

PAES de Invierno por región: brechas territoriales y concentración de puntajes altos

Un estudio descriptivo de la Admisión 2027 con remuestreo por establecimientos

Nicolás Pinto P. Observatorio PAES - Preuniversitario Gauss

Resumen ejecutivo. Este estudio describe la distribución territorial de los resultados de la PAES de Invierno del proceso de Admisión 2027. De 33.881 inscritos, 24.448 rindieron Competencia Lectora y Matemática 1 y cuentan con región del establecimiento de egreso, constituyendo la base territorial del índice obligatorio. La Región Metropolitana alcanza 719,1 puntos, frente a 681,7 en el resto del país: una diferencia de 37,3 puntos.

Parte de esa distancia se asocia a la composición por dependencia. Sobre los 24.245 casos con dependencia conocida, la brecha Metropolitana-resto es de 37,5 puntos y se reduce a 23,2 al igualar la mezcla de establecimientos particulares pagados, subvencionados y públicos. Sin embargo, la brecha permanece especialmente amplia dentro del sector público: estandarizada a una composición municipal/SLEP común, alcanza 45,8 puntos, frente a 14,9 entre egresados de particulares pagados. El contraste entre ambas brechas es de 31,0 puntos (IC 95 % [7,4; 50,6]).

La Metropolitana también concentra puntajes altos: la probabilidad de superar 900 puntos en el índice es 1,98 veces la del resto del país. No obstante, las diferencias de cuantiles muestran que la ventaja es máxima cerca del centro de la distribución y disminuye hacia el extremo superior. Los resultados describen a la población autoseleccionada que rindió en Invierno y no admiten interpretación causal ni extrapolación directa a la cohorte escolar de cada región.

Palabras clave: PAES; admisión universitaria; desigualdad territorial; regiones; dependencia escolar; bootstrap; estandarización.

1 Introducción

La PAES de Invierno ofrece una segunda instancia de rendición a quienes desean ingresar a la educación superior, pero la distribución territorial de sus puntajes o postulantes no necesariamente equivale a la del proceso Regular. La participación es voluntaria, las regiones difieren en tamaño y composición, y las pruebas electivas son rendidas por cohortes de distintas características. Por ello, comparar directamente promedios regionales requiere distinguir entre tres preguntas: cuánto difiere el nivel observado, cuánto de esa diferencia se asocia a la composición de quienes rinden y en qué parte de la distribución se concentra la ventaja de las regiones con mejores resultados.

El análisis territorial presenta además un problema de referencia. Una región grande, como la Metropolitana, influye de manera importante en el promedio nacional. Compararla contra ese promedio significa compararla parcialmente consigo misma. Para evaluar una ventaja territorial, la referencia natural es el resto del país, es decir, al evaluar la ventaja de una región, se utiliza como referencia el promedio nacional, excluyendo a dicha región. Esta distinción es especialmente relevante en Invierno: la Región Metropolitana representa el 41,8 % de la base territorial utilizada para construir el índice de pruebas obligatorias, definido como el promedio entre los puntajes de Competencia Lectora y Matemática 1.

Un segundo problema es composicional. La proporción de egresados de establecimientos particulares pagados, subvencionados, municipales y de Servicios Locales de Educación Pública (SLEP) varía entre regiones. Si estos grupos exhiben niveles medios distintos, una parte de la brecha territorial puede surgir de la mezcla de dependencias y no de diferencias dentro de cada categoría. El presente estudio utiliza

estandarización directa para separar descriptivamente ambos componentes, sin interpretar el residuo como un efecto causal del territorio.

Finalmente, una media no informa en qué sector de la distribución se produce la diferencia. Un territorio puede aventajar a otro de forma aproximadamente uniforme en cada cuantil, concentrar la diferencia cerca de la mediana o separarse especialmente en los puntajes más altos de cada territorio. Este punto adquiere particular importancia porque Matemática 1 presenta saturación en el extremo alto de la escala. En consecuencia, el estudio combina promedios, razón de proporciones de postulantes que superan umbrales (esta cantidad es llamada *riesgo relativo*) y diferencias de cuantiles.

