

---

## Valor agregado en la educación primaria y secundaria: siguiendo cohortes en el tiempo<sup>1</sup>

Value-added in primary and secondary education: Following cohorts over time

Jorge Duarte<sup>a</sup>  
jduarte@contratista.icfes.gov.co

Silvana Godoy Mateus<sup>b</sup>  
sgodoy@icfes.gov.co

Ximena Dueñas Herrera<sup>c</sup>  
xduenas@icfes.gov.co

---

### Resumen

Usando comparaciones de cohortes en las pruebas Saber 3°, 5° y 9° que realiza el Instituto Colombiano para la Evaluación de la Educación (Icfes) se estimaron dos modelos de valor agregado. El primero, que solo se basa en los resultados del año inicial para estimar el crecimiento en el año final, brinda una medida en términos absolutos de la ganancia educativa de los estudiantes dentro de un colegio. El segundo, que controla por varias características demográficas, sociales y económicas, limpia el resultado de condiciones que el colegio no puede controlar, para dar una estimación del efecto exclusivo del colegio sobre el aprendizaje de sus estudiantes. Los resultados del análisis muestran que las condiciones económicas de los estudiantes se relacionan fuertemente con el progreso educativo, pero que hay colegios que a pesar de tener bajo rendimiento tienen estudiantes en niveles de desempeño alto. Estos colegios logran mantener a sus estudiantes de desempeño alto en estos niveles, pero fallan en nivelar a sus estudiantes de bajo desempeño con el resto de sus estudiantes.

**Palabras clave:** valor agregado, educación, pruebas estandarizadas.

### Abstract

Two Value-Added models were fitted using Cohort comparisons on the Saber 3°, 5° y 9° test that are applied by the Icfes. The first one, which is only based on the

---

<sup>1</sup>Duarte, J., Godoy, S., Dueñas, X. (2016) Valor agregado en la educación primaria y secundaria: siguiendo cohortes en el tiempo. *Comunicaciones en Estadística*, 9(1), 11-41.

<sup>a</sup>Economista, Subdirección de Análisis y Divulgación, Icfes, Colombia.

<sup>b</sup>Subdirectora, Subdirección de análisis y divulgación, Icfes, Colombia.

<sup>c</sup>Directora General, Dirección General, Icfes, Colombia.

results of the initial year to estimate growth in the final year, provides a measure in absolute terms of educational gains of students within a school. The second, controlling for several demographic, social and economic characteristics, cleans the result of conditions that the school can not control, to give an estimate of the exclusive effect of the school on student learning. The analysis results show that the economic conditions of students are strongly related to educational progress, but there are schools that despite having underperforming students have high performance levels. These schools are able to maintain their high-achieving students at these levels, but fail to leverage its low-performing students with other students.

**Keywords:** value-added, education , standardized tests.

## 1. Introducción

El concepto de valor agregado (VA) en un sistema educativo se refiere al logro de los estudiantes, expresado como el crecimiento en conocimiento, habilidades, capacidades y otros atributos que estos han ganado como resultado de sus experiencias dentro de un sistema educativo en el tiempo (OECD 2008). También puede ser definido como la magnitud del efecto de un determinado colegio en los resultados educativos de sus estudiantes, que van más allá de sus calificaciones y afectan sus oportunidades futuras (Saunders 1998).

Del concepto de VA se desprende el concepto de modelos de VA. Estos pueden ser definidos en una categoría de modelos estadísticos que usan datos del logro educativo de los estudiantes en el tiempo, con el fin de medir ganancias en el aprendizaje (Doran & Lockwood 2006). Análogamente, las medidas de VA pueden ser definidas como el valor que agrega un colegio a sus estudiantes por encima de lo que se esperaría de acuerdo con sus antecedentes y logros previos (Hill 1995).

Hay diferencias importantes entre el objetivo de los modelos de VA y los de análisis transversales con datos de pruebas estandarizadas. Varios estudios usan datos en un periodo para comparar estudiantes, colegios, regiones y otros agregados. Estos análisis brindan información valiosa para conocer en términos absolutos el nivel de los estudiantes. Otro tipo de análisis usa datos de un grado para hacer comparaciones en el tiempo de ese mismo grado. Siempre y cuando se mantengan los estándares de aprendizaje y el diseño de las pruebas, estos análisis temporales permiten hacer comparaciones entre diferentes cohortes. Por el contrario, los modelos de VA siguen una misma cohorte en dos momentos del tiempo con el objetivo de conocer el cambio desde el punto de partida hasta el punto de llegada.

La OECD identifica dos beneficios principales de los modelos de VA. En primer lugar, las medidas de VA pueden ser “más justas” que otros análisis debido a que el resultado en VA tiene en cuenta el nivel inicial de los estudiantes (OECD 2008, Doran & Izumi 2004). Puede ser injusto evaluar la contribución de un colegio al logro de sus estudiantes centrándose únicamente en las condiciones de salida de estos o en el porcentaje de estudiantes que cumplen ciertos estándares, debido a que las

condiciones de entrada a los colegios no son iguales (Reardon & Raudenbush 2009). En Colombia, el desempeño de los estudiantes varía fuertemente según distintas condiciones socioeconómicas. Las medidas de VA, más que en el desempeño de salida del estudiante, pueden mostrar la magnitud de los cambios en el tiempo, dadas las condiciones de entrada.

El segundo beneficio de las mediciones de VA, que menciona la OECD, es que pueden ser “más precisas” debido a que incorporan características contextuales de los colegios y sus estudiantes. Las calificaciones transversales de las pruebas, aunque brindan información muy importante para un sinnúmero de objetivos de interés general, no permiten diferenciar la contribución puntual de los colegios de otras características que también afectan el desempeño escolar. Los modelos de VA pueden tener en cuenta condiciones sociales, económicas, demográficas y otros factores de contexto de los colegios y estudiantes, con el objetivo de generar una medida de los colegios que reconozca estas características.

Las pruebas Saber 3°, 5° y 9° del Icfes brindan la posibilidad de hacer estimaciones de VA. Estas pruebas evalúan a todos los estudiantes de los colegios de Colombia, con una periodicidad anual desde 2012. Entre otros resultados, cada colegio recibe una medida de tendencia central de su desempeño y cuatro medidas de la distribución de sus estudiantes en términos de su rendimiento. Esta distribución se basa en cuatro niveles de desempeño, definidos de acuerdo con los estándares educativos del Ministerio de Educación Nacional (MEN) para cada área y grado.

A la fecha es posible seguir a dos cohortes:

1. Estudiantes que en 2012 cursaban el grado tercero de primaria y que en 2014 estaban en el grado quinto del mismo ciclo educativo.
2. Estudiantes que en 2009 cursaban quinto de primaria y en 2013 estaban en noveno de bachillerato. Para cada una de las dos cohortes es posible obtener resultados para las áreas de matemáticas y lenguaje.

El ejercicio empírico de este documento estima dos modelos de VA. El primero, que solo se basa en los resultados del año inicial para estimar el crecimiento en el año final, brinda una medida en términos absolutos de la ganancia educativa de los estudiantes dentro de un colegio. El segundo, que controla por varias características demográficas, sociales y económicas, busca dar una estimación del efecto exclusivo del colegio sobre el aprendizaje de sus estudiantes. Consideramos que ambos resultados son de interés para las políticas educativas. El primer modelo puede ser usado para identificar los colegios con bajos desempeños, con el objetivo de focalizar ayudas o tomar planes de acción. Es importante saber cuáles son estos colegios aunque ellos no sean los únicos responsables de los resultados. El segundo modelo puede ayudar a recompensar o reconocer los colegios que, a pesar de las características de sus estudiantes, logran obtener buen desempeño.

Tras revisar la literatura, podemos afirmar que podría ser la primera aproximación a los modelos de VA para las pruebas Saber 3°, 5° y 9°. Los resultados deben ser interpretados con precaución debido a que no contamos con resultados a nivel de

estudiante; no podemos garantizar que los estudiantes al comienzo de un periodo sean exactamente los mismos que al final (problemas como deserción, repitencia, entre otros). Sin embargo, consideramos que la información de VA, acompañada de otra información de los colegios, puede ser un buen instrumento de política pública.

El documento se divide en cinco secciones incluyendo esta introducción. En la siguiente sección describimos brevemente las implicaciones de política pública de las estimaciones de VA. En las siguientes dos secciones exponemos la metodología y mostramos los resultados. Y, por último, concluimos con una discusión y con aspectos para mejorar en futuras investigaciones.

## 2. Implicaciones de política pública:

La escuela de la efectividad escolar nace por la necesidad de identificar el aporte directo de los colegios al desempeño de sus estudiantes (OECD 2011, Downes & Vindurampulle 2007, Kim & Lalancette 2013). Los primeros análisis estadísticos que buscaban este objetivo usaban el promedio de los estudiantes de un colegio para hacer comparaciones directas. Investigaciones posteriores, sobre los determinantes de los logros educativos, mostraron que el desempeño de los colegios estaba relacionado con las condiciones socioeconómicas de los estudiantes y, por tanto, que la comparación de estas medidas brutas no reconocía de forma precisa la contribución de los colegios (Hill 1995, Haveman & Wolfe 1995, Kim & Lalancette 2013). Raudenbush & Bryk (1986) hacen un estudio pionero de VA con modelos lineales jerárquicos para analizar la efectividad escolar, teniendo en cuenta no sólo las características socioeconómicas de los estudiantes, sino también los logros previos obtenidos por Haveman & Wolfe (1995) y Kim & Lalancette (2013).

Actualmente, los modelos de VA pueden ser definidos como herramientas que describen el proceso de aprendizaje. Estos modelos intentan separar las contribuciones del colegio en el desarrollo de los estudiantes, diferenciándolas de los aportes del resto de los factores asociados al desempeño (Willms 2009). A pesar de las múltiples definiciones teóricas y prácticas, todos los modelos de VA se aseguran de que la evaluación de la efectividad de los colegios tenga en cuenta todas las influencias externas a este, además de las condiciones socioeconómicas de los estudiantes, antecedentes familiares y logros previos (Wyatt 1996).

Los modelos de VA han sido usados para diferentes objetivos. En general, han sido usados para evaluar, monitorear o mejorar diferentes aspectos de un sistema educativo (OECD 2008). Estos modelos permiten identificar a los colegios que más aportan a la experiencia educativa de sus estudiantes, con la precaución de no penalizarlos o premiarlos por las condiciones y antecedentes de sus estudiantes (Dury & Doran 2003). Los modelos de VA pueden ser usados como una herramienta para la mejora escolar debido a que indican cómo está progresando el desempeño académico de los estudiantes respecto del desempeño de estudiantes comparables (OECD 2011, Downes & Vindurampulle 2007). A través de los modelos de VA

es posible identificar los colegios con desempeño por encima y por debajo de lo proyectado, con el fin de que aquellos colegios con dificultades sigan las prácticas de aquellos que tienen desempeños relativos altos (OECD 2011, Downes & Vindurampulle 2007).

