FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

UNIVERSIDAD DE MÁLAGA

FACULTAD DE COMERCIO Y GESTIÓN

Doctoral Thesis

Building composite indicators from a multicriteria approach: An empirical application for the performance appraisal and efficiency of the Spanish public higher education system

Author:
Samira El Gibari Ben Said

Supervisors:
Dr. Trinidad Gómez Núñez
Dr. Francisco Ruiz de la Rúa

Programa de Doctorado en Economía y Empresa
University of Málaga

Málaga 2020

AUTOR: Samira El Gibari Ben Said

http://orcid.org/0000-0003-0763-7314

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga

Statement on authorship and originality of the doctoral thesis

EL GIB
BEN SA
SAMIR
79360

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña SAMIRA EL GIBARI BEN SAID
Estudiante del programa de doctorado EN ECONOMÍA Y EMPRESA de la Universidad de Málaga, autor/a de la tesis, presentada para la obtención del título de doctor por la Universidad de Málaga, titulada: BUILDING COMPOSITE INDICATORS FROM A MULTICRITERIA APPROACH: AN EMPIRICAL APPLICATION FOR THE PERFORMANCE APPRAISAL AND EFFICIENCY OF THE SPANISH PUBLIC HIGHER EDUCATION SYSTEM

Realizada bajo la tutorización de FRANCISCO RUIZ DE LA RÚA y dirección de TRINIDAD GÓMEZ NÚÑEZ Y FRANCISCO RUIZ DE LA RÚA

DECLARO QUE:

La tesis presentada es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, conforme al ordenamiento jurídico vigente (Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), modificado por la Ley 2/2019, de 1 de marzo.

Igualmente asumo, ante a la Universidad de Málaga y ante cualquier otra instancia, la responsabilidad que pudiera derivarse en caso de plagio de contenidos en la tesis presentada, conforme al ordenamiento jurídico vigente.

En Málaga, a 23 de MAYO de 2020

Fdo.: SAMIRA EL GIBARI

Authorizations for the dissertation of the doctoral thesis

UNIVERSIDAD DE MÁLAGA

INFORME DE LOS DIRECTORES PARA LA AUTORIZACIÓN DE DEFENSA DE TESIS DOCTORAL

D.ª Trinidad Gómez Núñez y D. Francisco Ruiz de la Rúa, directores de la Tesis Doctoral de D.ª Samira El Gibari Ben Said, **informan favorablemente** la solicitud de autorización de defensa de la tesis con el título "*Building composite indicators from a multicriteria approach: An empirical application for the performance appraisal and efficiency of the Spanish public higher education system", presentada por dicha doctoranda.*

La citada tesis, que opta por la mención internacional, ha sido realizada por compendio de publicaciones y reúne todos los requisitos exigidos por el Programa de Doctorado en Economía y Empresa de la Universidad de Málaga. A este respecto, consideramos que la mencionada tesis supone una contribución de indudable relevancia científica, tanto en el ámbito Multicriterio, como en el campo de la Economía y de la Educación.

RUIZ DE
FRANCIS
2509347

Y para que así conste y tenga los efectos oportunos, en Málaga a 18 de Mayo de 2020.

Fdo. Trinidad Gómez
Catedrática del Departamento de
Economía Aplicada (Matemáticas)
Universidad de Málaga

Fdo. Francisco Ruiz
Catedrático del Departamento de
Economía Aplicada (Matemáticas)
Universidad de Málaga

UNIVERSIDAD DE MÁLAGA

INFORME DEL TUTOR PARA LA AUTORIZACIÓN DE DEFENSA DE TESIS DOCTORAL

D. Francisco Ruiz de la Rúa, tutor de la Tesis Doctoral de D.ª Samira El Gibari Ben Said, en el programa de doctorado en Economía y Empresa de la Universidad de Málaga, informa de que:

Dado que codirige la tesis de la doctoranda, con el título "*Building composite indicators from a multicriteria approach: An empirical application for the performance appraisal and efficiency of the Spanish public higher education system*", junto con D.ª Trinidad Gómez Núñez, se remite al informe elaborado por los dos, que se adjunta también a la documentación.

Y para que así conste y tenga los efectos oportunos, en Málaga a 18 de Mayo de 2020.

Fdo. Francisco Ruiz
Catedrático del Departamento de
Economía Aplicada (Matemáticas)
Universidad de Málaga

Certificate of stay at a foreign University

CERTIFICADO DE ESTANCIA EN UNIVERSIDAD O CENTRO DE INVESTIGACIÓN EXTRANJERO PARA LA OBTENCIÓN DE LA MENCIÓN INTERNACIONAL y PARA TESIS REALIZADAS EN RÉGIMEN DE COTUTELA

A CUMPLIMENTAR POR LA UNIVERSIDAD O CENTRO DE INVESTIGACIÓN EXTRANJERO DONDE EL DOCTORANDO REALIZÓ LA ESTANCIA:

APELLIDOS: de Witte	NOMBRE: Kristof
CORREO ELECTRÓNICO: kristof.dewitte@kuleuven.be	
UNIVERSIDAD / CENTRO DE INVESTIGACIÓN: Katholieke Universiteit Leuven	
VINCULACIÓN CON LA UNIVERSIDAD / CENTRO DE INVESTIGACIÓN: Professor at the Faculty of Economics and Business at KU Leuven and director of the research group "Leuven Economics of Education Research"	

CERTIFICA QUE EL DOCTORANDO CUYOS DATOS FIGURAN A CONTINUACIÓN HA REALIZADO UNA ESTANCIA EN LA UNIVERSIDAD O CENTRO DE INVESTIGACIÓN MENCIONADO ANTERIORMENTE, REALIZANDO ACTIVIDADES DE FORMACIÓN E INVESTIGACIÓN RELACIONADOS CON LA ELABORACIÓN DE SU TESIS DOCTORAL:

APELLIDOS: El Gibari	NOMBRE: Samira
D.N.I./PASAPORTE: 79360488-T	
CORREO ELECTRÓNICO: elgsamira@uma.es	
FECHA DE INICIO DE LA ESTANCIA: 01/03/2020	
FECHA DE FIN DE LA ESTANCIA: 31/05/2020	

En Leuven, a 29 de Mayo de 2020

KU LEUVEN
FACULTY OF ECONOMICS AND BUSINESS
Naamsestraat 69 - Box 3500
3000 LEUVEN, BELGIUM

Fdo:Kristof de Witte

Authorization of co-authors for the inclusion of articles

UNIVERSIDAD DE MÁLAGA

Authorization of co-authors for the inclusion of articles in a doctoral thesis made by publications.

