

Adjudicación Simplificada N° 010-2016-PRONABEC

SERVICIO PARA LA ELABORACIÓN DE UN ANÁLISIS DE RESULTADOS DEL EXAMEN ÚNICO DE BECA 18 DE PREGRADO Y BECAS ESPECIALES DELEGADAS EN ADMINISTRACIÓN – CONVOCATORIA 2016

Producto N° 3 Informe Final

Presentado para:



EVACBP SAC
EVALÚA CALIDAD BP SAC

Elaborado por:

Consortio: Centro de Estudios y Promoción del Desarrollo

Diciembre, 2016

Índice

1. Descripción del estudio sobre el Examen Único de Beca 18	3
2. Revisión de evidencia nacional e internacional	6
3. Análisis estadístico	28
4. Modelo econométrico	39
5. Conclusiones y recomendaciones de política	56
6. Bibliografía	58
Anexo	61

1. DESCRIPCIÓN DEL ESTUDIO SOBRE EL EXAMEN ÚNICO DE BECA 18

El Programa Nacional de Becas y Crédito Educativo - PRONABEC, dependiente del Ministerio de Educación, gestiona, a través de Beca 18, subsidios en educación superior a jóvenes de alto rendimiento y en condición de pobreza o pobreza extrema según el Sistema de Focalización de Hogares (SISFOH).

De esta forma, cada año se implementa una nueva Convocatoria de Becas, estableciendo requisitos de postulación según las necesidades del país, otorgando más de 50 mil becas en todo el país a la fecha, desde su primera convocatoria en el año 2012.

En la Convocatoria 2016 de Beca 18 se vio la necesidad de uniformizar o estandarizar lo que en este contexto se denomina “talento académico” y consiste en la medida de las capacidades académicas adquiridas por el postulante a lo largo de la secundaria. Si bien sigue siendo un requisito indispensable el haber alcanzado un puntaje mínimo como promedio general en dicho nivel, la implementación del Examen Único Nacional (EUN) de Beca 18 Convocatoria 2016 representa un esfuerzo único entre el Estado y el sector privado de uniformizar la medida de los logros académicos, en función a promover la óptima inversión del Estado y el buen desempeño de los becarios en la educación superior.

Por lo anterior, el PRONABEC ha considerado la necesidad de implementar un estudio sobre los resultados del EUN en esta su primera implementación en una Convocatoria de Beca 18, de tal manera que se mida la pertinencia de aplicar una prueba como esta al final de la secundaria.

En este sentido, se han establecido los objetivos del estudio de la siguiente manera:

Objetivo General

- Contar con un informe analítico de los resultados del Examen Único de Beca 18 de Pregrado y Becas Especiales delegadas en administración – Convocatoria 2016, con el fin de fortalecer la gestión del programa.

Objetivos Específicos

- Contar con una revisión de la evidencia mundial sobre la pertinencia de aplicar, al final de la secundaria, una evaluación estandarizada.
- Contar con un análisis estadístico de los resultados del Examen Único de Beca 18 de Pregrado y Becas Especiales delegadas en administración – Convocatoria 2016.
- Contar con un modelo econométrico relacionado a factores asociados a mejores resultados en el Examen Único.
- Contar con recomendaciones de política a partir del análisis de resultados del Examen Único de Beca 18 de Pregrado y Becas Especiales delegadas en administración – Convocatoria 2016.

Sobre las actividades y los plazos de entrega correspondientes por cada una, la Tabla 1 brinda esta información en detalle.

Tabla 1. Actividades y plazos de entrega

Producto	Descripción	Plazo
Firma de contrato: 11 de octubre de 2016		
Primer producto: Plan de trabajo general	Contiene todos los puntos a considerar en el análisis de resultados del Examen Único, detallando los procesos de recolección de información, un calendario de cada acción a implementar y una propuesta de bibliografía a utilizar.	A los 30 días calendario, los cuales empiezan a regir a partir del día siguiente de la suscripción del contrato
Segundo producto: Plan de análisis de la información	Contiene la elaboración de un análisis descriptivo de las variables que se tomarán en cuenta en el estudio, en un modelo multivariado, así como un modelo econométrico pertinente al análisis de resultados del Examen Único.	A los 60 días calendario, los cuales empiezan a regir a partir del día siguiente de la suscripción del contrato
Tercer producto: Informe final	Contiene: una revisión de evidencia nacional e internacional que justifique la pertinencia de aplicar una evaluación estandarizada al finalizar los estudios de educación secundaria; un análisis estadístico de los resultados en la evaluación única nacional diferenciando por condiciones de interés; un modelo econométrico que permita identificar qué factores se encuentran asociados a un mejor resultado en el examen único nacional, recomendaciones de política pública y una presentación (en formato ppt) de los hallazgos más importantes de las recomendaciones de política.	A los 75 días calendario, los cuales empiezan a regir a partir del día siguiente de la suscripción del contrato

2. REVISIÓN DE EVIDENCIA NACIONAL E INTERNACIONAL

Sobre evaluaciones estandarizadas a nivel nacional e internacional

Las evaluaciones estandarizadas a nivel mundial tienen larga data y su práctica se ha extendido por el planeta. Desde las evaluaciones internacionales como por ejemplo el *Programme for International Student Assessment* (Programa Internacional para la Evaluación de Estudiantes, o prueba PISA) realizada cada tres años en varios países del orbe, y que toma en cuenta campos de matemáticas, lectura, y ciencias, determinando así una suerte de valoración de los estudiantes por país, teniendo gran impacto a nivel social en los distintos sistemas educativos, hasta evaluaciones nacionales como el caso emblemático en América Latina de la Prueba de Selección Universitaria (PSU) en Chile, que es la puerta de entrada a la educación superior, y que ha dado pie a varios análisis de su impacto y necesidad de mejora continua, su lógica es cada vez más constante y replicada en otras realidades.

El impacto de las evaluaciones estandarizadas en muchos países pasan por varios aspectos, entre los cuales está el reconocimiento a la labor del sistema educativo, la medida de los logros, y en su parte más constructiva: un termómetro que permite conocer las deficiencias y suficiencias de cada política implementada desde el sector Educación. Es pertinente señalar que no se conoce aún con exactitud el impacto que se tiene respecto a la implementación, pues son varios aspectos a evaluar, desde el sentido que tiene homogeneizar la evaluación al final de la secundaria, siendo heterogénea la formación durante la misma, hasta el rol predictivo que tiene o debería tener dicha evaluación sobre las habilidades que necesita la persona cuando se inserte al mercado laboral.

Existe evidencia que muestra aspectos positivos, y una serie de críticas a su lógica que merecen atención. Desde su perspectiva, lo que se ha buscado desde un

principio es establecer una medición que sea válida para conocer el rendimiento cognitivo de cada estudiante respecto a los demás. Sin embargo, hay quienes arguyen sobre la insostenibilidad de tomar una evaluación estandarizada como mediador de la posibilidad de un estudiante de ingresar a la educación superior. También se cuestiona el hecho de homogenizar la formación académica partiendo del hecho de que entre los alumnos existen diferencias individuales, además de proceder de colegios distintos con currículos distintos.

Por lo anterior, en el presente capítulo se busca presentar lo que se conoce sobre las evaluaciones estandarizadas, su impacto, su lógica y su utilidad, prestando atención a los diversos actores y autores que indican tanto resultados positivos como negativos.

En este sentido, se tomará en cuenta dos aspectos generales:

- 1.- Conocer la experiencia de evaluaciones estandarizadas principalmente en países de Europa, en Estados Unidos, y en América Latina, resaltando los logros y profundizando en los análisis que de ellos han hecho diversos investigadores.
- 2.- Una reflexión sobre la pertinencia de implementar una evaluación estandarizada en un programa social como Beca 18, a la luz de las críticas a este tipo de evaluaciones que realizan expertos en distintas partes del mundo. A este respecto, lo que se busca es ver qué tan relacionadas están a lo que se viene implementando desde el Estado a través de Beca 18.

Los resultados de las evaluaciones durante la educación básica como al final de la misma están relacionadas en lógica y sirven ambas a la gestión de Beca 18, en tanto los postulantes al Programa proceden de una educación básica heterogénea, aparte de sus diferencias individuales, por un lado, y la necesidad de cuidar los recursos del Estado invirtiéndolo eficientemente, por el otro.

El análisis, desde esta perspectiva, pretende dar las razones por las que puede continuar replicándose en convocatorias posteriores de Beca 18, el Examen único

Nacional, siempre que consideren la mejora continua de las estrategias implementadas.

¿Qué evaluaciones estandarizadas existen a nivel nacional e internacional?

Es precisamente a partir del uso y relevancia que pruebas como la del Programa Internacional para la Evaluación de estudiantes (PISA) han adquirido sobre las políticas educativas mundiales que se ha empezado a pensar a nivel de países en la estandarización de las evaluaciones como medida válida, impulsando de esta forma el control de la calidad y la promoción de la reforma educacional (UNESCO, 2012). Por su parte, Ravela et al. (2007) resaltan las diversas perspectivas que los países tienen respecto al rol de la educación superior y las evaluaciones a gran escala, que se pueden resumir en: una base para adoptar políticas educativas, la mejora de los sistemas educativos, y como instrumento para el mejoramiento continuo al interior de los estos. Lo que sigue a continuación es una síntesis de las evaluaciones estandarizadas en diversas partes del mundo.

Evaluaciones estandarizadas en América Latina

En América Latina el caso más emblemático en cuanto a evaluaciones estandarizadas es el de la experiencia chilena con la Prueba de Selección Universitaria (PSU) creada a partir del año 2001 con el aval del Consejo de Rectores de las Universidades Chilenas (CRUCH) y administrada por primera vez en el año 2004. Esta prueba basa su homogeneización en los contenidos académicos considerados mínimos a impartir en la enseñanza media (Contenidos mínimos Obligatorios, CMO), y en los Objetivos Fundamentales (OF) aprobados mediante Decreto Supremo N° 220 del año 1998. Es justo en el contexto de las nuevas metas educativas de ese tiempo, en que se ve la necesidad de tomar medidas que busquen ir en la dirección planteada.

Después de más de una década de ser creada, Pearson (2013) fue encargado de realizar una evaluación de la PSU con varios objetivos, entre los cuales se encontraba la validez de contenido, validez predictiva respecto al rendimiento académico en la educación superior, su estructura interna y los procedimientos de puntuación, entre otros.

Anteriormente la PSU había sido evaluada en su validez y confiabilidad respecto a las pruebas de Lenguaje y Comunicación, y Matemática, obteniendo varias observaciones, siendo algunas de las más relevantes las relacionadas al ajuste del nivel de dificultad de las pruebas al nivel de habilidad de los estudiantes, en especial en la prueba de Matemática; también se recomendó investigar el sesgo generado a partir de las diferencias en el desempeño de los grupos poblacionales en la prueba, y realizar un análisis de confiabilidad a todas las poblaciones.

La PSU consta de dos pruebas obligatorias: Lenguaje y Comunicación, Matemática. Y dos pruebas opcionales: Historia y Ciencias Sociales, Ciencias: Biología, Física y Química. Las respuestas son de selección múltiple: 5 alternativas en todos los casos. Para la calificación, se suman las respuestas correctas y se restan un cuarto de punto por cada respuesta incorrecta. El máximo puntaje son 850 puntos y el mínimo 150 puntos.

Para el proceso de selección, cada universidad decide los requerimientos por cada una de sus carreras, estableciendo pesos ponderados mínimos para las mismas. El CRUCH establece los criterios generales para los límites inferiores y superiores de ponderación, y las instituciones educativas deciden dentro de aquellos límites.

Los resultados de la PSU se utilizan también para otorgar becas y créditos de estudio, considerando puntajes mínimos para calificar como beneficiarios, además de otros requisitos como su historial académico y condición socioeconómica, por ejemplo.

Para la evaluación de la PSU, Pearson (2013) formuló una serie de hallazgos y recomendaciones en función a los mismos, a partir de los objetivos de evaluación, que a su vez contienen objetivos. En esta parte nos centraremos en mencionar los

objetivos de evaluación y un resumen general de lo encontrado por cada objetivo contenido en los mismos:

Objetivo de evaluación: Construcción y usos de la PSU.

- *Estándares de calidad, seguridad y confidencialidad respecto del desarrollo de los ítems y pruebas:* se recomienda, entre otras cosas, optar por una revisión y validación externa, es decir, con personal distinto al que participó en la construcción de los ítems. Además, capacitar y certificar a quienes redactan los ítems. Por otro lado, debe encontrarse estrategias que promuevan la motivación de los estudiantes para rendir la prueba. También, se sugiere incluir un revisor de la prueba procedente tanto del nivel educativo secundario (medio) como de la educación superior.

Una observación de especial interés es la del NEM (nota de enseñanza media) que se incluye en la postulación a la universidad. Es pues, observable para los investigadores que no se cuente con información suficiente sobre las propiedades psicométricas del puntaje NEM, dado que dicha puntuación puede o no ser comparable entre instituciones educativas y modalidades curriculares.

- *Análisis de la suficiencia de un puntaje único para la prueba de Ciencias y de los procedimientos para calcular dicho puntaje, considerando los bloques electivos de Biología, Física y Química:* se recomienda usar pruebas separadas para cada curso, con distintos propósitos. También, documentar las razones por las que se optó por un puntaje único para dicha prueba.
- *Evaluación de modelos TRI para la calibración de ítems, el desarrollo de pruebas y los propósitos de equiparación:* la Teoría de Respuesta al Ítem (TRI), que se fija más en el valor de los ítems individuales que en el resultado global de la prueba se propone como reemplazo del que usualmente se considera en la PSU para la calibración de ítems.

- *Evaluación de software y de procesos usados en el análisis estadístico y banco de ítems:* se recomienda automatizar los procesos y documentar todo el proceso de modificaciones a cada ítem.
- *Evaluación del proceso de entrega y de la claridad de la información e cuanto a los examinados y los diferentes usuarios del sistema de selección:* se recomienda clarificar la información brindada a los estudiantes respecto a los informes de entrega, así como cuidar la información respecto a los usos de resultados de la prueba, dejando en claro los fines específicos de la misma.

Objetivo de evaluación: Los estudios de validez de la PSU.

