de 19 puntos en Lectura.

A nivel de instrumentos (A3), el portafolio y el informe referencia de terceros, consistentemente muestran efectos significativos y positivos tanto en Matemática como en Lectura. Sin embargo, la autoevaluación y la entrevista por un evaluador par parecen ser más débiles e incluso no significativos. Se observa que un aumento promedio de un punto en el portafolio promedio de todos los docentes de primer ciclo (de un máximo de 4), se asocia con un aumento de 18 puntos en el resultado promedio del establecimiento en la prueba de Matemática y de 12 puntos en Lectura. Por otro lado, el informe referencia de terceros tiene una asociación similar entre Matemática y Lectura, cercano a los 5 puntos promedio a nivel de establecimiento.

5.4 Selección de modelos

La selección de los modelos presentados se basó en criterios de mejor ajuste a los datos. Es decir, las especificaciones presentadas son aquellas con los mayores valores para los estadísticos de R^2 . La robustez de los resultados fue verificada por medio de un conjunto de pruebas con diferentes especificaciones.

En el Apéndice D se presentan un conjunto de especificaciones, tanto para el modelo individual como para el modelo a nivel de establecimientos. La lista de variables utilizadas es extensa, por lo que por simplicidad únicamente se presentan aquellas que son estadísticamente significativas. En el caso del modelo a nivel de estudiantes, las Tablas D.3 y D.4 presentan las regresiones realizadas para estudiantes de establecimientos con uno y más de un curso para el nivel de 4° básico. En tanto, la Tabla D.5 presenta los resultados de especificaciones realizadas sin los controles comunales para las estimaciones a nivel de establecimientos.

6. Síntesis y conclusiones

Los establecimientos subvencionados por el Estado concentran el mayor número de docentes del país, lo que se relaciona estrechamente con la mayor concentración de matrícula. En total los establecimientos municipales alcanzan a concentrar 43% de los docentes en ejercicio, así como también la mayor cantidad de docentes en edad de jubilar. Esto es, en torno a 16.500 docentes estarían próximos a dejar su plaza, lo que corresponde a 19% de la dotación municipal total. Cabe mencionar que estos docentes presentan un desempeño más bajo en la Evaluación Docente que aquellos más jóvenes.

Desde el año 2004 hasta 2010 se han realizado 79.763 evaluaciones, siendo educación básica el primer nivel educacional en incorporarse al sistema y educación de párvulos y educación especial la más recientes. Aun así, se observa que los docentes que ejercen en el nivel de educación de párvulos son los que proporcionalmente más se han evaluado y los que mejores resultados presentan. Por otro lado, en educación básica hay una gran porción de profesores evaluados, pero en el primer ciclo se encuentra la menor proporción de docentes calificados en el nivel Competente y Destacado.

Respecto a los docentes evaluados en la categoría de Insatisfactorio, ninguno logra llegar a un nivel de Destacado en evaluaciones posteriores, mientras que aproximadamente 60% transita hacia un desempeño Insatisfactorio y Básico en sus evaluaciones posteriores. Para el caso de docentes evaluados dos veces

consecutivas en nivel Insatisfactorio, sólo 16,5% logra un nivel de Competente en su tercera evaluación. Todo lo anterior es de interés si se consideran las nuevas consecuencias que pueden enfrentar los docentes con resultados en nivel Insatisfactorio o Básico en la Evaluación Docente, que establecen que ellos pueden ser desvinculados del establecimiento donde ejercen a partir de la primera vez que obtengan estos resultados.

A nivel de instrumentos, se identifica que los resultados de la pauta de autoevaluación, el informe de referencia de terceros y la entrevista por el evaluador par se encuentran sesgados hacia los puntajes superiores, mientras que los del portafolio de enseñanza pedagógica presentan una distribución más bien normal. A su vez, mientras los resultados de este último indican que en promedio los docentes no superan el punto de corte del nivel Básico (2,5 puntos), el instrumento de autoevaluación en promedio no baja del punto de corte del nivel Destacado (3,5 puntos). Esto deja la interrogante respecto de si estos puntajes son contradictorios entre sí. La respuesta podría ser negativa si los instrumentos abordan diferentes dominios, criterios o descriptores fijados por el Marco para la Buena Enseñanza. Pero podría ser positiva si se considera que finalmente los cuatro instrumentos buscan medir la práctica pedagógica del docente sin grandes diferencias en la escala de puntaje, sólo que se realiza a partir de diferentes miradas. Con todo, aun cuando se cumpliera que existen diferencias en lo que se intenta medir con cada instrumento, es cuestionable que éstas sean tan marcadas como para que un docente tenga una práctica en la categoría de Destacado en algunos aspectos y en nivel de Básico o Insatisfactorio en otros.

Adicionalmente, se observa que el efecto de los otros instrumentos consiste en desplazar la distribución de resultados hacia mejores niveles de desempeño, pero a nivel individual no todos los docentes se mantienen en el mismo cuartil de resultados. Es decir, la incorporación de uno o más de los instrumentos afecta el ordenamiento de los docentes en términos de resultados. Esto sugiere que los instrumentos de evaluación podrían estar midiendo diferentes constructos del desempeño docente, los que habría que determinar apropiadamente.

Al analizar la correlación no condicional entre los instrumentos de la Evaluación Docente y los resultados en las pruebas del SIMCE de Matemática y Lectura, se observa que a nivel agregado ésta es positiva y significativa. Al desagregar por instrumento se observa que la mayor magnitud se da con el informe de referencia de terceros y el portafolio, mientras que la más baja con la autoevaluación.

Por el lado del análisis de componentes principales, éste indica que la combinación ponderada de los instrumentos de la Evaluación Docente explica alrededor del 50% de la varianza total observada entre los resultados de los docentes. Adicionalmente, se desprende que el portafolio es el instrumento que debería tener la mayor ponderación en la combinación lineal, mientras que la autoevaluación debiera tener la menor, lo que es consistente con la fórmula de cálculo actual. Este análisis también sugiere que los instrumentos de entrevista por un evaluador par y el informe de referencia de terceros efectivamente añaden información acerca del desempeño docente, probablemente porque miden algún aspecto de las prácticas pedagógicas que no es capturado por los otros instrumentos. Los resultados de este ejercicio sugieren estudiar la posibilidad de darle mayor peso en el puntaje final al informe de referencia de terceros, actualmente en 10%, para incrementar el poder discriminativo de la Evaluación Docente. Cabe mencionar que en los casos de educación media y educación de párvulos el ponderador de la autoevaluación toma el mayor valor respecto a los otros ciclos de enseñanza, lo que es consistente con el análisis de varianza realizado al instrumento.

Finalmente, respecto al análisis realizado a los instrumentos, se observa que tanto el instrumento portafolio como el informe de referencia de terceros se asocian

siempre de manera positiva y significativa con el resultado de los estudiantes, tanto en las especificaciones a nivel individual como aquellas a nivel de establecimiento. La evaluación de los pares en general se asocia positiva y significativamente con el desempeño de los estudiantes, pero en menor magnitud que los instrumentos anteriores.

Estos análisis sobre los instrumentos de la Evaluación Docente son consistentes entre sí, posicionando al portafolio como el instrumento que más discrimina entre docentes y explica en mayor magnitud el desempeño de los estudiantes, seguido por el informe de referencia de terceros y finalmente por la entrevista de un evaluador par. El instrumento de autoevaluación parece ser el instrumento que menos aporta a la hora de identificar la calidad de un docente. Los dos primeros instrumentos son los que deberían tener el mayor peso en el puntaje final de la Evaluación Docente respecto a los otros dos instrumentos.

Por otro lado, se observa que existe una relación positiva entre el desempeño de los docentes en la Evaluación Docente y el resultado de sus estudiantes en el SIMCE de 4° básico. En general, las tres especificaciones realizadas en este documento, tanto para los aprendizajes de Matemática como de Lectura, son consistentes en mostrar que los alumnos vinculados en el tiempo con profesores de buen desempeño se asocian con un mejor resultado en las pruebas del SIMCE de 4° básico, que aquellos vinculados a docentes de peor desempeño, aun después de controlar por características observables de los estudiantes, docentes y del establecimiento. Cabe mencionar que los efectos son de mayor magnitud en los resultados de la prueba de Matemática que en la de Lectura, lo que se puede deber a que, como han sugerido diversas investigaciones, los resultados de Lectura se encuentran en mayor medida condicionados por factores relativos al capital cultural de las familias.

