

**UNIVERSIDAD DEL CEMA
Buenos Aires
Argentina**

Serie
DOCUMENTOS DE TRABAJO

Área: Economía

**ENTENDIENDO MEJOR LOS RANKINGS GLOBALES
DE UNIVERSIDADES**

Gustavo Ferro y Juan Dip

**Noviembre 2025
Nro. 911**

https://ucema.edu.ar/publicaciones/doc_trabajo.php
UCEMA: Av. Córdoba 374, C1054AAP Buenos Aires, Argentina
ISSN 1668-4575 (impreso), ISSN 1668-4583 (en línea)
Editor: Jorge M. Streb; Coordinador del Departamento de Investigaciones: Maximiliano Ivickas

Entendiendo mejor los rankings globales de universidades

Gustavo Ferro¹ y Juan Dip²

Resumen

Se utilizan modelos avanzados de análisis de eficiencia para decodificar los pesos implícitos de los tres índices compuestos globales que rankean universidades: *Academic Ranking of World Universities* (ARWU), *Quacquarelli Symonds* (QS) y *Times Higher Education* (THE), tras una discusión de la literatura sobre los rankings de universidades, sus aportes y sus problemas. La metodología central empleada por Ferro & Dip (2025), cuyos resultados se reseñan aquí, se basa en los modelos Benefit-of-the-Doubt (BoD), que, a su vez, derivan del Análisis Envolvente de Datos (DEA). Las principales contribuciones analíticas del estudio giran en torno a la aplicación de metodologías avanzadas de eficiencia para decodificar las estructuras de peso implícitas en los rankings (ARWU, QS y THE) y generar diagnósticos detallados y útiles para la mejora institucional. El análisis revela discrepancias entre los pesos oficiales asignados por las agencias de ranking y los pesos derivados de los modelos BoD (pesos implícitos de importancia).

Abstract

Advanced efficiency analysis models are used to decode the implicit weights of the three global composite indices that rank universities: the Academic Ranking of World Universities (ARWU), Quacquarelli Symonds (QS), and Times Higher Education (THE), following a discussion of the literature on university rankings, their contributions, and their problems. The methodology employed by Ferro & Dip (2025), whose results are summarized here, is based on Benefit-of-the-Doubt (BoD) models. These are derived from Data Envelopment Analysis (DEA). The study applies efficiency methodologies to decipher the implicit weight structures in the rankings (ARWU, QS, and THE) and provides detailed, actionable diagnoses for institutional improvement. The analysis reveals discrepancies between the official weights assigned by the ranking agencies and the weights derived from the BoD models (which are implicit importance weights).

1. Introducción

El primer antecedente de clasificación de universidades se produjo en Estados Unidos en 1888, basándose en la afiliación de científicos destacados entre sus profesores y en el estatus social de su alumnado (Webster, 1986). Sin embargo, la proliferación de clasificaciones globales en forma de rankings es un fenómeno más reciente (Stack, 2021). El ranking de Shanghai (ARWU) fue el primero en establecerse, impulsado por la Universidad Jiao Tong de China. Su objetivo inicial fue estratégico: identificar universidades de clase mundial y medir la brecha con las instituciones chinas, priorizando el rendimiento académico y de investigación (Wu & Liu, 2018). Times Higher Education Supplement (THE) se asoció con Quacquarelli Symonds (QS) en 2004 para lanzar su clasificación, motivada en parte por la necesidad del gobierno del Reino Unido de impulsar los vínculos entre la universidad y la industria y de crear una tabla de élite mundial (Stack, 2016; Baty, 2009). THE y QS se separaron en 2009/2010. THE desarrolló una metodología propia, destacando la relación universidad-industria (Ross, 2018; Cantú-Ortiz & Fangmeyer Jr., 2018). QS mantuvo y refinó la metodología original, con

¹ Universidad del CEMA (UCEMA) y CONICET. Email: gferro05@yahoo.com.ar; gaf97@ucema.edu.ar. Avenida Córdoba 374 (CP 1044), Buenos Aires (Argentina), teléfono: +54-11-63143000, extensión 780. Los puntos de vista de los autores son personales, no representan la posición de la Universidad del CEMA.

² Universidad Nacional de Misiones (UNaM). Facultad de Ciencias Económicas. Email: juan.dip@fce.unam.edu.ar. Av. Fernando Elías Llamosas 9458 Posadas (CP N3304) Misiones, teléfono: +54 (376) 4480006

énfasis en los indicadores de empleabilidad, y adoptó una filosofía "orientada al estudiante", en la que, en sus rankings, cobran gran importancia las encuestas dirigidas a referentes sectoriales y empleadores (Sowter et al., 2018; Cantú-Ortiz & Fangmeyer Jr., 2018).

Con el objetivo de decodificar los pesos implícitos de los tres índices compuestos globales que rankean universidades: *Academic Ranking of World Universities* (ARWU), *Quacquarelli Symonds* (QS) y *Times Higher Education* (THE), se emplean métodos avanzados basados en indicadores de eficiencia. Tras una discusión sobre la literatura sobre los rankings de universidades, sus aportes y sus problemas, se reseñan los resultados obtenidos por Ferro y Dip (2025). Su metodología central se basa en los modelos Benefit-of-the-Doubt (BoD), que, a su vez, derivan del Análisis Envolvente de Datos (DEA).