The doctoral thesis developed by *Samira El Gibari Ben Said*, with title *Building composite indicators from a multicriteria approach: An empirical application for the performance appraisal and efficiency of the Spanish public higher education system*, is presented by compendium of articles.

Mr. *Trinidad Gómez Núñez appears as co-author of the publications listed below:*

- Building composite indicators using multicriteria methods: a review.
- MRP-WSCI: Multiple reference point based weak and strong composite indicators.
- Evaluating university performance using reference point based composite indicators.

And so declares his acceptance to the inclusion of these articles in the aforementioned doctoral thesis, together with his resignation to use the same articles as part of another thesis.

Málaga, on 13th May, 2020.

Signed: *Trinidad Gómez Núñez*

UNIVERSIDAD DE MÁLAGA

Authorization of co-authors for the inclusion of articles in a doctoral thesis made by publications.

The doctoral thesis developed by *Samira El Gibari Ben Said*, with title *Building composite indicators from a multicriteria approach: An empirical application for the performance appraisal and efficiency of the Spanish public higher education system*, is presented by compendium of articles.

Mr. *Francisco Ruiz de la Rúa appears as co-author of the publications listed below:*

- Building composite indicators using multicriteria methods: a review.
- MRP-WSCI: Multiple reference point based weak and strong composite indicators.
- Evaluating university performance using reference point based composite indicators.

And so declares his acceptance to the inclusion of these articles in the aforementioned doctoral thesis, together with his resignation to use the same articles as part of another thesis.

Málaga, on 17th May, 2020.

Signed: *Francisco Ruiz de la Rúa*

UNIVERSIDAD DE MÁLAGA

Authorization of co-authors for the inclusion of articles in a doctoral thesis made by publications.

The doctoral thesis developed by *Samira El Gibari Ben Said*, with title *Building composite indicators from a multicriteria approach: An empirical application for the performance appraisal and efficiency of the Spanish public higher education system*, is presented by compendium of articles.

Mr. *José Manuel Cabello González appears as co-author of the publications listed below:*

- MRP-WSCI: Multiple reference point based weak and strong composite indicators.

And so declares his acceptance to the inclusion of these articles in the aforementioned doctoral thesis, together with his resignation to use the same articles as part of another thesis.

Málaga, on 10th May, 2020.

Signed: *José Manuel Cabello*

Acknowledgments

Undertaking this PhD would not have been possible to do without the support and encouragement that I received from many people. First, I wish to express my deepest acknowledgement of gratitude to my supervisors, *Dr. Trinidad Gómez* and *Dr. Francisco Ruiz*. Without their support, encouragement, guidance and constant feedback, this doctoral thesis would not have been achievable. It has been a real pleasure to learn from their extensive knowledge and experiences, not only professionally but personally as well.

I am also very grateful to all my colleagues at the Department of Applied Economics (Mathematics), especially to the director, *Dr. Rafael Caballero* for his invaluable advice and feedback on my work and, of course, for always being so supportive and helpful. Also, I am especially grateful to *Laura Delgado* and *Francisco Martos*, who helped me in numerous ways during various stages of my thesis (from restarting my "hung" computer for so many experiments to drowning our sorrows about how hard is carrying out a PhD). And of course a big thank you to *Sandra González* for the unforgettable MCDM2019 experience, in Istanbul.

Many thanks also to *Dr. Kristof de Witte* and *Dr. Giovanna d'Inverno* from the *Katholieke Universiteit Leuven*, who despite the bizarre and difficult situation of the COVID-19, they were always available to help me and support me in whatever it takes during this challenging period.

I would also like to say a thank you to the external reviewers and the PhD committee members: *Dr. Jian-Bo Yang*, *Dr. María Amparo León*, *Dr. Dylan Jones*, *Dr. Carlos Romero*, *Dr. Carmen Pérez*, *Dr. María Molinos*, *Dr. Miguel Ángel Hinojosa*, *Dr. Giovanna d'Inverno* and *Dr. María del Mar Muñoz*. I am also grateful to Dr. *Dr. Juan Hernández*, who has put his valuable wisdom at my disposal.

Last but not least, I can not let this opportunity pass without saying a very special thank you to my husband, *Othman Britel*, for his patience, and for always believing in me and encouraging me to follow my dreams. Also, special appreciations to the rest of my family: my parents, my sisters and brothers, my aunts and, of course, my nieces, nephews and my cousin.

الحمد لله الذي علم بالقلم، علم الإنسان ما لم يعلم، الحمد لله الذي أعانني على إنجاز رسالة الدكتوراه هذه. و بعد حمد الله تعالى و شكره على إنهائي لهذا العمل، أتقدم بخالص الشكر و عظيم الامتنان إلى عائلتي. خاصة لأبي الحنون، السيد عبد الله الجباري، و لأمي الغالية، السيدة صفية بنسعيد، اللذين غرسا في حب العلم من الصغر، و قدما لي كل غالي و نفيس، و كان لهما الفضل بعد الله فيما وصلت إليه الآن. فلا أملك إلا الدعاء لهما بطول العمر و حسن العمل و بلوغ الجنان.

Funding acknowledgments

Finally, I would like to thank the financial support received from the Spanish Ministry of Economy and Competitiveness (Project ECO2016-76567-C4-4-R and the corresponding predoctoral contract BES-2017-079743), from the Spanish Ministry of Science and Innovation (Project PID2019-104263RB-C42), from the Regional Government of Andalucía (Research group SEJ-417), and from the FEDER-UMA (Project UMA18-FEDERJA-065).

To all of you, a huge heartfelt thanks!