Respecto a la magnitud de la relación entre el desempeño de los docentes y el desempeño de los estudiantes, se observa en las estimaciones a nivel individual que tener profesores entre 1° y 4° básico con un punto adicional promedio en la Evaluación Docente se relaciona con un aumento promedio de 15,8 puntos en el resultado de la prueba de Matemática y de 9,5 puntos en la de Lectura del SIMCE. Es decir, cerca de un cuarto y un quinto de sus respectivas desviaciones estándar. Por otro lado, se observa que no hay un efecto estadísticamente significativo en el resultado de la prueba SIMCE de los estudiantes si éstos estuvieron vinculados a profesores con un desempeño en la categoría Básico, respecto de haber estado vinculados a profesores clasificados en nivel Insatisfactorio. Sin embargo, se observa un efecto positivo y creciente, tanto en Matemática como en Lectura, si es que los alumnos estuvieron vinculados con educadores en promedio calificados en nivel Competente o Destacado respecto de profesores en promedio en el nivel Insatisfactorio. A nivel de instrumentos, se observa que, en orden de magnitud el portafolio muestra la mayor contribución en el aumento del puntaje en el SIMCE, donde un punto adicional promedio se vincula con un aumento de 8 puntos en los resultados de la prueba SIMCE de Matemática y 6 en los de Lectura. En segundo lugar, el informe referencia de terceros con 4,2 y 2,2 puntos respectivamente y en tercer lugar, la entrevista por un evaluador par con 2,6 puntos en Matemática y 1 punto en Lectura. Finalmente, el instrumento de autoevaluación muestra ser no significativo en ambas pruebas.

Análisis similares pero a nivel de establecimientos son consistentes con los resultados anteriores, pero con mayores magnitudes. Un profesorado con un punto adicional promedio en la Evaluación Docente se asocia con un aumento del puntaje promedio de 27 puntos en la prueba de Matemática y de aproximadamente 19 puntos en la de Lectura. A nivel de instrumentos, el portafolio y el informe referencia de terceros, muestran efectos significativos y positivos tanto en Matemática como en Lectura. Sin embargo, la autoevaluación y la entrevista por un

evaluador par muestran ser más débiles e incluso no significativos, según sea la prueba.

Dadas las limitaciones de la información, los resultados deben ser tomados con cautela, pero considerando que las tendencias y magnitudes pueden ser informativas para el diseño de políticas públicas en esta materia.

7. Referencias bibliográficas

- Bravo, D., Falck, D., Gonzáles, R., Manzi, J., Peirano, C. "Do Good Teachers Make a Difference? The link between teacher certification programs and student achievement in Chile". LACEA 2009.
- Eisenberg, N. V. "The performance of teachers in Chilean public elementary schools: exploring its relationship with teacher backgrounds and student achievement, and its distribution across school and municipalities" Tesis doctoral, University of California Los Angeles. 2008.
- Goe, L. "The link between teacher quality and student outcomes: A research synthesis." Washington, D.C.: National Comprehensive Center for Teacher Quality. 2007.
- Goe, L., & Stickler, L. "Research and policy brief: Teacher quality and student achievement: Making the most of recent research." Washington, DC: National Comprehensive Center for Teacher Quality. 2008.
- Hanushek, E. "The Trade-Off Between Child Quantity and Quality.", *Journal of Political Economy* Vol. 100, No. 1, pp. 84-117. 1992.
- Ley N° 19.961, sobre Evaluación Docente. Promulgada el 9 de agosto de 2004. Ver www.docentemas.cl
- Ley N° 20.158, sobre Beneficios para Profesionales de la Educación. Promulgada el 27 de diciembre de 2006. Ver www.docentemas.cl
- Ley N° 20.501, sobre Calidad y Equidad de la Educación. Promulgada el 8 de febrero de 2011. Ver www.docentemas.cl
- Ramírez, M. "Diferencias Dentro de las Salas de Clases: Distribución del Rendimiento en Matemática". Centro de Estudios Públicos. 2007.
- Reglamento Evaluación Docente. Decreto N° 192 del Ministerio de Educación del 30 de agosto de 2004. Ver www.docentemas.cl
- Rivers, J. and Sanders, W. "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement". Research Progress Report, University of Tennessee Value-Added Research and Assessment Center, Knoxville, TN. 1996.
- Rivkin, S.; Hanushek, E.; Kain, J. "Teachers, Schools, and Academic Achievement". *Econometrica*, Vol. 73, No. 2. 2005.
- OECD. "Evaluating and Rewarding the Quality of Teachers International Practices". 2010.

8. Apéndices

Apéndice A

A.1: Planes de Superación.

Los profesionales de la educación que resulten evaluados con nivel de desempeño básico o insatisfactorio deben someterse a Planes de Superación Profesional diseñados a partir de las necesidades de desarrollo profesional derivadas de los informes de resultados de la Evaluación Docente. Dichos planes deben traducirse en acciones de aprendizaje y reaprendizaje respecto de las competencias, conocimientos, habilidades, dominios y criterios establecidos en el Marco para la Buena Enseñanza, aprobado por el Ministerio de Educación.

Corresponde al Ministerio de Educación, a través del Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP), aprobar tanto el diseño como la rendición de los Planes presentados por las comunas y transferir a los municipios y corporaciones municipales los recursos económicos para financiarlos, atendiendo al número de docentes evaluados con nivel Básico o Insatisfactorio.

Para tal efecto, durante los años 2004 – 2009 el CPEIP ha desarrollado un modelo de aplicación de los PSP que funciona a través de tres procesos reforzadores del diseño, monitoreo y evaluación de dichos planes.

- Un primer proceso está referido al diseño, seguimiento y rendición de los planes por parte de los encargados comunales y a la aprobación del diseño y rendición por parte del CPEIP. Este proceso se realiza mediante el software dedicado PSP Online, especialmente diseñado para tal efecto, que opera a través de internet.
- El segundo proceso está referido al monitoreo y evaluación de la gestión comunal de los PSP a nivel local, el cual se realiza a través de indicadores de gestión incluidos en el software, que se traducen en un informe anual de gestión comunal que elabora el CPEIP.
- El tercer proceso está referido al monitoreo y evaluación del programa Planes de Superación Profesional, con el propósito de mejorar su implementación a través de la optimización del software, de las competencias de los encargados comunales y el establecimiento de los principales aspectos a ser mejorados en siguientes aplicaciones. Este último proceso, durante los primeros años de la aplicación del programa, fue realizado a través de entrevistas en profundidad a los actores involucrados en la realización de los PSP, en un número determinado de comunas; en los últimos años el estudio focalizado ha sido remplazado por un análisis sobre el conjunto de los planes implementados, a partir de la información generada a través de las evidencias e indicadores contenidos en el software.

Con objeto de asegurar que las acciones de superación profesional cumplan con la condición de ser adecuadas, pertinentes, oportunas y efectivas, tal como lo estipula el reglamento que lo rige, le corresponde al CPEIP, en su rol de responsable técnico del proceso de Evaluación Docente, desarrollar las siguientes actividades:

- Aprobar la pertinencia de los Planes de Superación Profesional presentados por las comunas.
- Elaborar materiales de apoyo pedagógico para los docentes insatisfactorios.
- Evaluar la calidad de la prestación y la rendición del gasto ejecutado por la comuna.
- Monitorear y evaluar los Planes de Superación Profesional en su ejecución.

- Transferir los recursos para financiar las acciones de formación.

El modelo de implementación de los Planes de Superación Profesional y de sus tres procesos reforzadores, puede apreciarse en la gráfica siguiente.

A.2: Indicadores evaluados en el portafolio.

El portafolio es evaluado a partir de diferentes indicadores. Cabe mencionar que en el año 2004 existía un sistema computacional actual y protocolos de uso de datos diferentes a los actuales. Debido a eso no se dispone de un puntaje final por indicador. Para el año 2005 aún no se empleaban los datos por indicador para los informes de resultados, por lo que ninguna consulta mostraba un puntaje final por indicador. Por lo que aun cuando mediante otros esfuerzos fue posible recuperar información:

- Hay casos en que no se registra puntaje por indicador debido a factores como Insatisfactorio por Falta de Evidencia u otras circunstancias especiales.
- Hay casos en que los puntajes por dimensión eran *afectados por marcas* (que cambian el resultado por dimensión, pero no así de los indicadores) por lo que podría haber discrepancias en esos casos entre los datos por indicador y aquel por dimensión.
- Hay casos en que, *por efecto del redondeo* en los procesos parciales de cálculo puede haber diferencias iguales o inferiores a 0,01 en el puntaje final por dimensión.

Adicionalmente, como se observa en la Tabla A, los indicadores presentan diferencias año a año, ya sea en términos de su definición, o bien, en las rúbricas empleadas para su corrección. Esto hace inviable poder realizar un análisis desagregado de los mismos.

Los indicadores y sus respectivas dimensiones año a año son los siguientes:

Tabla A.1: Indicadores evaluados en el portafolio.