- *Estructura interna de los exámenes de la PSU:* se recomienda usar el marco teórico de la TRI y seleccionar ítems operacionales para la construcción de las pruebas.
- *Validez de contenidos:* el marco curricular sobre el que se basa la PSU no sería íntegramente considerado para dicha evaluación, de tal manera que se recomienda revisar la política que lo elige como base. Se recomienda, por tanto, tomar en cuenta variables no cognitivas y aptitudes que los estudiantes necesitarán en la universidad.
- *Análisis de trayectorias de puntajes de la PSU para las subpoblaciones en el tiempo, considerando dependencia, modo y género:* se recomienda comparar los puntajes de la prueba a través de los años y analizar la invarianza a nivel de grupos poblacionales.
- *Validez predictiva de la PSU:* se ha encontrado cierta validez predictiva de la PSU sobre el rendimiento académico en la universidad, y en especial en los primeros años. Al respecto, los puntajes en la PSU de Matemática y Ciencias Sociales y el NEM (y también el *ranking* en la secundaria) se mostraron con mayor valor predictivo que los puntajes en la PSU de Lenguaje y Comunicación e Historia y Ciencias Sociales.

Respecto a si los puntajes de la PSU deberían usarse para asignar becas, los resultados de la evaluación indican que hace falta una validación de los puntos de corte que toman como requisito para ser calificado como beneficiario.

Las recomendaciones generales para una prueba estandarizada como la PSU giran en torno a la necesidad de considerar la TRI como base para la construcción de la prueba, además de la investigación continua sobre la validación de usos e interpretaciones de la misma.

En Brasil, el Instituto Nacional de Estudios e Investigaciones Educativas Anísio Teixeira (INEP, por sus siglas en portugués) coordina, entre otros, el Examen Nacional de la Secundaria (ENEM), creado en 1998 para evaluar el estado académico de los alumnos al final de la secundaria, aplicándose a millones de alumnos, hasta establecerse como criterio de selección para universidades.

La Tabla 2 muestra cómo ha sido la evolución histórica del ENEM desde su creación.

Tabla 2. Evolución histórica del ENEM desde su implementación

Año	Participantes	Cambios
1998	157 221	Primer ENEM: voluntario, solo para autoevaluación.
2001	1 624 131	Introducción del número de identificación (ID) del colegio, igual al del censo pero diferente cada año.
2003	1 882 393	Grandes cambios en el equipo responsable debido a las elecciones presidenciales del año anterior.
2005	3 004 491	El ENEM empieza a utilizarse como forma de acceder a la enseñanza superior y como criterio para recibir becas del gobierno según el programa Universidad para Todos (ProUni) en las instituciones privadas.
2006	3 742 827	Publicación de la primera clasificación de colegios del ENEM.
2007	3 584 569	El número de identificación de los colegios (ID) pasa a ser el mismo a lo largo de los años, haciendo posible el análisis longitudinal.
2009	4 148 721	Cambios estructurales en el examen, incluida la matriz de competencias, el formato de la prueba y su duración.
2010	4 626 094	El ENEM empieza a utilizarse como certificado de finalización de la enseñanza media para cualquier ciudadano con más de 18 años. La mayoría de las universidades federales utiliza el ENEM como criterio de selección.

Fuente: Travitzki, Calero & Boto, 2014, a partir de los microdatos del Examen Nacional de Enseñanza Media (ENEM).

Si bien, como se observa en la Tabla 2, las universidades federales utilizan este examen como criterio de selección, dada su relevancia oficial como medida de rendimiento en la educación secundaria, existe evidencia de que esto último no tendría el sustento suficiente para ser considerado como tal. Es así que Travitzki, Calero & Boto (2014) señalan que, dado que no todos los colegios brasileños rinden en ENEM, al analizar los que participan, se encontró condiciones de infraestructura superiores a la media del país. También se encontró influencia del contexto en el rendimiento del examen. En sí, concluyen que los resultados de esta prueba estandarizada representan más las condiciones socioeconómicas que el mérito de los colegios, siendo un indicador más o menos válido para los más favorecidos, y no para las personas en condición de pobreza, dado su limitado valor informativo.

En Colombia existe el Instituto Colombiano para la Evaluación de la Educación – ICFES, que se encarga de velar por la calidad educativa en el país, administrando distintas pruebas durante la educación básica y media, entre ellas la llamada prueba Saber 11, como requisito para acceder a la educación superior.

El Examen de Estado de Educación Media, también conocida como prueba Saber 11 se administra desde 1965, y está dirigida a estudiantes que cursan el 11° grado de media. Incluye las pruebas de: Lectura crítica, Matemática, Ciencias Sociales y Ciencias ciudadanas, Ciencias Naturales e Inglés; y las sub-pruebas de: Razonamiento cuantitativo y Competencias ciudadanas.

Por su parte, Chica-Gómez, Galvis-Gutiérrez & Ramírez-Hassan (2010) resaltan el importante rol que tienen las variables socioeconómicas en los resultados alcanzados por los estudiantes, específicamente las referidas al ingreso económico y al nivel educativo de los padres: a medida que el nivel educativo de los padres aumenta, la probabilidad de ubicarse en un nivel bajo disminuye. Su estudio se centró en las áreas de Matemática y Lenguaje, y se encontró también que aquellos individuos con mejores condiciones socioeconómicas presentaban mayores probabilidades de alcanzar puntuaciones óptimas en la prueba.

Es importante señalar que en Colombia existen modalidades de becas que toman en cuenta aquellos estudiantes con mejores resultados en la prueba Saber 11 y recursos económicos insuficientes para acceder a la educación superior.

En República Dominicana, con el objetivo de obtener diagnósticos que sirvan para tomar decisiones a nivel macro en el plano educativo, se administran las Pruebas Nacionales a todos los alumnos de la educación básica y media para medir el logro académico adquirido. Comprenden las áreas de Lengua, Matemática, Ciencias Sociales y Ciencias de la Naturaleza. Su importancia radica en que su puntuación influye en la promoción de grado (la nota del colegio representa el 70% de la nota final, mientras que el puntaje en la Prueba Nacional, el 30%).

En México, el Instituto Nacional para la Evaluación de la Educación se encarga de la aplicación de la prueba PISA en el país. Además de esto, diseñó una metodología para crear el Examen de Calidad y Logro Educativo – EXCALE contiene las áreas de Matemática, Lenguaje, Ciencias Naturales y Ciencias Sociales. Este examen, sin embargo, al igual que la prueba PISA, se aplica sobre una muestra de estudiantes (en el caso de EXCALE, cada cuatro años) para medir el estado académico en general de la población a partir de una muestra representativa.

También existe en México la experiencia de la aplicación de una prueba estandarizada en educación media superior, llamada ENLACE, y que mide las habilidades en lectura y matemática de los estudiantes al final de la educación secundaria, aunque con resultados (en cuanto a la prueba en sí) que análisis posteriores revelan preocupantes, revelando la necesidad de realizar ajustes metodológicos en la misma ya que, como señalan de Hoyos, Espino & García (2012), existe evidencia que muestra que variables relacionadas al hogar, la institución y la escuela explican en gran medida los puntajes alcanzados por los estudiantes.

En el caso de Perú, el Examen Único Nacional realizado por primera vez en el 2016, analizado en el presente documento, corresponde a una iniciativa inédita en el país, y acaso en América Latina, dado que se da entre la culminación de la secundaria y el ingreso a la educación superior. Su implementación no corresponde a una finalidad directa de la medición de la calidad de las escuelas del país, sino como estrategia de selección de los estudiantes con mejor rendimiento académico, y recursos insuficientes, además de otras condiciones de vulnerabilidad, para acceder a una educación superior de calidad, a un programa de becas (Beca 18) basado en una visión de desarrollo basado en la formación de capital humano.

Anteriormente ha habido otras experiencias de evaluación, aunque con la diferencia de que no han surgido dentro de un programa social enfocado en educación, es decir, no han sido filtros para ser beneficiario, sino que han significado evaluaciones estandarizadas para conocer el rendimiento escolar. De esta manera, se tiene desde 1996 lo que podría denominarse “los inicios” de la evaluación estandarizada; en un segundo momento es donde se posiciona la prueba PISA como una suerte de diagnóstico de las destrezas académicas de los estudiantes de la escuela básica; por último, en un tercer momento están las evaluaciones censales desde el año 2006, aunque con problemas metodológicos que resultan en resultados poco confiables (CNE, 2013).

Aun siendo muestrales, la Tabla 3 nos da una idea de los resultados de pruebas estandarizadas durante los años 2001 y 2004, en función al porcentaje de estudiantes con nivel “suficiente”.

Tabla 3: Porcentaje de estudiantes con nivel suficiente en evaluaciones nacionales (2001-2004).

Grados	Evaluación Nacional 2001		Evaluación Nacional 2004	
	Comunicación	Matemática	Comunicación	Matemática
2.º de primaria			15,1	9,6
4.º de primaria	34,8	26,6		
6.º de primaria	9,1	7,1	12,1	7,9
3.º de secundaria			15,1	6,0
4.º de secundaria	21,3	5,2		
5.º de secundaria			9,8	2,9

Fuente: CNE (2013).

Evaluaciones estandarizadas en Estados Unidos

Desde sus inicios, la implementación de evaluaciones estandarizadas en Estados Unidos ha buscado medir el nivel académico (destrezas) alcanzado por los estudiantes.

La Evaluación Nacional del Progreso Educativo (NAEP, por sus siglas en inglés) administrada por el Centro Nacional para Estadísticas de la Educación del Departamento de Educación de Estados Unidos, desde 1969 es la más representativa al respecto, abarcando principalmente los cursos de ciencias, matemáticas, lectura y escritura, constando de dos partes: la primera es la evaluación principal y consta de preguntas primordialmente sobre los cursos anteriormente mencionados; la segunda evalúa preguntas utilizadas en anteriores evaluaciones para analizar el progreso a lo largo del tiempo. Sus resultados son comparables con los demás estados dado que se toma la misma prueba en todo el país, lo que hace de su análisis más rico en materia de propuesta de políticas educativas basadas en el rendimiento escolar, por ejemplo.

Por su parte, el examen estandarizado del Desarrollo de Educación General (GED, por sus siglas en inglés) permite obtener un certificado de suficiencia en el manejo de conocimientos a nivel de educación secundaria, siendo válido tanto en Estados Unidos como en Canadá. Este examen corresponde a aquellos ciudadanos que, al no haber completado por algún motivo la preparatoria, quieren certificar sus conocimientos de tal manera que puedan acceder a un trabajo o una carrera de educación superior o capacitación académica.

La implementación de evaluaciones estandarizadas en Estados Unidos tuvo un giro en la década de los ochenta, buscando introducir un enfoque de teoría económica basado en la rendición de cuentas, o de “gerenciar la educación”, como señalan Martínez & Prada (2012), lo que devino con el correr de los años en iniciativas llevadas a cabo como la aprobación de la ley denominada “*No Child Left Behind*” - NCLB (en español, “Ningún Niño Atrás”) en el 2002.

Para los autores, existe evidencia que pone en agenda de los decisores de políticas las consecuencias negativas que genera el evaluar mediante pruebas estandarizadas como único criterio, el progreso educativo de un estudiante. Estas se basarían en primer lugar en el exceso de confianza en la estandarización por parte del gobierno. A partir de esto, las principales consecuencias serían:

Lo que miden las pruebas presentaría un sesgo de distinta naturaleza, dado que, en primer lugar, hay limitaciones en las áreas que miden, luego está el hecho de no contar con una metodología sólida, según algunos autores, y además no toma en cuenta la diversidad cultural. También está el riesgo de manipulación de los resultados en algunas instituciones educativas. Luego, está la diferencia de estándares entre un Estado y otro, lo que implica no sólo diferencias interestatales sino en el mismo Estado, al cambiar su nivel y por ende los contenidos que implica su propio estándar de calidad.

Está el hecho además de los casos en que estudiantes de bajo rendimiento no acuden a rendir el examen. Un escenario negativo adicional pasa por el tema

económico, pues estandarizar una evaluación tiene menor costo que realizar pruebas más profundas, de mayor alcance cualitativo.

Pero la NCLB como evaluación estandarizada tiene aspectos positivos de reconocer. Una de ellas puede ser considerada respecto a la comparación entre escuelas que, mirándola de manera constructiva, puede ser un espacio para conocer las debilidades en temas académicos específicos y monitorearlas. Esto específicamente referente a las escuelas públicas. De esto se derivan lineamientos de investigación que busquen resultados óptimos a gran escala en función además a la inversión del Estado para dicha mejora.

Evaluaciones estandarizadas en Europa

Sobre las pruebas estandarizadas en países europeos, se tiene los niveles CINE 1 (educación primaria) y CINE 2 (educación secundaria) que buscan evaluar el avance en un momento clave del proceso educativo.

Por su parte, Eurydice (2010) señala que la implementación de pruebas estandarizadas difiere respecto al año de su implementación en los países europeos, siendo la década del noventa del siglo pasado en donde se incrementa significativamente su uso. Su ideación respondió a la necesidad de los países de tomar decisiones respecto a los avances académicos de los estudiantes, siendo en sus inicios implementadas para permitir la promoción de grado o para orientarlos al final de cada etapa educativa, siendo Islandia y Portugal de los primeros países en implementarlas (años 1946 y 1947, respectivamente).

Los primeros países europeos en implementar pruebas estandarizadas a gran escala (nacional) fueron Irlanda, Francia, Hungría, Suecia y Reino Unido, las mismas que se basaban ya no en resultados individuales que supeditaban el paso de un nivel a otro, sino a nivel de políticas educativas nacionales. En la década del noventa del siglo pasado las evaluaciones estandarizadas a nivel nacional fueron implementadas en España, Bélgica, Letonia, Estonia y Rumania. En la actualidad,

existen pruebas estandarizadas también en países como Dinamarca, Francia, Hungría, Portugal, Lituania, Eslovenia. Actualmente no se realizan este tipo de pruebas nacionales en Bélgica, República Checa, Grecia, Gales y Liechtenstein.

Los objetivos de las pruebas nacionales en Europa difieren en alguna medida según el país. Al respecto, Eurydice (2010) hace un recuento de las mismas, basado en tres grandes grupos:

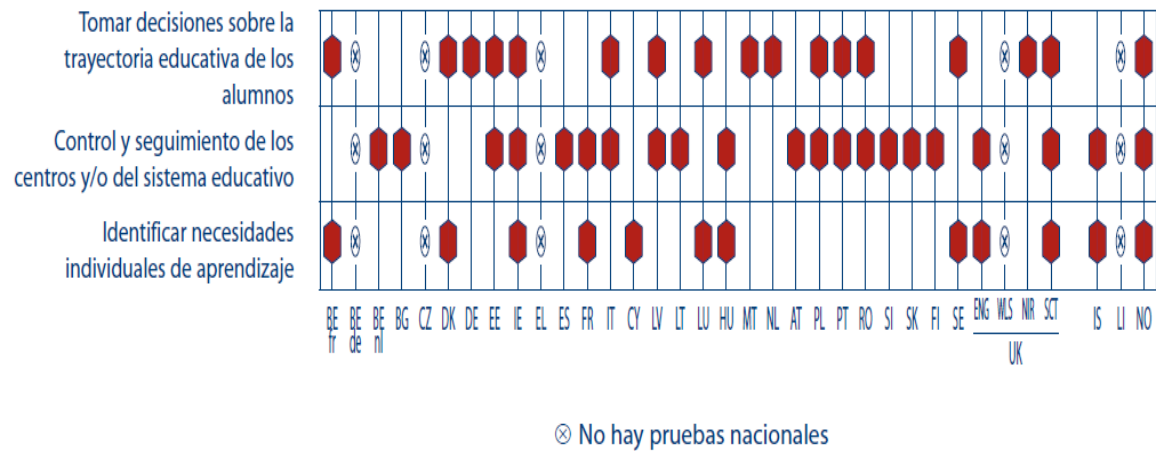
1.- Pruebas que dan cuenta del avance del alumno al final de una etapa educativa o un curso escolar, denominadas “pruebas sumativas o de evaluación del aprendizaje”, y sus resultados indican el camino a seguir a continuación durante el proceso académico.