List of acronyms and abbreviations

- AHP: Analytic Hierarchy Process
- ARWU: Academic Ranking of World Universities
- BoD: Benefit of the Doubt
- CA: Conjoint Analysis
- CI: Composite indicator
- CRUE: The Spanish University Rectors' Conference
- CYD Ranking: Conocimiento y Desarrollo Ranking
- DEA: Data Envelopment Analysis
- DMU: Decision Making Unit
- EHEA: European Higher Education Area
- ELECTRE: Elimination and Choice Expressing Reality
- EOCI; End of Childhood Index
- ERA: European Research Area
- EU-SPI: EU-Regional Social Progress Index
- FTE: Full-Time Equivalent
- GBOARD: Government Budget Appropriations or Outlays for R&D
- GRA: Grey Relational Analysis
- GPCI: Goal Programming Composite Indicator
- HE: Higher Education
- HEIs: Higher Education Institutions
- IUNE: The Observatory of Spanish University Research
- JCR: Journal Citation Reports
- MACBETH: Measuring Attractiveness by a Categorical Based Evaluation Technique
- MADM: Multi-Attribute Decision-Making
- MAUT: Multi-Attribute Utility Theory
- MAVT: Multi-Attribute Value Theory
- MCDM: Multiple Criteria Decision Making
- MCI: Mixed Composite Indicator
- MECD: Ministry of Education, Culture and Sports
- MCGP: Multi-Choice Goal Programming
- MODM: Multi-Objective Decision-Making
- MPI: Mazziotta-Pareto Index
- MRP-WSCI: Multiple Reference Point Weak-Strong Composite Indicator
- NGPCI: Net Goal Programming Composite Indicator
- OECD: Organization for Economic Co-operation and Development
- PCA: Principal Component Analysis
- PROMETHEE: Preference Ranking Organization Method for Enrichment Evaluations
- QS Ranking: Quacquarelli Symonds World University Ranking
- RGPCI: Restrictive Goal Programming Composite Indicator
- R&D: Research and Development
- SAW: Simple Additive Weighting
- SCI: Strong Composite Indicator
- SMART: Simple Multi-Attribute Rating Technique
- THE Ranking: Times Higher Education World University Ranking
- TOPSIS: Technique for Order Preferences by Similarity to Ideal Solutions
- USCI: Unweighted Strong Composite Indicator
- UTA: Utility Theory Additive
- WCI: Weak Composite Indicator
- WP: Weighted Product

Resumen de la tesis doctoral

El importante papel que muestra la educación superior en la competitividad global de los países confirma a las universidades como un factor fundamental para el desarrollo socio-económico de una nación. No cabe duda de que las universidades juegan un rol cada vez más crucial en la provisión de conocimiento, facilitando el desarrollo y siendo los principales actores de la innovación regional (OECD, 2007). Dehon *et al.* (2009) y Sánchez-Barrioluengo (2012) señalan que dicha transmisión de conocimiento se realiza principalmente a través de las actividades de docencia, su generación mediante la investigación científica, mientras que las actividades de transferencia se agrupan en la llamada "tercera misión" de la universidad.

En los últimos años, el uso los rankings universitarios se ha convertido en una de las herramientas más extendidas para la evaluación de las instituciones de educación superior (Sanz-Casado and De-Filippo, 2011; Glanzel and Debackere, 2009). De hecho, desde la publicación, en 2003, de la primera edición del ranking de Shanghai (*ARWU*) [1], y sucesivas iniciativas, tales como el *Times World University Ranking*[2], el *QS World University Ranking*[3], el *CWTS Leiden Ranking*[4], el *Ranking de Webometrics*[5], etc., la valoración del prestigio y la reputación de las universidades y los sistemas universitarios ha sido uno de los focos de atención del interés público y político. Cabe mencionar que en el ámbito nacional, destacan el proyecto *U-Ranking*[6] de la Fundación BBVA y el Instituto Valenciano de Investigaciones Económicas (IVIE) y *el ranking de las universidades españolas a tu medida*[7] de la Fundación Conocimiento y Desarrollo.

Autores como Marope *et al.* (2013) señalan que la existencia de los conocidos rankings mundiales de universidades pone de manifiesto que las universidades realmente están viviendo "en una época de mediciones y comparaciones", lo cual ha dado lugar a que los rankings pasen de ser "un producto de consumo" a "un instrumento estratégico global" (Hazelkorn, 2013). Esto se debe principalmente a la manera simple y sintética en la que ofrecen la información, lo cual facilita las comparaciones entre las distintas universidades al tiempo que las simplifica. Lamentablemente, según la última edición (2019) del *Ranking de Shanghai*, tan solo el 15 % del total de las universidades españolas está entre las 500 primeras instituciones universitarias del mundo, y de ellas tan solo una, la *Universidad de Barcelona*, está entre las 200 mejores. Este dato se menciona con frecuencia como prueba de la limitada calidad y escasa proyección internacional del Sistema Universitario Español.

No obstante, a pesar de su amplia popularidad, a los rankings de universidades se les suele criticar por los fines que persiguen y por sus metodologías (Climent *et al.*, 2013). Por su parte, Rauhvargers (2011) señala que las críticas que reciben los rankings universitarios son principalmente su limitada cobertura y su opacidad. En esta línea, la mayoría de iniciativas que comparan instituciones de educación superior recurren a técnicas de agregación simples, en especial los rankings comerciales. Esto implica un carácter compensatorio entre los in-

[1] http://www.shanghairanking.com/es/
[2] https://www.timeshighereducation.com/world-university-rankings
[3] http://www. topuniversities.com/university-rankings
[4] http://www.leidenranking.com/
[5] http://www.webometrics.info/es
[6] http://www.u-ranking.es/index2.php
[7] http://www.rankingcyd.org/

dicadores, y consecuentemente la imposibilidad de detectar las diferencias entre universidades según sus rendimientos en los diferentes indicadores. Así, algunos informes recientes (Rauhvargers, 2011), ponen de manifiesto la importancia que tienen los criterios con los que se elaboran los rankings para acreditar su relevancia e interpretar sus resultados. De esta manera, el valor otorgado a los rankings está ligado a cómo se elaboran, así como a la métrica usada para ello. En definitiva, la mayoría de los rankings ofrecen unos resultados bien ordinales, o numéricos de difícil interpretación, por lo que ofrecen una información limitada.

Partiendo de lo anterior, esta tesis doctoral persigue ir más allá de la mera elaboración de un ranking del Sistema Universitario Público Español. Para ello, se propone una nueva metodología, basada en el método de múltiples puntos de referencia (*MRP-WSCI*; Ruiz *et al.*, 2020), para la evaluación del rendimiento de las universidades públicas españolas y, además, proponemos su combinación con el análisis envolvente de datos (DEA; Charnes *et al.*, 1978).