El objetivo es caracterizar la distribución regional de la PAES de Invierno de 2026, correspondiente a la Admisión 2027, respondiendo cuatro preguntas: (i) qué regiones presentan mayores niveles observados; (ii) cuánto cambia la brecha entre la Región Metropolitana y el resto de las regiones al estandarizar por dependencia; (iii) cómo varía esa brecha dentro de cada dependencia; y (iv) si la ventaja metropolitana se intensifica o se atenúa hacia la cola superior de la distribución.

2 Metodología

2.1 Fuente, población y estimando

Se utilizó el Archivo C publicado por DEMRE para el proceso de Invierno de la Admisión 2027. El archivo contiene datos de 33.881 postulantes inscritos; 30.984 rindieron al menos una prueba y 24.640 rindieron las dos pruebas obligatorias. Entre estos últimos, 192 no cuentan con región

identificada, por lo que la base territorial del índice queda en 24.448 casos.

La región corresponde al establecimiento de egreso y no a la residencia del postulante. Dentro de la base de datos, se consideraron válidos los puntajes entre 100 y 1000; los ceros y campos vacíos fueron tratados como prueba no rendida. El índice obligatorio se definió como

$$IND_i = \frac{CLEC_i + M1_i}{2}, \quad (1)$$

únicamente para quienes rindieron ambas pruebas, donde $CLEC_i$ corresponde al puntaje de Competencia lectora y $M1_i$, al de Matemáticas 1, e i es un índice para distinguir entre distintos postulantes. Por tanto, bajo un enfoque de superpoblación, el estimando principal es el valor esperado de IND condicionado a inscripción en Invierno, rendición de Competencia Lectora y Matemática 1 y región identificada. La proporción de inscritos que satisface esta condición varía entre regiones desde 69,2 % hasta 79,1 %.

Los promedios de pruebas individuales utilizan a quienes rindieron cada prueba y tienen región conocida; por ello, la cantidad de casos de las pruebas individuales difiere del índice. La composición descriptiva por dependencia utiliza a los inscritos con región, mientras que las estandarizaciones se restringen a los 24.245 casos con índice, región y dependencia conocidas.

2.2 Marco inferencial e intervalos

Los datos constituyen un registro completo del proceso de Invierno analizado, no una muestra probabilística. En sentido de diseño, las medias observadas son parámetros de esa población y no existe error muestral. Los intervalos se interpretan bajo un modelo de superpoblación: cuantifican la variación que se obtiene al remuestrear la distribución empírica observada, pero no incorporan otros factores como cambios entre cohortes, formas de prueba, procesos de equating entre otros que podrían diferir en procesos de selección previos o posteriores.

La incertidumbre de las medias se estimó mediante bootstrap BCa al 95 %, con $B = 1000$ réplicas y semilla fija. El intervalo principal remuestrea establecimientos de egreso completos; el intervalo que remuestrea estudiantes se mantiene solo como referencia. Los 203 registros sin RBD forman conglomerados unitarios. El cociente entre las varianzas bootstrap por establecimientos e individuos se utilizó como análogo del efecto de diseño.

Como análisis de sensibilidad, se aplicó una contracción de efectos aleatorios a los promedios regionales. Este procedimiento acerca en mayor medida al centro común las estimaciones regionales menos precisas y permite evaluar hasta qué punto el orden observado depende de la distinta precisión con que se estima cada región. Los valores contraídos no sustituyen los promedios observados. La heterogeneidad se estimó mediante DerSimonian–Laird a partir de los errores estándar por conglomerados. El centro de contracción se estimó mediante una media ponderada de efectos aleatorios, con pesos $1/(s_r^2 + \tau^2)$, donde s_r^2 representa la varianza del estimador regional y τ^2 la heterogeneidad entre regiones.

2.3 Comparación territorial y estandarización

La comparación principal de nivel utiliza la diferencia entre la Región Metropolitana y el resto del país. Para estudiar la composición, municipal y SLEP se agruparon en una categoría pública, obteniendo tres estratos de dependencia: particular pagado, particular subvencionado y público. Los pesos nacionales de estos estratos en la base analítica son 22,3 %, 55,3 % y 22,4 %, respectivamente.