Las principales contribuciones analíticas del estudio giran en torno a la aplicación de metodologías avanzadas de eficiencia para decodificar las estructuras de peso implícitas en los rankings (ARWU, QS y THE) y generar diagnósticos detallados y útiles para la mejora institucional. El análisis revela discrepancias entre los pesos oficiales asignados por las agencias de ranking y los pesos derivados de los modelos BoD (pesos implícitos de importancia).

El Análisis Envolvente de Datos (DEA) es una técnica no paramétrica utilizada para medir la eficiencia relativa de "Unidades de Toma de Decisiones" (DMUs), que en este estudio son las universidades. No asume una forma funcional predefinida entre insumos y productos. En su lugar, DEA construye una "frontera de eficiencia" a partir de las DMUs de mejor desempeño y la eficiencia de todas las demás unidades se mide en relación con dicha frontera.

El Modelo Benefit-of-the-Doubt (BoD) es una variante de DEA diseñada para la construcción de indicadores compuestos. El principio fundamental del BoD otorga a cada universidad el "beneficio de la duda". Esto significa que, al calcular su puntaje de desempeño general (o índice compuesto) para cada universidad, se le permite elegir el conjunto de ponderaciones (pesos) que maximice su propia eficiencia relativa, siempre y cuando dichas ponderaciones sean factibles para todas las demás unidades de la muestra. Matemáticamente, el BoD busca maximizar la suma ponderada de los indicadores (productos deseables, como publicaciones o la internacionalización), manteniendo fijos los insumos. El resultado es un índice compuesto de eficiencia que se sitúa entre 0 (peor desempeño) y 1 (mejor desempeño). Este enfoque evita la imposición de ponderaciones arbitrarias y refleja las fortalezas específicas de cada unidad.

Un problema común en el BoD estándar es el fenómeno de "peso cero", en el que ciertos indicadores se ignoran si no contribuyen a la máxima eficiencia de la unidad evaluada. Para contrarrestar esto, se utiliza el modelo BoD CONST (BoD con restricciones sobre los pesos). En BoD CONST se aplican restricciones a las ponderaciones para asegurar que reflejen prioridades sustantivas y resulten más realistas. Por ejemplo, se puede establecer un valor mínimo (p. ej., 0.05) y un valor máximo (p. ej., 0.5) para el peso de un subindicador, reduciendo la flexibilidad, pero mejorando la validez del índice compuesto.

El Multi-directional Robust Benefit-of-the-Doubt (MDir_RBoD) es la extensión más avanzada empleada. El modelo MDir_RBoD extiende el marco BoD (Benefit-of-the-Doubt) mediante un enfoque multidireccional y robusto. La característica de no compensación evita que un bajo rendimiento en una métrica (p. ej., calidad de investigación) sea perfectamente compensado por un

alto rendimiento en otra (p. ej., cantidad de publicaciones). La robustez garantiza la confiabilidad y la estabilidad de los resultados ante la presencia de valores atípicos (*outliers*) (Vidoli et al., 2024). La robustez se obtiene mediante un procedimiento de remuestreo (*resampling*) para reducir la influencia de valores atípicos (*outliers*) o de unidades anormales. Esto asegura que la frontera de eficiencia y las mediciones de desempeño sean más estables y confiables. MDir_RBoD proporciona un puntaje de eficiencia general e identifica la "trayectoria de mejora" para cada indicador simple. Mide la proporción en la que debe aumentar cada indicador para que la universidad alcance la frontera de eficiencia³.

Tras esta introducción, la sección 2 resume los antecedentes, la sección 3 formaliza el método, la sección 4 presenta los resultados, la sección 5 concluye.

2. Antecedentes

ARWU ha mantenido una construcción basada casi exclusivamente en indicadores cuantitativos y "datos duros", con relativamente escasos cambios en el método y en las ponderaciones, sin encuestas de reputación (Wu & Liu, 2018). Esta elección, sin embargo, se critica por excluir la literatura abierta o gris de relevancia regional y por limitar la cobertura (Hazelkorn & Mihut, 2021; Olcay & Bulu, 2016). QS y THE utilizan una combinación de datos "duros" y "blandos" (*soft data*), estos últimos basados en encuestas de opinión (Olcay & Bulu, 2017). Estas encuestas son fuertemente criticadas por su alta subjetividad y el riesgo de efecto de halo (en el que la reputación influye en las métricas; Martínez Rizo, 2011; Bowman & Bastedo, 2011). QS depende en un 45% de las encuestas de reputación (Stack, 2021).

La evolución del análisis de las clasificaciones universitarias globales (Rankings) puede entenderse como un debate continuo y cronológico que ha transitado desde la crítica teórica de la jerarquía, pasando por el cuestionamiento metodológico, hasta la decodificación de los pesos reales mediante herramientas avanzadas de eficiencia. Hazelkorn (2015) divide la literatura sobre rankings en preocupaciones metodológicas y de comprensión teórica y subraya que los rankings son útiles para las partes interesadas al simplificar la información y guiar la asignación de recursos y la atracción de talento. Los rankings han reafirmado un entendimiento común sobre la posición jerárquica internacional de las universidades, lo que incide en la posición geopolítica de sus respectivos países (Hazelkorn & Mihut, 2021). Stack (2016) afirma que parte del atractivo de los rankings radica en que refuerzan narrativas de jerarquías de larga data. Conecta los rankings con la gestión de la reputación y el financiamiento gubernamental e industrial. En analogía con otros gigantes de servicios en el mundo corporativo, las tres agencias más grandes (ARWU, QS, THE) son conocidas como "The Big Three". Como consecuencia, se utilizan los rankings para motivar el cambio organizacional, definir políticas y asignar recursos. Locke (2021) conceptualiza seis formas principales en las que las universidades son influenciadas y reaccionan ante los rankings, incluyendo su uso para el