No cabe duda de que la literatura ofrece una gran variedad de trabajos que han hecho uso del método no paramétrico DEA para medir el desempeño de las instituciones de educación superior. De hecho, desde la publicación de la tesis doctoral de Rhodes (1978), donde el método DEA tiene su origen, la educación superior ha sido uno de los focos de atención de los estudios de evaluación de la eficiencia a través de DEA. Sin embargo, autores como Martínez-Campillo and Fernandez-Santos (2019) señalan que, en el ámbito nacional, hay un número limitado de investigaciones que analizan la eficiencia técnica de las universidades públicas españolas (Berbegal-Mirabent *et al.*, 2013; de la Torre *et al.*, 2017a,b; Berbegal-Mirabent, 2018; de Jorge Moreno *et al.*, 2019; Salas Velasco, 2019; Martínez-Campillo and Fernandez-Santos, 2019; Salas Velasco, 2020).

En esta línea, el número de outputs considerados en la mayoría de trabajos en la literatura oscila entre dos y cuatro, siendo los más utilizados el número de alumnos graduados y el número de publicaciones. Sin embargo, autores como González-Garay *et al.* (2019) señalan que la calidad de la educación superior es intrínsecamente multidimensional y, por lo tanto, difícil de evaluar utilizando solo unos pocos indicadores. Es por ello que la evaluación de fenómenos tan complejos, como es el caso del campo de la educación superior, requiere tener en cuenta dos aspectos importantes. Primero, el sistema de educación superior es bastante complejo, diverso e incluye diferentes tareas interconectadas entre sí. Segundo, en el actual entorno altamente competitivo, la evaluación de un sistema universitario requiere el uso de técnicas capaces de gestionar su complejidad (Attardi *et al.*, 2017). Para ello, en esta tesis doctoral, utilizamos los resultados obtenidos del método multicriterio (*MRP-WSCI*) como outputs, lo cual, hasta donde sabemos, no ha sido aplicado hasta el momento en el análisis envolvente de datos DEA. En definitiva, la combinación del método *MRP-WSCI* y el análisis DEA nos permite tener en cuenta todos los aspectos considerados relevantes para evaluar el desempeño universitario.

Además, el uso del método *MRP-WSCI* en el análisis del Sistema Universitario Público Español, resulta interesante, principalmente, por dos razones. En primer lugar, se utilizan niveles de referencia para cada indicador, y por lo tanto, las puntuaciones se interpretan fácilmente como la posición actual de la universidad con respecto a estos niveles. Asimismo, dado que la valoración de una universidad es en esencia un concepto multidimensional, se hace necesario el desarrollo de unos indicadores sintéticos para el análisis del mismo. De

acuerdo con Saisana and Tarantola (2002), un indicador sintético es aquel que resume la información que proporcionan un conjunto de indicadores parciales que no tienen una unidad común de medida y que no pueden ser ponderados de manera obvia. Así pues, un indicador sintético es una combinación matemática de un conjunto de indicadores parciales, con la principal ventaja de ofrecer una visión simplificada de un problema mucho más complejo. En esta línea, se van a construir unos indicadores sintéticos que permitan evaluar el funcionamiento de cada una de las universidades, ofreciéndonos una visión general del Sistema Universitario Público Español, teniendo en cuenta el carácter multidimensional que presenta, e identificando de esta manera las posibles áreas de mejora de cada una de las universidades que lo integran. En definitiva, el método *MRP-WSCI* permite obtener unos indicadores sintéticos para diferentes grados de compensación, de tal manera que las puntuaciones, además de dar una medida del rendimiento general de las universidades, también proporcionan señales de alarma que permiten al usuario detectar áreas de mejora.

Para la consecución de nuestro objetivo, esta tesis doctoral es presentada como compendio de publicaciones y se organiza de la siguiente manera:

- La primera parte contiene las tres publicaciones que conforman el compendio. Concretamente, el marco teórico y metodológico es abordado en los dos primeros artículos, mientras que la tercera y última publicación contiene la aplicación empírica. Seguidamente, se proporciona un resumen de cada uno de estos artículos.

    - Dado que hoy en día, existe un amplio abanico de metodologías para la construcción de los indicadores sintéticos de desempeño universitario, el primer artículo, El Gibari *et al.* (2019), abarca una revisión de la literatura de los diferentes métodos multicriterio utilizados para la construcción de indicadores. Además, en este artículo, hemos clasificado dichos métodos en cinco categorías: los métodos elementales, los métodos basados en valor y utilidad, el enfoque outranking, los métodos basados en DEA y los métodos basados en funciones de distancia. Asimismo, se discuten dos aspectos relevantes relacionados con la construcción de los indicadores sintéticos haciendo uso de los métodos MCDM, que son la ponderación y la compensación. Primero, de acuerdo con Nardo *et al.* (2005), los pesos de los diferentes criterios pueden tener un significativo efecto sobre el indicador sintético global y los resultados obtenidos. En esta línea, los métodos de ponderación se pueden clasificar en tres principales categorías: la ponderación igualitaria significa que todos los indicadores tienen el mismo peso, las técnicas positivas (endógenas o estadísticas), que derivan los pesos de las características de los datos y las técnicas normativas (exógenas o participativas), donde se tienen en cuenta las opiniones subjetivas de los expertos y/o los decisores. En el caso particular de los métodos MCDM usados para la construcción de indicadores sintéticos, nuestra revisión ha mostrado que las dos últimas técnicas han sido los enfoques más utilizados, siendo, en el primer caso, la técnica de ponderación basada en el análisis envolvente de datos DEA la más frecuente. Segundo, en función del método MCDM de agregación elegido, se asume que la compensación entre

los indicadores puede ser total (e.g. los métodos de la ponderación aditiva simple, SAW, la teoría de la utilidad aditiva, UTA, la técnica simple de calificación de múltiples atributos, SMART, el análisis envolvente de datos, DEA, y TOPSIS), parcial (e.g. los enfoques de la teoría de utilidades múltiples, MAUT, la teoría de múltiples atributos, MAVT, o el producto ponderado, WP) o nula (e.g. ELECTRE y PROMETHEE). En general, de esta revisión se desprende una clara tendencia hacia un número creciente, desde 2014, de artículos que recurren al uso de los métodos multicriterio para construir indicadores sintéticos. En concreto, nuestros resultados revelan que, desde 2012, ha habido una creciente tendencia a adoptar los métodos basados en las funciones de distancia. No obstante, a pesar de la amplia variedad de aplicaciones de los métodos MCDM para la construcción de los indicadores sintéticos, tales como sostenibilidad (29 %), medio ambiente (12 %), entorno empresarial (8 %), energía (6 %), turismo (5 %), desarrollo humano (5 %), servicio de agua (4 %) e inversión (4 %), esta publicación nos ha permitido identificar una escasez en el uso de dichas técnicas dentro del ámbito de la educación superior. Además, la mayoría de los artículos revisados se han publicado en revistas pertenecientes a la categoría de "Environmental Sciences"' (48 %), lo cual, parece lógico, ya que, bajo esta vertiente, la sostenibilidad y el medio ambiente han sido los campos más analizados. Otros artículos han elegido revistas que se centran en la categoría de "Operations Research & Management Science" (27 %). Sin embargo, hay un escaso número de publicaciones que recurren al uso de los métodos MCDM para la construcción de indicadores sintéticos en la categoría de "'Education & Educational Research". Este artículo está publicado en la revista *Journal of Business Economics*, que tiene un índice de impacto de 0.378, y está indexada en el segundo cuartil de *Scopus*. Su índice de citas es de 33 en *Google Scholar* y su impacto de citas ponderadas en el campo, según la base de datos *Scopus*, es de 8.36 (fecha de referencia: 31 de mayo de 2020).