Sea r una región y k uno de los tres estratos de dependencia. Se denota por $\bar{Y}_{rk}$ el promedio del índice obligatorio en la región r y el estrato k , y por w_k la proporción nacional correspondiente a dicho estrato. La media estandarizada de la región r se calculó como

$$\bar{Y}_r^* = \sum_k w_k \bar{Y}_{rk}. \quad (2)$$

De este modo, todas las regiones se evalúan bajo una misma composición por dependencia. Para incluir un estrato en el cálculo se exigió un mínimo de 30 casos por celda región–dependencia. Ñuble y Aysén no alcanzan cobertura completa de los tres estratos, por lo que sus valores estandarizados se reportan como parciales.

Para cuantificar cuánto cambia la brecha entre la Región Metropolitana y el resto del país al igualar esta composición, se calculó

$$\Delta_{RM-resto}^* = \sum_k w_k (\bar{Y}_{RM,k} - \bar{Y}_{resto,k}), \quad (3)$$

donde $\bar{Y}_{RM,k}$ y $\bar{Y}_{resto,k}$ corresponden a los promedios del índice obligatorio en el estrato k para la Región Metropolitana y el resto del país, respectivamente.

Todas las combinaciones lineales de brechas se estimaron utilizando el estadístico observado como valor puntual. Los intervalos se obtuvieron mediante un bootstrap conjunto, estratificado e independiente entre celdas.

Dentro del sector público se realizó una segunda estandarización, ahora con pesos comunes para municipal y SLEP: 58,2 % y 41,8 %. Esto permite comparar la brecha pública con la de particulares pagados sin arrastrar diferencias en la mezcla administrativa interna del sector público.

2.4 Concentración y forma de la distribución

La concentración de puntajes altos se midió mediante riesgo relativo (RR). Para una región r y un umbral u ,

$$RR_{r,u} = \frac{P(IND \geq u \mid r)}{P(IND \geq u \mid \text{resto del país})}. \quad (4)$$

Se utilizaron los umbrales 700, 800 y 900. La referencia excluye siempre a la propia región. Se reportan intervalos de Miettinen–Nurminen. El intervalo de confianza principal se construyó a partir de una distribución bootstrap obtenida mediante remuestreo por establecimientos, separando la región analizada del resto del país. Como análisis de sensibilidad, las proporciones regionales se ajustaron mediante un modelo beta-binomial empírico de Bayes, que aproxima en mayor medida hacia el promedio general las estimaciones basadas en pocos casos. Este ajuste permite evaluar la estabilidad de los patrones regionales frente a tamaños de grupos muy diferentes. el mapa utiliza siempre las proporciones sin ajustar.

Para estudiar si la brecha entre la Región Metropolitana y el resto del país es similar entre estudiantes con puntajes bajos, medios y altos, se calcularon las brechas entre cuantiles metropolitanos y del resto del país en los percentiles 5, 10, 25, 50, 75, 90, 95 y 99, con intervalos bootstrap por establecimientos. Como escenario contrafactual descriptivo, se desplazó en bloque la distribución del resto del país, una vez por la diferencia de medianas y otra por la diferencia de medias, y se comparó el RR resultante con el observado.

2.5 Variación interna y reproducibilidad

La variación del índice se descompuso exactamente en sumas de cuadrados entre regiones, entre comunas dentro de cada región y entre estudiantes dentro de cada comuna. Las proporciones resultantes corresponden a una partición descriptiva de la variación observada.

El análisis fue implementado en Python mediante las librerías `numpy`, `pandas`, `scipy` y `openpyxl`. El cuaderno genera las tablas regional, comunal y de concentración, junto con un archivo JSON que alimenta el reporte y algunas figuras.

El desarrollo del código fue generado con apoyo de modelos de lenguaje a gran escala, ChatGPT y Claude.

3 Resultados

3.1 La Región Metropolitana se separa del resto

La Región Metropolitana obtiene 719,1 puntos en el índice obligatorio, frente a 697,4 del conjunto territorial y 681,7 del resto del país. La diferencia relevante para una afirmación territorial es, por tanto, de 37,3 puntos. La comparación contra el promedio nacional entrega solo 21,7 puntos porque ese promedio incluye a la propia Metropolitana.