³ El método MDir_RBoD es como un entrenador personalizado y robusto para atletas olímpicos. En lugar de simplemente asignarle al atleta una puntuación general (como un ranking) o un peso fijo para cada disciplina, el MDir_RBoD observa cómo rinden los atletas mejor rankeados ("situados en la frontera de eficiencia"). Luego, identifica exactamente dónde está desequilibrado el rendimiento de un atleta específico y proporciona una dirección de mejora precisa que no permite que su fortaleza compense su debilidad, lo que lo obliga a mejorar en aquello donde está más rezagado. Además, al ser robusto, el entrenador ignora los resultados anómalos o "fuera de escala" de otros atletas, lo que garantiza que el plan de entrenamiento se base en el rendimiento más estable y real de los situados en la frontera.

posicionamiento estratégico, la redefinición de actividades, el impacto en la moral del personal y la promoción de cambios internos.

Con la aparición de los tres rankings principales (ARWU en 2003 y THE/QS, que se separaron en 2009), resultó necesario comprender qué buscaba cada agencia. Los rankings internacionales existentes han sido clasificados en "orientados al estudiante" (incluyendo inicialmente THE-QS y, posteriormente, QS) y "orientados a la investigación" (ARWU) (Aguillo et al., 2010). Cantú-Ortiz & Fangmeyer (2018) señalan que la competencia entre agencias ha llevado a cada una a desarrollar factores distintivos: QS se especializa en la experiencia de los estudiantes y la empleabilidad de los graduados; THE, en la relación universidad-industria; y ARWU, en su construcción puramente cuantitativa. Para Ross (2018), el ranking THE se centra en universidades intensivas en investigación, con un amplio currículo y un rol de enseñanza de grado. La filosofía central del ranking QS es "orientada al estudiante", con el objetivo de fomentar la movilidad internacional y el desarrollo profesional (Sowter et al., 2018). En tanto, el ranking ARWU se originó para identificar universidades de clase mundial y medir la brecha entre las universidades chinas y estas instituciones, priorizando el rendimiento académico y de investigación, así como la disponibilidad de datos comparables (Wu & Liu, 2018).

Se los ha criticado por distintas razones. Si bien los rankings sirven como herramientas de rendición de cuentas y guías para las partes interesadas (estudiantes, padres, responsables políticos), tienden a centrarse en la productividad investigadora, privilegiando los campos STEM y las publicaciones en inglés frente a las humanidades y la literatura gris. También se argumenta que, debido a la dificultad de encontrar métricas más adecuadas, la enseñanza se mide indirectamente mediante *proxies* (Altbach, 2021). Hazelkorn & Mihut (2021) también señalan que los rankings se centran en publicaciones en inglés y excluyen la literatura abierta o gris relevante a nivel regional. Barsky (2012) alerta sobre posibles sesgos geográficos o lingüísticos en los rankings, que pueden favorecer a ciertas universidades o disciplinas científicas en detrimento de otras. Bowman & Bastedo (2011) argumentan que los rankings se están convirtiendo cada vez más en sinónimo de reputación, en lugar de que la reputación sea un indicador independiente de la calidad. También sugieren que los efectos de anclaje (*anchoring effects*), en los que los rankings iniciales influyen en las evaluaciones futuras de la reputación. También Kim & Bastedo (2021) documentan que la relación entre rankings, recursos y reputación a veces es autocorrelacionada. Sobre el particular, Merton (1968, 1988) acuñó el concepto del "Efecto Mateo" (Matthew Effect) o ventaja acumulativa, que describe la tendencia a acumular éxito en proporción directa al nivel inicial, un fenómeno observado en las clasificaciones universitarias. También se ha argumentado que los rankings promueven un modelo único de la universidad de investigación anglosajona, rica y de élite (Ordorika & Lloyd, 2013, 2015). Giacalone (2009) critica los rankings por plantear preocupaciones sobre la construcción de una realidad basada en métricas o en una "metrabilidad", una lógica que podría distorsionar eventualmente la vida académica. En una línea similar a las anteriores, Hamann & Ringel (2023) citan las críticas comunes a los rankings por falta de transparencia, ponderación arbitraria y sesgos a favor de instituciones orientadas a la investigación. Se critica que los pesos asignados a los indicadores sean arbitrarios y que el procedimiento para agregar los resultados parciales se realice sin respetar las reglas metodológicas básicas (Martínez Rizo, 2011; Salmi & Saroyan, 2007). Por ejemplo, en ARWU, ignorar el tamaño institucional en los primeros cinco factores confiere una ventaja considerable a las universidades más grandes (Martínez Rizo, 2011).