- La segunda publicación, Ruiz *et al.* (2020), presenta el método de múltiples puntos de referencia (*MRP-WSCI*), que es una generalización del método de doble punto de referencia (Ruiz *et al.*, 2011). El método de punto de referencia fue propuesto originalmente por Wierzbicki (1980), y se diseñó en un primer momento, para resolver problemas de programación multiobjetivo, mediante la generación de soluciones eficientes que estaban más cerca de ciertos niveles de referencia (deseados) para los objetivos planteados. Más tarde, Wierzbicki *et al.* (2000) extendieron esta metodología a un esquema de doble punto de referencia, sugiriendo su utilización para la confección de unos rankings objetivos, donde se pueden especificar para cada objetivo un valor de reserva (nivel mínimo considerado como admisible) y un valor de aspiración (nivel considerado como óptimo o deseable). Ruiz *et al.* (2011) y Cabello *et al.* (2014) adaptaron y desarrollaron aún más esta idea para la construcción de los indicadores sintéticos según el grado de compensación entre los indicadores y basándose en los pesos otorgados por el usuario. No obstante, a pesar de las

ventajas del método de doble punto de referencia, nos hemos dado cuenta de que, tal y como se había planteado en un primer momento, dicho método tiene dos posibles inconvenientes. En primer lugar, la utilización de únicamente dos puntos de referencia, en determinadas situaciones, puede suponer una limitación para el decisor. Además, los resultados se daban en una escala fija (de -1 a 2), lo cual puede ser una limitación a la hora de combinar esta metodología con otras, tales como el análisis envolvente de datos DEA, donde la implementación del modelo establece la restricción de no negatividad de las variables. En esta línea, los decisores pueden desear utilizar una escala con diferentes rangos. De esta manera, el método de múltiples puntos de referencia (*MRP-WSCI*) se puede aplicar en un contexto más general, donde el usuario pueda establecer cualquier número de niveles de referencia, que pueden suponer distintos niveles de rendimiento (muy malo, malo, ..., bueno, muy bueno), y pueda adoptar cualquier escala. Con el objetivo de ilustrar el comportamiento y las ventajas del método *MRP-WSCI*, se ha aplicado a la construcción del Índice de Progreso Social Regional de la Unión Europea (*EU-SPI*), en el contexto español. En concreto, hemos evaluado el comportamiento de las regiones españolas en comparación con todas las regiones europeas, dado que los niveles de referencia se han establecido teniendo en cuenta todas las regiones que aparecen en el análisis *EU-SPI*. Lógicamente, en el proceso de construcción de los indicadores sintéticos *MRP-WSCI*, los resultados obtenidos dependen de los niveles de referencia elegidos, que a su vez deben establecerse de acuerdo con el objetivo final del estudio. En esta línea, con el fin de analizar los efectos de un cambio de los niveles de referencia en los resultados, las regiones españolas han sido también analizadas, comparándolas únicamente entre sí. Como resultado, cuando los niveles de referencia utilizados son los españoles, las diferencias entre las regiones evaluadas se acentúan. Asimismo, se han llevado a cabo diferentes estudios, tales como el análisis de robustez de los indicadores sintéticos *MRP-WSCI* frente a cambios en los niveles de referencia y en los pesos otorgados, así como el efecto de la construcción de los indicadores *MRP-WSCI* en diferentes etapas de agregación. Por último, cabe mencionar que, en este artículo, se ha demostrado que los indicadores sintéticos proporcionados por el método *MRP-WSCI* satisfacen una serie de propiedades que la literatura científica considera altamente deseables (existencia y unicidad, simetría, transitividad, monotonicidad, invarianza y propiedades de exhaustividad). Este trabajo ha sido publicado en *OMEGA*, que cuenta con un factor de impacto *JCR* del 5.341, de hecho, es la segunda revista mejor clasificada dentro de su categoría "Operations Research & Management Science". Su índice de citas es de 3 en *Google Scholar* y su impacto de citas ponderadas en el campo, según la base de datos *Scopus*, es de 0.52 (fecha de referencia: 31 de mayo de 2020).

- El tercer artículo, El Gibari *et al.* (2018), trata de una versión preliminar del método de doble punto de referencia aplicado a las universidades públicas andaluzas. Dentro del ámbito multicriterio y educa-

ción superior, la metodología adoptada en este artículo, supone un paso más allá de la mera construcción de indicadores sintéticos, ya que analizamos los indicadores individualmente, mediante los niveles de referencia y planteamos la construcción de indicadores sintéticos para diferentes grados de compensación en la confección de rankings universitarios. Además, con el objetivo de llevar a cabo un análisis más exhaustivo, analizamos por separado las tres misiones básicas de un sistema universitario (investigación, docencia y transferencia de tecnología). Nuestros resultados afirman que, en general, el rendimiento de las universidades públicas andaluzas en la misión docente es mejor que en las misiones investigadora y de transferencia de tecnología, aunque los valores de los indicadores sugieren que todas las universidades todavía tienen mucho margen de mejora. De hecho, la internacionalización de las universidades se presenta como uno de los focos claves de transformación en el campo de la educación superior, especialmente en el contexto europeo (OECD, 2012). Sin embargo, nuestros resultados muestran que la "*atracción internacional*" y las "*prácticas en empresa*" suponen una debilidad para la mayoría de las universidades públicas andaluzas. Asimismo, además de proporcionar una información valiosa del desempeño universitario andaluz mediante el análisis de los escenarios compensatorio y no compensatorio, hemos elaborado un ranking, según diferentes grados de compensación, para cada una de las misiones analizadas. De esta manera, estudiamos el posicionamiento de cada una de las universidades en el ranking, a medida que cambia el grado de compensación entre los indicadores, desde una compensación total hasta una compensación nula entre los indicadores. En este caso, nuestros resultados revelan que las clasificaciones en el ranking de las universidades públicas andaluzas son más estables en la misión investigadora, no viéndose muy afectadas por ligeros cambios en la compensación permitida. En definitiva, de este artículo, podemos concluir que la metodología multicriterio ofrece un sistema de advertencia para ayudar en la toma de decisiones estratégicas. Este trabajo está publicado en la revista *Journal of Informetrics*, que cuenta con un factor de impacto *JCR* del 3.879 y está indexada en el primer cuartil en la categoría "Computer Science & Interdisciplinary Applications". Su índice de citas es de 7 en *Google Scholar* y su impacto de citas ponderadas en el campo, según la base de datos *Scopus*, es de 0.90 (fecha de referencia: 31 de mayo de 2020).