Los Lagos ocupa el segundo lugar con 699,9 puntos y es además segunda en las cinco pruebas por separado. La diferencia observada entre la Región Metropolitana y Los Lagos es de 19,2 puntos, con IC 95, % por conglomerados [2,3; 35,6]. Este intervalo corresponde específicamente a la comparación entre ambas regiones.

Las dos mayores diferencias de la Región Metropolitana respecto del promedio nacional por prueba aparecen en Historia y Ciencias Sociales (34,2) y Matemática 1 (25,9); siguen Ciencias (22,3), Matemática 2 (20,5) y Competencia Lectora (18,1).

La contracción modifica poco esta estructura. La heterogeneidad entre regiones es $\tau = 19,9$ puntos y los errores estándar por conglomerados varían entre 2,8 y 16,4. El centro de efectos aleatorios es 682,3; el desplazamiento máximo es de 4,6 puntos, en Atacama. Cambian de posición cuatro de las dieciséis regiones, siempre por un lugar, y se mantienen Metropolitana, Los Lagos y Biobío en los tres primeros puestos.

3.2 La composición explica una parte, pero no la mayor parte, de la brecha

La composición de la base analítica es distinta en la Metropolitana: 31,2 % de sus casos proviene de establecimientos particulares pagados, frente a 22,3 % del conjunto. Para comparar brechas sobre la misma base, se utilizaron los

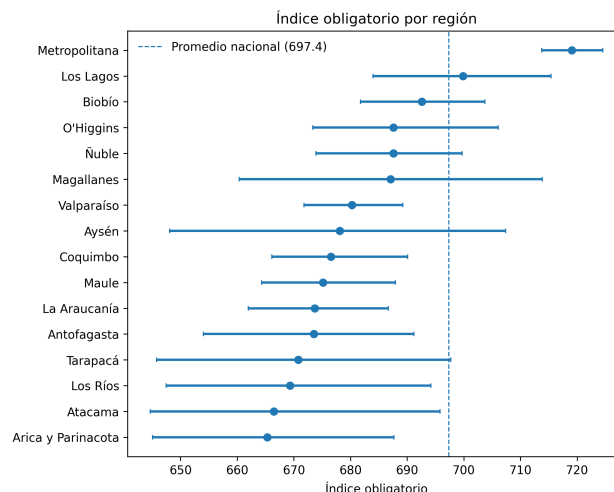


Figura 1: Índice obligatorio por región. Las barras gruesas corresponden al intervalo bootstrap por establecimientos; el remuestreo individual se mantiene como referencia.

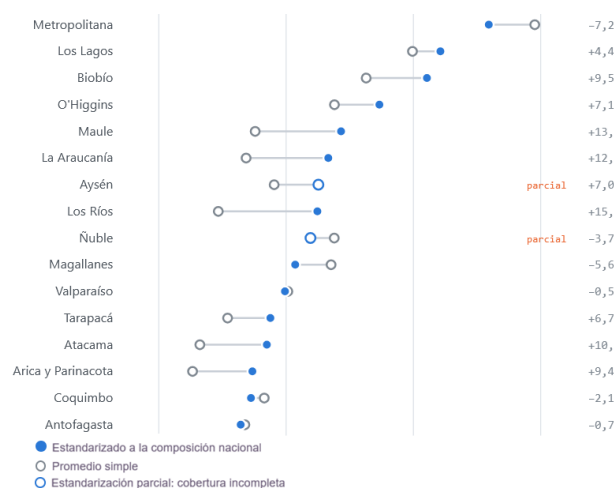


Figura 2: Promedio regional observado y estandarizado a la composición nacional por dependencia. Ñuble y Aysén tienen cobertura incompleta y se consideran estandarizaciones parciales.

24.245 casos con dependencia conocida. Allí, la diferencia sin ajustar entre la Región Metropolitana y el resto es de 37,5 puntos (IC 95 % [31,4; 43,9]). Al igualar la composición por dependencia en ambos territorios, la brecha disminuye a 23,2 puntos ([16,9; 29,0]), una reducción descriptiva de 14,3 puntos.

El resultado indica que la distinta mezcla de dependencias está asociada a una fracción apreciable de la ventaja metropolitana, pero deja en pie más de veinte puntos de diferencia.