2.- Pruebas nacionales que buscan medir la “calidad” de un centro educativo o de un sistema educativo. Tiene dos vertientes de análisis de resultados: orientado a calibrar la eficiencia de los profesores y su actividad docente, y la eficiencia de los sistemas educativos en concreto.

3.- Pruebas nacionales orientadas a los resultados individuales del alumno, centrado en conseguir un panorama acertado de sus necesidades específicas para tomar medidas respecto a las mismas, es decir una “evaluación para el aprendizaje”.

Por supuesto, estas categorías no son exclusivas, dado que los resultados de las evaluaciones en cada país pueden tener varios usos, encuadrándose muchas veces en más de una categoría, como muestra el Gráfico 1.

Gráfico 1: Objetivos de las pruebas estandarizadas nacionales en Europa.



Nota: códigos de países: ver Anexo 1.

Fuente: Eurydice (2010).

Como se observa, son pocos los países europeos en los que no hay pruebas nacionales. Como se indicó anteriormente, éstos son: Bélgica (comunidad germanófono), República Checa, Grecia, Gales y Liechtenstein. Por otro lado, son pocos los países donde se toman en cuenta los tres objetivos principales de las evaluaciones estandarizadas nacionales: Irlanda, Escocia y Noruega. También, el objetivo que se toma en cuenta mayormente es el de control y seguimiento de los centros y/o del sistema educativo, y el de menor atención, el de identificar las necesidades individuales de aprendizaje. Respecto a los elementos que componen las pruebas (materias o cursos), se tiene un antecedente: en la Recomendación del Parlamento Europeo y del Consejo, de 18 de diciembre de 2006, sobre las competencias clave para el aprendizaje permanente, se consideran ocho de éstas: comunicación en lengua materna; comunicación en lenguas extranjeras; competencia matemática y competencias básicas en ciencia y tecnología, competencia digital; aprender a aprender; competencias sociales y cívicas; sentido de la iniciativa y el espíritu de empresa; conciencia y expresión culturales. De ésta, sólo tres se relacionan a lo evaluado en las pruebas estandarizadas. El uso de los resultados también se cuida en relación a la publicación de los mismos. No todos los países publican estos por cada centro de estudios.

Así, en gran parte de los países los resultados sirven para conocer en qué puntos necesita refuerzo el alumno. Como puede verse por lo anterior, existiría un sesgo en este “diagnóstico” dado que no mediría la verdadera magnitud de la necesidad. Es importante señalar el cuidado que debe tenerse sobre el uso de un examen estandarizado a partir de esta última consideración.

Evaluaciones estandarizadas en Asia

En el caso de China, el *gaokao* o examen nacional estandarizado para acceder a la educación superior es criticado no sólo por la presión que se ejerce a través de las familias sobre los postulantes, dado que se ha construido sobre éste una imagen de gran importancia, sino por la diferencia en la selectividad. Al respecto, Liu (2013) observa que aunque la selección no se encuentra significativamente relacionada a un criterio socioeconómico, sí existen otros factores en relación, como sociodemográficos, esto es, de orden cultural que la afectan.

¿Es pertinente una evaluación estandarizada como el Examen Único Nacional en un programa social como Beca 18?

La estandarización de las habilidades cognitivas, o dígame académicas, de los estudiantes ha recibido muchas críticas, algunas de las cuales se han mencionado en párrafos precedentes. En lo que sigue, se hará una síntesis de las principales y se dará cuenta de qué tanto están relacionadas las mismas a la realidad de un programa social como Beca 18, cuyos estandartes de talento académico e inclusión social han ido dirigiendo su gestión a los largo de los años desde su creación. Lo que se pretende es analizar si el lado negativo de la estandarización puede aplicarse a la lógica de Beca 18, en tanto ha buscado en este primer examen, uniformizar la medida del talento académico en vías de realizar una selección más idónea en beneficio de los postulantes (mejor adaptación académica a la educación superior) y el Estado Peruano (garantizar la eficiencia de la inversión).

Si bien en sociedades homogéneas la estandarización de una prueba puede dar resultados, en sociedades cuyo sistema educativo se caracteriza por su heterogeneidad, el uso de una evaluación estandarizada puede ampliar la discriminación y, entre otras cosas, puede surgir el “reduccionismo instrumental”, es decir, el entender la calidad sólo a partir de los aspectos que las pruebas miden (UNESCO & LLECE, 2008). Por esto es que se ha visto necesario, al idear una evaluación estandarizada, que sus propósitos vayan en función a mejorar las oportunidades de aprendizaje (UNESCO, 2012).

Por su parte, Treviño-Villarreal (2006) observa que existe una tensión entre lo que mide una prueba estandarizada y la diversidad cultural al interior de un país, en especial en las comunidades indígenas, y en especial, recalcamos, una evaluación que tiene importantes consecuencias para el futuro del estudiante.

Arancibia (2015) incide en la no pertinencia de la estandarización de las pruebas académicas nacionales (e internacionales, además) dado que no toman en cuenta la diversidad de los sujetos individuales, de las inteligencias, la diversidad geográfica, las diversidades étnico culturales, por lo que debe enfocarse en sistemas de evaluación contextualizados.

Por otro lado, Benavidez-Ormaza (2010) considera que las evaluaciones nacionales e internacionales sí reflejan la calidad de la educación y se pueden sacar conclusiones válidas a partir de sus resultados. La autora sostiene, además, que la inversión en la medición del logro académico de los estudiantes, en países de bajos recursos, debe hacer especial énfasis en la calidad técnica de la propuesta y los ciudadanos deben conocer que dichas evaluaciones pueden tener impacto política a nivel de objetivos educativos nacionales.

En el Perú, el Consejo Nacional de Educación (2013) manifiesta que es mucho lo que se ha avanzado en cuanto a medición del logro académico en la escuela básica en el país desde la implementación de pruebas estandarizadas desde la última década del siglo pasado, sin embargo, esto debe ser basado en un modelo de bajas

consecuencias, en el que a partir de un currículo nacional se mide lo que el estudiante ha adquirido en materia de aprendizaje.

Anteriormente, se han presentado documentos que dan cuenta de la crítica a las evaluaciones estandarizadas, las mismas que no tendrían sustento suficiente para ser tomadas en consideración como evaluadoras de la calidad, como se muestra en la Revista Docencia (2003). Allí, se comentan algunas posturas críticas sobre las evaluaciones estandarizadas en la realidad chilena. Comenta Juan Cassassus, especialista de la UNESCO, que uno de los problemas en un sistema educativo de este tipo es que se pretende “integrar homogeneizando”, generando un currículo único que no es sensible a las diferencias de etnia, por ejemplo. Otro punto débil, a decir de Cassassus, es la idea de “controlar los privilegios” asumiendo una postura de emitir y dirigir los criterios de acceso a los mismos. Destacan además que los alumnos de escuelas de bajos recursos económicos representan una realidad concreta de necesidad, que va más allá de lo económico, pues resultan marginadas e incluso estigmatizadas por los mismos pobladores.

También, otros especialistas como Calderón, Mardones, Muñoz & Valenzuela (2003, citados en Revista Docencia, 2003) son de la opinión de lo inviable del supuesto de la homogeneidad de la educación en las escuelas que subyace al enfoque de estandarización de la calidad, dado que los estudiantes tienen diferencias individuales en contexto cultural, social, geográfico, además de ritmos de aprendizaje distintos.

Por su parte, Barrenechea (2010) presenta lo que llama “seis limitaciones clave que tienen las pruebas estandarizadas”, que son:

- 1.- Las diferencias entre las inteligencias múltiples y lo que miden las pruebas estandarizadas, indicando que “no pareciera existir una relación coherente entre el tipo de enseñanza que se pretende que los alumnos reciban y el tipo de evaluación que se les impone”, y resaltando la necesidad de enfocarse en las fortalezas del educando, que van más allá de una homogeneización de la educación, siendo que

los test estandarizados sólo miden dos tipos de inteligencia: lógico-matemática y lingüística.

2.- Las diferencias entre el currículum oculto y el real, que incide en la diferencia entre el currículo tal como se presenta, y el que se enseña en clase, es decir la enseñanza en concreto. Es decir hay un margen de diferencia donde reside la adaptación del currículo a la clase y el método de enseñanza. En relación a esto, una evaluación estandarizada no contemplaría las diferencias ni en lo que saben ni en lo que se les ha enseñado en la escuela.

3.- El riesgo de enseñar para la evaluación, que implica varias cosas en juego, como el prestigio de la escuela, el financiamiento que recibe desde distintas instituciones, sea pública o privada, etcétera. A partir de esto, la escuela busca implementar estrategias que den resultados positivos frente a la sociedad, como consecuencia de la presión de alcanzar el estándar de calidad esperado. En consecuencia, se enseña al estudiante a preocuparse más por la nota alcanzada que por lo que se aprende, siguiendo la misma lógica de las evaluaciones estandarizadas.

4.- Incentivos para hacer trampas en los exámenes, relacionado al punto anterior y referido específicamente a los medios ilícitos para lograr resultados positivos en las evaluaciones estandarizadas, surgidos en alguna medida de la presión de la publicidad y la consiguiente necesidad de reconocimiento de la institución (o mas bien de las personas que las guían, es decir, las encargadas de tomar las decisiones).

5.- Las evaluaciones estandarizadas en relación a los contextos sociales, siendo el norte de esta crítica las diferencias de los distintos grupos sociales, en lengua materna, por ejemplo, donde aquellos que tienen una lengua distinta a la usada en la evaluación estandarizada, tendrían menores posibilidades de logro respecto a los demás. Dado esto, se resalta la falta de una metodología que tome en cuenta dichas diferencias.

6.- Predecir el éxito en el mercado laboral de los estudiantes, cuya crítica toma en consideración el hecho de la validez de las evaluaciones estandarizadas como

determinantes de la aptitud y capacidad de los alumnos durante los estudios superiores y luego, en el mercado laboral, pues dichas evaluaciones no toman en cuenta la medición de destrezas específica que requiere un empleo, entrando a este respecto la discusión del tiempo requerido para resolver un problema (distinto al de un empleo en la vida real) y la rigidez de los estándares que caracterizan la lógica de la evaluación.

Es de resaltar la diferencia curricular de las diversas escuelas, estableciendo una diferencia en contenidos a los que estuvieron expuestos los alumnos. Dicha lógica curricular no tuvo en cuenta una estandarización inter escolar, por lo que afrontar una pretensión de homogenización mediante una evaluación al final de la secundaria responde a una escisión entre las dos lógicas: durante y al final de la escuela.

Las críticas, a decir de Barrenechea (2010) no deben verse como una propuesta de eliminación de las evaluaciones estandarizadas, porque en buena medida, al fin de cuentas, han dado datos sobre las deficiencias y fortalezas del sistema educativo. Más bien, lo que debería hacerse es no tomarlas como la única medida que hace posible el acceso a la educación superior.

Como puede observarse, las evaluaciones estandarizadas han recibido diversas críticas por la misma naturaleza “estandarizadora” que promueve. Al respecto, se ha dicho, o más bien recordado, que la crítica a nivel internacional indica que las evaluaciones estandarizadas no toman en cuenta la diversidad cultural del evaluado, ni sus capacidades diferenciadas, lo que resulta en una suerte de discriminación. Por otro lado, dan mala imagen a la educación pública y no consideran las necesidades de los alumnos. Por último, en su afán de promover la competencia olvidan todo lo referente a la colaboración interinstitucional. Es más, se ha visto en una prueba estandarizada al final de la secundaria como la PSU en Chile que el 43% de los más pobres no alcanzan el puntaje mínimo para ingresar a

una universidad, mientras que sólo al 5% de los jóvenes de los estratos socioeconómicos más altos les sucede lo mismo (Revista Docencia, 2009).

En el caso del Programa Beca 18, sin embargo, dicha evaluación estandarizada sustenta su pertinencia en función a:

1.- La población evaluada, en su totalidad, presente condiciones de exclusión, lo que corresponde a un enfoque de necesidades, llamado también inclusión social. De esta forma, es justo a aquellos postulantes más necesitados que se les evalúa.

2.- La diferencia que se hace entre los resultados de la población, esto es, consideran los puntajes alcanzados dentro de cada grupo objetivo, según modalidad.

3.- Dicha estandarización responde a una necesidad de proteger la inversión del Estado, seleccionando a los postulantes con mejores resultados en la evaluación, la misma que está relacionada, en contenidos, a los cursos de la carrera de educación superior elegida.

4.- En el caso de la lengua materna, la educación superior ofrecida en el país es predominantemente en español, por lo que los postulantes a dichas instituciones son evaluados en dicho idioma. Resultaría contraproducente diseñar una prueba estandarizada para evaluarlos en una lengua distinta al español, para luego estudiar en esta última. El tema de este punto es más de fondo. Existe sí la necesidad de establecer políticas de inclusión educativa en materia de lengua, dado que nuestro país es muy rico en diversidad de etnias, aunque esta problemática excede el alcance del Programa.

A decir de Barrenechea (2010), “al atar el resultado de las evaluaciones estandarizadas a cuestiones como el financiamiento que reciben las escuelas (tal como sucede en los Estados Unidos) o la posibilidad de acceder al circuito de educación superior (como sucede en países como Chile, Holanda y Nueva Zelanda) lo que se está haciendo es perpetuar esas inequidades”, refiriéndose al contexto social propia de cada población que, dentro de un país, es diversa.

En el caso del Perú, esta diversidad es especialmente significativa, y en lo que corresponde a Beca 18, la evaluación estandarizada presente un diseño metodológico afín a las necesidades de cada población objetivo, esto es, obtiene su presupuesto basado en una lógica clara, precisa, cuya dependencia pasa de conseguir resultados dirigidos específicamente a tender la diversidad poblacional, como es por ejemplo modalidades de beca específicas para postulantes procedentes de Comunidades Nativas Amazónicas, para adolescentes en situación de abandono o tutelados por el Estado, para pobladores del Valle de los ríos Apurímac, Ene y Mantaro (VRAEM), para licenciados de las Fuerzas Armadas, para víctimas de la violencia habida en el país desde el año 1980, para pobladores procedentes del Huallaga, para quienes desean formarse en la carrera de Educación Intercultural Bilingüe, y para la población en general en condición de pobreza y con un rendimiento académico mínimo, establecido por el Programa.