Dimensión	Indicador	2004	2005	2006	2007	2008	2009	2010
A	Relación entre el OFV-CMO y los objetivos de la unidad	si	No	no	no	no	no	no
	Formulación de objetivos	si	Si	si	si	no	no	no
	Formulación y coherencia entre los objetivos	no	No	no	no	no	si	si
	Secuencia de clases en la unidad	si	Si	si	si	no	no	no
	Secuencia de clases en función de los objetivos de aprendizaje de la unidad	no	No	no	no	si	no	si
	Calidad de la secuencia entre las clases la unidad	no	no	no	no	no	si	no
	Fundamentación del criterio para secuencias las actividades	si	si	no	no	no	no	no
	Relación entre las características de los alumnos y el OAU	si	no	no	no	no	no	no
	Análisis crítico de la unidad implementada	no	si	si	no	no	no	no
	Relación entre las actividades y los objetivos	no	no	si	no	si	si	si
	Análisis de la unidad implementada	no	no	si	si	si	no	no
B	Contribución de las actividades a los objetivos de la clase inicial	si	si	no	no	no	no	no
	Contribución de las actividades a los objetivos de la clase intermedia	si	no	no	no	no	no	no
	Contribución de las actividades a los objetivos de la clase final	si	si	no	no	no	no	no
	Fundamentación de las actividades y recursos	no	si	no	no	no	no	no
	Análisis de una clase implementada	no	no	si	si	si	si	si
	Desarrollo de las fortalezas académicas de los alumnos	no	no	si	no	si	no	no
	Variedad de estrategias didácticas	no	no	si	no	si	no	no
	Clima para el Aprendizaje	no	no	no	si	no	no	no
	Relación entre las actividades y los objetivos	no	no	no	si	no	no	no
	Análisis de los aspectos por mejorar de la unidad implementada	no	no	no	no	no	si	si
Análisis de las Fortalezas de la Unidad	no	no	no	no	no	si	si	
C	Fundamentación del instrumento o procedimiento de evaluación	si	si	no	no	no	no	no
	Adecuación de los ítems a las habilidades y contenidos declarados	si	no	no	no	no	no	no
	Adecuación de los ítems a los aprendizajes declarados	no	si	no	si	no	no	no
	Coherencia entre a evaluación y los objetivos de aprendizaje de la unidad	no	si	no	no	no	no	no
	Calidad del instrumento de evaluación	no	no	si	no	si	si	si
	Adecuación de la evaluación a los objetivos de aprendizaje de la unidad	no	no	si	si	si	si	si
	Calidad de la retroalimentación	no	no	si	no	si	no	no
	Calidad de las preguntas o instrucciones de la evaluación	no	no	no	si	no	no	no

	Calidad de la pauta de corrección de la evaluación	no	no	no	no	no	si	si
D	Tipo de comentarios frente a los aciertos y errores de los alumnos	si	no	no	no	no	no	no
	Información a partir de los resultados de la evaluación	si	si	no	no	no	no	no
	Utilización de la información obtenida en la evaluación	si	no	no	no	no	no	no
	Fundamentación del instrumento de evaluación	no	si	no	no	no	no	no
	Diseño de la evaluación	no	no	si	no	si	no	no
	Atribución de resultados de aprendizaje	no	no	si	si	si	no	no
	Análisis de la evaluación	no	no	si	si	si	no	no
	Variedad de estrategias didácticas	no	no	no	si	no	no	no
	Calidad de la retroalimentación	no	no	no	no	no	si	si
	Análisis de los resultados de la evaluación	no	no	no	no	no	si	si
	E	Identificación de aspectos más efectivos de la unidad y argumentos mencionados.	si	no	no	no	no	no
Identificación de aspectos menos efectivos de la unidad y argumentos mencionados		si	no	no	no	no	no	no
Responsabilidad por los resultados de aprendizaje de los alumnos		si	si	no	no	no	no	no
Detección y satisfacción de las propias necesidades de aprendizaje		si	si	no	si	no	no	no
Riqueza pedagógica en el diseño de unidades de aprendizaje		no	si	no	no	no	no	no
Incorporación de las características de los alumnos		no	no	si	no	si	no	no
Estrategias para detectar el nivel de aprendizaje de los alumnos		no	no	si	no	si	no	no
Detección de las propias necesidades de aprendizaje		no	no	si	no	si	no	no
Desarrollo de las fortalezas de los alumnos		no	no	no	si	no	no	no
Promoción de la formación de valores		no	no	no	si	no	no	no
Variedad de actividades para trabajar dificultades de los alumnos		no	no	no	no	no	si	si
Trabajo colaborativo con otros docentes		no	no	no	no	no	si	si
Adecuación de las actividades a las distintas características de los alumnos		no	no	no	no	no	si	si
F	Grado en que el profesor logra que sus estudiantes se involucren en la clase	si	no	no	no	no	no	no
	Grado en que el profesor logra que se focalicen en las actividades propuestas	no	si	no	no	no	no	no
	Focalización de los alumnos en las actividades propuestas	no	no	si	si	si	si	si
	Tipo de actitudes que el profesor promueve entre los alumnos	si	no	no	no	no	no	no
	Promoción de la participación de todos los alumnos	no	si	si	si	si	si	si
	Tipo de relación que el profesor establece con sus alumnos	si	no	no	no	no	no	no
	Manejo del grupo de curso	no	si	si	si	si	no	no

	Supervisión de las actividades	no	no	no	no	no	si	si
G	Calidad de la estructura de la clase	si	si	si	si	si	no	no
	Utilización del tiempo instruccional	si	si	si	si	si	no	no
	Contribución de las actividades al logro de los objetivos y contenidos de la clase	si	si	si	si	si	si	si
	Calidad del Inicio de la Clase	no	no	no	no	no	si	si
	Calidad del Cierre de la Clase	no	no	no	no	no	si	si
H	Relaciones que establece el profesor entre los aprendizajes de la clase y los conocimientos de los alumnos	si	si	no	no	no	no	no
	Calidad de la incorporación de las intervenciones de los alumnos	si	no	no	no	no	no	no
	Calidad de las explicaciones que hace el profesor	si	no	si	si	si	si	si
	Acompañamiento del profesor en el desarrollo de las actividades en función del aprendizaje	si	si	si	si	si	no	no
	Calidad de las preguntas que formula el profesor e interacción a partir de ellas	si	si	no	no	no	no	no
	Calidad de las preguntas	no	no	no	no	no	si	si
	Calidad de la interacción que promueve el profesor	no	no	si	si	si	no	no
	Calidad de la expresión oral en Inglés*	no	no	no	no	si	no	no
	Calidad de la retroalimentación a los alumnos	no	no	no	no	no	si	si

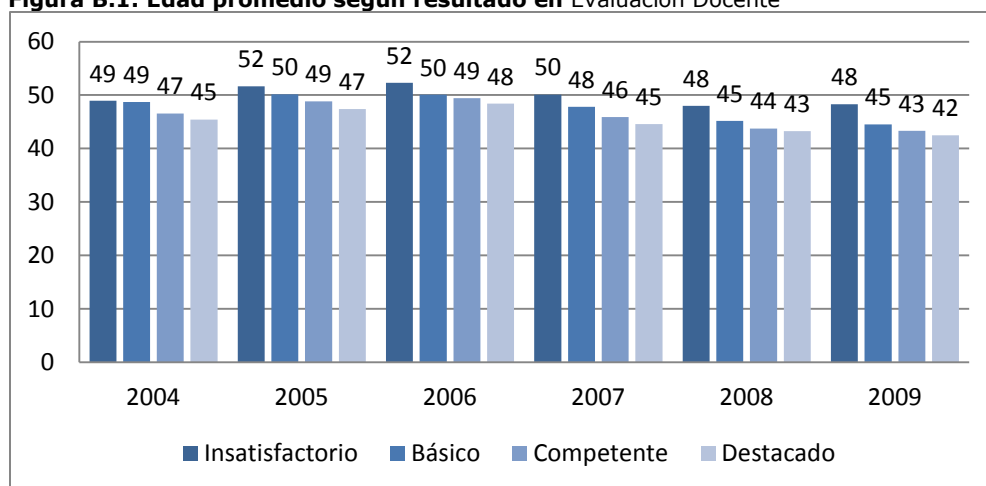
Fuente: Elaboración propia con registros del CPEIP.

*: Este indicador sólo se aplica para los docentes de inglés.

Nota: Cabe mencionar que indicadores pueden cambiar de nombre agregando o sacando alguna especificación, por lo que se consideran como diferentes. Así como mantienen el nombre pero cambian respecto a la dimensión en la cual son considerados.

Apéndice B: Análisis de datos

Figura B.1: Edad promedio según resultado en Evaluación Docente



Fuente: Elaboración propia con registros del CPEIP.

Tabla B.1: Análisis de varianza de la entrevista del evaluador par por ciclo.