Chankseliani (2023) y Marginson (2021) abordan la diversificación de la producción científica, el crecimiento explosivo de las publicaciones globales y el aumento de la colaboración transfronteriza y de la citación. Se han documentado cambios importantes en la dinámica mundial del rendimiento de la investigación, entre los que destaca el aumento significativo de la proporción de universidades chinas en el grupo de élite (Marhl et al., 2025). No obstante, a pesar del cambio en el origen de la producción científica, que pasa de Occidente a Asia, el impacto de las citas se ve influido no solo por la calidad (relevancia, evidencia), sino también por características no científicas como la longitud del resumen, las referencias y el prestigio o la universidad del autor (Mammola et al., 2022). Por otra parte, la búsqueda de mejoras en los rankings produce resultados indeseables. Se han documentado prácticas fraudulentas de investigación en sistemas científicos de alto crecimiento, originadas en regímenes de evaluación institucional que ponen presión sobre los académicos, así como en culturas epistémicas (tolerancia hacia ciertas prácticas) y en cierta influencia de la industria editorial comercial, que incentivan el volumen por encima de la calidad (Wang et al., 2024).

Con respecto a estudios que han indagado sobre la metodología aquí utilizada (Benefit-of-the-Doubt (BoD), basada en el Análisis Envolvente de Datos (DEA)), Cherchye et al. (2007) formalizaron el enfoque Benefit-of-the-Doubt (BoD), que utiliza el marco del DEA para construir índices compuestos al permitir que cada unidad elija las ponderaciones que maximicen su propia eficiencia relativa, evitando ponderaciones arbitrarias. Dip (2023) utilizó el método BoD para analizar qué información proporciona U-Multirank sobre la transferencia de conocimiento e investigación. El Gibari et al. (2022) contribuyeron al uso del BoD en la evaluación universitaria combinando indicadores compuestos basados en puntos de referencia con el análisis DEA. Ferro & D'Elia (2020) identificaron las variables a considerar en el análisis de fronteras de eficiencia en la educación superior. Jain & Gulati (2024) aplicaron el enfoque BoD para analizar el desarrollo de la educación superior en India mediante un índice multidimensional. Johnes (2018) exploró, utilizando rankings universitarios y el método BoD, qué muestran realmente dichas clasificaciones. Matulová (2023) aplicó el método BoD para clasificar universidades europeas mediante un indicador de sostenibilidad basado en DEA. Szuwarzyński (2019) y De Witte & Hudrlikova (2013) aplicaron el enfoque BoD para evaluar el desempeño de la investigación o del docente en universidades. Van Puyenbroeck & Rogge (2017) propusieron un método para la construcción de índices compuestos multiplicativos, como números de índice de cantidad de media geométrica, con ponderaciones BoD, un enfoque utilizado para las restricciones de pesos. Vidoli et al. (2024) introdujeron el modelo Multi-directional Robust Benefit-of-the-Doubt (MDir_RBoD), diseñado para proporcionar una medida de calidad compuesta resiliente a valores atípicos (outliers) e identificar las direcciones de mejora potenciales para cada indicador componente. Los resultados aquí discutidos provienen de Ferro y Dip (2025).

3. Método

Matemáticamente, el indicador compuesto de desempeño I_{c0} , (se obtiene resolviendo el siguiente problema de optimización):

$$I_{c0} = \text{Max}_{w_{0,i}} \sum_{i=1}^m w_{0,i} y_{0,i}$$

$$\text{s.t. } \sum_{i=1}^m w_{0,i} y_{j,i} \leq 1 \quad \text{con } j = 1, \dots, n; \quad w_{0,i} \geq 0, \quad \text{con } i = 1, \dots, m \quad (1)$$

La solución óptima de la ecuación (1) nos proporciona un indicador compuesto que representa una medida de eficiencia para cada DMU (unidad de decisión), donde $0 \leq I_{c0} \leq 1$. Un valor de 0 indica el peor desempeño de la DMU en comparación con otras DMUs de la muestra; por el contrario, un valor de 1 indica que la DMU es completamente eficiente. Por otro lado $w_{0,i}$ es el ponderador al indicador individual i de cada DMU, $y_{0,i}$ es el valor de cada indicador i para la DMU₀, n es el número de DMUs y m es el número de indicadores individuales. La ecuación (1) se resuelve maximizando n veces por cada DMU para obtener un conjunto de I_{c0} para la muestra completa. A este modelo se le llama BOD.

Si cada ponderación óptima se multiplica por el valor de cada indicador, se obtienen las "Participaciones de torta" (*Pie Shares*) $w_{0,i}y_{j,i}$. Estas indican en qué medida la dimensión "i" contribuye al puntaje compuesto de cada universidad. Si se ejecuta el modelo (1) en su estado actual, se debe anticipar con cierta frecuencia la posibilidad de un fenómeno de ponderación cero. Para evitarlo, se aplica la restricción a dichas "participaciones de torta", siguiendo a Van Puyenbroeck & Rogge (2016). Proponen construir CIs a partir de la media geométrica de los índices de cantidad. Ello introduce un procedimiento indirecto en dos etapas para combinar la agregación multiplicativa con los ponderadores BOD. El fundamento es el indicador de cantidad media geométrica, que agrupa razones de subindicadores (valores de DMUs relativos a un estándar de referencia, y_{ri}/y_{rB} usando como pesos de importancia los exponentes (ω_{ri}^*), donde y_{rB} son los valores del indicador base y el valor del subindicador respectivo es y_{ri} para $r = 1, \dots, s$.

$$\text{Entonces, } CI_i = \prod [(y_{ri}/y_{rB})^{\omega_{ri}^*}] \quad (2).$$

Se comienza con un modelo lineal BOD (1) para derivar los ponderadores, resolviendo un modelo lineal estándar BOD para cada universidad. Ello identifica los precios sombra óptimos que maximizan el valor de CI para la unidad evaluada, sujeta a un umbral máximo (habitualmente 100%). Luego, para calcular las participaciones de los subindicadores BOD ("Budget Shares") (ω_{ri}^*), se los estima a partir de los valores de los pesos óptimos (w_{ri}) obtenidos en el primer paso y se completa la analogía con los índices económicos de cantidades, donde los subindicadores se usan como ponderadores:

$$\omega_{ri}^* = (w_{ri}^* * y_{ri}) / \sum (w_{ri}^* * y_{ri}) \quad (3).$$

La suma de dichos subindicadores (ω_{ri}^*) para cada universidad i es igual a 1. El modelo también permite incorporar restricciones sobre los pesos de los subindicadores (por ejemplo, $0.005 \leq \omega_{ri}^*$). Ese modelo restringido se llama BODCONST.