3. ANÁLISIS ESTADÍSTICO

Las siguientes tablas resumen describen la naturaleza socioeconómica de los postulantes.

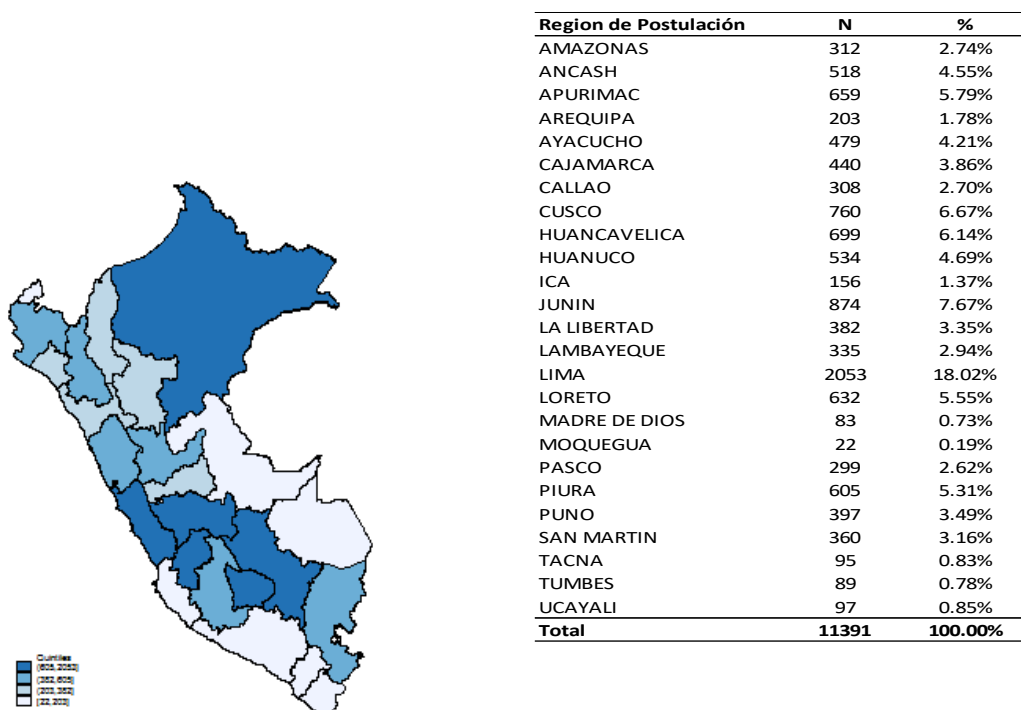
Tabla 4. Postulantes por modalidad y sexo

Beca	Mujeres	Hombres	% Mujeres	% Hombres
Ordinaria	4399	3317	39%	29%
Albergue	136	89	1%	1%
Repared	210	254	2%	2%
VRAEM	476	420	4%	4%
Huallaga	168	169	1%	1%
FFAA	173	635	2%	6%
EIB	276	143	2%	1%
CNA	220	306	2%	3%
Total	6058	5333	53%	47%

Fuente: PRONABEC. **Elaboración propia.**

La Tabla 4 muestra la preponderancia de postulantes a la modalidad Ordinaria con el 68% del total de postulantes que rindieron el examen estandarizado. Por su parte, las modalidades con menos postulantes son las de Albergue y Huallaga con solo 4% de los postulantes. Esto sucede debido a que las demás clasificaciones abarcan a un público objetivo mucho más específico (razón por la cual se otorgan cuotas específicas para dichas modalidades).

Gráfico 2. Distribución de postulantes por departamento



Fuente: PRONABEC. Elaboración propia.

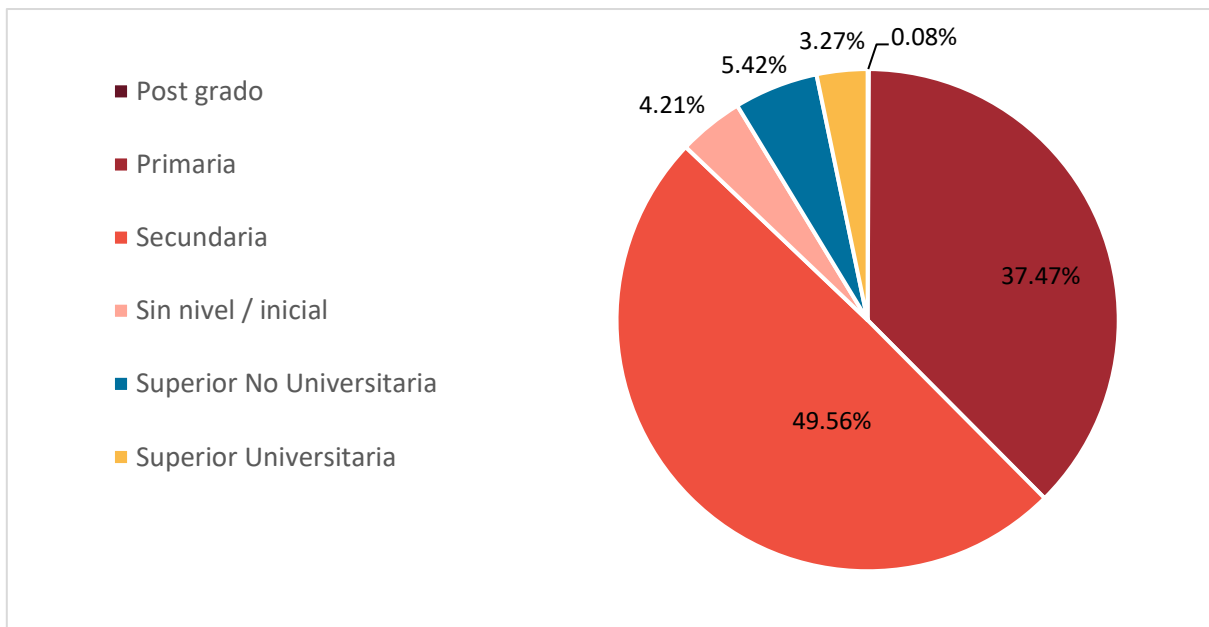
Tabla 5. Postulantes por zona de residencia

Zona	N	%
Lima Metropolitana	1730	15.19%
Resto Urbano	6372	55.94%
Rural	3289	28.87%
Total	11391	100.00%

Fuente: PRONABEC. Elaboración propia.

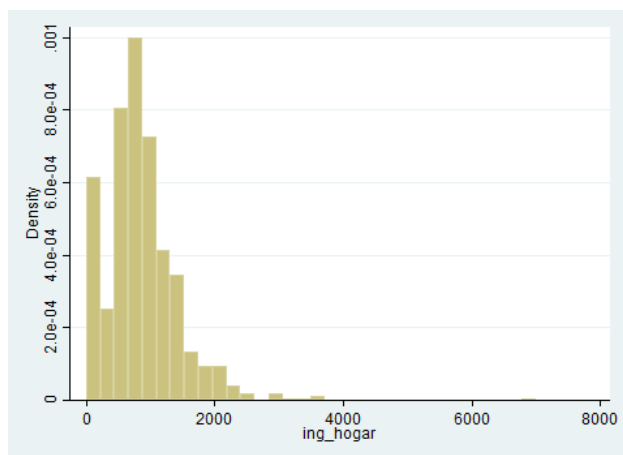
A nivel departamental, la mayoría de las postulantes proviene desde la zona central (costa y sierra). Lima, Junín, Huancavelica, Cuzco, Apurímac y Loreto tienen la mayor concentración de postulantes (casi el 50% de todos los postulantes residen en esos departamentos). Lima tiene el 18% del total de la muestra, en contraposición con Moquegua que solo tiene el 0.19% (22 postulantes). Adicionalmente, la Tabla 5 expone la concentración de postulantes en el ámbito urbano, con lo que podría decirse que la cobertura a la zona rural y la estrategia de descentralización del examen necesita ser mejorada.

Gráfico 3. Nivel educativo del jefe de hogar

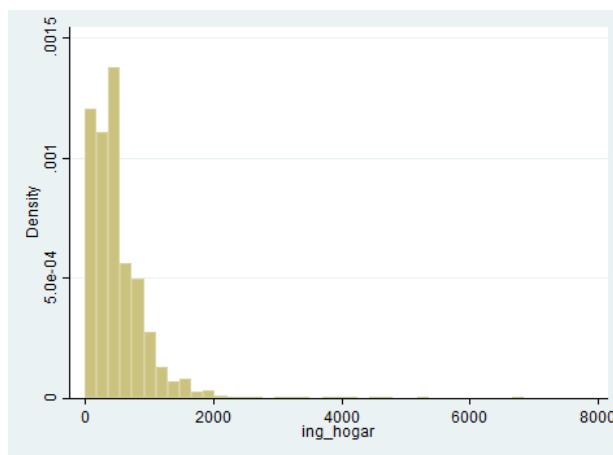


Fuente: PRONABEC. Elaboración propia.

Histograma 1. Ingresos en Lima Metropolitana

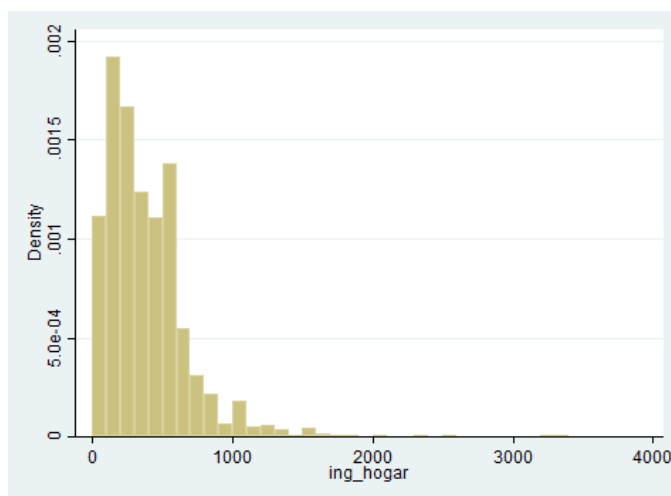


Histograma 2. Ingresos en el Resto Urbano



Fuente: PRONABEC. Elaboración propia.

Histograma 3. Ingresos en el ámbito Rural



Fuente: PRONABEC. Elaboración propia.

El nivel educativo del jefe del hogar se concentra en primaria y secundaria. El 49.6% de los hogares de los postulantes tienen un jefe de hogar con secundaria como máximo nivel educativo alcanzado y el 37.47% solo alcanzó educación primaria. Como describen Glewwe y Miguel (2008) el set de activos acumulados (en

este caso el capital humano) puede influir directamente sobre el desarrollo cognitivo de los miembros dependientes en el hogar. El Gráfico 3 muestra una focalización, aparentemente, adecuada para aquellos hogares cuyos jefes no tuvieron la oportunidad de acceder a educación superior. Como característica complementaria al nivel educativo del jefe del hogar, se tienen los Histogramas 1, 2 y 3 que muestran la distribución del nivel de ingreso conjunto (ingreso de padres e ingreso de madres). Los Histogramas de frecuencia muestran concentraciones distintas para cada ámbito geográfico. El ingreso del hogar en Lima metropolitana tiene una media mayor y un sesgo menos pronunciado hacia la izquierda. Lo cual se encuentra correlacionado muy fuertemente con el nivel educativo (Saavedra, 2007; Yamada, 2010).

Esto puede otorgar una primera idea de las posibles privaciones a las que el postulante puede estar expuesto ya que el poder adquisitivo de la población con bajo nivel educativo tiende a ser bastante insuficiente para proveer ambientes adecuados para el desarrollo cognitivo (reflejado en el resultado del examen). Estas privaciones, como se mostrara más adelante, pueden incidir negativamente en el desarrollo cognitivo del individuo condicionando su desempeño en la educación superior. Después de este resumen de las características personales de los estudiantes y de los jefes, se muestra la caracterización de la vivienda de los postulantes mediante las variables de combustible de cocina, provisión de servicios básicos, hacinamiento y calidad de los materiales de construcción de la vivienda.

Gráfico 4. Combustible de cocina

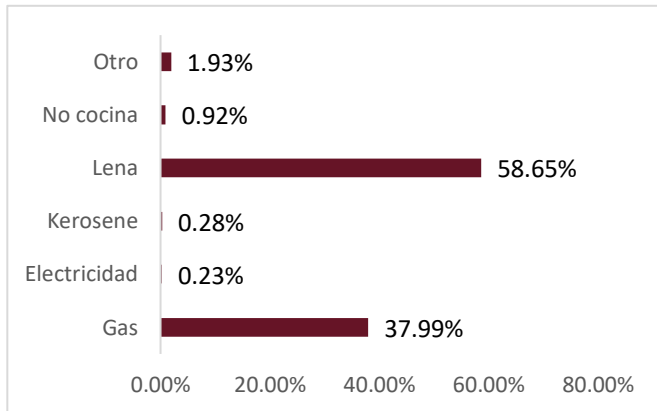
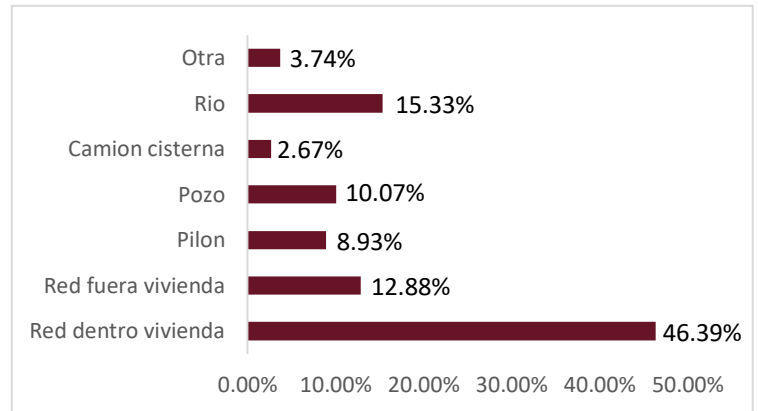


Gráfico 5. Abastecimiento de agua



Fuente: PRONABEC. Elaboración propia.