Varianza	2005	2006	2007	2008	2009	2010	Total
E. Parvularia				0,37	0,38	0,58	0,38
1er Ciclo	0,38	0,49	0,48	0,42	0,35	0,45	0,44
2do Ciclo	0,38	0,48	0,48	0,42	0,38	0,49	0,46
E. Media	0,31	0,47	0,39	0,40	0,36	0,45	0,40
E. Especial					0,41	0,50	0,43

Fuente: Elaboración propia con registros del MINEDUC.

Tabla B.2: Análisis de varianza del portafolio por ciclo.

Varianza	2005	2006	2007	2008	2009	2010	Total
E. Parvularia				0,08	0,08	0,06	0,10
1er Ciclo	0,12	0,07	0,06	0,08	0,08	0,07	0,09
2do Ciclo	0,14	0,10	0,12	0,10	0,08	0,08	0,10
E. Media	0,14	0,07	0,10	0,09	0,09	0,09	0,10
E. Especial					0,09	0,11	0,10

Fuente: Elaboración propia con registros del MINEDUC.

Tabla B.3: Ponderador de cada indicador del portafolio según análisis de Componentes Principales. 2006 – 2010. Todos los ciclos.

	Ponderador de componentes principales				
	2006	2007*	2008**	2009***	2010
A: Organización de los elementos de la unidad	0,436	0,408	0,379	0,359	0,367
B: Análisis de las actividades de las clases	0,447	0,410	0,420	0,387	0,388
C: Calidad de la evaluación de la unidad	0,357	0,358	0,389	0,339	0,327
D: Reflexión a partir de los resultados de la evaluación	0,413	0,394	0,402	0,361	0,389
E: Reflexión pedagógica	0,410	0,392	0,426	0,432	0,419
F: Ambiente de la clase para el aprendizaje	0,231	0,203	0,243	0,284	0,297
G: Estructura de la clase	0,250	0,274	0,260	0,342	0,333
H: Interacción pedagógica	0,170	0,334	0,242	0,304	0,286

Fuente: Elaboración propia con registros del MINEDUC.

*: Educación media comienza desde este año a ser evaluada.

** : Educación de párvulos comienza desde este año a ser evaluada.

***: Educación especial comienza desde este año a ser evaluada.

Tabla B.4: Ponderador de cada indicador del portafolio según análisis de Componentes Principales. 2006 – 2010. Educación de párvulos..

	Ponderador de componentes principales		
	2008	2009	2010
A: Organización de los elementos de la unidad	0,369	0,356	0,449
B: Análisis de las actividades de las clases	0,406	0,449	0,441
C: Calidad de la evaluación de la unidad	0,393	0,351	0,341
D: Reflexión a partir de los resultados de la evaluación	0,398	0,382	0,399
E: Reflexión pedagógica	0,426	0,438	0,480
F: Ambiente de la clase para el aprendizaje	0,230	0,236	0,133
G: Estructura de la clase	0,285	0,293	0,231
H: Interacción pedagógica	0,266	0,265	0,169

Fuente: Elaboración propia con registros del MINEDUC.

Tabla B.5: Ponderador de cada indicador del portafolio según análisis de Componentes Principales. 2006 – 2010. 1er ciclo.

	Ponderador de componentes principales				
	2006	2007	2008	2009	2010
A: Organización de los elementos de la unidad	0,427	0,425	0,432	0,412	0,384
B: Análisis de las actividades de las clases	0,442	0,442	0,444	0,444	0,416
C: Calidad de la evaluación de la unidad	0,393	0,399	0,295	0,360	0,371
D: Reflexión a partir de los resultados de la evaluación	0,406	0,354	0,409	0,412	0,399
E: Reflexión pedagógica	0,414	0,394	0,413	0,384	0,432
F: Ambiente de la clase para el aprendizaje	0,151	0,221	0,266	0,206	0,245
G: Estructura de la clase	0,230	0,267	0,293	0,280	0,268
H: Interacción pedagógica	0,239	0,253	0,186	0,256	0,255

Fuente: Elaboración propia con registros del MINEDUC.

Tabla B.6: Ponderador de cada indicador del portafolio según análisis de Componentes Principales. 2006 – 2010. 2do ciclo.

	Ponderador de componentes principales				
	2006	2007	2008	2009	2010
A: Organización de los elementos de la unidad	0,427	0,386	0,445	0,421	0,403
B: Análisis de las actividades de las clases	0,413	0,432	0,444	0,421	0,435
C: Calidad de la evaluación de la unidad	0,399	0,407	0,367	0,363	0,404
D: Reflexión a partir de los resultados de la evaluación	0,416	0,341	0,409	0,380	0,390
E: Reflexión pedagógica	0,365	0,362	0,404	0,379	0,425
F: Ambiente de la clase para el aprendizaje	0,205	0,260	0,224	0,222	0,233
G: Estructura de la clase	0,290	0,314	0,269	0,274	0,243
H: Interacción pedagógica	0,237	0,291	0,134	0,318	0,199

Fuente: Elaboración propia con registros del MINEDUC.

Tabla B.7: Ponderador de cada indicador del portafolio según análisis de Componentes Principales. 2006 – 2010. Educación media.

	Ponderador de componentes principales			
	2007	2008	2009	2010
A: Organización de los elementos de la unidad	0,400	0,451	0,419	0,380
B: Análisis de las actividades de las clases	0,396	0,452	0,428	0,410
C: Calidad de la evaluación de la unidad	0,349	0,334	0,294	0,389
D: Reflexión a partir de los resultados de la evaluación	0,345	0,412	0,411	0,400
E: Reflexión pedagógica	0,399	0,415	0,421	0,391
F: Ambiente de la clase para el aprendizaje	0,287	0,179	0,216	0,257
G: Estructura de la clase	0,314	0,257	0,250	0,287
H: Interacción pedagógica	0,320	0,201	0,316	0,273

Fuente: Elaboración propia con registros del MINEDUC.

Tabla B.8: Ponderador de cada indicador del portafolio según análisis de Componentes Principales. 2006 – 2010. Educación especial.

	Ponderador de componentes principales	
	2009	2010
A: Organización de los elementos de la unidad	0,350	0,356
B: Análisis de las actividades de las clases	0,374	0,349
C: Calidad de la evaluación de la unidad	0,305	0,349
D: Reflexión a partir de los resultados de la evaluación	0,343	0,381
E: Reflexión pedagógica	0,385	0,394
F: Ambiente de la clase para el aprendizaje	0,324	0,339
G: Estructura de la clase	0,384	0,344
H: Interacción pedagógica	0,355	0,310

Fuente: Elaboración propia con registros del MINEDUC

Apéndice C: Bases de datos para estimaciones

En el proceso de construcción de la base de datos se fueron perdiendo observaciones. En la base de SIMCE hay 99.690 estudiantes en el sector municipal, una vez terminada la vinculación de esta con las otras bases, el registro indica 81.707 observaciones. Esta diferencia se explica fundamentalmente por la imposibilidad de vincular a alumnos de distintos años, es decir, que de un año a otro hay un conjunto de estudiantes que no es posible encontrar el año anterior. Posteriormente, la base fue filtrada para aquellos estudiantes que tienen para todos los años a un profesor con un resultado en la Evaluación Docente. Finalmente, el número de observaciones total es de 55.811.

Sin embargo, esta pérdida es relevante si es que no es aleatoria. En la Tabla C.1, las comparaciones de medias de las principales variables a nivel de alumnos indican que la pérdida de observaciones es aleatoria, ya que ambos grupos tienen medias similares en las variables analizadas.

Tabla C.1: Comparación de medias entre base inicial y base final construida

Variables	Base Simce		Base construida	
	Obs	Media	Obs	Media
Alumno hombre	99.690	0,505	55.811	0,4904
Ptje Lectura	98.881	248,09	55.610	252,79
Ptje Matemática	99.231	231,95	55.811	236,92
Educación Madre				
Sin Escolaridad	99.690	0,0063	55.811	0,0050
Ed. Básica Incompleta	99.690	0,2138	55.811	0,2132
Ed. Básica Completa	99.690	0,1758	55.811	0,1779
Ed. Media Incompleta	99.690	0,1999	55.811	0,1983
Ed. Media Completa	99.690	0,3044	55.811	0,3130
Ed. CFT o IP Incompleta	99.690	0,0237	55.811	0,0215
Ed. CFT o IP Completa	99.690	0,0496	55.811	0,048
Ed. Universitaria Incompleta	99.690	0,0139	55.811	0,0122
Ed. Universitaria Completa	99.690	0,0167	55.811	0,0162
Ed. Postgrado	99.690	0,0008	55.811	0,0005

Fuente: Elaboración propia con registros del MINEDUC.