Finalmente, para evaluar la presencia de valores atípicos (*outliers*) y determinar qué tan lejos están las universidades de sus valores objetivo, se sigue a Vidoli et al. (2024). Ellos introducen el modelo Multi-directional Robust Benefit-of-the-Doubt (MDir_RBoD). Su objetivo principal es proporcionar una medida de calidad compuesta resistente a valores atípicos e identificar posibles direcciones de mejora para cada indicador, a nivel individual.

MDir_BoD busca el mayor aumento posible en un solo indicador y_{ri} requerido para alcanzar la frontera de eficiencia, mientras se mantienen fijos otros indicadores. Además, obtenemos el vector direccional, que representa la disparidad entre el valor ideal (frontera) y el valor observado de cada indicador. Esta métrica mide la proporción en la que debe aumentar cada indicador simple para alcanzar la frontera.

Por lo tanto, el modelo MDir_RBoD emplea un esquema de ponderación que favorece a cada unidad y propone una dirección de mejora al identificar subas potenciales, en lugar de simplemente seguir la producción pasada o una dirección específica. Además, mejora este enfoque al compararlo con una frontera robusta frente a valores atípicos y unidades anormales. Se llama a este modelo BODRDIR.

4. Resultados

Las Tablas 1 a 3 muestran los resultados para los tres rankings, cotejándolos con los ponderadores oficiales de cada ranking (para una explicación más detallada, véase Dip & Ferro, 2025)

En cuanto a los resultados de ARWU, los modelos BoD (sin restricciones) indican un dominio significativo de la productividad científica en el ranking. Las ponderaciones implícitas (BoD) muestran que Publications representa el 38,3 % en el Top 10 y alcanza el 59,9 % en el Top 50. A pesar de que los indicadores de calidad de investigación, como Highly Cited Publications y Nature & Science, tienen un peso oficial del 20%, su contribución real en el modelo BoD es marginal, a menudo inferior al 0,5%. Esto sugiere que ARWU sobreestima oficialmente estos indicadores y subestima las publicaciones. El análisis de eficiencia (BoDRMDIR) por indicador muestra que Alumni y Awards presentan las brechas más notables entre la élite y el promedio. La eficiencia promedio de Alumni es del 65,8 % en el Top 10, pero cae al 6,48 % en la muestra completa. Las Publications mantienen un desempeño más equilibrado a lo largo de los estratos.

Tabla 1: Resultados para ARWU (ponderadores oficiales informados de 2024)

Importance Weights – BoD-	Top 10	Top 50	Top 100	ALL	Oficiales
Alumni	8,80%	2,00%	0,90%	0,20%	20,00%
Awards	31,90%	10,20%	4,60%	0,60%	20,00%
Highly Cited Publication	0,00%	0,00%	0,10%	0,00%	20,00%
Nature & Science	0,40%	0,20%	0,10%	0,00%	20,00%
Publications	38,30%	59,90%	55,50%	31,80%	20,00%
Per Capita Academic Performance	13,20%	4,00%	4,50%	3,60%	10,00%
Importance Weights – BoDCONST	Top 10	Top 50	Top 100	ALL	Oficiales
Alumni	17,00%	8,40%	7,45%	6,75%	20,00%
Awards	44,00%	15,20%	12,30%	6,41%	20,00%
Highly Cited Publication	5,00%	12,00%	10,25%	9,85%	20,00%
Nature & Science	7,50%	12,40%	11,90%	8,56%	20,00%
Publications	12,00%	36,80%	39,90%	45,11%	20,00%
Per Capita Academic Performance	14,50%	15,20%	18,20%	23,33%	10,00%
Efficiency scores for each indicator. BoDRMDIR	Top 10	Top 50	Top 100	ALL	Oficiales
Alumni	65,80%	21,40%	14,48%	6,48%	20,00%
Awards	73,02%	23,28%	15,80%	5,39%	20,00%
Highly Cited Publication	62,56%	21,96%	18,39%	13,11%	20,00%
Nature & Science	68,53%	24,04%	20,11%	12,96%	20,00%
Publications	72,05%	40,27%	36,42%	32,96%	20,00%
Per Capita Academic Performance	67,26%	25,46%	21,38%	15,48%	10,00%

Fuente: Dip & Ferro (2025), sobre ARWU.

Los resultados de THE (Times Higher Education) muestran, en sus ponderaciones implícitas (BoD), que Industry Impact (Ingresos de la industria) es el indicador más prominente, contribuyendo con un 79,3 % en el Top 10, una cifra significativamente mayor que su peso oficial del 4 %. Se observa una subestimación extrema del Industry Impact y una sobreestimación de la Teaching (Enseñanza), que tiene un peso oficial del 29,5 %, pero una contribución real inferior al 1 % en el modelo BoD no restringido. En cuanto a la eficiencia (BoDRMDIR), las universidades de élite (Top 10) alcanzan una eficiencia cercana a la perfección (>0.97) en todas las categorías. Destacan en Research Quality (98.59%) e Industry Impact (97.00%).