Gráfico 6. Sistema de desagüe

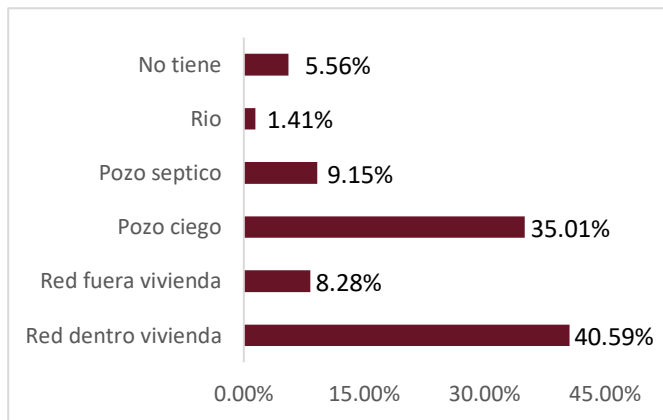
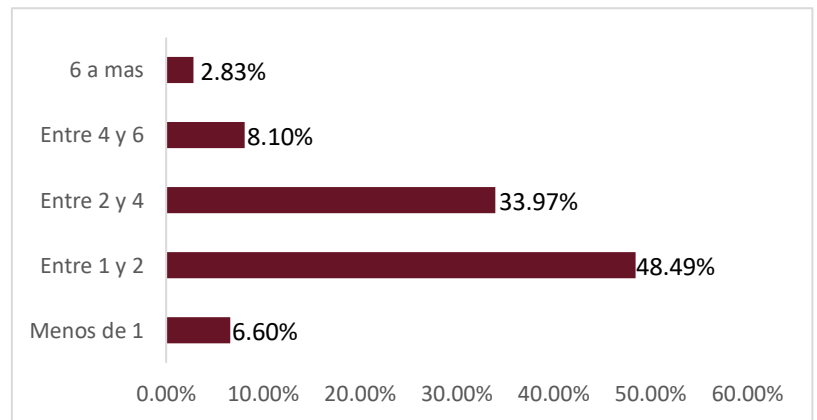


Gráfico 7. Hacinamiento



Fuente: PRONABEC. Elaboración propia.

Gráfico 8. Material de los pisos

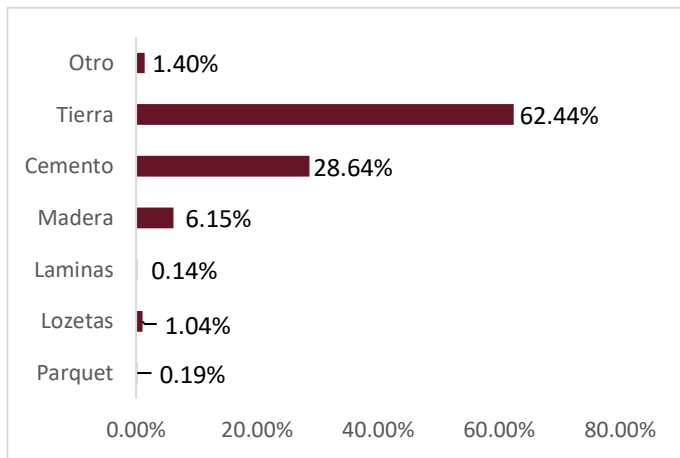
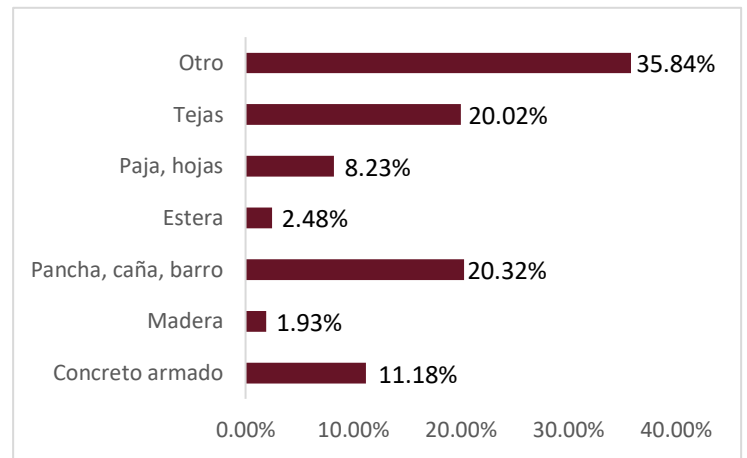
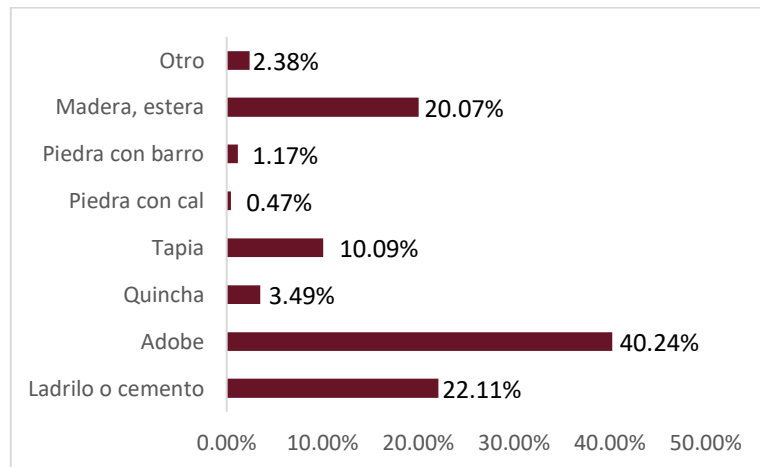


Gráfico 9. Material de los techos



Fuente: PRONABEC. **Elaboración propia.**

Gráfico 10. Material de las paredes

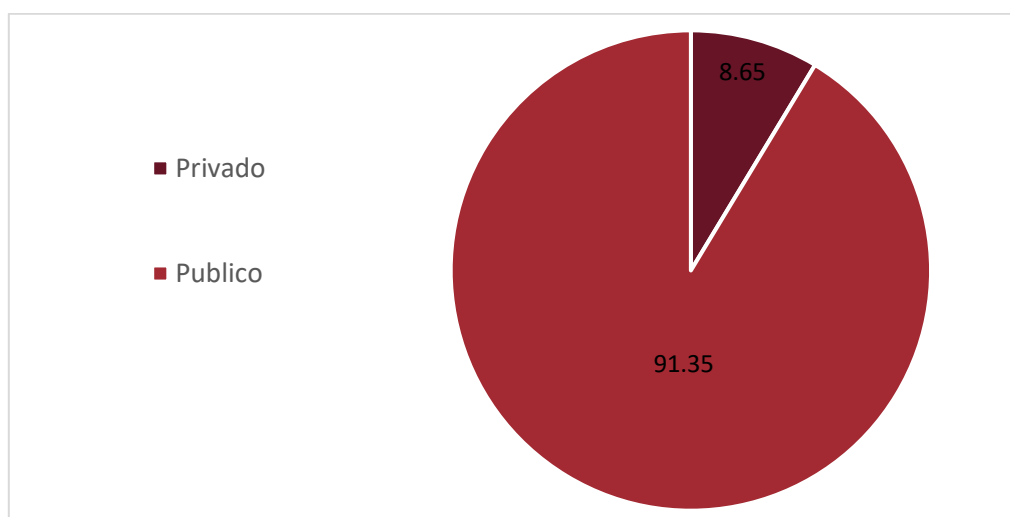


Fuente: PRONABEC. **Elaboración propia.**

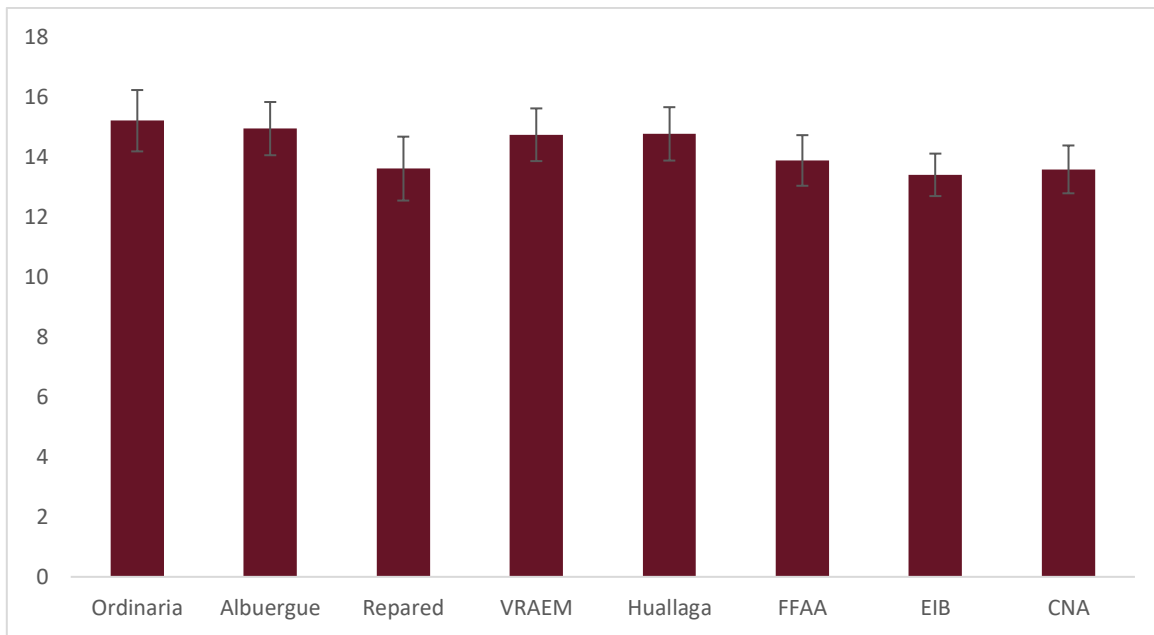
Las posibles deficiencias que los becarios pueden experimentar dentro de la vivienda se concentran en el sistema de eliminación de excretas, las características de la construcción de la misma y en el combustible que utilizan para cocinar. Como

el Gráfico 6 señala, casi el 40% reporta tener pozo ciego como sistema de desagüe, el Gráfico 4 muestra que más del 60% de los hogares utiliza leña para cocinar los alimentos. Por su parte los Gráficos 8, 9 y 10 evidencian que los materiales predominantes para la construcción de los pisos, techos y paredes son tierra (62.44%), otro (35.84%) y adobe (40.24%) respectivamente, aunque la concentración no es tan pronunciada para el caso del material de los techos, lo cual también puede ser indicio de una estrategia de identificación adecuada en cuanto a la población objetivo. Sin embargo, esto también señala que, a pesar de la evolución en los indicadores de pobreza, todavía existe parte de la población que experimenta privaciones bastante fuertes. Esto se da aun cuando la mayoría de los postulantes proviene desde zonas urbanas donde se esperaría que esta clase de privaciones (combustible de cocina inadecuado, materiales de construcción deficientes, entre otros) se encuentren en rangos mínimos, lo cual refleja que incluso dentro de un ámbito geográfico definido como “urbano” todavía se encuentra heterogeneidad y esta heterogeneidad es capturada por la definición de “población objetivo” del programa. La última parte de esta sección se concentra en un breve análisis de la oferta educativa concerniente a la muestra de becarios.

Gráfico 11. Gestión del colegio del postulante



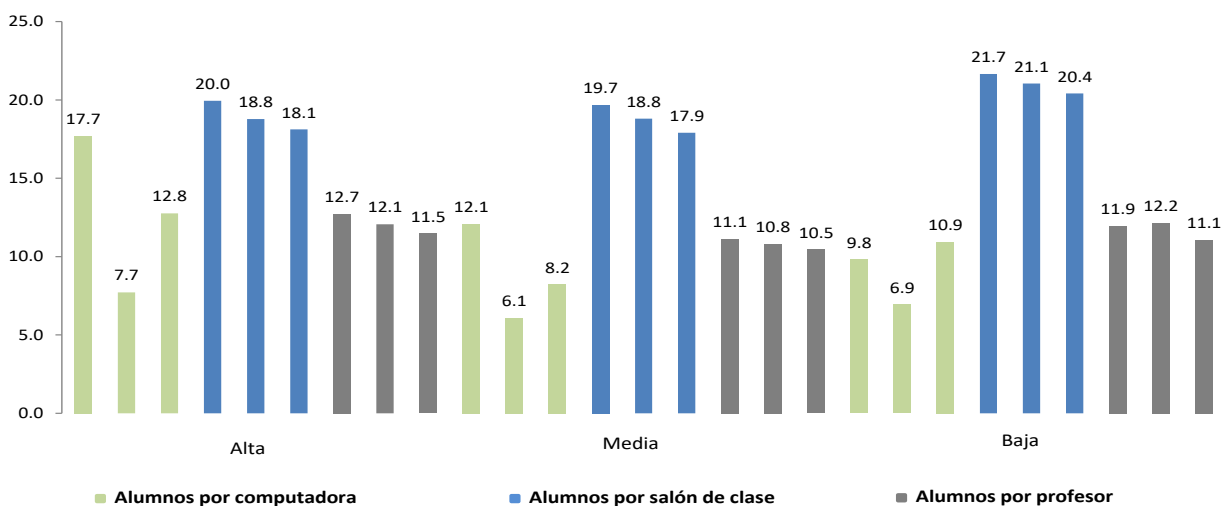
Fuente: PRONABEC. Elaboración propia.

Gráfico 12. Promedio de secundaria por modalidad

Fuente: PRONABEC. **Elaboración propia.**

Más del 90% de los postulantes provienen de colegios públicos y el promedio más alto de secundaria se encuentra en los postulantes de la Modalidad Ordinaria. Como se mencionó líneas arriba, todas estas características pueden ser determinantes sobre el desarrollo cognitivo, resultado que se ha medido mediante la calificación que se obtuvo al momento de rendir la evaluación estandarizada. Como gráficos complementarios se presentan las características de los colegios divididas por la prioridad distrital en donde se encuentran.