La Tabla C.2 indica que alrededor del 60% de las madres no completó la enseñanza media y sólo 6,4% tiene educación terciaria completa. El 49% de los estudiantes son hombres. La media de los resultados en Matemática es de 236 puntos y la de Lectura de 252 puntos. El tamaño de los cohorte en donde se encuentran los estudiantes es en promedio de 24 estudiantes y la clasificación socioeconómica de sus establecimientos se concentra en el Grupo B con un 17% que pertenece al grupo más bajo. Los docentes de Matemática y Lectura no presentan diferencias en los puntajes promedio de la Evaluación Docente, con un puntaje promedio del portafolio de 2,2.

La Tabla C.3 presenta la estadística descriptiva de la base de datos a nivel de establecimientos. Cerca de 80% de los establecimientos pertenecen a los dos grupos socioeconómicos más bajos (GSE A y B). Respecto a sus resultados en la prueba SIMCE, el promedio de Matemática es de 228 puntos y el de Lectura de 246 puntos.

Tabla C.2: Estadísticas descriptivas. Alumnos 1° básico 2005 a 4° básico 2008

Variable	N° Observaciones	Promedio	Desviación Estándar
Alumnos 2008			
Madre: Sin Escolaridad	55.811	0,01	0,07
Madre Ed. Básica Incompleta	55.811	0,21	0,41
Madre: Ed. Básica Completa	55.811	0,18	0,38
Madre: Ed. Media Incompleta	55.811	0,20	0,40
Madre: Ed. Media Completa	55.811	0,31	0,46
Madre: Ed. CFT o IP Incompleta	55.811	0,02	0,15
Madre: Ed. CFT o IP Completa	55.811	0,05	0,21
Madre: Ed. Universitaria Incompleta	55.811	0,01	0,11
Madre: Ed. Universitaria Completa	55.811	0,02	0,13
Madre: Ed. Postgrado	55.811	0,00	0,02
Hombre	55.811	0,49	0,50
Preescolar	55.811	0,77	0,41
Ranking promedio	55.811	11,78	7,75
Ptje Matemática	55.811	236,92	51,79
Ptje Lectura	55.610	252,79	50,93
Establecimientos			
Ptje Prom Estab Matemática 2005 – 2007	53.894	234,73	19,21
Ptje Prom Estab Lectura 2005 - 2007	53.893	243,52	17,16
N° de Alumnos por cohorte 2008	55.811	24,24	41,30
N° de Alumnos ² por cohorte 2008	55.811	1.512,62	7.140,34
GSE A 2008	55.811	0,17	0,38
GSE B 2008	55.811	0,52	0,50
GSE C 2008	55.811	0,27	0,44
GSE D 2008	55.811	0,04	0,19
GSE E 2008	55.811	0,00	0,00
Profesores			
Matemática			
Hombre 2008	53.551	0,16	0,37
Hombre 2007	53.518	0,15	0,35
Hombre 2006	54.407	0,10	0,31
Hombre 2005	53.725	0,09	0,29
Prom Ptje Portafolio 2004 - 2010	55.811	2,20	0,23
Prom Eval Docente 2004 – 2010	55.811	2,74	0,48
Años de experiencia	53.551	22,83	11,13
Edad promedio	55.778	49,03	7,33
Prom Insatisfactorio	55.401	0,004	0,06
Prom Básico	55.401	0,30	0,46
Prom Competente	55.401	0,64	0,47
Prom Destacado	55.401	0,03	0,17
Lectura			
Hombre 2008	53.499	0,13	0,34
Hombre 2007	53.273	0,13	0,33
Hombre 2006	54.224	0,10	0,30
Hombre 2005	53.505	0,09	0,29
Prom Ptje Portafolio 2004 – 2010	55.610	2,20	0,23

Prom Eval Docente 2004 - 2010	55.610	2,73	0,48
Años de experiencia	53.499	22,83	11,09
Edad promedio	55.577	49,05	7,37
Prom Insatisfactorio	55.610	0,003	0,06
Prom Básico	55.610	0,30	0,46
Prom Competente	55.610	0,64	0,47
Prom Destacado	55.610	0,03	0,17

Fuente: Elaboración propia, en base a registros del MINEDUC.

Tabla C.3: Estadísticas descriptivas. Todos los docentes de primer ciclo. Base de establecimientos.

Establecimientos 2008	N° Observaciones	Promedio	Desviación Estándar
Prom Matemática 2008	2.763	228,99	23,71
Prom Lectura y comunicación 2008	2.765	246,65	21,77
GSE A 2008	2.765	0,44	0,497
GSE B 2008	2.765	0,43	0,494
GSE C 2008	2.765	0,11	0,318
GSE D 2008	2.765	0,01	0,103
GSE E 2008	2.765	0,0002	0,015
Horas de contrato promedio docentes 2008	2.765	34,94	6,63
Años de servicio promedio docente 2008	2.765	20,45	8,16
Porcentaje de docentes hombres 2008	2.765	31,93	33,55
N° de Alumnos por cohorte 2008	2.765	24,15	29,73
N° de Alumnos^2 por cohorte 2008	2.765	1.467,80	3.535,36
Ptje Prom final ED	4.241	2,58	0,18
% docentes Insatisfactorios	4.310	0,72	6,01
% docentes Básicos	4.310	33,81	33,12
% docentes Competentes	4.310	57,67	33,48
% docentes Destacados	4.310	6,17	14,81
Ptje Prom Portafolio	4.241	2,17	0,19
Ptje Prom Autoevaluación	4.241	3,83	0,28
Ptje Prom IRT	4.207	3,09	0,48
Ptje Prom Evaluador par	4.240	2,98	0,47
Tamaño comuna	3.520	65.055	86.888
% competencia comuna	3.549	1,15	6,17
DAEM	4.310	0,79	0,40

Fuente: Elaboración propia, en base a registros del MINEDUC

Apéndice D: Estimaciones

Tabla D.1: Estimación SIMCE 4° Básico 2008 a nivel de alumno, controlando por características de los estudiantes, docentes y establecimientos.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
P. Básico	-1.245			-0.221		
	(-0.94)			(-0.16)		
P. Competente	3.801**			2.947*		
	(2.90)			(2.15)		
P. Destacado	9.305***			6.601***		
	(5.61)			(3.84)		
Promedio Puntajes ED		15.85***			9.531***	
		(17.06)			(10.04)	
Puntaje promedio PF			8.051***			6.193***
			(9.38)			(7.03)
Puntaje promedio IRT			4.205***			2.210***
			(11.94)			(6.07)
Puntaje promedio AE			-0.763			-0.459
			(-1.32)			(-0.78)
Puntaje promedio EP			2.674***			1.021**
			(7.42)			(2.76)
Edad promedio docentes	-0.123***	-0.0908**	-0.0966**	-0.0500	-0.0306	-0.0339
	(-3.68)	(-2.72)	(-2.89)	(-1.47)	(-0.90)	(-0.99)
Años de experiencia	0.307***	0.265***	0.276***	0.216**	0.200**	0.209**
	(4.53)	(3.91)	(4.06)	(3.03)	(2.81)	(2.92)
Años de experiencia^2	-0.00490**	-0.00392*	-0.00433**	-0.00366*	-0.00336	-0.00364*
	(-2.94)	(-2.36)	(-2.60)	(-2.10)	(-1.93)	(-2.09)
Docente 4° B. hombre	-3.512***	-3.292***	-3.108***	-4.013***	-3.850***	-3.734***
	(-6.58)	(-6.17)	(-5.82)	(-6.66)	(-6.38)	(-6.18)
Alumno hombre (d)	11.38***	11.39***	11.35***	-6.093***	-6.080***	-6.099***
	(32.48)	(32.54)	(32.45)	(-16.79)	(-16.77)	(-16.82)
Escolaridad madre: Básica incompleta	-0.958	-0.933	-0.958	-0.176	-0.174	-0.176
	(-0.99)	(-0.96)	(-0.99)	(-0.18)	(-0.17)	(-0.18)
Escolaridad madre: Básica completa	-0.933	-0.894	-0.898	-0.990	-0.989	-0.986
	(-0.95)	(-0.91)	(-0.91)	(-0.98)	(-0.98)	(-0.97)
Escolaridad madre: Media incompleta	0.785	0.802	0.778	0.501	0.513	0.513
	(0.81)	(0.82)	(0.80)	(0.50)	(0.51)	(0.51)
Escolaridad madre: Media completa	2.658**	2.692**	2.635**	1.697	1.723	1.708
	(2.80)	(2.84)	(2.78)	(1.73)	(1.76)	(1.75)
Escolaridad madre: CFT o IP incompleta	4.651**	4.613**	4.547**	2.659	2.640	2.634
	(3.27)	(3.25)	(3.20)	(1.80)	(1.78)	(1.78)