En E.E.U.U. los sistemas de información, cada vez más, permiten identificar y relacionar a los estudiantes con sus profesores. Esta posibilidad ha hecho que los modelos de VA midan, o intenten medir, cuál es la contribución específica de los profesores al progreso de sus estudiantes y, en consecuencia, que los resultados puedan tener implicaciones importantes para ellos. Esta es la práctica más controversial del uso de los modelos de VA (Koedel et al. n.d., Aaronson et al. 2007, Rivkin 2005). En Luisiana, Nueva York y en algunos condados de Tennessee, los modelos se usan para evaluar la efectividad de los programas de preparación docente, para ayudar al mejoramiento de estos y para estudiar la distribución de profesores en colegios pobres (Armour-Garb 2009). Sin embargo, las mayores críticas provienen de Florida, Denver y otros condados de Tennessee, donde VA se usa para definir la permanencia y el pago de los profesores. Sindicatos de profesores en Nueva York y California han prohibido cualquier clase de decisión de alto impacto para profesores que se base en VA y, particularmente, en pruebas estandarizadas a los estudiantes.

Los grupos que están a favor de las decisiones de alto impacto basadas en VA afirman que estas medidas son menos costosas que otros mecanismos (como evaluaciones observacionales del desempeño de los profesores) y que, si los modelos están especificados correctamente, pueden generar mejores resultados. Los críticos de estos métodos piensan que las estimaciones no son lo suficientemente confiables como para categorizar a los profesores. Nosotros pensamos que ejercicios empíricos, como el de este documento, permiten dar a conocer las ventajas y desventajas de estos procedimientos y que abren la discusión de para qué y cómo pueden ser usados los resultados en el contexto colombiano.

A pesar de las múltiples aplicaciones de los modelos de VA, el uso de modelos estadísticos para medir las ganancias marginales en el aprendizaje generan varias dificultades prácticas que, aunque entendidas teóricamente, son difíciles de responder empíricamente. Para obtener estimaciones precisas deben ser usados los modelos especificaciones apropiadas dadas las propiedades de los datos disponibles y los objetivos de política pública (OECD 2008). En la siguiente sección exponemos nuestra metodología.

### 3. Metodología:

Para el ejercicio empírico usamos las pruebas Saber del Icfes de los grados 3°, 5° y 9°, de los años 2009, 2012, 2013 y 2014, para las áreas de matemáticas y lenguaje. Estas pruebas evalúan a todos los colegios del país. Para cada una de las pruebas en estos años, grados y áreas, usamos el puntaje promedio del colegio y el porcentaje de estudiantes en los cuatro niveles de desempeño (insuficiente,

mínimo, satisfactorio y avanzado).

Los niveles de desempeño son definidos de acuerdo con los estándares educativos del MEN y, por tanto, brindan una medida de la distribución en términos de aprendizaje de los estudiantes dentro del colegio. Con el fin de reducir el número de variables de respuesta, sin perder información de la distribución, unimos el porcentaje de estudiantes en los niveles de desempeño insuficiente y mínimo para tener una medida de la cola de la distribución (es decir, de los estudiantes con las calificaciones menos deseables del colegio), excluimos el nivel satisfactorio y usamos el nivel avanzado como medida de la punta de la distribución. Esta decisión la tomamos porque, en promedio, la distribución es asimétrica hacia la derecha. De esta forma contamos con tres medidas de interés: el puntaje promedio, que brinda una medida del desempeño promedio de los estudiantes del colegio; la suma de los porcentajes de los estudiantes en los niveles insuficiente y mínimo, que muestra el comportamiento de los estudiantes en la cola de la distribución del colegio; y, por último, el porcentaje en avanzado, que muestra el comportamiento en la punta de la distribución.

Las pruebas Saber 3°, 5° y 9° no brindan una calificación específica para los estudiantes. La razón es que cada estudiante presenta solo una parte de la prueba para alguna de las áreas evaluadas con el fin de estimar una calificación para el colegio. Esto implica que no hay un registro que identifique a los estudiantes en el tiempo. Por lo tanto, la unidad de observación de nuestro análisis son los colegios. Este hecho hace que, a pesar de que podemos seguir a una cohorte específica en el tiempo, no podemos garantizar que los estudiantes evaluados en dos periodos sean exactamente los mismos. Esta restricción puede sesgar las estimaciones de los cálculos en ambas direcciones. Por ejemplo, si un colegio en el grado quinto de primaria tenía estudiantes muy buenos y cuatro años después en el grado noveno de bachillerato tenía a otros estudiantes no tan buenos, puede parecer que no hubo aumentos en el rendimiento escolar. Si lo contrario le pasara a un colegio, el sesgo sería positivo. Por lo tanto, la validez de nuestros modelos de VA con observaciones a nivel de colegio depende del supuesto fuerte de que la gran mayoría de estudiantes no cambian de colegio. Este supuesto no se puede comprobar debido a que aun sabiendo la proporción de estudiantes en el año final respecto del inicial, no sabemos si son los mismos estudiantes. El supuesto se podría relajar si se supone que los estudiantes que un colegio recibe tienen condiciones similares a los estudiantes que pierde.

Con los datos que contamos es posible seguir a dos cohortes en el tiempo (ver tabla 1) La primera cohorte que estudiamos es la de los estudiantes que en 2012 estaban en el grado tercero de primaria y que dos años después, en 2014, estaban en quinto de primaria. A esta cohorte la llamamos Tercero 2012 a Quinto 2014. La segunda cohorte es la de los estudiantes que en 2009 estaban en el grado quinto de primaria y cuatro años después, en 2013, estaban en noveno de bachillerato. A esta cohorte la llamamos Quinto 2009 a Noveno 2013. La cohorte Tercero 2012 a Quinto 2014 es de gran interés porque es la que más garantiza el supuesto de estabilidad estudiantil en los colegios, debido a que ambas calificaciones pertenecen al mismo

ciclo escolar. La cohorte Quinto 2009 a Noveno 2013 es la que menos garantiza este supuesto, pero es de gran interés debido a que hay cambio de ciclo escolar. Desde el año 2012 las pruebas Saber 3° , 5° y 9° se vienen aplicando con una periodicidad anual y desde el próximo año se aplicará la prueba al grado séptimo, así que se ampliará considerablemente la posibilidad de seguir cohortes en el tiempo.

Tabla 1: Cohortes. Fuente: elaboración propia.

2009	2010	2011	2012	2013	2014
					Tercero
				Tercero	Cuarto
			Tercero	Cuarto	Quinto
		Tercero	Cuarto	Quinto	Sexto
	Tercero	Cuarto	Quinto	Sexto	Séptimo
Tercero	Cuarto	Quinto	Sexto	Séptimo	Octavo
Cuarto	Quinto	Sexto	Séptimo	Octavo	Noveno
Quinto	Sexto	Séptimo	Octavo	Noveno	Décimo
Sexto	Séptimo	Octavo	Noveno	Décimo	Once
Séptimo	Octavo	Noveno	Décimo	Once	
Octavo	Noveno	Décimo	Once		
Noveno	Décimo	Once			
Décimo	Once				
Once					

Usamos dos modelos diferentes, para cada una de las dos cohortes, cada una de las dos áreas y cada una de las tres calificaciones. El primero, que llamamos *modelo de crecimiento*, es una regresión lineal simple donde la variable dependiente es la calificación del año final y la independiente es la calificación del año inicial. El segundo modelo, que llamamos modelo de VA, es una regresión lineal de efectos fijos e interacciones en dos niveles que controla por varias características contextuales.

### 3.1. Modelo de crecimiento

Si los exámenes que se aplican en el año inicial y en el año final fueran exactamente los mismos, la diferencia entre ambas calificaciones sería una medida del cambio en el logro educativo, o crecimiento, en el periodo de tiempo comprendido entre los dos exámenes. Como las pruebas Saber 3° , 5° y 9° evalúan competencias específicas en cada grado escolar, las pruebas entre grados no son comparables. Por ejemplo, si un colegio en el año inicial tiene muy buenos resultados y en el año final tiene unos no tan buenos, no podemos afirmar que desmejoró debido a que posiblemente mejoró respecto del primer examen, pero no tanto como para tener un buen desempeño en el segundo. Para poder saber si en realidad empeoró, habría que aplicarles a los estudiantes del año final el examen del año inicial por segunda vez.

Por esta razón, la diferencia no puede hacerse entre el resultado del año final y

el del año inicial, sino entre el del año final y una estimación del puntaje del año final basada en la calificación del año inicial. En otras palabras, el crecimiento es la diferencia entre el resultado alcanzado por el colegio y el resultado esperado. Estimamos por mínimos cuadrados ordinarios (MCO) el siguiente modelo

$$F_i = \alpha_0 + \alpha_1 I_i + \epsilon_i \quad (1)$$

Donde  $i = 1, 2, \dots, N_i$  denota el número de colegios.  $F_i$  denota la calificación de  $i$  en el año final, donde la calificación es el puntaje promedio, la suma del porcentaje de estudiantes en los niveles de desempeño insuficiente y mínimo o el porcentaje de estudiantes en el nivel de desempeño avanzado. Es decir, se estiman tres modelos diferentes, uno para cada variable dependiente.  $I_i$  denota la calificación inicial de  $i$ .  $\alpha_0$  y  $\alpha_1$  son los coeficientes de regresión y  $\epsilon_i$  es el término de error.

La predicción de  $F_i$ ,  $\hat{F}_i$ , es el resultado esperado del colegio para el año final dado el año inicial y dado el comportamiento de los colegios de todo el país. La diferencia  $F_i - \hat{F}_i$  es el crecimiento del colegio  $i$ .

El modelo de crecimiento muestra el aporte marginal de un colegio en el tiempo sin ajustar por otras variables contextuales, es decir que muestra una medida absoluta del cambio en el logro educativo de los estudiantes de un colegio. Consideramos que este modelo es pertinente por dos razones principales. La primera es que su sencillez es una ventaja a la luz de la opinión pública, la segunda es que debido a que el resultado no se afecta por otras variables diferentes a las de las calificaciones, se pueden hacer comparaciones entre colegios sin otras consideraciones y basadas únicamente en el cambio real.

### 3.2. Modelo de valor agregado

Una de las ventajas de VA es que es posible incluir información contextual de los estudiantes y el colegio, para hacer medidas que tengan en cuenta inequidades o que logren separar el efecto de los colegios de otras características de los estudiantes. Aun controlando por el nivel inicial de los estudiantes, es razonable pensar que los cambios en el logro educativo se afecten por otras variables exógenas al colegio. Por ejemplo, padres de los estudiantes más comprometidos con la educación de sus hijos, estudiantes con más recursos alimenticios o tecnológicos, pertinencia de los estándares educativos en ciertas regiones del país, actividades extracurriculares de los estudiantes, grupos minoritarios, creencias religiosas, discapacidades y motivaciones de los estudiantes, entre otras. Con este objetivo, y con la restricción de los datos disponibles, el modelo de VA incluye otros regresores con características contextuales del colegio y sus estudiantes.