Tabla 2: Resultados para THE (ponderadores oficiales informados de 2024)

Importance Weights – BoD-	Top 10	Top50	Top 100	All	Oficiales
Teaching	0.15%	0.10%	0.06%	0.72%	29.50%
Research Environment	7.63%	4.34%	3.05%	0.27%	29.00%
Research Quality	6.45%	7.09%	14.28%	22.72%	30.00%
Industry Impact	79.29%	71.94%	57.13%	19.30%	4.00%
International Outlook	6.43%	16.07%	24.22%	20.94%	7.50%
Importance Weights – BoDCONST-	Top 10	Top 50	Top 100	All	Oficiales
Teaching	7.18%	6.10%	5.71%	7.08%	29.50%
Research Environment	10.08%	7.57%	6.66%	5.23%	29.00%
Research Quality	29.50%	28.40%	30.11%	38.65%	30.00%
Industry Impact	34.93%	40.29%	39.89%	22.70%	4.00%
International Outlook	18.31%	17.65%	17.63%	26.34%	7.50%
Efficiency scores for each indicator. BoDRMDIR	TOP 10	TOP 50	TOP 100	ALL	Oficiales
Teaching	98.18%	78.97%	63.68%	25.07%	29.50%
Research Environment	98.42%	81.29%	67.30%	22.36%	29.00%
Research Quality	98.59%	94.82%	90.83%	52.57%	30.00%
Industry Impact	97.00%	91.98%	88.04%	38.87%	4.00%
International Outlook	97.63%	94.98%	89.10%	42.98%	7.50%

Fuente: Dip & Ferro (2025), sobre THE.

Los resultados de QS (Quacquarelli Symonds), un ranking centrado en el estudiante, sobrevaloran la internacionalización y las nuevas dimensiones, como la sostenibilidad. Las ponderaciones implícitas (BoD) indican que los indicadores de internacionalización (International Faculty e International Students) son los principales motores de eficiencia y, colectivamente, contribuyen con un 60,16 % en el Top 10, a pesar de tener un peso oficial modesto (10 % en total). El modelo BoD concluye que QS subestima sistemáticamente la internacionalización, mientras que sobreestima componentes tradicionales como la Academic Reputation (45% oficial vs. 7,84% real en la muestra completa). El modelo BoDCONST resalta la importancia creciente de Sustainability (15,38 % vs. 5 % oficial) y de International Research Network (16,04 % vs. 5 % oficial) para la élite. En cuanto a eficiencia (BoDRMDIR), International Faculty (99.87%) y Employer Reputation (99.87%) son los indicadores más sólidos para diferenciar a las universidades de primer nivel en el ranking de QS.

Tabla 3: Resultados para QS (ponderadores oficiales de 2024)

Importance Weights – BoD-	Top 10	Top 50	Top 100	All	Oficiales
Academic Reputation	1.77%	3.36%	2.39%	1.66%	30.00%
Employer Reputation	0.92%	2.64%	4.51%	6.18%	15.00%
Faculty / Student	12.87%	13.88%	12.54%	11.84%	10.00%
Citations	7.35%	11.14%	11.61%	11.52%	20.00%
International Faculty	42.36%	34.57%	30.47%	18.95%	5.00%
International Students	17.80%	14.65%	13.92%	9.50%	5.00%
International Research	5.17%	6.51%	8.24%	11.78%	5.00%
Employers	2.12%	1.98%	2.16%	1.85%	5.00%
Sustainability	9.60%	10.90%	13.05%	12.98%	5.00%
Importance Weights – BoDCONST-	Top 10	Top 50	Top 100	All	Oficiales
Academic Reputation	14.15%	9.92%	8.02%	6.35%	30.00%
Employer Reputation	5.00%	5.57%	6.38%	9.00%	15.00%
Faculty / Student	5.77%	8.77%	8.45%	11.90%	10.00%
Citations	13.82%	12.81%	12.13%	13.51%	20.00%
International Faculty	10.86%	14.44%	16.81%	15.07%	5.00%
International Students	12.16%	16.23%	13.99%	11.25%	5.00%
International Research	16.04%	10.51%	11.41%	12.73%	5.00%
Employers	6.81%	5.88%	5.55%	6.22%	5.00%
Sustainability	15.38%	15.88%	17.27%	13.98%	5.00%
Efficiency scores for each indicator. BoDRMDIR	Top 10	Top 50	To 100	All	Oficiales
Academic Reputation	99.84%	98.32%	83.37%	40.38%	30.00%
Employer Reputation	99.26%	95.95%	79.39%	40.29%	15.00%
Faculty / Student	98.60%	93.75%	76.60%	44.70%	10.00%
Citations	99.51%	96.87%	86.80%	49.41%	20.00%
International Faculty	99.87%	94.38%	92.19%	53.14%	5.00%
International Students	99.34%	95.09%	88.48%	44.38%	5.00%
International Research	98.75%	94.52%	89.83%	50.74%	5.00%
Employers	99.87%	95.12%	77.49%	34.10%	5.00%
Sustainability	99.44%	96.34%	92.37%	52.78%	5.00%

Fuente: Dip & Ferro (2025), sobre QS.