Escolaridad madre: CFT o IP completa	4.171***	4.067***	4.039***	3.895**	3.824**	3.823**
	(3.63)	(3.54)	(3.52)	(3.28)	(3.22)	(3.22)
Escolaridad madre: Universitaria incompleta	8.313***	8.257***	8.142***	9.161***	9.128***	9.081***
	(4.80)	(4.77)	(4.71)	(5.13)	(5.11)	(5.09)
Escolaridad madre: Universitaria completa	7.992***	7.959***	7.930***	6.297***	6.277***	6.247***
	(4.97)	(4.96)	(4.94)	(3.77)	(3.76)	(3.74)
Escolaridad madre: Postgrado	8.499	8.398	8.530	-6.473	-6.449	-6.334
	(1.18)	(1.16)	(1.18)	(-0.85)	(-0.85)	(-0.83)
Alumno: Preescolar	0.214	0.215	0.198	0.642	0.641	0.641
	(0.49)	(0.50)	(0.46)	(1.43)	(1.42)	(1.42)
Alumnos: Ranking promedio	4.137***	4.133***	4.126***	3.815***	3.813***	3.809***
	(160.00)	(159.99)	(159.75)	(141.96)	(141.93)	(141.75)
Tamaño cohorte	-0.572***	-0.574***	-0.573***	-0.519***	-0.520***	-0.520***
	(-38.81)	(-39.04)	(-38.98)	(-34.79)	(-34.90)	(-34.85)
Tamaño cohorte^2	0.00220** *	0.00221** *	0.00221** *	0.00188** *	0.00188** *	0.00189** *
	(27.87)	(28.06)	(28.06)	(24.09)	(24.14)	(24.15)
Promedio SIMCE 4°B. 2005-2007. Matemática	0.424***	0.418***	0.414***			
	(36.28)	(35.83)	(35.43)			
Promedio SIMCE 4°B. 2005- 2007. Lectura				0.372***	0.369***	0.366***
				(27.19)	(26.98)	(26.71)
Constante	106.1***	67.22***	74.72***	152.5***	129.8***	134.0***
	(19.13)	(11.33)	(12.07)	(26.77)	(21.45)	(21.15)
N	51809	51809	51809	51757	51757	51757
R² ajustado	0.423	0.424	0.424	0.362	0.362	0.362

Fuente: Elaboración propia, en base a registros del MINEDUC.

Nota 1: ED: Evaluación Docente. PF: portafolio. IRT: informe de referencia de terceros. AE: autoevaluación. EP: evaluador par.

Nota 2: *** nivel de significancia 0.1%, ** 1% y * 5%.

Tabla D2: Estimación SIMCE 4° Básico 2008 a nivel de establecimiento, controlando por características de los estudiantes, docentes, establecimientos y comuna.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
% Básico	0.252*			0.283*		
	(2.05)			(2.49)		
% Competente	0.385**			0.360**		
	(3.14)			(3.19)		
% Destacado	0.502***			0.439***		
	(3.95)			(3.74)		
Puntaje promedio final ED		26.85***			18.72***	
		(7.59)			(5.77)	
Puntaje promedio PF			17.89***			11.91***

			(5.44)			(3.95)
Puntaje promedio AE			4.712*			0.946
			(1.98)			(0.43)
Puntaje promedio IRT			4.641***			4.703***
			(3.90)			(4.32)
Puntaje promedio EP			3.333*			2.031
			(2.37)			(1.57)
Tamaño cohorte	-0.106**	-0.105**	-0.111**	-0.171***	-0.170***	-0.170***
	(-2.55)	(-2.56)	(-2.68)	(-4.50)	(-4.47)	(-4.46)
Tamaño cohorte^2	0.000811*	0.000820*	0.000840*	0.00104**	0.00104**	0.00104**
	(3.01)	(3.05)	(3.11)	(4.19)	(4.19)	(4.19)
GSE B	7.617***	7.568***	7.596***	6.689***	6.690***	6.625***
	(6.59)	(6.56)	(6.57)	(6.29)	(6.29)	(6.24)
GSE C	22.06***	21.87***	21.83***	21.58***	21.50***	21.22***
	(12.54)	(12.46)	(12.40)	(13.34)	(13.29)	(13.11)
GSE D	44.72***	44.79***	44.36***	41.11***	41.17***	40.51***
	(11.42)	(11.47)	(11.33)	(11.42)	(11.44)	(11.25)
Años de servicio	0.415	0.399	0.402	0.391	0.391	0.370
	(1.24)	(1.19)	(1.19)	(1.26)	(1.26)	(1.19)
Años de servicio^2	-0.00962	-0.00937	-0.00938	-0.0101	-0.00998	-0.00966
	(-1.18)	(-1.11)	(-1.11)	(-1.29)	(-1.28)	(-1.24)
Hrs contrato promedio	0.111	0.0872	0.114	0.223**	0.202**	0.218**
	(1.01)	(0.80)	(1.04)	(2.21)	(2.00)	(2.17)
% hombre	-0.0765**	-0.0749**	-0.0747**	-0.0706**	-0.0721**	-0.0685**
	(-3.23)	(-3.16)	(-3.16)	(-3.23)	(-3.29)	(-3.13)
% Docentes Evaluado	0.0402	0.0523	0.0350	0.0194	0.0283	0.0145
	(1.06)	(1.38)	(0.92)	(0.55)	(0.81)	(0.41)
DAEM	6.111	7.053	6.184	12.72	13.13	12.85
	(0.41)	(0.48)	(0.42)	(0.94)	(0.97)	(0.95)
Tamaño Comuna	-	-	-	-	-	-
	0.0000252	0.0000236	0.0000251	0.0000213	0.0000205	0.0000207
	(-1.33)	(-1.25)	(-1.33)	(-1.23)	(-1.18)	(-1.19)
% Competencia	0.110	0.147	0.104	-0.0913	-0.0567	-0.106
	(0.30)	(0.40)	(0.28)	(-0.27)	(-0.17)	(-0.31)
Constante	146.9***	179.9***	135.7***	192.8***	205.9***	192.3***
	(6.12)	(7.08)	(5.41)	(8.74)	(8.80)	(8.35)
N	2378	2378	2377	2379	2379	2378
R² ajustado	0.266	0.269	0.268	0.272	0.272	0.275

Fuente: Elaboración propia, en base a registros del MINEDUC.

Nota 1: ED: Evaluación Docente. PF: portafolio. IRT: informe de referencia de terceros. AE: autoevaluación. EP: evaluador par.

Nota 2: *** nivel de significancia 0.1%, ** 1% y * 5%.

Tabla D.3: Estimación SIMCE 4° Básico 2008 a nivel de alumno en establecimientos con un solo curso por nivel.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
P. Básico	-7.774**			-3.225		
	(-2.95)			(-1.27)		
P. Competente	-0.724			1.741		
	(-0.28)			(0.69)		
P. Destacado	-0.0175			5.191		
	(-0.01)			(1.57)		
Promedio Puntajes ED		20.32***			13.12***	
		(12.08)			(7.86)	
Puntaje promedio PF			12.48***			10.78***
			(8.03)			(7.02)
Puntaje promedio IRT			3.543***			1.800**
			(5.88)			(2.93)
Puntaje promedio AE			-0.344			0.732
			(-0.35)			(0.75)
Puntaje promedio EP			3.838***			1.187
			(5.76)			(1.77)
Edad promedio docentes	-0.102	-0.0665	-0.0722	-0.0513	-0.0376	-0.0369
	(-1.84)	(-1.20)	(-1.30)	(-0.92)	(-0.67)	(-0.66)
Años de experiencia	0.211	0.143	0.153	0.137	0.108	0.116
	(1.81)	(1.23)	(1.31)	(1.14)	(0.90)	(0.97)
Años de experiencia^2	-0.00290	-0.00117	-0.00141	-0.00280	-0.00205	-0.00232
	(-1.01)	(-0.41)	(-0.49)	(-0.95)	(-0.70)	(-0.79)
Docente 4° B. hombre	-4.182***	-3.801***	-3.860***	-6.674***	-6.389***	-6.270***
	(-4.87)	(-4.44)	(-4.50)	(-6.96)	(-6.67)	(-6.54)
Alumno hombre (d)	10.25***	10.28***	10.27***	-7.022***	-7.001***	-6.996***
	(17.28)	(17.36)	(17.36)	(-11.65)	(-11.62)	(-11.62)
Escolaridad madre: Básica incompleta	0.323	0.400	0.342	-1.788	-1.804	-1.799
	(0.20)	(0.24)	(0.21)	(-1.07)	(-1.08)	(-1.08)
Escolaridad madre: Básica completa	0.257	0.346	0.308	-2.600	-2.600	-2.620
	(0.15)	(0.21)	(0.19)	(-1.53)	(-1.53)	(-1.54)
Escolaridad madre: Media incompleta	1.736	1.860	1.810	-1.061	-1.043	-1.052
	(1.04)	(1.11)	(1.08)	(-0.62)	(-0.61)	(-0.62)
Escolaridad madre: Media completa	3.489*	3.527*	3.435*	-0.378	-0.364	-0.393
	(2.11)	(2.14)	(2.09)	(-0.22)	(-0.22)	(-0.23)
Escolaridad madre: CFT o IP incompleta	7.717**	7.881**	7.774**	5.862*	6.000*	6.020*
	(2.68)	(2.74)	(2.71)	(1.99)	(2.03)	(2.04)
Escolaridad madre: CFT o IP completa	7.383***	7.300***	7.288***	3.066	3.035	2.958