3.3 La brecha es considerablemente mayor dentro del sector público

Al comparar dentro de cada dependencia, la Región Metropolitana supera al resto del país en 14,9 puntos entre egresados de particulares pagados, 17,3 entre subvencionados, 39,3 entre municipales y 55,0 entre SLEP. Los intervalos individuales son amplios y se solapan, por lo que el orden exacto de las cuatro estimaciones no queda establecido.

La comparación principal combina municipal y SLEP. La brecha observada entre regiones en establecimientos públicos es de 46,2 puntos; estandarizada a una mezcla común de

Cuadro 1: Índice obligatorio por región. El intervalo principal remuestrea establecimientos de egreso. La columna contraída corresponde al análisis de sensibilidad de efectos aleatorios.

Región	n	Índice	IC 95 %	Contraída
Metropolitana	10.222	719,1	[713,8; 724,5]	718,4
Los Lagos	904	699,9	[684,0; 715,4]	697,4
Biobío	2.476	692,6	[681,8; 703,7]	691,8
O'Higgins	1.447	687,6	[673,3; 706,1]	686,9
Ñuble	739	687,6	[673,9; 699,7]	687,0
Magallanes	212	687,1	[660,3; 713,9]	685,5
Valparaíso	1.935	680,3	[671,8; 689,2]	680,4
Aysén	139	678,1	[648,1; 707,4]	679,8
Coquimbo	1.160	676,6	[666,1; 690,1]	677,1
Maule	1.701	675,1	[664,3; 688,0]	675,7
La Araucanía	1.283	673,7	[661,9; 686,7]	674,5
Antofagasta	720	673,5	[654,0; 691,2]	675,2
Tarapacá	438	670,8	[645,8; 697,7]	674,4
Los Ríos	429	669,3	[647,4; 694,2]	672,7
Atacama	343	666,4	[644,7; 695,8]	671,0
Arica y Parinacota	300	665,3	[645,1; 687,7]	669,1

Cuadro 2: Ventaja de la Región Metropolitana sobre el resto del país dentro de cada dependencia.

Dependencia	Diferencia	IC 95 %	d
Part. pagado	+14,9	[4,9; 24,9]	0,13
Part. subv.	+17,3	[9,8; 24,2]	0,14
Municipal	+39,3	[12,7; 62,2]	0,32
SLEP	+55,0	[21,0; 79,6]	0,44

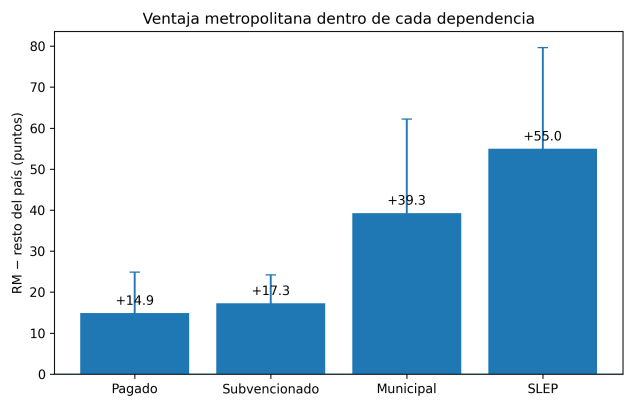


Figura 3: Ventaja metropolitana dentro de cada dependencia, con intervalo por conglomerados. La comparación entre categorías es descriptiva; municipal y SLEP son categorías administrativas, no niveles ordinales de una escala socioeconómica.

58,2 % municipal y 41,8 % SLEP, queda en 45,8 puntos (IC 95 % [23,5; 61,2]). La brecha entre establecimientos particulares pagados es 14,9. El contraste directo entre ambas brechas alcanza 31,0 puntos, con IC 95 % [7,4; 50,6].

En unidades de desviación típica, el patrón se conserva: $d = 0,13$ para particulares pagados, $0,14$ para subvencionados, $0,32$ para municipales y $0,44$ para SLEP. El patrón se mantiene al expresar las brechas en unidades de desviación estándar. Por tanto, las diferencias de dispersión entre los grupos no bastan para explicar el resultado. Este análisis no permite descartar posibles efectos de saturación de la escala.