Para evaluar la consistencia entre los rankings oficiales y los derivados de las metodologías de eficiencia, se calculó la correlación de rangos de Spearman. El modelo BoDCONST muestra correlaciones casi perfectas, especialmente con THE (0,98 %), lo que sugiere que las restricciones de pesos ayudan al modelo a ajustarse a los rankings basados en la reputación y la calidad de la investigación. Por otro lado, MDir_RBoD presenta una correlación muy débil con ARWU (0,18 %), lo que podría indicar que esta metodología, al centrarse en direcciones de mejora y no en la compensación, refleja mejor la diversidad de misiones universitarias más allá de un enfoque puramente bibliométrico.

Tabla 4: Correlaciones entre resultados

Método	ARWU	THE	QS
BoD	0,48	0,91	0,55
BoDCONST	0,91	0,98	0,86
BoDRMULTI	0,18	0,96	0,91

Fuente: Dip & Ferro (2025), sobre ARWU.

El MDir_RBoD identifica cuánto debe mejorar cada indicador para alcanzar la frontera de eficiencia. Un valor de 0.30 indica que se requiere una mejora del 30%. El uso de la metodología MDir_RBoD permite un diagnóstico granular que trasciende la puntuación compuesta, proporcionando a las universidades información precisa sobre los indicadores específicos en los que deben enfocarse para lograr mejoras estratégicas y medibles en su clasificación global.

Por ejemplo, en el caso de ARWU, universidades como Oxford y Stanford muestran la necesidad de mejoras significativas (30–45 %) en indicadores como *Alumni* y *Per Capita Performance (PCPD)*. Las universidades de América Latina, como la Universidad de Buenos Aires (UBA) o la Pontificia Universidad Católica de Chile, presentan brechas sustanciales en Awards, Alumni e Highly Cited Researchers, lo que a menudo requiere duplicar sus puntajes. En el caso de THE, por ejemplo, las universidades latinoamericanas (como la UNAM o la USP) presentan mayores ineficiencias en *Teaching*, *Research Environment* y *Research Quality*, con mejoras requeridas que oscilan entre el 40% y el 90%. En contraste, las instituciones de élite global muestran ineficiencias mínimas (cerca de cero).

5. Conclusiones

En esencia, la discrepancia entre el peso oficial y el peso implícito (BoD) demuestra que las agencias de ranking a menudo subestiman los verdaderos pesos que definen la frontera de eficiencia. El uso del MDir_RBoD, al proporcionar un diagnóstico detallado que cuantifica la proporción en la que debe mejorar cada indicador, ofrece una herramienta de acción para las administraciones universitarias que deseen mejorar sus posiciones, trascendiendo la simple puntuación compuesta.

El análisis proporciona benchmarks prácticos y recomendaciones estratégicas para universidades y responsables políticos, facilitando la comprensión de la eficiencia y la mejora estratégica en las clasificaciones. Los resultados sugieren que las estrategias de mejora deben diferenciarse según el estrato institucional. Las universidades de élite deben enfocarse en mantener ventajas en el impacto industrial y en la calidad de la investigación (para THE) y en la reputación e internacionalización (para QS), ya que muestran ineficiencias mínimas. En tanto, las universidades que están más abajo en los rankings se beneficiarían al centrarse en la internacionalización y en mejoras selectivas en la calidad de la investigación, ya que estas ofrecen ganancias de eficiencia más alcanzables.

El BoDRMDIR identificó brechas significativas en los indicadores tradicionales de ARWU (Alumni, Awards) fuera del Top 50. En el ranking THE, las universidades latinoamericanas (UNAM, USP, etc.) mostraron brechas considerables que requieren mejoras de entre el 40% y el 90% en áreas como Teaching, Research Environment y Research Quality.

Referencias

- Aguillo, I., Bar-Ilan, J., Levene, M., & Ortega, J. (2010). Comparing university rankings. *Scientometrics*, 85(1), 243–256. <https://doi.org/10.1007/s11192-010-0190-z>
- Altbach, P. G. (2009). China and India: A steep climb to world-class universities. In J. A. Douglass, C. J. King & I. Feller (Eds.), *Globalization's muse: Universities and higher education systems in a changing world* (pp. 205–226). Berkeley Public Policy Press.
- Altbach, P. G. (2021). Foreword: reflections on rankings. In E. Hazelkorn & G. Mihut (Eds.), *Research Handbook on University Rankings. Theory, Methodology, Influence and Impact*. Elgar Handbook in Education. Edward Elgar Publishing.
- Baty, P. (2009). Rankings 09: Talking Points. The 2009 world ranking methodology. *The Times Higher Education*.
- Baty, P. (2014). Matter of opinion? Try matter of fact. *Times Higher Education*, (2142), 6–8.
- Barsky, O. (2012). Acerca de los rankings internacionales de las universidades y su repercusión en Argentina. *Debate Universitario*, 11, ISSN: 2314–1530.
- Bowman, N. A. & Bastedo, M. N. (2011). Anchoring effects in world university rankings: exploring biases in reputation scores. *Higher Education*, 61(4), 431–444. <https://doi.org/10.1007/s10734-010-9339-1>
- Cantú-Ortiz, F. J. & Fangmeyer Jr., J. (2018). University Performance in the Age of Research Analytics. In F. J. Cantú-Ortiz (Ed.), *Research Analytics: Boosting University Productivity and Competitiveness through Scientometrics*. CRC Press Taylor & Francis Group.
- Cherchye, L., Moesen, W., Rogge, N., & Van Puyenbroeck, T. (2007). An introduction to 'Benefit of the doubt' composite indicators. *Social Indicators Research*, 82, 111–145.
- Dip, J.A. (2021). What does U-Multirank tell us about knowledge transfer and research? *Scientometrics*, 126, 3011–3039.
- Dip, J. A. & Ferro, G. (2025). *The three global rankings of universities. The why, the how, and the what are they informing?* LX Reunión Anual de la Asociación Argentina de Economía Política (AAEP), UNRC.
- Espeland, W. N. & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, 113(1), 1–40. <https://doi.org/10.1086/517897>
- Espeland, W. N. & Sauder, M. (2009). The discipline of rankings: Tight coupling and organizational change. *American Sociological Review*, 74, 63–82.
- Giacalone, R. A. (2009). Academic rankings in research institutions: A case of skewed mind-sets and professional amnesia. *Academy of Management Learning & Education*, 8(1), 122–126.
- Hamann, J. & Ringel, L. (2023). The discursive resilience of university rankings. *Higher Education*, 86, 845–863.
- Hazelkorn, E. (2015). *Rankings and the Reshaping of Higher Education. The Battle for World-Class Excellence* (2nd ed.). Palgrave Macmillan.