	(3.38)	(3.34)	(3.34)	(1.38)	(1.36)	(1.33)
Escolaridad madre: Universitaria incompleta	13.09***	13.24***	13.00***	9.991**	10.08**	10.04**
	(3.61)	(3.65)	(3.59)	(2.71)	(2.74)	(2.73)
Escolaridad madre: Universitaria completa	8.872**	8.952**	8.828**	2.580	2.616	2.429
	(2.91)	(2.94)	(2.90)	(0.84)	(0.85)	(0.79)
Escolaridad madre: Postgrado	-22.76	-23.40	-23.76	-16.79	-16.99	-17.16
	(-1.40)	(-1.44)	(-1.47)	(-1.02)	(-1.03)	(-1.04)
Alumno: Preescolar	1.133	1.125	1.088	0.625	0.642	0.608
	(1.57)	(1.56)	(1.51)	(0.85)	(0.87)	(0.83)
Alumnos: Ranking promedio	5.162***	5.152***	5.151***	4.779***	4.774***	4.775***
	(92.37)	(92.36)	(92.36)	(83.49)	(83.45)	(83.48)
Tamaño cohorte	-2.420***	-2.443***	-2.431***	-2.512***	-2.528***	-2.524***
	(-17.18)	(-17.37)	(-17.27)	(-17.43)	(-17.54)	(-17.51)
Tamaño cohorte^2	0.0141***	0.0144***	0.0140***	0.0172***	0.0175***	0.0174***
	(4.74)	(4.84)	(4.71)	(5.61)	(5.71)	(5.67)
Promedio SIMCE 4°B. 2005-2007. Matemática	0.457***	0.450***	0.448***			
	(24.28)	(23.93)	(23.78)			
Promedio SIMCE 4°B. 2005- 2007. Lectura				0.413***	0.405***	0.403***
				(18.93)	(18.59)	(18.47)
Constante	108.7***	54.88***	60.76***	170.7***	139.2***	138.3***
	(11.25)	(5.33)	(5.64)	(19.10)	(14.68)	(13.80)
N	18233	18233	18233	18267	18267	18267
R² ajustado	0.406	0.408	0.409	0.360	0.360	0.360

Fuente: Elaboración propia, en base a registros del MINEDUC

Nota 1: ED: Evaluación Docente. PF: portafolio. IRT: informe de referencia de terceros. AE: autoevaluación. EP: evaluador par.

Nota 2: *** nivel de significancia 0,1%, ** 1% y * 5%.

Tabla D.4: Estimación SIMCE 4° Básico 2008 a nivel de alumno en establecimientos con más de un curso por nivel.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
P. Básico	1.504			1.065		
	(0.97)			(0.64)		
P. Competente	6.258***			3.676*		
	(4.11)			(2.22)		
P. Destacado	11.64***			6.634**		
	(6.04)			(3.22)		
Promedio Puntajes ED		15.04***			8.625***	
		(13.09)			(7.18)	
Puntaje promedio PF			8.057***			5.662***
			(7.58)			(5.07)
Puntaje promedio IRT			4.208***			2.026***

			(9.33)			(4.26)
Puntaje promedio AE			-0.814			-0.614
			(-1.09)			(-0.79)
Puntaje promedio EP			2.220***			0.713
			(5.03)			(1.55)
Edad promedio docentes	-0.118**	-0.0841	-0.0858*	-0.0563	-0.0326	-0.0371
	(-2.75)	(-1.95)	(-1.98)	(-1.26)	(-0.73)	(-0.82)
Años de experiencia	0.339***	0.313***	0.322***	0.283**	0.274**	0.287**
	(3.89)	(3.60)	(3.69)	(3.06)	(2.96)	(3.09)
Años de experiencia^2	-0.00496*	-0.00451*	-0.00501*	-0.00406	-0.00397	-0.00435
	(-2.31)	(-2.11)	(-2.34)	(-1.79)	(-1.75)	(-1.92)
Docente 4° B. hombre	-6.449***	-6.247***	-5.872***	-6.174***	-6.004***	-5.815***
	(-8.85)	(-8.58)	(-8.02)	(-7.51)	(-7.30)	(-7.04)
Alumno hombre	12.66***	12.67***	12.60***	-5.021***	-5.008***	-5.042***
	(29.91)	(29.95)	(29.80)	(-11.29)	(-11.27)	(-11.34)
Escolaridad madre: Básica incompleta	-2.422*	-2.399*	-2.370*	-0.117	-0.112	-0.0782
	(-2.06)	(-2.04)	(-2.01)	(-0.10)	(-0.09)	(-0.06)
Escolaridad madre: Básica completa	-2.583*	-2.535*	-2.517*	-1.309	-1.291	-1.271
	(-2.17)	(-2.13)	(-2.11)	(-1.06)	(-1.04)	(-1.03)
Escolaridad madre: Media incompleta	-0.524	-0.516	-0.523	0.501	0.504	0.522
	(-0.45)	(-0.44)	(-0.45)	(0.41)	(0.42)	(0.43)
Escolaridad madre: Media completa	0.890	0.930	0.905	1.439	1.463	1.477
	(0.79)	(0.82)	(0.80)	(1.22)	(1.24)	(1.25)
Escolaridad madre: CFT o IP incompleta	2.489	2.446	2.417	1.246	1.219	1.228
	(1.55)	(1.52)	(1.51)	(0.74)	(0.72)	(0.73)
Escolaridad madre: CFT o IP completa	1.850	1.779	1.755	3.542*	3.490*	3.511*
	(1.39)	(1.34)	(1.32)	(2.55)	(2.51)	(2.53)
Escolaridad madre: Universitaria incompleta	5.772**	5.679**	5.602**	8.464***	8.406***	8.384***
	(2.98)	(2.94)	(2.90)	(4.19)	(4.16)	(4.15)
Escolaridad madre: Universitaria completa	5.849**	5.789**	5.760**	6.699***	6.670***	6.670***
	(3.16)	(3.13)	(3.12)	(3.42)	(3.41)	(3.41)
Escolaridad madre: Postgrado	16.43*	16.36*	16.72*	-3.596	-3.584	-3.407
	(2.10)	(2.09)	(2.14)	(-0.43)	(-0.43)	(-0.41)
Alumno: Preescolar	0.268	0.285	0.293	1.366*	1.372*	1.383*
	(0.50)	(0.53)	(0.55)	(2.43)	(2.44)	(2.46)
Alumnos: Ranking promedio	4.055***	4.054***	4.045***	3.760***	3.759***	3.755***
	(139.80)	(139.86)	(139.52)	(123.07)	(123.08)	(122.85)
Tamaño cohorte	-0.328***	-0.328***	-0.329***	-0.329***	-0.330***	-0.328***

	(-11.23)	(-11.25)	(-11.29)	(-10.91)	(-10.95)	(-10.88)
Tamaño cohorte^2	0.00109** *	0.00108** *	0.00109** *	0.00103** *	0.00103** *	0.00102** *
	(8.53)	(8.52)	(8.59)	(7.98)	(7.98)	(7.95)
Promedio SIMCE 4°B. 2005-2007. Matemática	0.427***	0.424***	0.419***			
	(25.34)	(25.18)	(24.81)			
Promedio SIMCE 4°B. 2005- 2007. Lectura				0.354***	0.352***	0.348***
				(17.88)	(17.82)	(17.54)
Constante	100.7***	65.32***	72.54***	144.9***	124.4***	129.9***
	(14.56)	(8.79)	(9.37)	(19.08)	(15.38)	(15.45)
N	33576	33576	33576	33490	33490	33490
R² ajustado	0.465	0.466	0.466	0.396	0.396	0.396

Fuente: Elaboración propia, en base a registros MINEDUC

Nota 1: ED: Evaluación Docente. PF: portafolio. IRT: informe de referencia de terceros. AE: autoevaluación. EP: evaluador par.

Nota 2: *** nivel de significancia 0.1%, ** 1% y * 5%.