Usamos la identificación de las Entidades Territoriales Certificadas (ETC) como uno de los efectos fijos del modelo. La Ley 715 de 2001 da la categoría de ETC a todos los departamentos y municipios que en 2002 contaran con más de cien mil habitantes. Los municipios con menor cantidad de habitantes que tengan la capacidad de administrar el servicio público de educación, pueden solicitar la cer-



tificación si cumplen con ciertos requisitos técnicos, administrativos y financieros. Estos requisitos son: que el plan de desarrollo municipal sea consistente con las políticas educativas nacionales, que cuenten con establecimientos educativos estatales organizados para ofrecer al menos educación básica completa, que tengan una planta docente y directiva definida de acuerdo con los parámetros nacionales y que tengan la capacidad institucional para operar el sistema de información del sector educativo. Los municipios con más de cien mil habitantes deben cumplir estos requisitos para seguir certificados (decretos 2700 de 2004 y 3940 de 2007). Consideramos que la ETC ofrece un control regional mejor que el municipio o el departamento porque mantiene la identificación departamental, que es una característica regional importante, pero diferencia dentro del departamento a los municipios que cuentan con las anteriores características. Para este ejercicio usamos las 94 ETC existentes hasta el 2014. Desde el 2015 el municipio de Yumbo, Valle del Cauca, está certificado.

Como segundo nivel de efectos fijos usamos la zona (rural o urbana) y el sector (oficial o no oficial). Además de las diferencias regionales, la zona y el sector capturan características determinantes para la educación. En Colombia, y en general en los países menos desarrollados, se mantienen grandes diferencias entre las zonas rurales y urbanas, en términos del nivel socioeconómico de las familias, la calidad de la educación y la deserción escolar (Colbert 1999, Kazeem & Stokes 1995, Zhang 2006). Respecto del sector, los colegios no oficiales suplen la demanda que los oficiales no alcanzan a cubrir. Esta demanda proviene de padres que buscan mejor educación para sus hijos o de padres que no encontraron cupo en los colegios oficiales. Por tal motivo, aunque en promedio los colegios no oficiales tienen mejor desempeño que los oficiales, esta característica no es un control por el desempeño escolar (hay colegios oficiales con muy buen desempeño y no oficiales con muy mal desempeño). El sector es relevante porque hay marcadas diferencias en incentivos docentes y planes curriculares (Nunez et al. 2002, Tobón et al. 2008, Gaviria & Barrientos 2001).

Como aproximación al nivel socioeconómico del colegio usamos el índice de nivel socioeconómico (INSE) del Icfes. El INSE se construye con información de la encuesta socio-demográfica que responden los estudiantes en el momento de la aplicación de las pruebas Saber. El cálculo tiene en cuenta la composición del hogar, la infraestructura y las posesiones o enseres de la vivienda, la interrelación con los padres (hablar sobre noticias, libros o programas de televisión) y con un componente cultural (en términos de su asistencia a actividades de teatro, exposiciones, ferias, carnavales, parques, circos, cuenteros o títeres, bibliotecas y cine). Esta clasificación es particular para el campo educativo y diferente de otras medidas socioeconómicas con otros objetivos, por lo tanto el INSE no es comparable con cálculos de pobreza del Departamento Administrativo Nacional de Estadística (DANE).

El INSE se calcula para cada estudiante, mientras que los colegios se clasifican en cuatro niveles socioeconómicos (NSE) basados en el INSE. Para nuestro ejercicio, con el fin de tener una variable continua, tomamos el promedio por colegio del INSE como variable explicativa. Las cuatro categorías del NSE las usamos para

presentar estadísticas descriptivas en la siguiente sección. Incluimos una medida de la estabilidad de la cohorte, definida como la diferencia entre el número de estudiantes que presentaron la prueba en el año final y el número de estudiantes que presentaron la prueba en el año inicial. Esta medida no garantiza que los estudiantes sean los mismos. Por último, usamos la desviación estándar del puntaje promedio de los estudiantes de un colegio como medida de la dispersión del aprendizaje dentro de este. Con estas variables estimamos el siguiente modelo lineal de efectos fijos e interacciones en dos niveles:

$$F_{ijk} = \beta_0 + \beta_1 I_{ijk} + \beta_2 INSE_{ijk} + \beta_3 EST_{ijk} + \beta_4 DE_{ijk} + v_{1j} I_{ijk} ETC_j + \pi_{1k} I_{ijk} ZS_k + v_{2j} + \pi_{2k} + \mu_{ijk}$$

Donde, además de lo descrito en el modelo de crecimiento,  $j = 1, 2, 3$ ,  $N_j$  denota el número de ETC y  $k =$  Oficial rural, oficial urbano, no oficial rural y no oficial urbano denota la zona y sector de  $i$ .  $INSE_{ijk}$  es el  $INSE$  del colegio  $i$  que pertenece a  $j$  y a  $k$ ,  $Est_{ijk}$  es la medida de estabilidad de la cohorte de  $i$  y  $DE_{ijk}$  es la desviación estándar del puntaje promedio de  $i$ .  $ETC_j$  identifica a cada  $ETC$  y  $ZS_k$  a la zona y el sector.  $\beta_0, \beta_1, \beta_2, \beta_3$  y  $\beta_4$  son los coeficientes de regresión comunes a todos los colegios.  $v_{1j}$  y  $v_{2j}$  son los coeficientes de pendiente y constante, respectivamente, comunes a las  $ETC$  y  $\pi_{1k}$  y  $\pi_{2k}$  son los coeficientes comunes a cada zona y sector. Empleamos interacciones debido a que las pendientes de regresión deben ser también específicas para los niveles de  $ETC$  y de zona y sector. En el anexo 1 se muestran todas las regresiones del modelo de crecimiento y del modelo de VA.

Se estimó el modelo por medio de efectos fijos con el objetivo de solucionar el problema de endogeneidad de  $I_{ijk}$ . El modelo no incluye muchas variables contextuales de los estudiantes del colegio no correlacionadas con  $I_{ijk}$  que quedan recogidas en  $\mu_{ijk}$ . Esto hace que  $\beta_1$  esté sesgado y que la medida de VA incluya efectos no atribuibles al colegio. Al usar efectos fijos incluimos en el modelo todas las características de la ETC y de la zona y el sector invariantes en el tiempo correlacionadas con  $I_{ijk}$ , corrigiendo así, al menos parcialmente, el problema de endogeneidad. El modelo no usa efectos aleatorios porque estos modelos se basan en el supuesto de que  $E[v_{2j} + \pi_{2k} + \mu_{ijk} \mid I_{ijk}, INSE_{ijk}, Est_{ijk}, DE_{ijk}] = 0$ , lo cual es aún más difícil de garantizar.

El supuesto que queda por garantizar para que las estimaciones sean insesgadas es que  $E[\mu_{ijk} \mid I_{ijk}, INSE_{ijk}, Est_{ijk}, DE_{ijk}] = 0$ . Este supuesto es difícil de cumplir debido a que casi que la única forma de garantizarlo es que los datos provengan de un experimento controlado. Es decir, para conocer el efecto de un colegio en el desempeño de sus estudiantes, habría que repartir aleatoriamente en diferentes colegios a estudiantes con características similares y evaluarlos al comienzo y al final de un periodo. A pesar de que no podemos garantizar que nuestro modelo cumpla este supuesto, consideramos que tenemos una buena aproximación.

## 4. Resultados

Las estimaciones de crecimiento y VA deben ser comparadas entre colegios o agregados (municipios, ETC y departamentos) similares en el desempeño del examen del año inicial, debido a dos razones. La primera es una razón matemática. Al menos teóricamente, los colegios con peor desempeño en el año inicial son los que tienen mayor posibilidad matemática de crecer en el periodo. Este hecho puede hacer que parezca, si no se separan los colegios por el resultado del año inicial, que un colegio con buen desempeño en ambos periodos agrega menos valor a sus estudiantes que un colegio que mejoró, pero tenía bajo desempeño en el año inicial. Como la posibilidad matemática de mejorar es diferente a la probabilidad real de hacerlo, si un colegio de bajo desempeño mejora considerablemente en el periodo, es apropiado resaltar este crecimiento entre los colegios con un rendimiento similar.

La segunda razón es que la dificultad del aprendizaje puede aumentar a medida que más se aprende. Por tal motivo no deberían valorarse de igual forma crecimientos en el aprendizaje en diferentes niveles de este. Por ejemplo, no genera el mismo valor un colegio de desempeño bajo que aumentó cien puntos del examen, que otro colegio que aumentó los mismos cien puntos, pero que está en un desempeño alto. Posiblemente fue más difícil para el colegio de desempeño alto conseguir este crecimiento. Nuestros modelos, en especial el de crecimiento, no tienen en cuenta estos cambios marginales. Aunque el modelo de VA basa sus comparaciones entre colegios similares, no deja de ser un modelo lineal. Por ambas razones dividimos el puntaje promedio del año inicial de los colegios en tres categorías de desempeño: bajo, medio y alto. Las categorías las escogimos de tal forma que en los años iniciales de ambas cohortes y ambas materias, en promedio, 40 % de los colegios quedarán en desempeño bajo, 40 % en medio y 20 % en alto. Esto equivale a que los colegios de desempeño bajo tienen puntajes entre 100 y 290, los de desempeño medio entre 291 y 350, y los de desempeño alto entre 351 y 500. Las estimaciones son comparables dentro de estas categorías y no son comparables entre ellas.

La figura 1 muestra un ejemplo de los seis cuadrantes resultantes del modelo de VA, para el puntaje promedio de los colegios de la cohorte Tercero 2012 a Quinto 2014 en el área de matemáticas. Pueden hacerse comparaciones verticales entre los colegios pertenecientes a una misma categoría de desempeño. Además, podría dársele nombre a los cuadrantes para dar una medida cualitativa de ambas dimensiones. Resaltan las categorías de bajo desempeño y VA decreciente, debido a que contiene a los colegios con mayores falencias para el aprendizaje, y la categoría de colegios con desempeño alto y VA creciente, debido a que son los colegios que tienen los mejores resultados.

La Figura 2 muestra el ejemplo del modelo de VA cuando la variable de interés es la suma del porcentaje de estudiantes en los niveles de desempeño insuficiente y mínimo. Como el corte para las tres categorías de desempeño se basa en el puntaje promedio, no hay una correspondencia perfecta entre el corte y el porcentaje de estudiantes en los niveles de desempeño. Por ejemplo, la figura 2 muestra que hay un colegio clasificado en desempeño alto a pesar de que tiene a más del 40 % de

sus estudiantes en mínimo e insuficiente. Este colegio debe tener el resto de sus estudiantes en un buen nivel de desempeño para que tenga de promedio un puntaje mayor que 350 y, entonces, quede clasificado en desempeño alto. La pertinencia de los niveles de desempeño radica en que si dos colegios con el mismo puntaje promedio tienen clasificados a sus estudiantes en diferentes niveles de desempeño, la dispersión del aprendizaje de sus estudiantes es diferente. Además de analizar el comportamiento promedio del colegio, es importante conocer la dispersión de los niveles del aprendizaje dentro de éste.