Figura 4: Riesgo relativo de superar 700, 800 y 900 puntos, región contra el resto del país. El mapa utiliza proporciones crudas; la contracción beta-binomial se reserva como sensibilidad.

3.4 La Metropolitana concentra puntajes altos, pero no ensancha la cola superior

La proporción metropolitana que supera 700 puntos es 1,27 veces la del resto del país; sobre 800, el RR aumenta a 1,41; sobre 900, a 1,98. Para este último umbral, el intervalo por establecimientos es [1,66; 2,34]. Si se utilizara el promedio nacional en lugar del resto del país como referencia, el RR sobre 900 bajaría a 1,40 por la inclusión de la propia Metropolitana en el denominador.

La razón de proporciones(RR) aumenta al elevar el umbral de 700 a 900 puntos. Sin embargo, este aumento no implica que la distancia entre la Región Metropolitana y el resto del país sea mayor entre los puntajes más altos. Una misma diferencia de nivel puede producir razones cada vez mayores cuando se comparan eventos menos frecuentes. La comparación de percentiles muestra un patrón distinto: la brecha aumenta desde 26,0 puntos en el percentil 5 hasta 46,5 en la mediana, pero luego disminuye a 23,5 en el percentil 90, 21,0 en el 95 y 12,2 en el 99.

El escenario contrafactual de desplazamiento puro confirma esta lectura. Si la distribución del resto del país se desplazara en bloque por la diferencia de medianas (46,5 puntos), produciría un RR sobre 900 de 3,32; si se desplazara por la diferencia de medias (37,3), produciría 2,83. Ambos superan ampliamente el 1,98 observado. En la escala

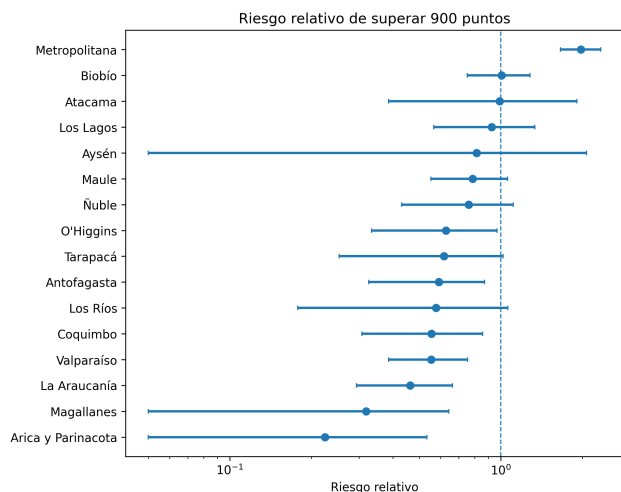


Figura 5: Riesgo relativo de superar 900 puntos. El intervalo principal remuestrea establecimientos; Miettinen–Nurminen se muestra como referencia bajo independencia individual.

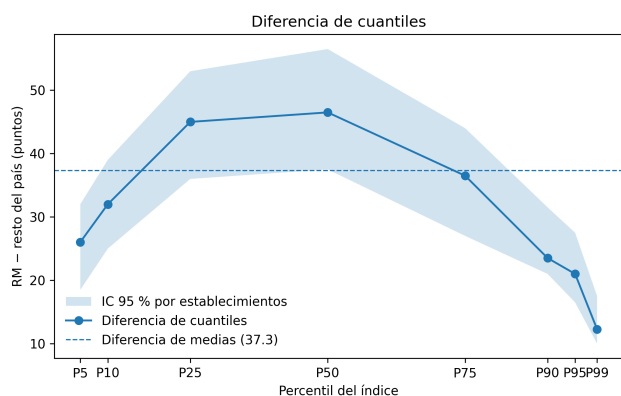


Figura 6: Diferencia de cuantiles entre la Región Metropolitana y el resto del país. La banda principal corresponde al bootstrap por establecimientos.

medida, por tanto, no aparece ensanchamiento de la cola superior: la ventaja metropolitana es comparativamente menor en el extremo alto de lo que produciría un desplazamiento uniforme.