- La segunda parte de la presente tesis doctoral engloba algunas líneas de investigación adicionales relacionadas con estas publicaciones:

    - Para la toma de decisiones en nuestro entorno dinámico, el uso de los indicadores sintéticos se ha convertido en indispensable, ya que son herramientas poderosas para resumir, enfocar y condensar la complejidad de dicho entorno. No obstante, la información proporcionada por el indicador sintético debe ser máxima, permitiendo detectar áreas de mejora, en lugar de solo clasificar las unidades. En esta línea, la literatura ofrece pocos enfoques que permiten la construcción de

indicadores sintéticos para diferentes grados de compensación. Partiendo de ello, en este capítulo, llevamos a cabo un análisis comparativo de los indicadores sintéticos proporcionados por algunos métodos que permiten la construcción de indicadores sintéticos para diferentes grados de compensación. En concreto, nuestro objetivo consiste en ilustrar el comportamiento de dichos métodos, analizando la información adicional que proporciona cada uno de los métodos, al considerar los escenarios compensatorios y no compensatorios, conjuntamente. En concreto, mediante la aplicación a un caso real, a través de la construcción del índice de vulnerabilidad infantil (*EOCI*), en el contexto africano, hemos comparado la información proporcionada por el uso conjunto de las medias geométrica y aritmética, la Función *Mean-Min* (Tarabusi and Guarini, 2013), el enfoque Goal-Programming (*GPCI;* Blancas *et al.*, 2010) y el método *MRP-WSCI* (Ruiz *et al.*, 2020). Todos estos enfoques siguen, hasta cierto punto, los patrones generales para la construcción de indicadores sintéticos. Sin embargo, la filosofía detrás de esta construcción y el significado de sus valores son diferentes para cada metodología. En particular, el procedimiento de normalización de la Función *Mean-Min* permite identificar las unidades con un rendimiento superior o inferior al promedio, mientras que los enfoques *GPCI* y *MRP-WSCI* llevan a cabo una normalización basada en la distancia, teniendo en cuenta la posición de cada unidad con respecto a un conjunto de niveles de referencia (uno, en el caso del primero, o más, en el segundo). Además, los valores mínimos y máximos de la escala de los indicadores sintéticos están predefinidos para las medias clásicas y el método *MRP-WSCI*, mientras que en el caso de la Función *Mean-Min* y el enfoque *GPCI*, los resultados pueden tomar cualquier valor real, ya que no hay niveles superiores ni inferiores. Desde el punto de vista práctico, nuestros resultados revelan que los escenarios no compensatorios de la media clásica y el *GPCI* permiten, en la práctica, una compensación parcial, mientras que los indicadores sintéticos no compensatorios proporcionados por la Función *Mean-Min* y el enfoque *MRP-WSCI* no permiten ningún tipo de compensación entre los indicadores. En esta línea, además de la interpretación de cada uno de los indicadores sintéticos propuestos por las metodologías empleadas para el análisis comparativo, en el artículo se hace hincapié en las características más importantes de los enfoques analizados, tales como por ejemplo el significado de la distancia entre el escenario compensatorio y no compensatorio o la tendencia de la correlación entre ambos escenarios, a medida que el número de simulaciones aumenta. Cabe mencionar que para este último caso, hemos llevado a cabo una serie de experimentos computacionales en los que todos los métodos se han aplicado a una serie de problemas que se han generado de manera aleatoria. Nuestros resultados revelan que la Función *Mean-Min* y el método *MRP-WSCI* son los menos correlacionados en casi un 50 % de las simulaciones cada uno, y solo el enfoque *GPCI* obtiene los resultados menos correlacionados en un número residual de casos. En definitiva, podemos concluir que la Función *Mean-Min* y el método *MRP-WSCI* proporcionan una información más rica cuando se consideran ambos es-

cenarios conjuntamente (compensatorio y no compensatorio). Cabe mencionar que una versión extendida de este trabajo se encuentra actualmente en fase de revisión en la revista *Internacional Journal of Information Technology & Decision Marking*, que tiene un factor de impacto *JCR* del 2.862 y está indexada en el segundo cuartil en las categorías "Operations Research & Management Science" y "Computer Science & Information Systems".

- Por último, la evaluación de las instituciones de educación superior debe estar de acuerdo con las nuevas demandas del complejo entorno actual. Para este fin, el uso de técnicas que sean capaces de manejar esta complejidad es altamente adecuado. En esta última línea de investigación, se propone una combinación del método *MRP-WSCI* con el análisis no paramétrico DEA para la evaluación del Sistema universitario público Español. Esta combinación nos permite tener en cuenta como resultados todos los aspectos considerados relevantes para evaluar el desempeño universitario e incorporar los recursos, de distinta naturaleza, que utilizan las universidades. Para ilustrar esta combinación, nuestra base de datos comprende a las 47 universidades presenciales públicas españolas para el año académico más reciente disponible, 2016-2017. Para ello, al igual que el artículo que recoge la aplicación empírica de la presente tesis doctoral, hemos analizado las diferentes misiones universitarias. En esta línea, la construcción de los indicadores sintéticos *MRP-WSCI* se ha llevado a cabo en dos etapas (agregación de los indicadores individuales en los diferentes sub-bloques correspondientes, y luego agregación de los sub-bloques a las correspondientes misiones). En la primera etapa de agregación, dado que la información proporcionada por los indicadores individuales dentro de los correspondientes sub-bloques es homogénea, se permite una compensación completa entre ellos. En la segunda etapa, se proporcionan diferentes grados de compensación entre los sub-bloques (compensatorio y no compensatorio). Los indicadores sintéticos *MRP-WSCI* son utilizados, a posteriori, como outputs en el análisis envolvente de datos DEA. Cabe mencionar que la construcción de los indicadores sintéticos *MRP-WSCI* se ha llevado a cabo en un entorno robusto con respecto a los pesos de los indicadores y sub-bloques. De esta manera, además de analizar el rendimiento de las universidades públicas españolas, desde los dos escenarios de compensación, mediante el método *MRP-WSCI*, también llevamos a cabo un doble análisis de eficiencia, en el que los indicadores de resultados (compensatorios y no compensatorios) son confrontados con los recursos disponibles por cada una de las universidades. En concreto, entre las 47 universidades públicas españolas analizadas, nuestros resultados muestran que 18 universidades son eficientes en el escenario compensatorio, y 17 en el no compensatorio, mientras que 14 de ellas son eficientes en ambos análisis. Además, en promedio, los resultados de las universidades públicas españolas son bastante similares en ambos análisis (1,11 y 1,12, respectivamente). Esto nos pone de manifiesto que, manteniendo sus inputs en los niveles actuales, para operar de manera eficiente, las universidades españolas deberían aumentar,