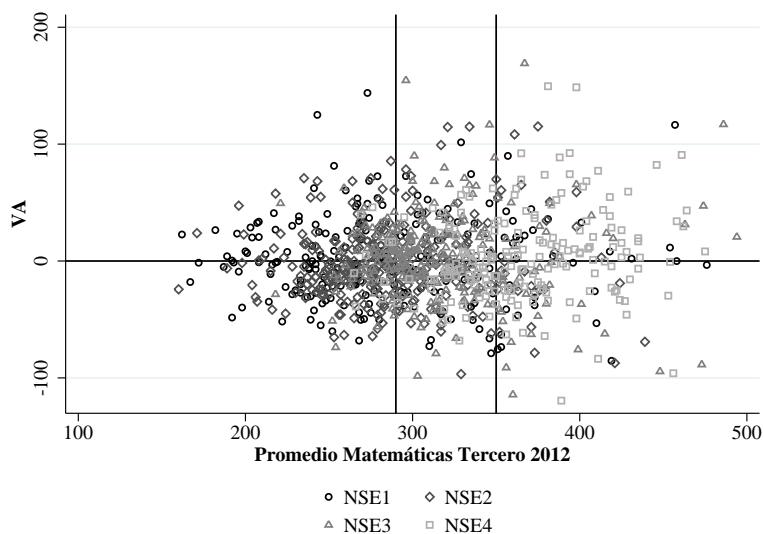


Figura 1: Cuadrantes de desempeño. Cohorte Tercero 2012 a Quinto 2014 matemáticas. Promedio. Fuente: elaboración propia.

La figura 3 continúa con el ejemplo del modelo de VA para la cohorte Tercero 2012 a Quinto 2014 en el área de matemáticas, pero el resultado de interés que muestra es el porcentaje de estudiantes en el nivel de desempeño avanzado; es decir, el análisis se centra en la punta de la distribución. En general, pero no por definición, para que un colegio quede clasificado en desempeño alto, debe tener más del 40% de sus estudiantes en el nivel de desempeño avanzado y los colegios en la categoría baja no tienen más del 20% de sus estudiantes en avanzado. Las figuras 1 a 3 muestran que a medida que se avanza en las categorías de desempeño, aumenta la dispersión del VA. El grupo de comparación compuesto por los colegios clasificados en desempeño alto, tanto para el puntaje promedio como para los extremos de la distribución, varía más que los otros grupos de comparación en el aporte en términos de aprendizaje que brindan a sus estudiantes.

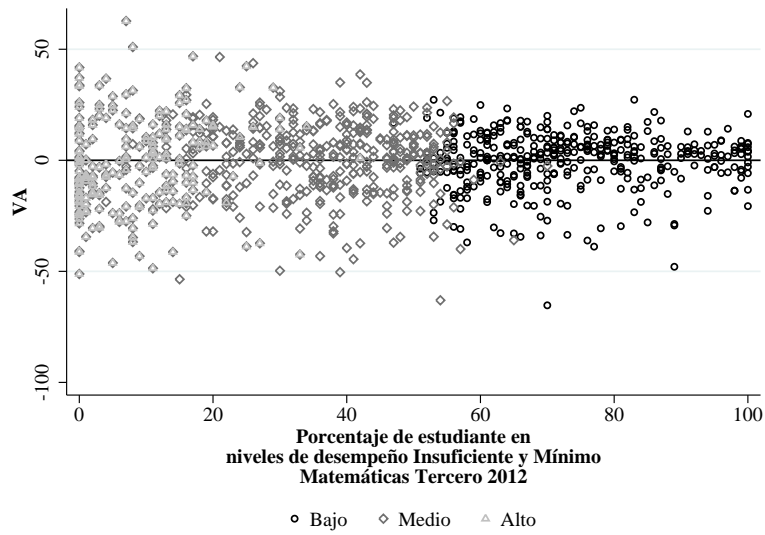


Figura 2: Cohorte Tercero 2012 a Quinto 2014 matemáticas. Insuficiente y Mínimo. Fuente: elaboración propia.

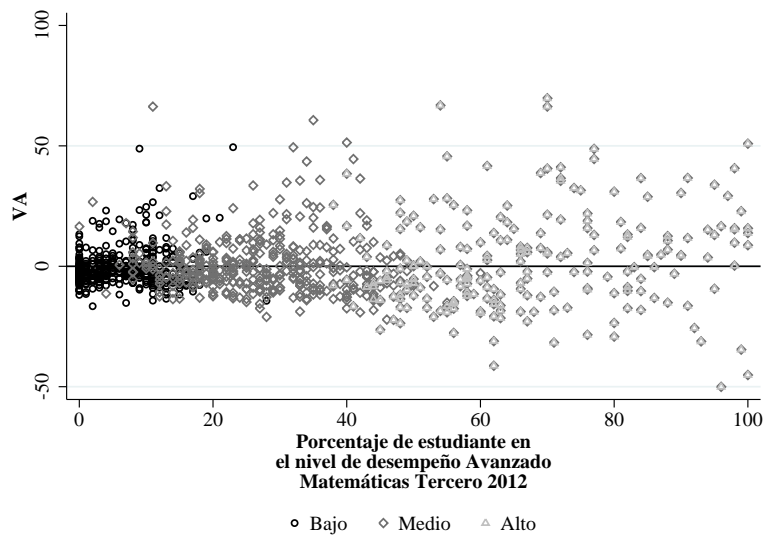


Figura 3: Cohorte Tercero 2012 a Quinto 2014 matemáticas. Avanzado. Fuente: elaboración propia.

A continuación mostramos los resultados de ambas cohortes agregados a nivel

nacional. El agregado nacional es el promedio de los colegios del país ponderado por el número de estudiantes que presentaron la prueba en ambos momentos del tiempo. Para la cohorte Tercero 2012 a Quinto 2014 presentamos los resultados del modelo de crecimiento y del modelo de VA solo para el área de lenguaje y para la cohorte Quinto 2009 a Noveno 2013 presentamos los resultados de ambos modelos solo para el área de matemáticas. Presentamos sólo un área para cada cohorte debido a que a nivel nacional los resultados entre áreas son muy similares. La correlación de las pruebas Saber a nivel de colegio para el agregado nacional entre áreas de un mismo grado nunca es menor a 0.7. Esta tendencia no se mantiene para otro tipo de comparaciones.

Los resultados nacionales son sólo una parte de los hallazgos del análisis debido a que, con ciertas precauciones y para diferentes objetivos, es posible hacer comparaciones entre colegios, municipios, ETC o departamentos. La base de datos con los resultados para estos agregados, para ambas cohortes y áreas y para los resultados de las tres variables dependientes, puede ser solicitada a los autores.

#### **4.1. Resultados nacionales de la cohorte Tercero 2012 a Quinto 2014**

La figura 4 muestra el agregado nacional para los modelos de crecimiento y VA en el área de lenguaje, para el puntaje promedio de los colegios por categoría de desempeño y NSE. Los resultados entre categorías no son comparables, pero dentro de cada categoría es posible comparar entre NSE. En el modelo de crecimiento se evidencia una correlación positiva marcada entre el cambio en el aprendizaje y el nivel socioeconómico de los estudiantes. El crecimiento es negativo en las tres categorías de desempeño para los dos niveles socioeconómicos más bajos. Esto muestra que, sin importar el desempeño del colegio, en promedio, los colegios de estudiantes con más carencias económicas no generan valor al terminar el ciclo de primaria, respecto del aprendizaje que tenían en el grado tercero. Análogamente, sin importar el desempeño, los colegios de estudiantes con más bienestar económico son los que más generan valor al terminar el ciclo de primaria.

La parte derecha de la figura 4 muestra los resultados del modelo de VA. Estos resultados no están en la misma escala del modelo de crecimiento porque los resultados de VA se “encogen” en comparación con los del modelo anterior. La razón es que en el modelo de VA la diferencia entre el resultado observado y esperado del colegio es, en promedio, menor que la del modelo de crecimiento, porque el valor esperado se calcula respecto a colegios similares. Por el contrario, el valor esperado en el modelo de crecimiento se calcula respecto a todos los colegios del país. El modelo de VA quita el efecto de varias características contextuales (variables que podrían estar afectando el desempeño académico de los colegios) para dar una estimación más “limpia”. Los resultados muestran que se pierde la correlación entre el valor agregado de los colegios y los niveles socioeconómicos.

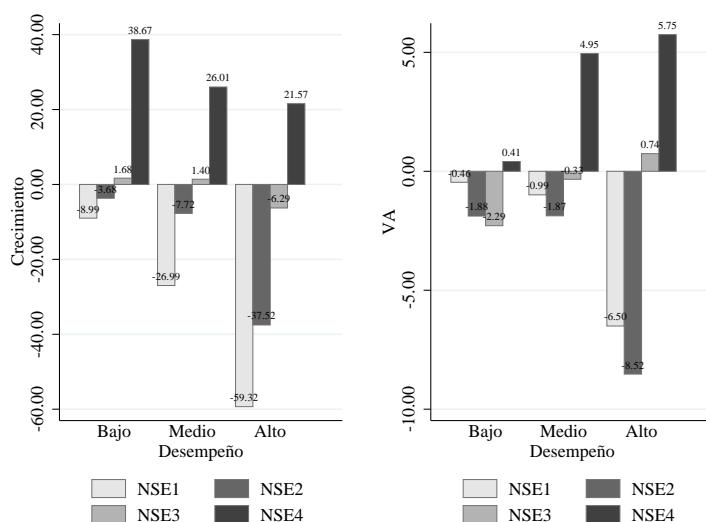


Figura 4: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño y nivel socioeconómico. Cohorte Tercero 2012 a Quinto 2014. Área de lenguaje. Fuente: elaboración propia.

Es importante resaltar que ambos modelos brindan información relevante para objetivos de política pública. El modelo de VA quita del efecto del colegio características de los estudiantes que él no puede controlar y, por tanto, el resultado es una estimación del efecto directo y exclusivo del colegio. Este resultado tiene efectos de equidad a la hora de comparar colegios. Por el contrario, el modelo de crecimiento brinda una medida absoluta de cómo están mejorando los estudiantes en el tiempo, independientemente de sus características demográficas y socioeconómicas. La diferencia de las estimaciones entre el modelo de crecimiento y el modelo de VA, puede ser interpretada como la diferencia en el aumento del aprendizaje de los estudiantes que no fue causada directamente por las variables que no fueron controladas en el modelo, entre ellas el colegio. Los resultados que más llaman la atención son en los que hay cambio de signo entre ambos modelos. En las estimaciones de la figura 4 pasa para los colegios de NSE 3. Por ejemplo, para los de desempeño bajo, el aumento promedio en términos absolutos es efecto de las características contextuales y para los de desempeño alto, el modelo de VA muestra que estos colegios, en promedio, están aportando valor a sus estudiantes. Cabe resaltar que en el modelo de VA los colegios de desempeño alto, categorizados en el NSE más deseable, son los que en promedio generan mayor valor agregado en el aprendizaje. Esto muestra que más allá de las características contextuales, los colegios que atienden a este tipo de estudiantes son los que mejor están haciendo el trabajo.