Gráfico 13. Alumnos por computadora, tamaño de clase y alumnos por profesor (2011-2013)



Fuente: División de Estadística MINEDU, ESCALE. **Elaboración propia.**

Gráfico 14. Profesores Titulados (%)

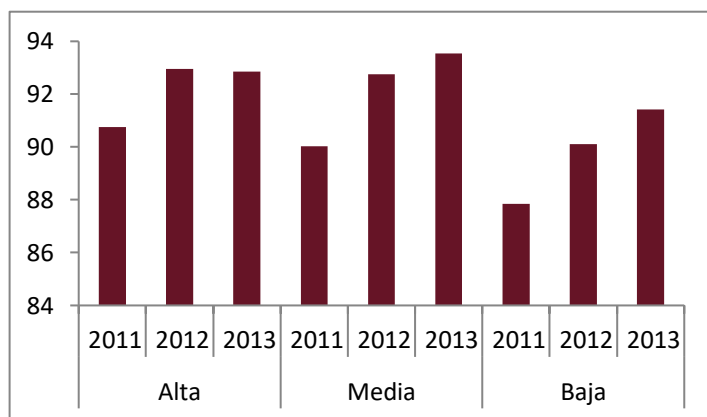
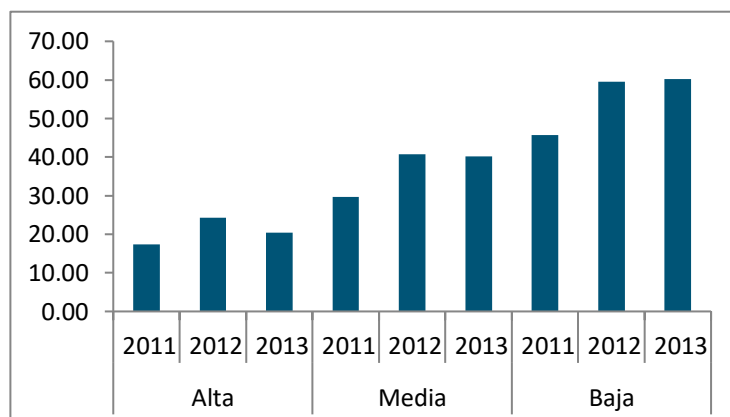


Gráfico 15. Escuelas con acceso a Internet (%)



Fuente: División de Estadística MINEDU, ESCALE. **Elaboración propia.**

Estos últimos gráficos revelan las diferencias entre la oferta educativa y las potenciales deficiencias estructurales de la oferta educativa peruana. Aunque el porcentaje de profesores con titulación sea mayor para los distritos de alta prioridad, el acceso a servicios como internet, tamaño de clase y alumnos por profesor empeora mientras la prioridad distrital aumenta. Lo cual, podría significar que la

heterogeneidad de los resultados en esta evaluación no solo se encuentran condicionados por los activos del hogar y el ambiente que pueden propiciar (óptimo o no) para el desarrollo cognitivo, si no por las diferencias en la calidad de la oferta educativa. La infraestructura escolar y la calidad del contexto en el cual se forma el estudiante puede ser determinante para el desempeño correcto en las siguientes fases educativas y, por consiguiente, en la etapa laboral mediante la cual se pueden definir los estándares de la calidad de vida futura.

4. MODELO ECONÓMÉRICO

Esta parte del documento apunta a estimar una función de producción de conocimiento de acuerdo a los trabajos de Glewwe y Miguel (2008), Artega y Glewwe (2014), Heckman (2006) y Cunha y Heckman (2007). La función de producción tiene la siguiente form

$$A_{i2} = A_2(H_{i2}, H_{i1}, S_{i2}, SY_{i2}, h_{i2}, h_{i1}, f_i, \mu_{i0})$$

El resultado cognitivo en el periodo 2 del individuo i es una función que depende de inputs contemporáneos y de inputs pasados. Así, H_{i2} y H_{i1} son el conjunto de factores educativos que provienen del hogar en el periodo 1 y 2 respectivamente (el modelo puede extenderse para más periodos), S_{i2} representa los inputs del colegio en el periodo 2, SY_{i2} los años de escolaridad h_{i2} , h_{i1} el estado de salud del individuo f_i las influencias externas predeterminadas y μ_{i0} la habilidad innata. Asumiendo un grado de separabilidad relativo en la función de producción de conocimiento, podemos expresar una función lineal en parámetros estimable.

$$A_{i2} = H_{i2}'\gamma_1 + H_{i1}'\gamma_2 + S_{i2}'\phi_1 + h_{i2}\varphi_1 + h_{i1}\varphi_2 + f_i'(2) + \mu_{i0}\beta(2)$$

De este modo, las variables capturadas por la base de datos pueden incluirse (otras deben ser incorporadas desde otras bases de datos como el censo escolar) y estimar un modelo multinivel (Hox, 2010) que considera efectos fijos a nivel de los distintos ámbitos en los que se encuentra el individuo.

Para el análisis inicial, se necesita evaluar si existen diferencias en el desarrollo cognitivo (medido por el resultado de la evaluación) que pueden estar correlacionadas con las diferencias entre las características de la vivienda, nivel socioeconómico y colegio de donde provienen los postulantes.

Tabla 6. Resultado del examen por modalidad

Modalidad beca	N	Media	Mediana	Desv. Est.
Ordinaria	7716	48.23	46.33	14.34
Albergue	225	45.66	45.67	11.90
Repared	464	44.45	42.00	13.35
VRAEM	896	46.34	44.33	12.60
Huallaga	337	43.28	41.00	11.88
FFAA	808	40.26	39.67	11.07
EIB	419	37.47	37.67	8.56
CNA	526	35.41	34.67	9.64
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 7. Resultado a nivel de ámbito geográfico.

Zona de residencia	N	Media	Mediana	Desv. Est.
Lima Metropolitana	1730	56.36	55.67	15.14
Resto Urbano	6372	46.12	44.67	13.36
Rural	3289	40.93	39.67	11.33
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 8. Resultado por sexo

Sexo	N	Media	Mediana	Desv. Est.
Hombre	5333	45.96	43.67	14.73
Mujer	6058	46.37	44.67	13.28
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 9. Resultado por nivel educativo del jefe

Nivel educativo jefe	N	Media	Mediana	Desv. Est.
Primaria	4196	42.56	41.33	11.89
Secundaria	5550	47.63	45.84	14.39
Sin nivel / inicial	471	41.93	41.33	11.55
Superior	No			
Universitaria	607	55.41	55.33	15.08
Superior Universitaria	366	55.84	56.00	15.78
Post grado	9	53.78	52.33	14.35
Total	11199	46.19	44.33	14.00

Fuente: PRONABEC. Elaboración propia.

La distribución de los resultados teniendo en cuenta cada una de las clasificaciones mostradas en las Tablas 6-9 muestra que podría haber heterogeneidad que viene

explicada por el sexo, el nivel educativo del jefe y el ámbito geográfico de donde proviene el postulante. A simple vista, las mujeres obtienen un mejor resultado en la nota global; del mismo modo, la Tabla 7 señala que aquellos postulantes que provienen de Lima Metropolitana obtienen entre 10 y 16 puntos más que los que provienen del resto urbano o del ámbito rural. Como es de esperar, el nivel educativo del jefe del hogar (Tabla 9) también muestra resultados consistentes con la literatura mencionada. Los desempeños más altos se concentran en aquellos postulantes que residen en un hogar cuyo jefe tiene educación superior (universitaria o no universitaria). Esto es consistente con los hallazgos de Artega y Glewwe (2014), ya que se produce un efecto traspaso del capital humano desde los padres (o jefes de hogar) a los hijos (o miembros dependientes del hogar).

Tabla 10. Resultado por provisión de agua

Abastecimiento de agua	N	Media	Mediana	Desv. Est.
Red dentro vivienda	5284	49.16	47.33	14.46
Red fuera vivienda	1467	45.18	43.33	13.33
Pilon	1017	44.61	42.67	13.18
Pozo	1147	41.89	40.67	12.21
Camion cisterna	304	50.96	49.33	14.24
Rio	1746	40.47	39.67	11.54
Otra	426	47.88	46.67	13.88
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 11. Resultado por provisión de desagüe

Desagüe	N	Media	Mediana	Desv. Est.
Red dentro vivienda	4624	50.73	49.00	14.68
Red fuera vivienda	943	46.33	44.33	13.13
Pozo ciego	3988	42.34	41.00	12.30
Pozo septico	1042	43.26	41.33	12.58
Rio	161	42.79	41.33	13.31
No tiene	633	42.56	41.67	12.40
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 12. Resultado por combustible de cocina

Combustible de cocina	N	Media	Mediana	Desv. Est.
Gas	4219	52.53	51.00	14.89
Electricidad	26	42.46	39.50	15.87
Kerosene	31	50.32	49.33	17.02
Lena	6513	42.03	41.00	11.62
No cocina	102	44.63	42.50	15.39
Otro	214	49.53	47.17	14.59
Total	11105	46.21	44.33	14.03

Fuente: PRONABEC. Elaboración propia.

Tabla 13. Resultado por hacinamiento

Hacinamiento	N	Media	Mediana	Desv. Est.
Menos de 1	752	44.91	43.00	14.08
Entre 1 y 2	5524	45.81	44.00	13.90
Entre 2 y 4	3870	47.19	45.33	14.23
Entre 4 y 6	923	46.53	44.00	13.58
6 a mas	322	42.23	41.67	11.75
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 14. Resultado por material de las paredes

Material de las paredes	N	Media	Mediana	Desv. Est.
Ladrillo o cemento	2518	52.59	50.67	15.22
Adobe	4584	44.52	43.00	12.73
Quincha	397	41.91	41.00	13.20
Tapia	1149	43.59	42.33	11.86
Piedra con cal	53	52.47	49.33	15.15
Piedra con barro	133	44.15	41.33	13.28
Madera, estera	2286	44.35	42.67	13.79
Otro	271	47.00	45.67	14.65
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 15. Resultado por material de los pisos

Material de los pisos	N	Media	Mediana	Desv. Est.
Parquet	22	48.97	43.84	14.76
Losetas	119	59.43	58.67	14.14
Laminas	16	50.62	49.17	12.85
Madera	701	43.99	42.00	14.70
Cemento	3262	52.29	50.67	15.16
Tierra	7112	43.28	42.00	12.20
Otro	159	49.32	48.33	13.47
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

Tabla 16. Resultado por material de los techos

Material de los techos	N	Media	Mediana	Desv. Est.
Concreto armado	1273	54.58	52.67	15.65
Madera	220	50.69	50.17	15.50
Pancha, caña, barro	2315	43.86	42.67	12.39
Estera	282	46.34	44.67	13.04
Paja, hojas	938	37.18	37.00	10.28
Tejas	2281	44.43	42.67	12.94
Otro	4082	47.66	46.00	13.82
Total	11391	46.18	44.33	13.98

Fuente: PRONABEC. Elaboración propia.

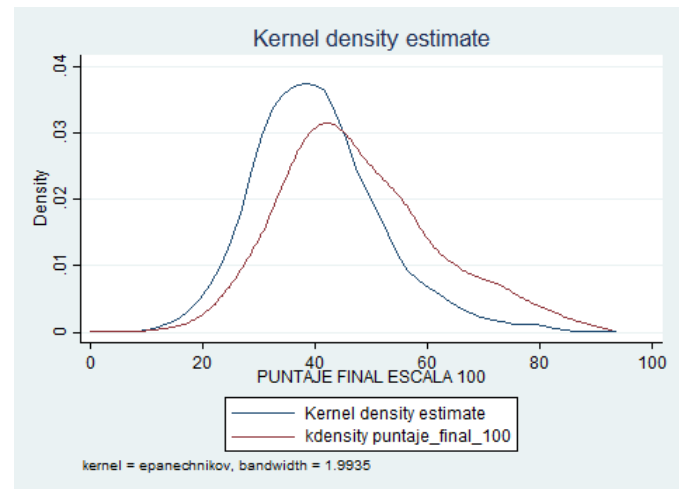
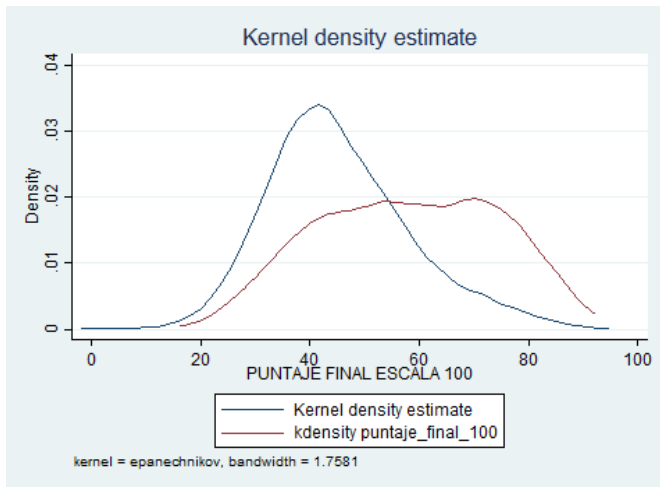
Las Tablas 10-16 también exponen resultados consistentes con la literatura citada ya que muestran diferencias bastante marcadas en los resultados de acuerdo a cada tipo de condición de la vivienda inadecuada. Los resultados alcanzados por aquellos postulantes que residen en viviendas con redes de agua y desagüe dentro del hogar (Tablas 10 y 11) son mucho mayores (entre 5 y 10 puntos) que los obtenidos por los postulantes con un establecimiento inadecuado de estas redes.

Por su parte, la Tabla 12 muestra que la deficiencia de cocinar con un combustible de cocina inadecuado puede incidir de manera bastante negativa (42 puntos en contraste con los 50 puntos alcanzados en donde el combustible de cocina es gas). Asimismo, la condición de hacinamiento también puede incidir en un resultado bastante más bajo debido a que no se cuenta con los espacios adecuados para poder destinar las horas de estudio necesarias.

Por otro lado, las Tablas 14, 15 y 16 muestran que un material de construcción adecuado para las paredes y los techos se relacionan a un resultado mayor hasta por 10 puntos con respecto al mínimo de sus respectivas categorías.

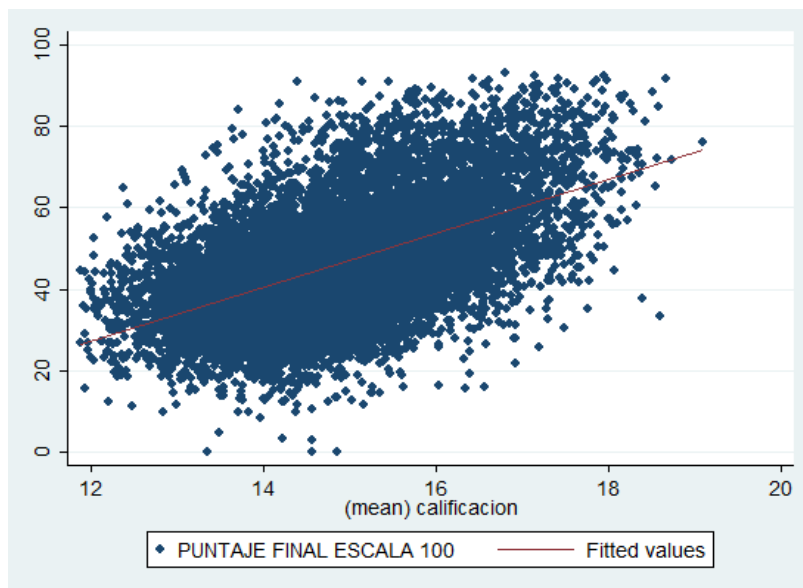
Estas tablas junto a los gráficos mostrados anteriormente, comienzan a darle forma al potencial modelo predictivo para poder evaluar aquellas características que se encuentran asociadas, de manera significativa, al rendimiento en un examen estandarizado (medida del desarrollo cognitivo). La exposición a privaciones tiene un impacto bastante negativo en el desempeño académico futuro (medido por el resultado de una evaluación que mide las capacidades básicas necesarias para la etapa de educación superior).