en promedio, sus outputs en un 11 % en el escenario compensatorio y un 12 % en el no compensatorio. En general, la mayoría de las universidades tienden a ser menos ineficientes cuando se analiza el rendimiento global (en concreto, 21 de 47). De hecho, nuestros resultados muestran que ciertas universidades son mucho más ineficientes en el escenario no compensatorio, mientras que en el compensatorio, sus niveles de eficiencia mejoran sustancialmente. Este hecho nos indica que dichas universidades tienen un mal rendimiento en al menos algún sub-bloque analizado, pero estos rendimientos desfavorables se compensan con buenos rendimientos en otros, lo que resulta en un mejor rendimiento global (escenario compensatorio). En definitiva, nuestro análisis resalta la necesidad de utilizar escenarios compensatorios y no compensatorios en este contexto. Además, la combinación del uso de indicadores sintéticos con el análisis DEA nos da una imagen más completa de las instituciones evaluadas, lo que permite a las universidades comprender mejor sus debilidades y fortalezas. Cabe resaltar que, actualmente, estamos en proceso de elaboración de una versión extendida de este trabajo de investigación, la cual prevemos enviar, en breve, a una revista indexada en la base de datos *JCR*, y de gran difusión internacional.

Una conclusión relevante de esta tesis doctoral es la notable variedad dentro de los resultados de cada una de las misiones analizadas. Los valores de los indicadores sugieren que, en general, todas las universidades públicas españolas todavía tienen mucho margen de mejora en todas las misiones. Asimismo, nuestro análisis muestra que, además de las universidades públicas españolas que siempre están presentes en los rankings universitarios más prestigiosos, hay otras universidades que aún disponiendo de unos recursos bastante "modestos", hacen un uso eficiente de dichos recursos, obteniendo así "buenos" rendimientos en algunos aspectos analizados.

En resumen, hoy en día, un mejor desempeño y compromiso para lograr una ventaja competitiva es esencial para el Sistema Universitario Público Español, dado que resulta evidente que en el futuro, las universidades tendrán una contribución aún más vital en nuestra sociedad. En esta tesis doctoral, respetando el carácter multidimensional y partiendo de la necesidad de considerar el aspecto complejo y diverso de un sistema universitario, se propone por un lado, la utilización de una nueva metodología para la evaluación del rendimiento y, por otro lado, la combinación de la misma con el análisis envolvente de datos DEA, para confrontar el rendimiento de cada universidad con los recursos disponibles, tanto en un escenario compensatorio como no compensatorio.

En esta línea, podemos concluir que, a diferencia de la mayoría de los rankings universitarios que existen hoy en día, una ventaja de esta tesis es la posibilidad de analizar las universidades públicas españolas desde diferentes perspectivas simultáneamente, desarrollando y usando indicadores sintéticos conjuntamente con el análisis de los indicadores individuales de rendimiento universitario. Además, la interpretación de los indicadores sintéticos propuestos resulta más intuitiva y fácil para el decisor, dado que representan la posición global de la universidad correspondiente con respecto a los niveles de referencia establecidos. De esta manera, la puntuación final no es solo un número, sino una medida más informativa. En segundo lugar, se ha demostrado que la consideración de

manera conjunta del escenario compensatorio y no compensatorio proporciona una información más rica para la toma de decisiones. De hecho, además de proporcionar el rendimiento global de las universidades públicas españolas, el método multicriterio *MRP-WSCI* proporciona señales de alerta que permiten al usuario detectar malos desempeños en ciertos indicadores que de lo contrario podrían pasar desapercibidos. Además, la combinación del método multicriterio de múltiples puntos de referencia con el análisis envolvente de datos proporciona una imagen más completa de las instituciones analizadas, ya que permite confrontar los indicadores de resultados con los recursos disponibles por cada una de las universidades, y en consecuencia las universidades públicas españolas pueden detectar mejor sus debilidades y fortalezas, facilitando así una toma de decisiones por parte de sus gestores acordes con ello.

Por último, tal y como define Sabino (1992), la investigación es "un esfuerzo que se emprende para resolver un problema, claro está, un problema de conocimiento". Por su parte, Cervo and Bervian (1980) la definen como "una actividad encaminada a la solución de problemas. Su objetivo consiste en hallar respuesta a preguntas mediante el empleo de procesos científicos". No obstante, cualquier trabajo de investigación elaborado con un mínimo de interés no sólo contribuye a resolver las preguntas que se plantean sobre el tema tratado, sino también genera nuevas preguntas, ideas y/o futuras líneas de investigación. Así, como futuras líneas de investigación de esta tesis doctoral tenemos pretensión de abordar las siguientes. En primer lugar, sería interesante estudiar la evolución de las universidades españolas a lo largo del tiempo, y no sólo limitarnos a analizar un año concreto. En particular, resulta llamativo estudiar si las universidades públicas españolas logran mantener sus rendimientos y las posiciones alcanzadas en los rankings elaborados para cada una de las misiones a largo plazo. Para ello, habría que estudiar la situación de las universidades públicas españolas frente a un año concreto. En este caso, los niveles de referencia se mantendrían fijos, lo que permite conocer la evolución de las instituciones universitarias frente a la situación que presentaban en el año base. Alternativamente, los valores de referencia pueden ser determinados por parte de expertos, según lo que consideren oportuno, atendiendo a sus juicios de valor, para cada indicador. De esta forma, obtendríamos una medida absoluta con respecto a estos valores. Asimismo, la presente tesis doctoral se ha centrado en el Sistema Universitario Público Español. No obstante, también sería conveniente hacer una comparación más amplia entre las universidades públicas y privadas españolas, donde los resultados darán una idea de la posición relativa de cada universidad española con respecto a todas las instituciones del Sistema Universitario Español. Además, con el fin de sacar el mayor provecho posible del Proceso de Bolonia, llevar a cabo una comparación más amplia entre distintas universidades europeas podría resultar interesante y útil para los propósitos del Gobierno español. Cabe destacar que esta comparación se podrá llevar a cabo siempre que la calidad y la homogeneidad de los datos empleados de las distintas universidades sean las adecuadas.