Las figura 5 y 6 muestran los resultados para los extremos de la distribución de los colegios. La figura 5, que muestra el progreso en términos de los niveles de desempeño insuficiente y mínimo debe interpretarse de tal forma que un decrecimiento se valore positivamente. Esto debido a que el aprendizaje se evidencia en la reducción del porcentaje de estudiantes en estos niveles de desempeño. Los resultados muestran que se mantiene la correlación entre desempeño y NSE. Los modelos de crecimiento revelan que los colegios de desempeño alto que atienden estudiantes del nivel socioeconómico menos deseable, son los que generan menores ganancias en el aprendizaje.

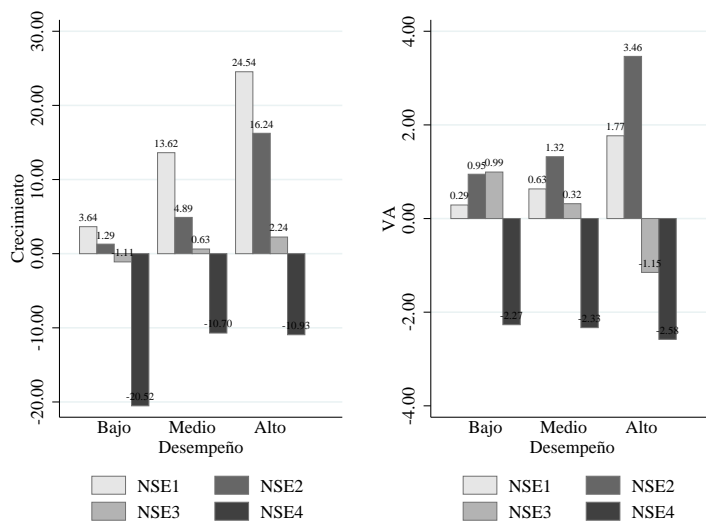


Figura 5: *Crecimiento y valor agregado para niveles de desempeño insuficiente y mínimo por categorías de desempeño y nivel socioeconómico. Cohorte Tercero 2012 a Quinto 2014. Área de lenguaje. Fuente: elaboración propia.*

Al comparar los paneles de cada figura, se evidencia que el decrecimiento en términos absolutos en las puntas de la distribución de los colegios del nivel socioeconómico menos deseable cambia drásticamente al ajustar por características contextuales. Los colegios del primer nivel socioeconómico logran incluir estudiantes en el nivel avanzado, pero fallan en reducir el porcentaje en insuficiente y mínimo.

La comparación entre los resultados del promedio del colegio y las puntas de la distribución (figuras 4, 5, 6), muestra que los colegios que agregaron valor lo hicieron principalmente por mover al centro de la distribución a los estudiantes con peores resultados, y no necesariamente por desplazarlos a la derecha. Muchos colegios crecieron en promedio, a pesar de disminuir el porcentaje de estudiantes en el nivel de desempeño avanzado. Análogamente, los colegios que no agregaron valor en promedio, lo hicieron principalmente por no sacar a sus estudiantes de los



dos niveles de desempeño menos deseables. Este hecho sugiere que, sin importar el desempeño o el nivel socioeconómico, los esfuerzos de política pública deben enfocarse en nivelar a los estudiantes que están debajo del promedio de sus respectivos colegios.

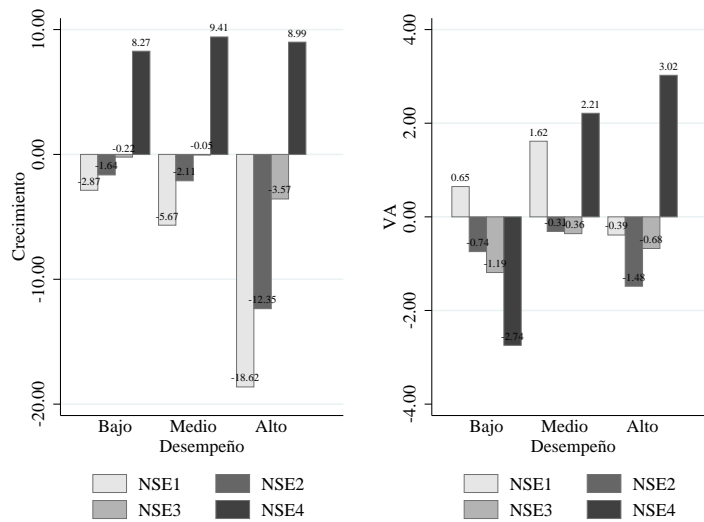


Figura 6: Crecimiento y valor agregado para el nivel de desempeño avanzado por categorías de desempeño y nivel socioeconómico. Cohorte Tercero 2012 a Quinto 2014. Área de lenguaje. Fuente: elaboración propia.

La figura 7 muestra los resultados para el promedio del área de lenguaje por categorías de desempeño y por zona y sector. Tanto en el modelo de crecimiento como en el de VA los colegios que generan más valor a sus estudiantes son los rurales no oficiales. Estos colegios son en general campestres y bilingües, con características totalmente diferentes a los rurales oficiales. Este resultado confirma, una vez más, que la educación que reciben los estudiantes de nivel socioeconómico alto es representativamente más alta que la que reciben otros estudiantes.

Hay dos conclusiones importantes para los colegios oficiales en los resultados por zona. Después de quitar el efecto de las características contextuales, los colegios rurales siguen sin generar valor a sus estudiantes, pero los colegios urbanos muestran mejor comportamiento, en especial los de desempeño alto que son los que generan más valor entre todos los grupos de comparación.

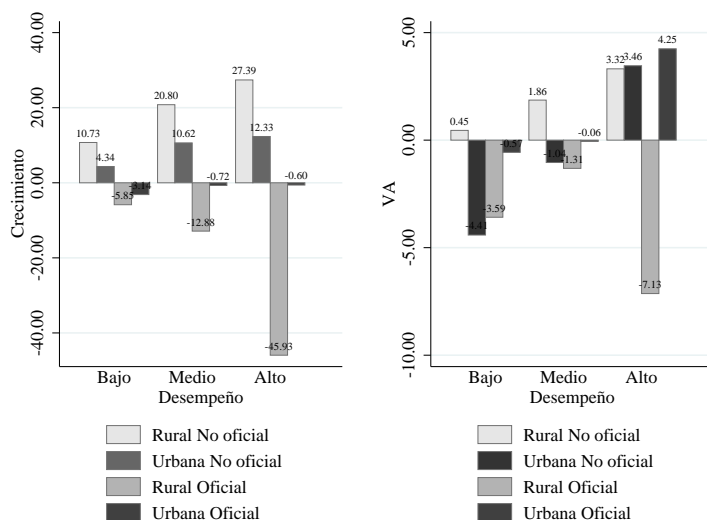


Figura 7: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño, Zona y Sector. Cohorte Tercero 2012 a Quinto 2014. Área de lenguaje. Fuente: elaboración propia.

#### 4.2. Resultados nacionales de la cohorte Quinto 2009 a Noveno 2013

Hay cuatro años de diferencia en el periodo de estudio de la cohorte Quinto 2009 a Noveno 2013. Además, quinto es el último grado del ciclo de primaria y noveno el último de secundaria. La magnitud del cambio en el aprendizaje de esta cohorte en este periodo es determinante para el futuro cercano de estos estudiantes. La figura 8 muestra que, al igual que para la cohorte Tercero 2012 a Quinto 2014, los colegios de desempeño alto de NSE 4 son los que más aportan valor a sus estudiantes. Además, en el modelo de VA el efecto del colegio es el doble que el de la anterior cohorte. Por tanto, se destaca que este tipo de colegios, aún después de quitar el efecto de características socioeconómicas, sigue reportando los mejores resultados. De igual forma se evidencia que la educación de los estudiantes de nivel socioeconómico alto es categóricamente mejor que la que reciben estudiantes menos afortunados.

La figura 9 muestra el análisis por zona y sector. Al igual que en la cohorte Tercero 2012 a Quinto 2014, los colegios que generan más valor a sus estudiantes son los rurales no oficiales y los que generan menos son los rurales oficiales. A diferencia de la anterior cohorte, no hay evidencia de buenos resultados para los colegios urbanos oficiales en comparación con los colegios rurales oficiales.

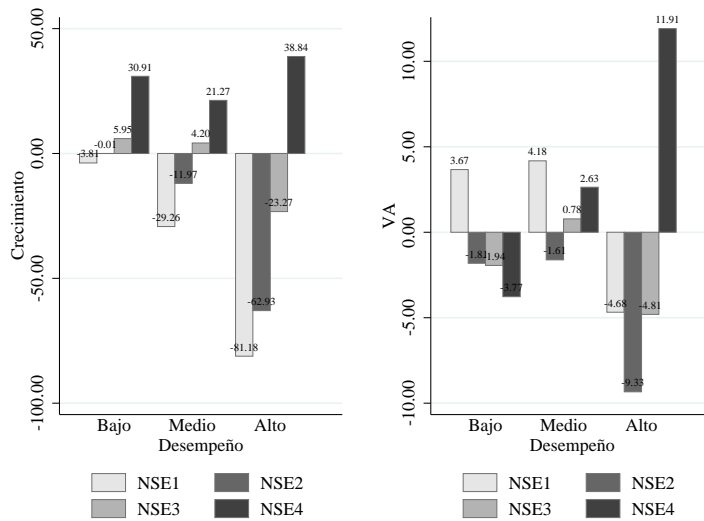


Figura 8: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño y nivel socioeconómico. Cohorte Quinto 2009 a Noveno 2013. Área de matemáticas. Fuente: elaboración propia.

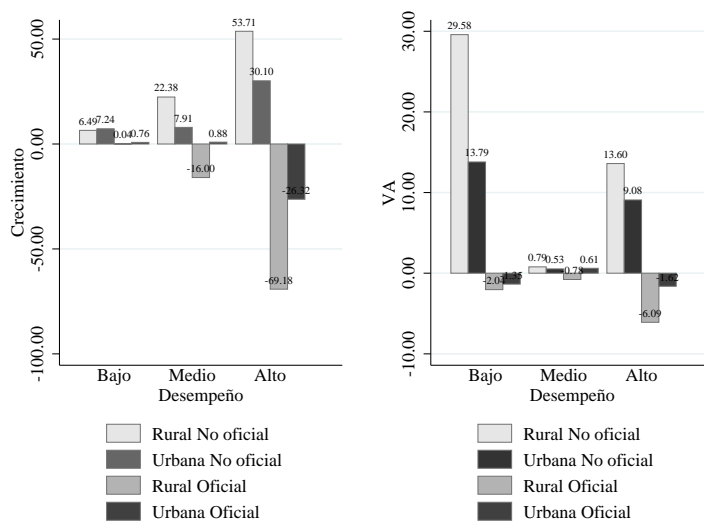


Figura 9: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño, zona y sector. Cohorte Quinto 2009 a Noveno 2013. Área de matemáticas. Fuente: elaboración propia.