Tabla D.5: Estimación SIMCE 4° Básico 2008 a nivel de establecimiento. Sin controles comunales.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
% Básico	0.280*			0.303**		
	(2.34)			(2.77)		
% Competente	0.408***			0.375***		
	(3.42)			(3.44)		
% Destacado	0.546***			0.478***		
	(4.43)			(4.24)		
Puntaje promedio final ED		27.97***			19.37***	
		(8.44)			(6.39)	
Puntaje prom. PF			18.61***			12.60***
			(6.08)			(4.52)
Puntaje prom. AE			3.522			0.420
			(1.58)			(0.21)
Puntaje prom. IRT			4.797***			4.523***
			(4.25)			(4.41)
Puntaje prom. EP			3.619**			2.070
			(2.73)			(1.71)
Tamaño cohorte	-0.0908*	-0.0929*	-0.0973*	-0.154***	-0.157***	-0.157***
	(-2.28)	(-2.33)	(-2.43)	(-4.24)	(-4.30)	(-4.28)
Tamaño cohorte^2	0.00074**	0.00073**	0.00076**	0.00096***	0.00096***	0.00097***
	(2.79)	(2.77)	(2.87)	(3.96)	(3.96)	(3.99)
GSE B	7.891***	7.938***	7.887***	7.164***	7.166***	7.088***
	(7.27)	(7.30)	(7.25)	(7.21)	(7.21)	(7.13)
GSE C	22.33***	22.56***	22.34***	21.97***	22.09***	21.79***
	(13.28)	(13.40)	(13.25)	(14.27)	(14.35)	(14.14)
GSE D	45.12***	45.08***	44.69***	41.51***	41.49***	40.94***

	(11.46)	(11.43)	(11.33)	(11.52)	(11.51)	(11.35)
Años de servicio	0.268	0.302	0.282	0.307	0.323	0.306
	(0.86)	(0.97)	(0.90)	(1.07)	(1.13)	(1.07)
Años de servicio^2	-0.00520	-0.00610	-0.00549	-0.00576	-0.00618	-0.00583
	(-0.66)	(-0.78)	(-0.70)	(-0.80)	(-0.86)	(-0.81)
Hrs contrato promedio	0.0776	0.0999	0.101	0.126	0.143	0.141
	(0.76)	(0.98)	(0.99)	(1.35)	(1.54)	(1.51)
% hombre	-0.0843***	0.0852***	0.0840***	-0.0731***	-0.0710***	-0.0694***
	(-3.81)	(-3.86)	(-3.80)	(-3.58)	(-3.49)	(-3.41)
% Docentes Evaluado	0.0480	0.0380	0.0344	0.0290	0.0221	0.0178
	(1.35)	(1.06)	(0.96)	(0.89)	(0.68)	(0.54)
Constante	187.1***	152.1***	145.5***	217.1***	203.1***	205.1***
	(12.46)	(12.34)	(10.26)	(15.79)	(18.04)	(15.81)
N	2761	2761	2760	2763	2763	2762
R² ajustado	0.275	0.272	0.274	0.279	0.279	0.281

Fuente: Elaboración propia, en base a registros MINEDUC

Nota 1: ED: Evaluación Docente. PF: portafolio. IRT: informe de referencia de terceros. AE: autoevaluación. EP: evaluador par.

Nota 2: *** nivel de significancia 0.1%, ** 1% y * 5%.

Tabla D.6: Estimación SIMCE 4° Básico 2008 a nivel de alumno con errores por cluster por establecimiento.

	Matemática			Lectura		
	A1	A2	A3	A1	A2	A3
P. Básico	-1.248			-0.221		
	(-0.51)			(-0.11)		
P. Competente	3.802			2.947		
	(1.58)			(1.46)		
P. Destacado	9.326**			6.601*		
	(2.86)			(2.40)		
Promedio Puntajes ED		15.83***			9.531***	
		(8.00)			(5.57)	
Puntaje promedio PF			8.090***			6.193***
			(4.37)			(3.76)
Puntaje promedio IRT			4.169***			2.210**
			(5.55)			(3.21)
Puntaje promedio AE			-0.794			-0.459
			(-0.68)			(-0.45)
Puntaje promedio EP			2.675***			1.021
			(3.41)			(1.57)
Edad promedio docentes	-0.122	-0.0902	-0.0961	-0.0500	-0.0306	-0.0339
	(-1.95)	(-1.46)	(-1.56)	(-0.90)	(-0.55)	(-0.60)
Años de experiencia	0.301*	0.260	0.270	0.216	0.200	0.209
	(2.04)	(1.76)	(1.84)	(1.71)	(1.59)	(1.65)

Años de experiencia^2	-0.00479	-0.00382	-0.00422	-0.00366	-0.00336	-0.00364
	(-1.33)	(-1.07)	(-1.18)	(-1.19)	(-1.09)	(-1.18)
Docente 4° B. hombre	-3.475**	-3.257**	-3.075**	-4.013***	-3.850***	-3.734***
	(-3.11)	(-2.93)	(-2.78)	(-3.65)	(-3.50)	(-3.39)
Alumno hombre	11.37***	11.37***	11.34***	-6.093***	-6.080***	-6.099***
	(26.87)	(26.88)	(26.91)	(-14.81)	(-14.79)	(-14.86)
Escolaridad madre: Básica incompleta	-1.202	-1.182	-1.207	-0.176	-0.174	-0.176
	(-1.17)	(-1.15)	(-1.17)	(-0.16)	(-0.16)	(-0.16)
Escolaridad madre: Básica completa	-1.161	-1.124	-1.129	-0.990	-0.989	-0.986
	(-1.12)	(-1.09)	(-1.09)	(-0.89)	(-0.88)	(-0.88)
Escolaridad madre: Media incompleta	0.599	0.611	0.587	0.501	0.513	0.513
	(0.59)	(0.60)	(0.58)	(0.45)	(0.46)	(0.46)
Escolaridad madre: Media completa	2.429*	2.458*	2.402*	1.697	1.723	1.708
	(2.47)	(2.49)	(2.44)	(1.57)	(1.59)	(1.57)
Escolaridad madre: CFT o IP incompleta	4.459**	4.419**	4.348**	2.659	2.640	2.634
	(3.09)	(3.06)	(3.01)	(1.69)	(1.68)	(1.68)
Escolaridad madre: CFT o IP completa	4.043***	3.936**	3.908**	3.895**	3.824**	3.823**
	(3.37)	(3.27)	(3.25)	(2.98)	(2.92)	(2.92)
Escolaridad madre: Universitaria incompleta	8.142***	8.082***	7.968***	9.161***	9.128***	9.081***
	(4.26)	(4.23)	(4.18)	(4.88)	(4.86)	(4.84)
Escolaridad madre: Universitaria completa	7.795***	7.753***	7.723***	6.297***	6.277***	6.247***
	(4.59)	(4.55)	(4.55)	(3.49)	(3.48)	(3.47)
Escolaridad madre: Postgrado	8.569	8.470	8.600	-6.473	-6.449	-6.334
	(1.09)	(1.07)	(1.08)	(-0.72)	(-0.72)	(-0.71)
Alumno: Preescolar	0.218	0.221	0.204	0.642	0.641	0.641
	(0.48)	(0.48)	(0.45)	(1.36)	(1.36)	(1.36)
Alumnos: Ranking promedio	4.135***	4.131***	4.125***	3.815***	3.813***	3.809***
	(93.83)	(93.60)	(93.36)	(90.16)	(90.04)	(89.71)
Tamaño cohorte	-0.572***	-0.574***	-0.573***	-0.519***	-0.520***	-0.520***
	(-12.99)	(-12.95)	(-13.13)	(-16.18)	(-16.18)	(-16.20)
Tamaño cohorte^2	0.00220** *	0.00221** *	0.00221** *	0.00188** *	0.00188** *	0.00189** *
	(7.55)	(7.52)	(7.67)	(9.45)	(9.40)	(9.47)
Promedio SIMCE 4°B. 2005-2007. Matemática	0.424***	0.418***	0.414***			
	(16.20)	(15.99)	(15.82)			
Promedio SIMCE 4°B. 2005- 2007. Lectura				0.372***	0.369***	0.366***
				(14.08)	(13.85)	(13.74)
Constante	106.4***	67.56***	75.15***	152.5***	129.8***	134.0***

	(7.98)	(4.93)	(5.25)	(14.26)	(11.46)	(11.33)
N	51809	51809	51809	51757	51757	51757
R² ajustado	0.422	0.424	0.424	0.362	0.362	0.362

Fuente: Elaboración propia, en base a registros del MINEDUC

Nota 1: ED: Evaluación Docente. PF: portafolio. IRT: informe de referencia de terceros. AE: autoevaluación. EP: evaluador par.

Nota 2: *** nivel de significancia 0.1%, ** 1% y * 5%.