La interpretación debe permanecer en la escala observada. Matemática 1 concentra 678 de los 697 puntajes de 1000 del proceso; el 2,50 % de quienes la rinden obtiene el máximo y el 18,8 % supera 900. El índice obligatorio reduce la dependencia de ese techo al promediar M1 con Competencia Lectora, pero no la elimina. Por ello, la compresión observada en la cola no permite inferir la forma de una distribución latente de desempeño.

3.5 La heterogeneidad interna supera a la interregional

Entre las 49 comunas con al menos 50 casos, Vitacura alcanza 797,8 puntos y La Pintana 637,8, un rango cercano a 160 puntos. El contraste de rangos es ilustrativo pero no es una comparación formal, porque las comunas constituyen una partición más fina que las regiones.

La descomposición de sumas de cuadrados muestra que el 2,4 % de la variación individual corresponde a diferencias entre regiones y el 5,2 % a diferencias entre comunas dentro de una misma región; el 92,4 % restante se encuentra entre estudiantes de una misma comuna. Dentro de la Región Metropolitana, la comuna de egreso representa el 7,2 % de

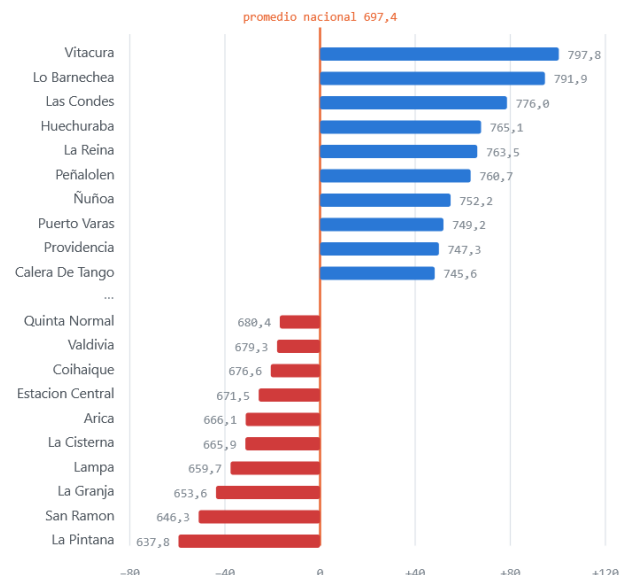


Figura 7: Comunas con mayor y menor promedio del índice entre aquellas con al menos 50 casos. La comparación gráfica se complementa con la descomposición formal de la variación.

la variación observada.

4 Discusión

El patrón territorial de la PAES de Invierno está dominado por una separación nítida de la Región Metropolitana. La diferencia de 37,3 puntos respecto del resto del país es mayor que la que sugiere la comparación contra el promedio nacional, precisamente porque la Metropolitana constituye una fracción muy grande de la base. Los Lagos aparece como un segundo caso particular: se ubica inmediatamente después de la Metropolitana tanto en el índice como en las cinco pruebas por separado, aunque con mayor incertidumbre debido a su menor tamaño.

La composición por dependencia explica descriptivamente una parte relevante de la brecha, pero no la mayor parte. Igualar la mezcla de particulares pagados, subvencionados y públicos reduce el diferencial de 37,5 a 23,2 puntos. Este resultado descarta una lectura simple según la cual el liderazgo metropolitano sería únicamente consecuencia de su mayor presencia de particulares pagados. Al mismo tiempo, el residuo no puede atribuirse al territorio en sentido causal: la estandarización controla una sola dimensión y no incorpora variables previas de desempeño ni mecanismos de selección hacia la rendición de Invierno.

El hallazgo más marcado aparece al comparar dependencias. La brecha Metropolitana–resto es relativamente pequeña entre egresados de establecimientos particulares pagados y considerablemente mayor entre egresados del sistema público. La estandarización municipal/SLEP muestra que este resultado no se debe a que la composición administrativa interna del sector público sea diferente en ambos territorios. No obstante, el análisis considera únicamente a quienes rindieron voluntariamente la PAES de Invierno. Si el perfil de quienes deciden rendirla difiere entre la Región Metropolitana y el resto del país, parte de la brecha observada podría deberse a esa selección. Con estos datos no es posible distinguir este efecto de una diferencia de rendimiento entre

territorios.