Gráfico 16. Densidad del resultado por tipo de gestión **Gráfico 17. Resultado por lengua materna**



Fuente: PRONABEC. Elaboración propia.

Gráfico 18. Relación entre el promedio de secundaria y resultado de la evaluación.



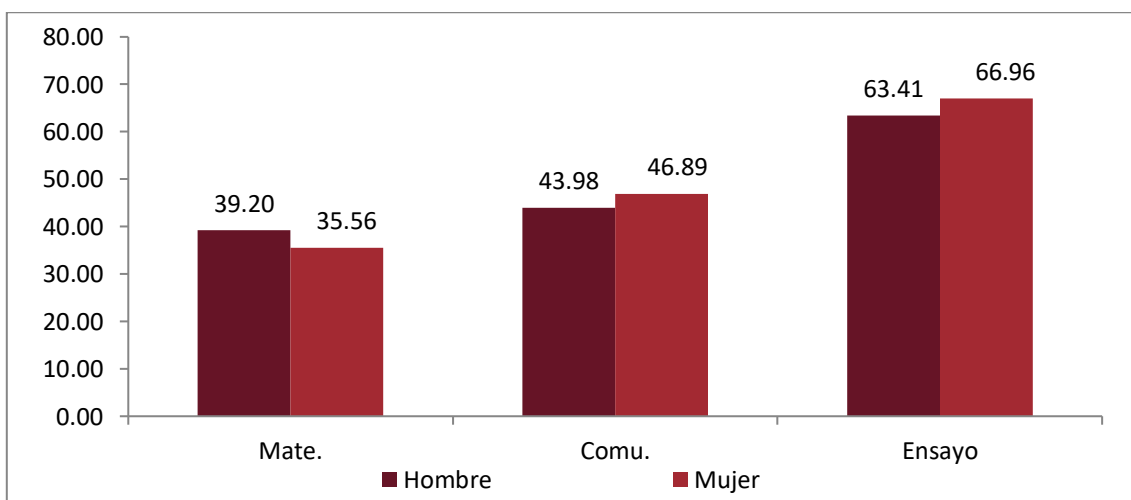
Fuente: PRONABEC. Elaboración propia.

Los Gráficos 16 y 17 señalan la relevancia de las variables del tipo de gestión del colegio donde estudió el postulante, la lengua materna y el promedio obtenido en los últimos 5 años de estudio con respecto a la nota obtenida en el examen. Para el Gráfico 16, la distribución estimada para los resultados obtenidos por los postulantes que provienen de colegios privados es la línea roja, distribución que tiene una media muy por encima de su contraparte pública. De manera más exacta, el promedio de aquellos en colegios de gestión privada se encuentra casi 12 puntos por encima. Para aquellos postulantes cuya lengua materna es el castellano (Gráfico 17, línea roja) el promedio que obtuvieron se encuentra por encima de aquellos cuya lengua materna es quechua u otra. Por último, el conocimiento acumulado (medido mediante el promedio de los últimos 5 años) presenta una correlación positiva y bastante marcada con el resultado obtenido en la evaluación.

Análisis de los componentes de la calificación global.

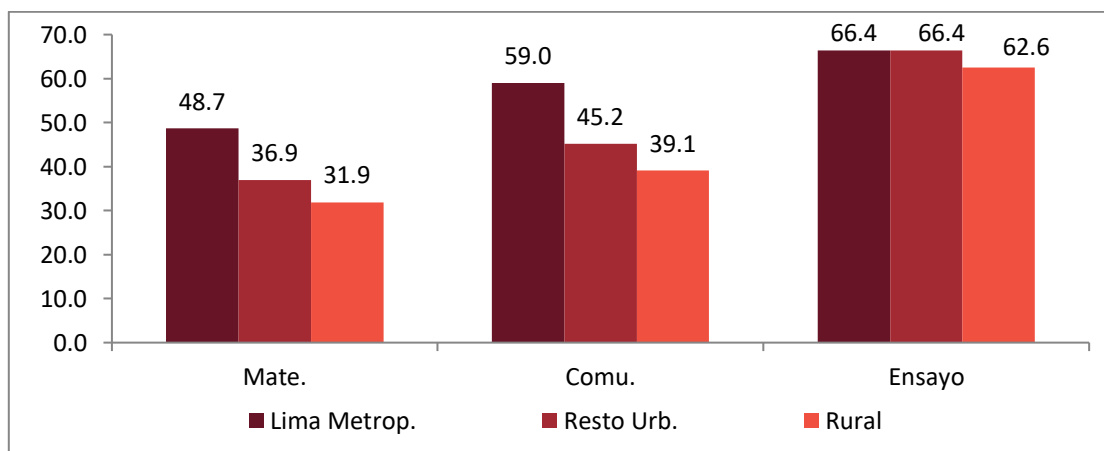
Para otorgar una idea acerca de las diferencias que existen dentro del logro académico de aquellos estudiantes que terminan la secundaria es necesario ver los resultados de cada componente de la nota global de la evaluación estandarizada.

Gráfico 19. Resultado de los componentes por sexo



Fuente: PRONABEC. **Elaboración propia.**

Gráfico 20. Resultado de los componentes por ámbito geográfico



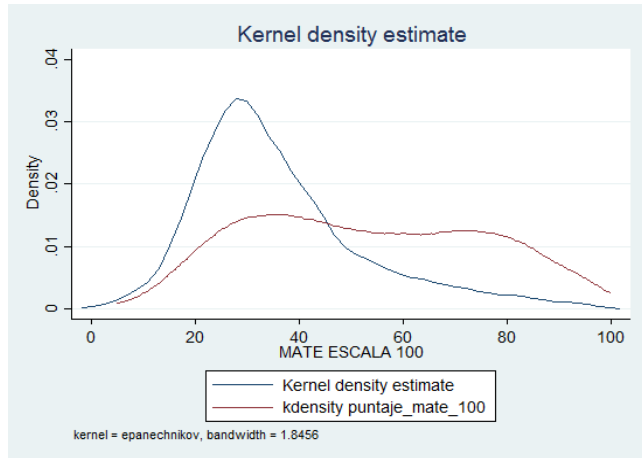
Fuente: PRONABEC. **Elaboración propia.**

Cuando se analizan los componentes de la calificación global pueden verse heterogeneidades que, en un principio, no aparecían como obvias. El Gráfico 19 se usa para el caso del análisis por sexo y se puede ver que, por un lado, las mujeres obtienen un resultado mayor para los componentes de comunicación y en la elaboración del ensayo. Por otro lado, los hombres obtienen un resultado mayor para la parte de matemáticas. Asimismo, el análisis de los componentes mediante el filtro del ámbito geográfico (Gráfico 20) también revela particularidades que no pueden verse con un análisis global.

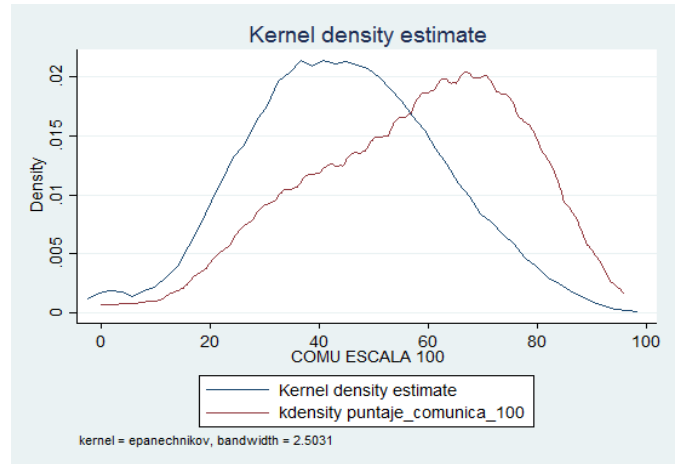
No existe una diferencia pronunciada en la evaluación del ensayo; sin embargo, los componentes de razonamiento verbal y comunicación si presentan una diferencia bastante notoria para los tres ámbitos. Los postulantes que vienen desde Lima metropolitana, como es de esperarse, tienen una calificación significativamente más alta que aquellos del resto urbano o de la zona rural.

Gráfico 21. Densidades de los resultados por tipo de gestión

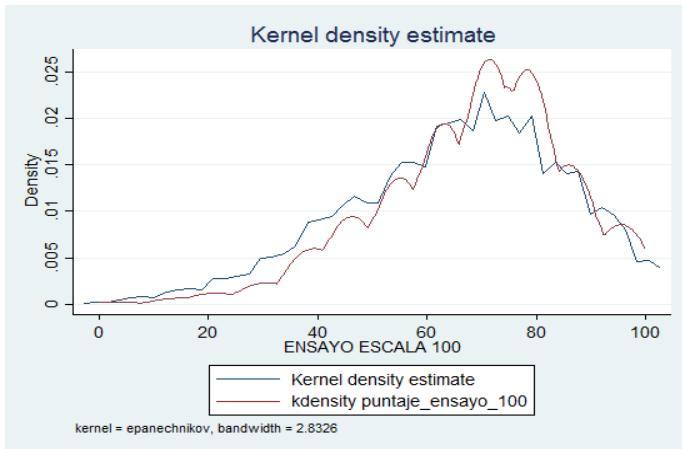
Matemática



Comunicación



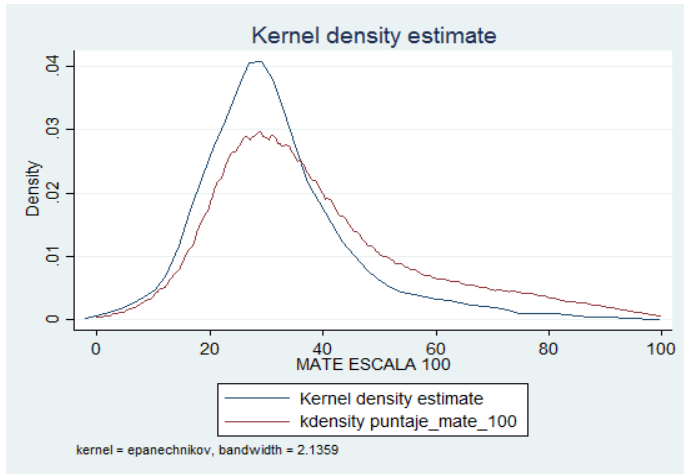
Ensayo



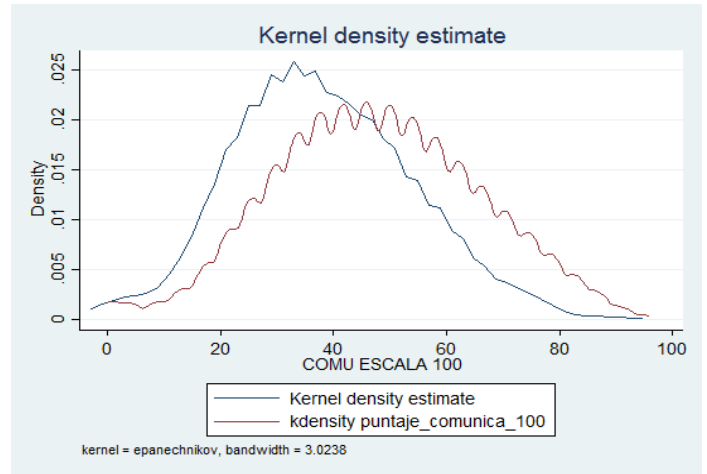
Fuente: PRONABEC. Elaboración propia.

Gráfico 22. Densidades de los resultados por lengua materna

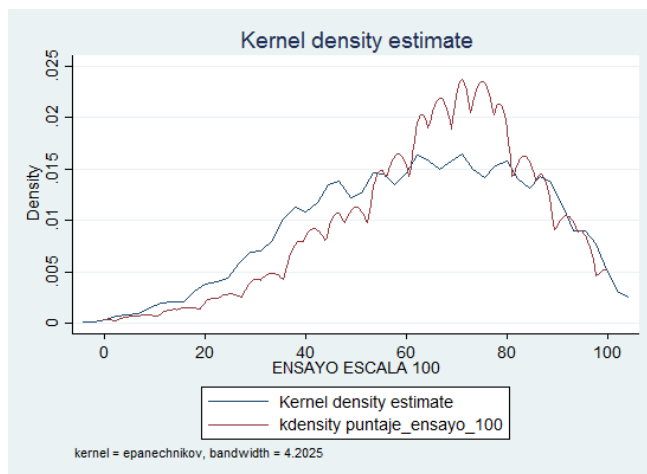
Matemática



Comunicación



Ensayo

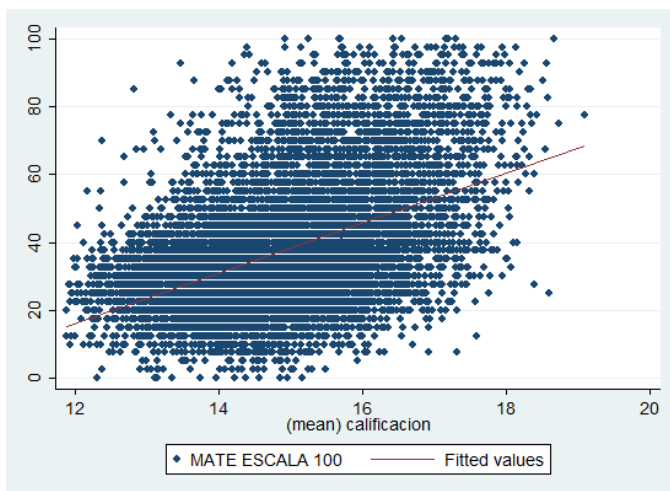


Fuente: PRONABEC. Elaboración propia.