En la figura 10, el modelo de crecimiento evidencia que hay un modesto aumento del porcentaje de estudiantes en avanzado en los colegios de bajo rendimiento de todos los niveles socioeconómicos. Esto muestra que, aunque la mayoría de estudiantes no alcanzan las metas esperadas, para el grado noveno, los colegios sí consiguen que las alcancen los estudiantes más aventajados académicamente.

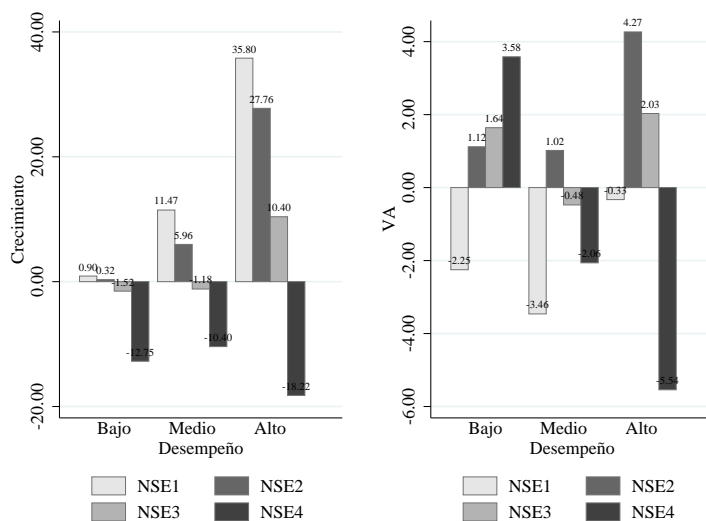


Figura 10: *Crecimiento y valor agregado para niveles de desempeño insuficiente y mínimo por categorías de desempeño y nivel socioeconómico. Cohorte Quinto 2009 a Noveno 2013. Área de matemáticas. Fuente: elaboración propia.*

El aspecto que más se destaca en las figuras 10 11 es que, a diferencia de la cohorte Tercero 2012 a Quinto 2014, los colegios de bajo NSE en el modelo de VA logran reducir el porcentaje de estudiantes en insuficiente y mínimo y aumentar el porcentaje en avanzado. El anexo 2 muestra los resultados nacionales del área de matemáticas para la cohorte Tercero 2012 a Quinto 2014 y del área de lenguaje para la cohorte Quinto 2009 a Noveno 2012. Las conclusiones en el análisis por NSE son muy similares para ambas áreas, pero hay una diferencia importante en el análisis por zona y sector. El modelo de VA muestra buenos resultados en ambas cohortes para los colegios oficiales urbanos en el área de lenguaje, pero no para el área de matemáticas.

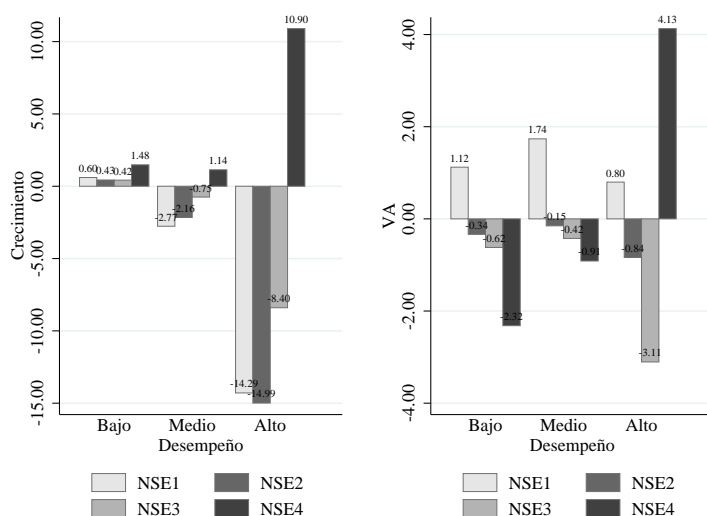


Figura 11: *Crecimiento y valor agregado para el nivel de desempeño avanzado por categorías de desempeño y nivel Socioeconómico. Cohorte Quinto 2009 a Noveno 2013. Área de matemáticas. Fuente: elaboración propia.*

## 5. Conclusión

Al menos hasta nuestro conocimiento, en este artículo se hace la primera aproximación a los modelos de VA para las pruebas Saber 3°, 5° y 9° en Colombia. Esta metodología nos permite seguir a dos cohortes en el tiempo. La primera cohorte es la de estudiantes que en 2012 cursaban el grado tercero de primaria y que en 2014 estaban en quinto de primaria. Como el Icfes implementará la prueba Saber 7°, en 2016 será posible volver a seguir a este mismo grupo de estudiantes. La segunda cohorte que seguimos es la de los estudiantes que en 2009 cursaban el grado quinto de primaria y que cuatro años después, en 2013, estaban en noveno de bachillerato. El próximo año, cuando el Icfes califique y publique los resultados de la prueba Saber 11 del año en curso, podrá seguirse de nuevo a esta cohorte. Esto implica que podrá estudiarse a este grupo de estudiantes en los tres ciclos que completan la educación básica y media: primaria, secundaria y media.

Las estimaciones de VA son pertinentes porque permiten seguir a una misma cohorte en el tiempo, con el objetivo de estimar el cambio en el aprendizaje de los estudiantes. Esta posibilidad permite conocer el resultado de un colegio en un punto final descontado por lo que sus estudiantes ya sabían en un punto inicial. Por el contrario, los estudios de corte transversal solo permiten hacer comparaciones en un momento del tiempo o comparaciones temporales entre diferentes cohortes. Los modelos de VA son un buen complemento para los análisis transversales, debido

a que saber cuánto han mejorado los estudiantes es tan importante como saber cómo están al finalizar un periodo.

Nuestro ejercicio empírico se basa en dos modelos: uno sin ajustes contextuales y otro que controla por diferentes características demográficas, sociales y económicas. El modelo sin ajustes brinda una medida del cambio en el aprendizaje que experimentan los estudiantes pertenecientes a un colegio en un periodo. El modelo con ajustes separa los efectos exógenos al colegio del efecto directo de este. Ambos modelos tienen implicaciones prácticas de política pública. El primer modelo puede ser usado para identificar colegios con resultados negativos y focalizar la atención hacia ellos y el segundo modelo para premiar o reconocer el esfuerzo de los mejores.

Los resultados del análisis nacional muestran que las condiciones económicas de los estudiantes se relacionan fuertemente con el progreso educativo. Los estudiantes pertenecientes a los niveles socioeconómicos más altos, y que atienden colegios con matrículas costosas, son los que mejoran sus competencias en mayor medida. Este resultado se evidencia en el modelo que separa las características de los estudiantes del efecto directo del colegio: no son solo las condiciones favorables de estos estudiantes las que les permiten mejor desempeño, sino que los colegios efectivamente brindan una mejor formación.

Nuestro análisis, además de ver el comportamiento promedio de los colegios, se centra en la distribución de los estudiantes dentro de este. Los resultados muestran que hay colegios que a pesar de tener bajo rendimiento tienen estudiantes en niveles de desempeño alto. Estos colegios logran mantener a sus estudiantes de desempeño alto en estos niveles. El problema radica en que fallan en nivelar a sus estudiantes de bajo desempeño con el resto de sus estudiantes.

La interpretación de los resultados de nuestros modelos de VA debe tener en cuenta el desempeño inicial de los colegios, debido a que los aumentos en el aprendizaje tienen diferentes significados dependiendo del nivel en que inicien. Por esta razón, presentamos nuestras estimaciones en tres categorías de desempeño basadas en los resultados del año inicial. Las estimaciones de los colegios u otros agregados son comparables solo dentro de cada categoría de desempeño y no son comparables entre estas categorías. Hay otras formas de mostrar los resultados de VA. Una de ellas es mostrar los resultados con medidas de aporte relativo. Estas medidas no muestran la estimación puntual de VA, sino una comparación de cada colegio respecto de otros con ciertas condiciones similares. En la página web del Icfes están publicados los resultados del estudio del aporte relativo de todas las universidades del país. (de Gestión de Proyectos de Investigación y Dirección de Evaluación. 2015) usa las pruebas Saber 11 y Saber Pro para estimar el VA de los grupos de referencia de las universidades. Los resultados presentan para cada institución-grupo de referencia, como el aporte relativo con respecto a las instituciones cuyos estudiantes hayan tenido resultados más similares en Saber 11.

Otra forma de mostrar los resultados de VA podría basarse en una única medida relativa al comportamiento de un grupo de colegios similares. En una investigación del (de Análisis y divulgación y Dirección de Evaluación. 2015) se están constru-

yendo grupos de comparación que no se basan en resultados cognitivos, sino en características contextuales. El objetivo de ese análisis es saber cómo está cada colegio respecto a colegios con condiciones socioeconómicas similares. Combinar este ejercicio con el de VA podría arrojar información valiosa.

Una falencia de nuestro trabajo es que, debido a que las pruebas Saber 3°, 5° y 9° no brindan una calificación particular para los estudiantes, la unidad de observación para el análisis son los colegios. Este hecho implica que las estimaciones son válidas bajo el supuesto de que hay estabilidad estudiantil en los colegios. Este supuesto se podría comprobar si existieran calificaciones a nivel de estudiante para las pruebas Saber 3°, 5° y 9°. Una opción menos costosa podría darse si desde otras fuentes se pudiera saber la proporción de estudiantes de una cohorte que continúan en cada colegio. De esta forma habría más confianza en la estimación para colegios con estabilidad estudiantil comprobada.

Nuestras medidas de VA pueden proveer a los encargados de las políticas educativas evidencias del aprendizaje de los estudiantes en los colegios. Estas evidencias pueden tener fines de rendición de cuentas, pueden ser usadas internamente por las instituciones para sus planes de mejoramiento o para identificar debilidades y fortalezas de los colegios (Steedle 2010). Sin embargo, pensamos que la complejidad del ambiente educativo requiere que las interpretaciones de las estimaciones de VA incluyan varias advertencias para una justa y correcta interpretación (OECD 2008). Aun corrigiendo los problemas de nuestro ejercicio empírico, varias dificultades se mantienen a la hora de estimar el efecto de los colegios en el progreso educativo de sus estudiantes. Todo lo que se aprende en un colegio se puede traducir en acumulación de conocimiento, habilidades, herramientas, costumbres y valores éticos, además del efecto en la forma en que los estudiantes piensan, sienten y actúan (Bennett 2001, Harvey & Green 1993). Los modelos de VA usan datos de pruebas estandarizadas que sólo pueden medir una parte de toda la educación que se brinda en un colegio.