La concentración de puntajes altos exige una interpretación similarmente cuidadosa. La Región Metropolitana casi duplica la tasa del resto del país sobre 900 puntos, pero las diferencias de cuantiles indican que su ventaja es mayor en el tramo intermedio de la distribución y se reduce hacia los percentiles superiores. El resultado corrige una posible lectura intuitiva del RR creciente con el umbral: El aumento del RR al elevar el umbral no implica necesariamente que la brecha entre la Región Metropolitana y el resto del país sea mayor entre los puntajes más altos. De hecho, las diferencias entre percentiles disminuyen en la parte superior de la distribución observada. Esta conclusión debe interpretarse considerando que en Matemática 1 existe saturación en el extremo superior de la escala, lo que puede comprimir las diferencias entre estudiantes de muy alto desempeño.

También resulta relevante la magnitud de la heterogeneidad dentro de las regiones. La componente entre comunas dentro de una misma región supera a la componente entre regiones, aun cuando la mayor parte de la variación permanece dentro de las comunas. En términos descriptivos, esto sugiere que el territorio regional es una escala de agregación demasiado gruesa para resumir por sí sola las diferencias asociadas al establecimiento de egreso.

Estos resultados tienen cuatro limitaciones principales. Primero, la población de Invierno es autoseleccionada y no representa a la cohorte escolar de cada región; el 73,9 % de los inscritos ya había rendido el proceso Regular anterior. Segundo, el estimando del índice está condicionado a rendir ambas obligatorias, condición cuya frecuencia varía entre territorios. Tercero, aunque el análisis ajusta descriptivamente por dependencia administrativa, no dispone de variables individuales que permitan igualar otras diferencias de composición socioeconómica entre territorios. Cuarto, los intervalos por conglomerados reproducen la variación entre establecimientos de esta distribución empírica, pero no cuantifican cambios de cohorte, forma de prueba, equating entre otras variables relevantes.

5 Conclusiones

La PAES de Invierno de la Admisión 2027 presenta una estructura territorial clara. Entre quienes rindieron Competencia Lectora y Matemática 1 y cuentan con región identificada, la Región Metropolitana promedia 719,1 puntos, 37,3 más que el resto del país. Los Lagos ocupa el segundo lugar y la mayoría de las restantes regiones forma un grupo relativamente compacto.

La composición por dependencia está asociada a una parte de la ventaja metropolitana: al igualar la mezcla de dependencias, la brecha territorial disminuye de 37,5 a 23,2 puntos. La diferencia no desaparece y, al estudiar los grupos por separado, es mucho mayor en el sector público que entre particulares pagados. La brecha pública estandarizada alcanza 45,8 puntos, frente a 14,9 en particulares pagados; el contraste entre ambas es de 31,0 puntos, con IC 95 % [7,4; 50,6].

La Metropolitana también presenta una mayor concentración de puntajes altos: su riesgo relativo de superar 900 puntos es 1,98 respecto del resto del país. Sin embargo, esa ventaja no se ensancha en la cola superior del índice obser-

vado. Las diferencias de cuantiles son máximas en torno a la mediana y disminuyen hacia P99, y los contrafactuales de desplazamiento uniforme producirían una separación superior a la observada sobre 900 puntos.

En conjunto, el análisis muestra que la ventaja metropolitana no puede reducirse a una sola explicación composicional y que su magnitud varía según el grupo y la zona de la distribución observada. Las conclusiones son descriptivas: caracterizan a quienes se inscribieron y rindieron la PAES de Invierno bajo estas condiciones, sin identificar efectos causales del territorio, del establecimiento o de la dependencia administrativa. La replicación del estudio en nuevos procesos permitirá evaluar qué componentes del patrón son estables y cuáles corresponden a particularidades de esta cohorte y esta aplicación.

Referencias

- [1] Datos Abiertos DEMRE. Archivo C, Datos de Rendición de las PAES de Invierno del Proceso de Admisión 2027. <https://portal-transparencia.demre.cl/portal-base-datos>
- [2] Efron, B. (1987). Better Bootstrap Confidence Intervals. *Journal of the American Statistical Association*, 82(397), 171–185.
- [3] Observatorio PAES - Preuniversitario Gauss. Cuaderno reproducible del análisis regional de la PAES de Invierno, Admisión 2027. Referencia web por completar en la versión de publicación.