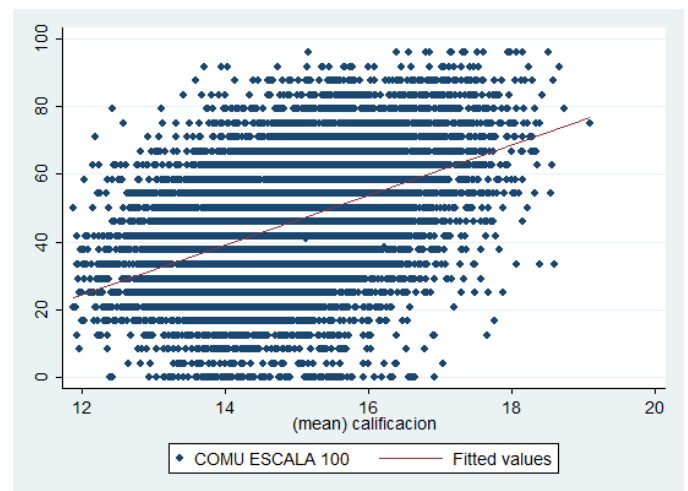
Nuevamente, los resultados, divididos por cada componente de la nota final, tienen una distribución distinta. El Gráfico 21 muestra que la distribución de los resultados de matemática no presenta una media tan diferente para aquellos cuya lengua materna es distinta al castellano (distribución en rojo). La calificación obtenida para comunicación y para el ensayo si presenta una media un poco mayor para los postulantes que identifican al castellano como su lengua materna. Dado esto, el análisis mediante la separación del tipo de gestión del último año del colegio (Gráfico 22) si presenta brechas muy marcadas. Aquellos estudiantes que provienen de colegios privados (a pesar de ser una minoría bastante reducida) tienen calificaciones bastante mayores en los componentes de comunicación y matemáticas. Toda la distribución (línea roja) se encuentra desplazada muy a la derecha y la media es bastante mayor, así como la varianza (que puede deberse al número reducido de observaciones disponibles).

Gráficos 23. Dispersión de los resultados y el promedio de secundaria

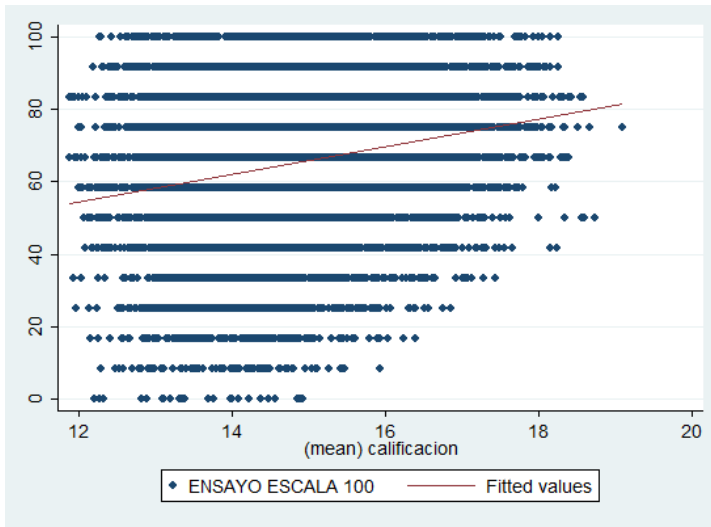
Matemática



Comunicación



Ensayo



Fuente: PRONABEC. **Elaboración propia.**

La relación entre el promedio de secundaria y el resultado dividido entre sus componentes sigue la misma tendencia que con el resultado global. Mientras se obtenga un promedio mayor (que puede interpretarse como la acumulación de conocimientos previos para el desenvolvimiento en la universidad) el desempeño en esta evaluación será mejor. Lo cual es más notorio en el componente de matemáticas, donde existe una varianza mayor en las notas obtenidas.

Para finalizar el análisis se presenta el modelo econométrico de la estimación de la función de producción de rendimiento académico (desarrollo cognitivo).

Tabla 17. Resultados de la estimación

Variables	Global	Verbal	Matemático	Ensayo
Promedio	5.473*** (0.110)	5.834*** (0.150)	6.337*** (0.152)	3.026*** (0.186)
Publica	-5.694*** (0.451)	-4.495*** (0.585)	-9.198*** (0.653)	-1.083* (0.648)
Mujer	-2.375*** (0.218)	-0.0432 (0.306)	-6.704*** (0.296)	1.618*** (0.389)
Edad	0.0595 (0.0609)	0.0553 (0.0948)	0.106 (0.0801)	-0.0256 (0.114)
Lengua	-0.470** (0.206)	-0.911*** (0.307)	0.678*** (0.250)	-1.882*** (0.445)
Años educ. jefe	0.243*** (0.0297)	0.308*** (0.0418)	0.284*** (0.0390)	0.0308 (0.0533)
Ing. Hogar	0.00362*** (0.000300)	0.00392*** (0.000413)	0.00466*** (0.000408)	0.000970** (0.000435)
Bienes	0.369*** (0.0600)	0.432*** (0.0825)	0.448*** (0.0840)	0.0829 (0.0961)
Privaciones	-1.049*** (0.0942)	-1.446*** (0.134)	-1.099*** (0.128)	-0.155 (0.165)
Rural	-0.742*** (0.247)	-1.093*** (0.360)	-0.505 (0.320)	-0.515 (0.499)
Horas capital	-0.0861*** (0.0315)	0.0150 (0.0448)	-0.115*** (0.0400)	-0.231*** (0.0683)
Constante	-30.79*** (2.379)	-38.48*** (3.386)	-50.30*** (3.244)	23.60*** (4.159)
N	10,246	10,246	10,246	10,246
R-cuadrado	0.426	0.315	0.368	0.058

Elaboración propia.

*** Significativo al 0.01, ** significativo al 0.05, * significativo al 0.1

Los resultados de la estimación muestran que las variables pertenecientes a las características cognitivas previas del postulante y al tipo de oferta educativa a la que estuvo expuesto son extremadamente relevantes. Puesto que 1 punto adicional en el promedio de notas de secundaria representa cinco puntos adicionales en el resultado final de la evaluación estandarizada (tal como se podía apreciar en el gráfico de puntos líneas arriba) y este coeficiente es significativo al 0.01.

Las mujeres obtienen, en promedio, 2 puntos menos que los varones y aquellos cuya lengua materna es el castellano pueden obtener casi 0.5 puntos de diferencia con respecto a aquellos cuya lengua materna es el quechua. Los activos del hogar como el capital humano del jefe (medido por los años de educación), el ingreso percibido por los padres y los bienes de riqueza con los que se cuenta (internet, televisor, cable, teléfono, entre otros) tienen una correlación positiva y significativa con el resultado obtenido. El coeficiente de privaciones muestra que por cada privación adicional a la que se encuentra expuesto el postulante (combustible de cocina inadecuado, servicios higiénicos insuficientes, condición de hacinamiento) se obtiene un punto menos en la evaluación estandarizada.

Además, los coeficientes de las estimaciones para cada componente de la nota global si presentan cierta heterogeneidad. Para la calificación del ensayo la significancia se reduce considerablemente y el coeficiente de bondad de ajuste baja hasta casi la octava parte cuando se le compara con el coeficiente del resultado global. Los signos de los coeficientes de estimación permanecen consistentes durante todas las estimaciones. Las privaciones tienen un efecto negativo y los activos del hogar (aproximados por los bienes de riqueza, la educación del jefe de hogar y el ingreso) se mantienen como provisorios de un ambiente de desarrollo cognitivo más propicio.

Lo más resaltante de estas estimaciones es el incremento entre la diferencia de los resultados obtenidos en la parte de matemáticas por aquellos estudiantes que provienen de escuelas públicas, el coeficiente pasa de 5 a casi 10 puntos de brecha.

5. CONCLUSIONES Y RECOMENDACIONES DE POLÍTICA

El análisis de resultados del Examen Único Nacional de Beca 18 – Convocatoria 2016 ha provisto de nociones bastante claras con respecto a los condicionantes del desarrollo cognitivo previo a la vida universitaria, entre las cuales tenemos las siguientes:

Las variables socioeconómicas son determinantes al momento de propiciar un desarrollo cognitivo óptimo.

Tanto el análisis univariado (tablas de resultados) como el análisis multivariado (regresiones) muestran que las privaciones (combustible de cocina inadecuado, malos materiales de construcción de paredes, piso y techo, condición de hacinamiento) tienen un impacto negativo en el logro cognitivo y potencial logro académico en el ámbito de educación superior.

En base a lo anterior, puede decirse que la definición de población objetivo por parte del programa es bastante adecuada. Es decir, que el foco de destino para una beca se encuentra bien formulado.

Aun con ello, a la luz de los resultados, pareciera necesario diseñar un ciclo de nivelación estandarizado que solvete las posibles deficiencias que presenten los potenciales becarios. Debido a que la distribución de los resultados presenta una media por debajo de 50 para todos los postulantes, debe reforzarse los conocimientos previos para que se asegure un resultado académico adecuado en la educación superior, teniendo en cuenta que el resultado académico obtenido en esta etapa puede determinar, casi por completo, el resultado laboral posterior.

Se recomienda, además:

- Asegurar la continuidad de una evaluación estandarizada complementaria a la evaluación del conocimiento acumulado en la educación secundaria.

- Propiciar, aun a riesgo de resultados, una mayor descentralización, centrando la cobertura a nivel de áreas rurales.
- Utilizar la evaluación estandarizada y los resultados de la educación secundaria para caracterizar un nivel de riesgo o por lo menos para identificar las condiciones en las que llegan los estudiantes a la educación superior. Profundizar en este análisis con miras a definir una estandarización de los procesos de educación superior de remediación (Ciclo de nivelación o “Ciclo 0”).
- Considerar la pertinencia de mantener los 3 niveles de evaluación en tanto la diferencia en sus resultados evidenciaría una intención de captura de dimensiones que pueden estar trascendiendo los límites cognitivos (académicos). En particular la consideración de un ensayo, que por otra parte, debería contar con una metodología de realización sólida y validada, representando así una forma de evaluación estandarizada sin precedentes en la región (en donde priman las evaluaciones de matemática y lenguaje solamente).
- Definir, en el marco de estos estudios, una línea de generación de evidencia pertinente con los objetivos de la Oficina, y del Programa en general, que gestiona Beca 18.

6. BIBLIOGRAFÍA

- Arancibia, J. (2015). *La estandarización de la evaluación. Las pruebas internacionales, ¿medición o evaluación?* Internacional de la Educación para América Latina.
- Arteaga, I., & Glewwe, P. (2014). *Achievement Gap between Indigenous and Non-Indigenous Children in Peru: An Analysis of Young Lives Survey Data*, Working Paper 130. Oxford: Young Lives.
- Barrenechea, I. (2010). Evaluaciones estandarizadas: seis reflexiones críticas. *Archivos Analíticos de Políticas Educativas*, 18(8).
- Benavidez-Ormaza, V. (2010). Las evaluaciones de logros educativos y su relación con la calidad de la educación. *Revista Iberoamericana de Educación*, 53, 83-96.
- Chica-Gómez, S. M., Galvis-Gutiérrez, D. M., & Ramírez-Hassan, A. (2010). Determinantes del rendimiento académico en Colombia. Pruebas ICFES – Saber 11°, 2009. *Revista Universidad EAFIT*, 46(160), 48-72.
- Consejo Nacional de Educación (2013). Evaluaciones estandarizadas del rendimiento escolar. *Boletín del Consejo Nacional de Educación*, 36.
- Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2), 31-47. doi: 10.1257/aer.97.2.31
- De Hoyos, R. E., Espino, J. M., & García, V. (2012). Determinantes del logro escolar en México. Primeros resultados utilizando la prueba ENLACE media superior. *El Trimestre Económico*, LXXIX(4), 783-811.

- Eurydice (2010). *Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados*. Bruselas: Agencia Ejecutiva en el Ámbito Educativo, Audiovisual y Cultural.
- Glewwe, P., & Miguel, E. (2008). The Impact of Child Health and Nutrition on Education in Less Developed Countries. En T.P. Schultz & J.A. Strauss (Eds.), *Handbook of Development Economics*, (Vol. 4), 3561-3606. doi: 10.1016/S1573-4471(07)04056-9
- Heckman, J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, 312(5782), 1900-1902. doi: 10.1126/science.1128898
- Hox, Joop (2010). *Multilevel Analysis, Techniques and Applications* (2da edición). London: Routledge.
- Liu, Y. (2013). Meritocracy and the Gaokao: a survey study of higher education selection and socio-economic participation in East China. *British Journal of Sociology of Education*, 34(5-06), 868-887. doi: 10.1080/01425692.2013.816237
- Martínez, L., & Prada, S. (2012). Consecuencias de medir progreso educativo con pruebas estandarizadas: el caso de los Estados Unidos. *Revista de Antropología y Sociología VIRAJES*, 14(1), 47-63.
- Pearson (2013). *Evaluación de la PSU en Chile. Informe Final*. Santiago: Ministerio de Educación de Chile.
- Treviño-Villarreal, E. (2006). Evaluación del aprendizaje de los estudiantes indígenas en América Latina. Desafíos de medición e interpretación en contextos de diversidad cultural y desigualdad social. *Revista Mexicana de Investigación Educativa*, 11(28), 225-268.
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez-Rizo, F., Aylwin, M., & Wolff, L. (2007). Las evaluaciones educativas que

América. Santiago: Programa de Promoción de la Reforma Educativa en América Latina y El Caribe-PREAL.

Revista Docencia (2003). Los sistemas estandarizados de medición: tema en debate. Revista Docencia N° 20, Santiago, Chile.

Revista Docencia (2009). Estandarización educativa en Chile: un peligroso hábito. Revista Docencia N° 39, Santiago, Chile.

Travitzki, R., Calero, J., & Boto, C. (2014). ¿Qué información proporciona el Examen Nacional de Enseñanza Media (ENEM) a la sociedad brasileña? *Revista Cepal*, 113, 163-181.

UNESCO (2012). Situación educativa de América Latina y el Caribe. Santiago: Oficina Regional de Educación de la UNESCO para América Latina y el Caribe - OREALC/UNESCO.

UNESCO & LLECE (2008). *Reflexiones en torno a la evaluación de la calidad educativa*. Santiago: Oficina Regional de Educación de la UNESCO para América Latina y el Caribe - OREALC/UNESCO, Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación - LLECE.

Anexo 1

Códigos de países europeos

EU-27	Unión Europea	NL	Países Bajos
BE	Bélgica	AT	Austria
BE fr	Bélgica - Comunidad francesa	PL	Polonia
BE de	Bélgica - Comunidad germanófono	PT	Portugal
BE nl	Bélgica - Comunidad flamenca	RO	Rumania
BG	Bulgaria	SI	Eslovenia
CZ	República Checa	SK	Eslovaquia
DK	Dinamarca	FI	Finlandia
DE	Alemania	SE	Suecia
EE	Estonia	UK	Reino Unido
IE	Irlanda	UK-ENG	Inglaterra
EL	Grecia	UK-WLS	Gales
ES	España	UK-NIR	Irlanda del Norte
FR	Francia	UK-SCT	Escocia
IT	Italia	Países AELC/AEE	Los tres países de la Asociación Europea de Libre Comercio que son miembros del Área Económica Europea
CY	Chipre	IS	Islandia
LV	Letonia	LI	Liechtenstein
LT	Lituania	NO	Noruega
LU	Luxemburgo		
HU	Hungría		
MT	Malta		

Fuente: Eurydice (2010).