**Recibido: 12 de febrero del 2016**

**Aceptado: 14 de abril del 2016**

## Referencias

- Aaronson, D., Barrow, L. & Sander, W. (2007), 'Teachers and student achievement in the Chicago public high schools.', *Journal of Labor Economics*. **25**(1), 95–135.
- Armour-Garb, A. (2009), 'Should "value-added" models be used to evaluate teachers?', *Journal of Policy Analysis and Management*. **28**(4), 692–712.
- Bennett, D. C. (2001), 'Assessing quality in higher education.', *Liberal Education*. **87**(2).

- Colbert, V. (1999), 'Mejorando el acceso y la calidad de la educación para el sector rural pobre: El caso de la escuela nueva en Colombia.', *Revista Iberoamericana de Educación*. .
- de Análisis y divulgación y Dirección de Evaluación., S. (2015), Interpretando con equidad: vecindades de comparación en saber 3°, 5° y 9°, Technical report, Instituto Colombiano para la Evaluación de la Educación. ICFES.
- de Gestión de Proyectos de Investigación y Dirección de Evaluación., O. A. (2015), Reporte de resultados en saber pro, medidas de aporte relativo y otros indicadores de calidad de las instituciones de educación superior 2013-2014., Technical report, Instituto Colombiano para la Evaluación de la Educación. ICFES.
- Doran, H. C. & Izumi, L. T. (2004), 'Putting education to the test: A value-added model for California.', *Pacific Research Institute*. .
- Doran, H. C. & Lockwood, J. R. (2006), 'Fitting value-added models in r.', *Journal of Educational and Behavioral Statistics*. **31**(2), 205–230.
- Downes, D. & Vindurampulle, O. (2007), 'Value added measures for school improvement.', *Office for Education Policy and Innovation*. (13).
- Dury, D. & Doran, H. (2003), 'The value of value-added analysis. policy research brief.', *National School Boards Association*. **3**(1), 1–4.
- Gaviria, A. & Barrientos, J. (2001), 'Características del plantel y calidad de la educación en Bogotá.', *Centro de Investigación Económica y Social, Fedesarrollo*. .
- Harvey, L. & Green, D. (1993), 'Defining quality.', *Assessment and Evaluation in Higher Education*. **18**(1), 9–34.
- Haveman, R. & Wolfe, B. (1995), 'The determinants of childrens attainments: A review of methods and findings.', *Journal of Economic Literature*. **33**(4), 1829–1878.
- Hill, P. (1995), 'Value-added measures of achievement.', *Incorporated Association of Registered Teachers of Victoria Seminar Series*. **44**.
- Kazeem, A. and Jensen, L. & Stokes, S. (1995), 'School attendance in Nigeria: Understanding the impact and intersection of gender, urbana rural residence, and socioeconomic status.', *Comparative Education Review* **54**(2), 295–319.
- Kim, H. & Lalancette, D. (2013), Literature review on the value added measurement in education., Technical report, OECD.
- Koedel, C., Mihaly, K. & Rockoff, J. (n.d.), Value added modelling: A review. Unpublished article. (2015).
- Nunez, J., Steiner, R., Cadena, X. & Pardo, R. (2002), '¿Cuáles colegios ofrecen mejor educación en Colombia? ', *CEDE, Universidad de los Andes* .



- OECD (2008), Measuring improvements in learning outcomes: Best practices to assess the value-added of schools., Technical report, OECD.
- OECD (2011), La medición del aprendizaje de los alumnos: Mejores prácticas para evaluar el valor agregado de las escuelas., Technical report, OECD.
- Raudenbush, S. W. & Bryk, A. S. (1986), 'A hierarchical model for studying school effects.', *Sociology of Education*. **59**, 1–17.
- Reardon, S. F. & Raudenbush, S. W. (2009), 'Assumptions of value added models for estimating school effects.', *Education Finance and Policy*. **4**(4), 492–519.
- Rivkin, S. (2005), 'Teachers, schools and academic achievement.', *Econometrica*. **73**(2), 17–58.
- Saunders, L. (1998), "“Value Added” measurement of school effectiveness: An overview.', *National Foundation of Economic Research* .
- Steedle, J. T. (2010), 'Improving the reliability and interpretability of value-added scores for post-secondary institutional assessment programs.', *Paper presented at the Annual Meeting of the American Educational Research Association*. .
- Tobón, D., Valencia, G., Ríos, P. & Bedoya, J. (2008), 'Organización jerárquica y logro escolar en Medellín. Lecturas de Economía.', *Universidad de Antioquia*. **68**, 145–173.
- Willms, D. (2009), 'Value added models in education. policy brief.', *Canadian Research Institute for Social Policy* .
- Wyatt, T. (1996), 'School effectiveness research: dead end, damp squib or smouldering fuse?', *Issues in Educational Research*. **6**(1), 79–112.
- Zhang, Y. (2006), 'Urban-rural literacy gaps in sub-saharan africa: The roles of socioeconomic status and school quality.', *Comparative Education Review*. **50**(4).

## Anexos:

## Anexo 1

Tabla A1.1:

Modelos de crecimiento y valor agregado. Variable dependiente: Puntaje promedio del colegio en el año final.\*

Variables independientes	Cohorte Tercero 2012 a Quinto 2014				Cohorte Quinto 2009 a Noviembre 2013			
	Matemáticas		Lenguaje		Matemáticas		Lenguaje	
	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado
Promedio año inicial	0.626	0.318	0.695	0.240	0.612	0.710	0.894	0.886
Desviación año inicial	0.007	0.194	0.007	0.162	0.009	0.281	0.009	0.248
		-0.054		0.014		-0.038		0.010
Nivel Socioeconómico		0.091		0.028		0.037		0.038
		<b>1.953</b>		<b>2.693</b>		<b>4.027</b>		<b>3.900</b>
DeserT12_Q14		0.102		0.094		0.112		0.108
		-0.178		-0.162		-0.053		-0.036
		0.090		0.026		0.024		0.023
FE ETC	No	Si	No	Si	No	Si	No	Si
FE Zona Sector	No	Si	No	Si	No	Si	No	Si
Int. (Promedio)(ETC)	No	Si	No	Si	No	Si	No	Si
Int. (Promedio)(ZonaSector)	No	Si	No	Si	No	Si	No	Si
Constante	109.337	64.749	92.614	50.260	51.945	-31.454	24.522	-80.280
	2.332	47.518	2.074	42.968	2.854	49.653	2.648	50.410
N	9876	9773	9888	9783	7612	7313	7764	7456
F	7081.46	5981.6	10936.07	92.85	7738.99	86.62	10940.07	10447
R2	0.42	0.54	0.53	0.65	0.50	0.70	0.58	0.73
R2 ajustado	0.42	0.53	0.53	0.64	0.50	0.69	0.58	0.73

\*Coeficiente en negrita y error estándar debajo. In=Interacción. Modelos de crecimiento estimados con MCO y modelos de valor agregado MCO con efectos fijos.

Tabla A1.2:

Modelos de crecimiento y valor agregado. Variable dependiente: Porcentaje de estudiantes en los niveles de desempeño Insuficiente y Mínimo en el año final.\*

Variables independientes	Cohorte Tercero 2012 a Quinto 2014				Cohorte Quinto 2009 a Noviembre 2013			
	Matemáticas		Lenguaje		Matemáticas		Lenguaje	
	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado
Nivel año inicial	0.549	0.194	0.645	0.214	0.691	0.589	0.800	0.995
Desviación año inicial	0.007	0.169	0.006	0.149	0.008	0.329	0.008	0.274
		-0.015		-0.035		-0.013		-0.031
Nivel Socioeconómico		0.013		0.013		0.017		0.019
		-0.797		-1.232		-1.744		-1.808
DeserT12_Q14		0.046		0.044		0.050		0.052
		0.072		0.070		0.023		0.009
		0.013		0.012		0.006		0.006
FE ETC	No	Si	No	Si	No	Si	No	Si
FE Zona Sector	No	Si	No	Si	No	Si	No	Si
Int. (Promedio)(ETC)	No	Si	No	Si	No	Si	No	Si
Int. (Promedio)(ZonaSector)	No	Si	No	Si	No	Si	No	Si
Constante	40.172	68.646	25.435	90.798	29.909	132.263	13.098	120.974
	0.357	5.374	0.359	5.063	0.581	7.158	0.513	7.277
N	9876	9773	9888	9783	7612	7313	7764	7456
F	6800.78	5576	10607.53	90.01	6831.99	77.59	9367.39	92.71
R2	0.41	0.53	0.52	0.64	0.47	0.67	0.55	0.71
R2 ajustado	0.41	0.52	0.52	0.63	0.47	0.67	0.55	0.70

\*Coeficiente en negrita y error estándar debajo. In=Interacción. Modelos de crecimiento estimados con MCO y modelos de valor agregado MCO con efectos fijos.

Tabla A1.3:

Modelos de crecimiento y valor agregado. Variable dependiente: Porcentaje de estudiantes en el nivel de desempeño Avanzado en el año final.\*

Variables independientes	Cohorte Tercero 2012 a Quinto 2014				Cohorte Quinto 2009 a Noveno 2013			
	Matemáticas		Lenguaje		Matemáticas		Lenguaje	
	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado	Crecimiento	Valor Agregado
Nivel año inicial	0.409	0.083	0.533	0.196	0.457	-0.329	0.508	-0.032
Desviación año inicial	0.006	0.199	0.006	0.246	0.006	0.267	0.005	0.254
		-0.017		-0.014		-0.038		-0.049
		0.011		0.011		0.009		0.009
Nivel Socioeconómico		0.521		0.908		0.662		0.647
		0.036		0.033		0.025		0.024
DeserT12_Q14		-0.059		-0.055		-0.022		-0.014
		0.011		0.009		0.003		0.003
FE ETC	No	Si	No	Si	No	Si	No	Si
FE Zona Sector	No	Si	No	Si	No	Si	No	Si
Int. (Promedio)(ETC)	No	Si	No	Si	No	Si	No	Si
Int. (Promedio)(ZonaSector)	No	Si	No	Si	No	Si	No	Si
Constante	4.422	-15.418	4.715	-34.644	-0.997	-32.990	-0.977	-28.580
	0.210	5.462	0.177	4.602	0.135	3.248	0.114	2.930
N	9876	9773	9888	9783	7612	7313	7764	7456
F	5282.53	41.61	8423.24	70.46	5161.99	58.62	8556.12	75.63
R2	0.35	0.45	0.46	0.58	0.40	0.61	0.52	0.66
R2 ajustado	0.35	0.44	0.46	0.57	0.40	0.60	0.52	0.66

\*Coeficiente en negrita y error estándar debajo. Int.=interacción. Modelos de crecimiento estimados con MCO y modelos de valor agregado MCO con efectos fijos.