Hazelkorn, E. & Mihut, G. (2021). Introduction: putting rankings in context -looking back, looking forward. In E. Hazelkorn & G. Mihut, *Research Handbook on University Rankings. Theory, Methodology, Influence and Impact*. Elgar Handbook in Education. Edward Elgar Publishing.

Johnes, J. (2018). University rankings: What do they really show? *Scientometrics*, 115, 585–606.

Locke, W. (2021). Researching and understanding the influence of rankings on higher education institutions: logics, methodologies and conceptualisations. In E. Hazelkorn & G. Mihut (Eds.), *Research Handbook on University Rankings. Theory, Methodology, Influence and Impact*. Elgar Handbook in Education. Edward Elgar Publishing.

Martínez Rizo, F. (2011). Los rankings de universidades: una visión crítica. *Revista de la Educación Superior*, 40(1, 157), 77–9

Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810), 56–63.,

Merton, R. K. (1988). The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. *ISIS*, 79(4), 606–623.

Moed, H. F. (2006). Bibliometric rankings of world universities. CWTS Report, 1. <https://pdfs.semanticscholar.org/352b/494e45c98e1faea4d5300a1aa0632517f213.pdf>.

Olcay, G. A., & Bulu, M. (2017). Is measuring the knowledge creation of universities possible? A review of university rankings. *Technological Forecasting and Social Change*, 123, 153–160. <https://doi.org/10.1016/j.techfore.2016.03.029>

Ordorika, I. & Lloyd, M. (2013). A decade of international university rankings: a critical perspective from Latin America. In P. T. M. Marope, P. J. Wells & E. Hazelkorn (Eds.), *Rankings and Accountability in Higher Education Uses and Misuses* (pp. 209–231). UNESCO Publishing.

Ordorika, I. & Lloyd, M. (2015). International rankings and the contest for university hegemony. *Journal of Education Policy*, 30(3), 385–405. <https://doi.org/10.1080/02680939.2014.979247>

Ordorika, I. & Rodríguez Gómez, R. (2010). El ranking Times en el mercado del prestigio universitario. *Perfiles Educativos*, 32(129), 8–25.

Ross, D. (2018). Times Higher Education World Universities Ranking. In F. J. Cantú-Ortiz (Editor), *Research Analytics Boosting University Productivity and Competitiveness through Scientometrics*. CRC Press Taylor & Francis Group.

Salmi, J. & Saroyan, A. (2007). League tables as policy instruments: Uses and misuses. *Higher Education Management and Policy*, 19, 1–39.

Sowter, B., Reggio, D., & Hijazi, S. (2018). QS World University Rankings. In F. J. Cantú-Ortiz (Editor), *Research Analytics Boosting University Productivity and Competitiveness through Scientometrics*. CRC Press Taylor & Francis Group.

Stack, M. (2016). *Global University Rankings and the Mediatization of Higher Education*. Palgrave Macmillan.

Szuwarzyński, A. (2019). Benefit of the doubt approach to assessing the research performance of Australian universities. *Higher Education Quarterly*, 73(2), 235–250. <https://doi.org/10.1111/hequ.12184>

THE (2018). World Academic Summit 2018 Programme. Recuperado de <http://web.archive.org/web/20190120213916/http://www.theworldsummitseries.com/events/the-world-academic-summit-2018/agenda-959e1e6cf1da41e0a093350cec4d7fdd.aspx>

Vessuri, H., Guédon, J., & Cetto, A. (2013). Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Current Sociology*, 62(5), 647–665.

Vidoli, F., Fusco, E., Pignataro, G., & Guccio, C. (2024). Multi-directional Robust Benefit of the Doubt model: An application to the measurement of the quality of acute care services in OECD countries. *Socio-Economic Planning Sciences*, 93, 101877.

Wu, Y. & Liu, N. C. (2018). Academic Ranking of World Universities (ARWU): Methodologies and Trends. In F. J. Cantú-Ortiz (Editor), *Research Analytics Boosting University Productivity and Competitiveness through Scientometrics*. CRC Press Taylor & Francis Group.