Anexo 2

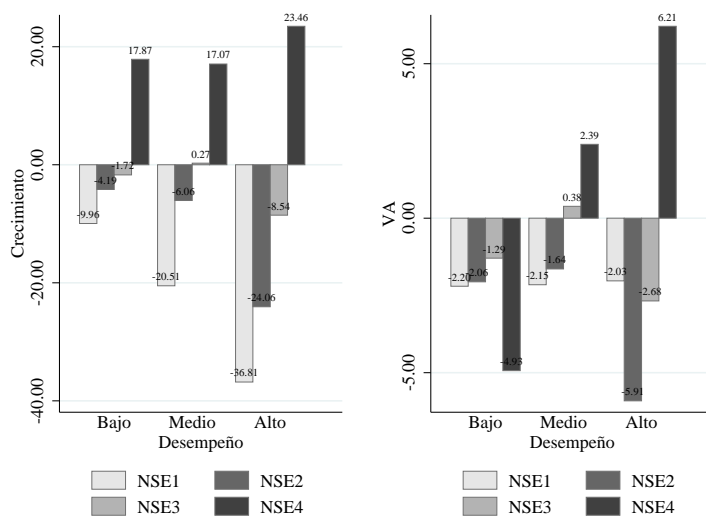


Figura 12: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño y nivel socioeconómico. Cohorte Tercero 2012 a Quinto 2014. Área de matemáticas. Fuente: elaboración propia.

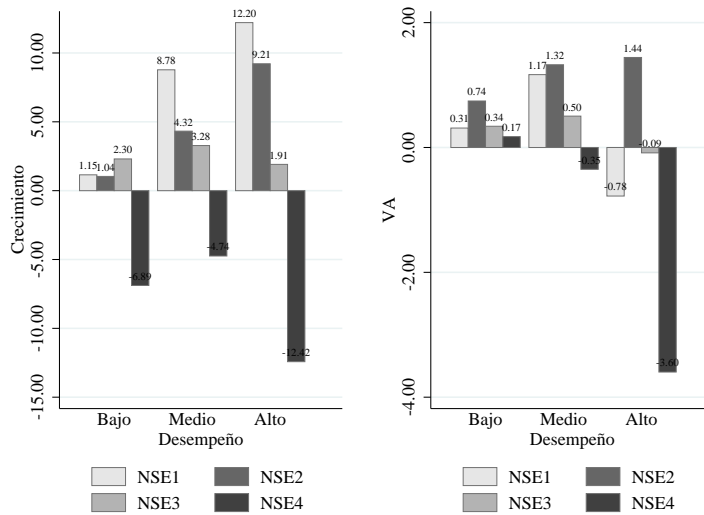


Figura 13: Crecimiento y valor agregado para niveles de desempeño Insuficiente y Mínimo por categorías de desempeño y nivel socioeconómico. Cohorte Tercero 2012 a Quinto 2014. Área de matemáticas. Fuente: elaboración propia.

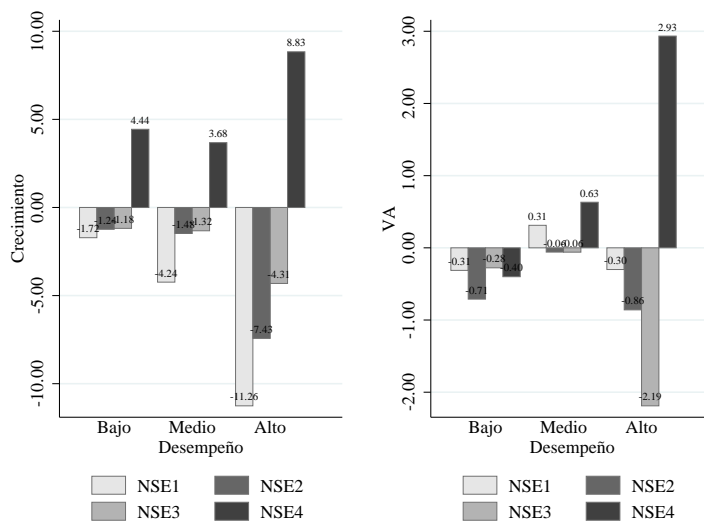


Figura 14: Crecimiento y valor agregado para el nivel de desempeño Avanzado por categorías de desempeño y nivel socioeconómico. Cohorte Tercero 2012 a Quinto 2014. Área de matemáticas. Fuente: elaboración propia.

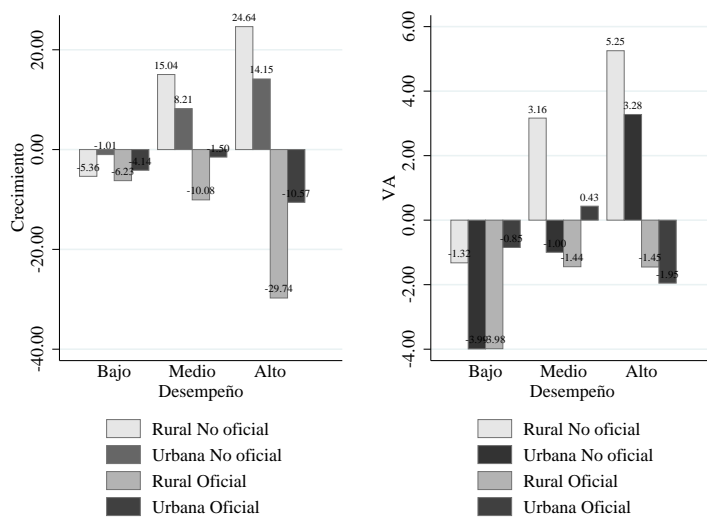


Figura 15: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño, Zona y Sector. Cohorte Tercero 2012 a Quinto 2014. Área de matemáticas. Fuente: elaboración propia.

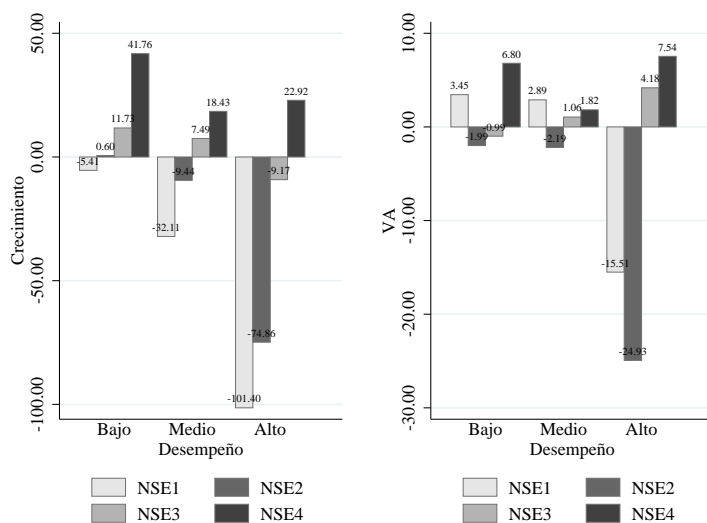


Figura 16: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño y nivel socioeconómico. Cohorte Quinto 2009 a Noveno 2013. Área de lenguaje. Fuente: elaboración propia.

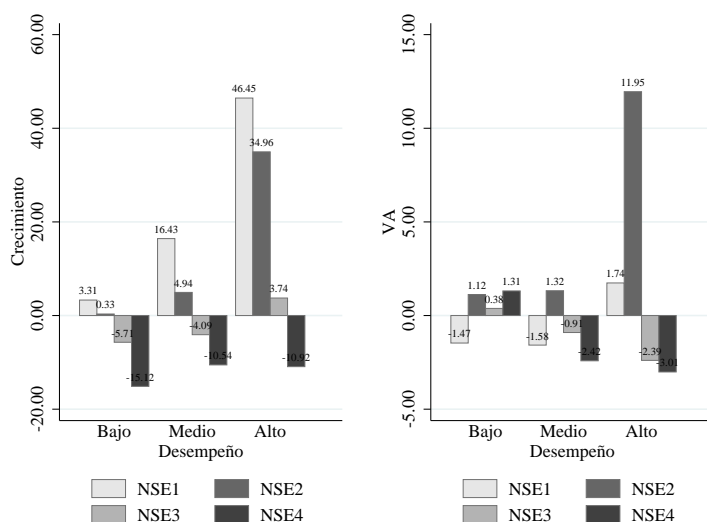


Figura 17: *Crecimiento y valor agregado para niveles de desempeño Insuficiente y Mínimo por categorías de desempeño y nivel socioeconómico. Cohorte Quinto 2009 a Noveno 2013. Área de lenguaje. Fuente: elaboración propia.*

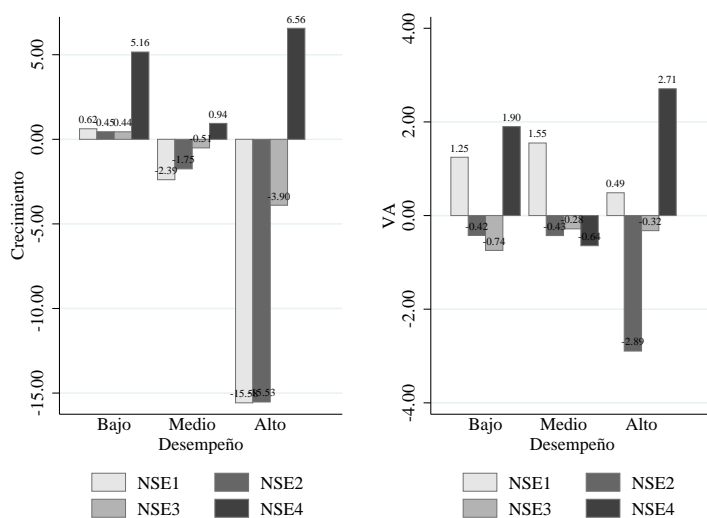


Figura 18: *Crecimiento y valor agregado para el nivel de desempeño Avanzado por categorías de desempeño y nivel socioeconómico. Cohorte Quinto 2009 a Noveno 2013. Área de lenguaje. Fuente: elaboración propia.*

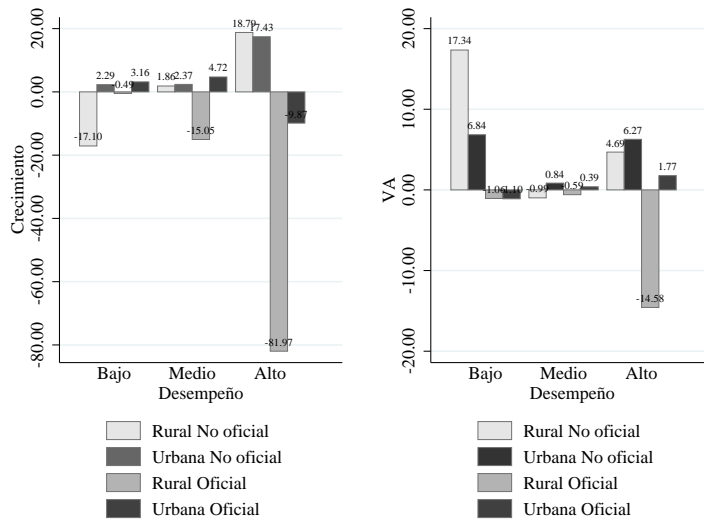


Figura 19: Crecimiento y valor agregado para el puntaje promedio por categorías de desempeño, zona y sector. Cohorte Quinto 2009 a Noveno 2013. Área de lenguaje. Fuente: elaboración propia.