Summary of the doctoral thesis

The increasing awareness around the world of the role that universities play in today's competitive environment highlights the need to assess them by using methods that are able to manage their diversity, complexity and multidimensional nature. On this premise, this doctoral thesis aims to go beyond simply providing a ranking of the Spanish public university system. To this end, this doctoral thesis is presented as a compendium of three research papers. Besides, some additional research questions related to these articles are presented.

First, the theoretical and methodological framework is addressed in two articles. On the one hand, given that in today's increasingly globalized and complex environment that is flooded with data, the number of composite indicators in existence around the world is constantly growing at a rapid pace, especially due to their aims of summarizing, focusing and condensing the complexity of our dynamic environment. For their construction, the functionality of Multiple Criteria Decision Making approaches is very high, as it has all the characteristics of a useful decision support tool: it plays an important role in designing and analyzing systems by considering various indicators and objectives. Therefore, our first goal has been the identification of the different MCDM methods used to construct composite indicators. In general, our results have shown a clear tendency towards an increasing number of papers in a wide variety of fields, that use MCDM methods to build composite indicators since 2014. However, a shortage in the field of higher education has been identified. On the other hand, when using MCDM methods to aggregate single indicators into a composite one, the compensatory character between indicators varies. In this PhD thesis, we propose a generalization of the double reference point method, which allows for different compensation degrees, depending on the aggregation scenario chosen. This generalization, called Multiple Reference Point based Weak and Strong Composite Indicators (*MRP-WSCI*), allows the decision maker give, for each indicator, a desired number of reference levels to assess the goodness of its possible values. Thus, the proposed approach can be applied in a more general context, where the decision makers can establish any number of reference levels, which somehow define the performance levels (e.g., very poor, poor, fair, good, very good,...), and any scale can be used.

Second, the empirical application is discussed in one article. To this end, we propose an application of multicriteria analysis techniques, based on the double reference point method, to the Andalusian public university system. In order to get a more accurate vision of each basic university mission, the three of them (research, teaching and technology transfer) are analyzed separately. Within the scopes of multicriteria and higher education field, the results obtained are particularly useful because they serve not only to measure the relative position of each university, but also to offer a warning system to assist in strategic decision making. Furthermore, aggregations for different compensation degrees are provided for each Andalusian university.

Furthermore, since in our dynamic environment, the information provided by any composite indicator should be as maximal as possible, allowing to detect improvement areas, rather than just ranking the units, in this PhD thesis, we extend the theoretical and methodological framework. Thus, we illustrate the behavior of some methods that build composite indicators allowing different compensation degrees. Specifically, we analyze the additional information provided by each of these methods when the compensatory and non-compensatory scenarios are jointly considered.

Finally, an extension of the empirical application to the Spanish public university system is proposed. To this end, we evaluate of the Spanish public universities performance by combining the Multiple reference point based weak and strong composite indicators proposed in this thesis and the well-known DEA analysis. This last section discusses how this combination can add value to traditional composite indicators, not only in terms of securing high performance of the Spanish public universities, but also in how to make an appropriate use of resources. Thus, we highlight the need to assess the university performance by using both compensatory and non-compensatory scenarios. Furthermore, combining the use of composite indicators with the DEA analysis provides warning signals that may assist user in strategic decision making for policy purposes, by identifying the weaknesses and strengths of each institution under assessment.

Summing up, recognizing the specific nature and complexity of a university system, this doctoral thesis has attempted to analyze the performance and efficiency of the Spanish public universities from a multicriteria approach. First, in order to evaluate the performance of a university system, a novel approach, based on the multicriteria reference point scheme, is proposed. Then, its combination with the well-known DEA method is carried out, by comparing the performance of each university (compensatory and non-compensatory) with their resources.

Contents

1. Introduction

The growing importance of universities in today's competitive environment highlights the need to assess them by using methods that are able to manage their diversity, complexity and multidimensional nature. Thus, this doctoral thesis aims to go beyond simply providing a ranking of the Spanish public university system. To this end, this research is presented as a compendium of articles.

On these premises, the structure of this Introduction is as follows: Subchapter 1.1 describes the main characteristics of the Spanish *HE* system; Subchapter 1.2 gives a short overview of the characteristics of the *CI*s in the *HE* field; the main MCDM methods applied to their construction are presented in Subchapter 1.3; Subchapter 1.4 highlights the need to assess the operational efficiency in the Spanish *HE* field; and finally, the introduction ends with the main purposes and structure of this PhD thesis in Subchapter 1.5.

1.1. Institutional context: The Spanish public higher education system

The Spanish higher education (*HE*) System comprises 82 universities, out of which 50 are public (47 on-site, 1 distance university and 2 special universities) and 32 are private (28 on-site and 4 distance universities). Traditionally, most of the Spanish universities focused their efforts on professional training and teaching, neglecting scientific research. However, the Government of Spain introduced "LOMLOU" (the Organic Law 4/2007[8] on universities), in order to redefine Spanish universities in the European Higher Education Area (*EHEA*). Since 2006, Spain has implemented structural changes in its *HE* system according to the Bologna Process, which ensures comparability in standards with the *EHEA*. In this direction, official university programs are adapted to the *EHEA*, and structured into three different degree levels: Grade degrees (undergraduate studies), University Master degrees (graduate studies) and Doctoral degrees (postgraduate studies), noticing a huge increase in the number of bilingual programs and in the international student mobility. Nowadays, most of the Spanish universities cover nearly all the subject areas offered by the Spanish official degrees listed by the Ministry of Education. Furthermore, the introduction of "LOMLOU" has encouraged the Spanish public *HE* system to adapt to the *HE* global competition, which has led to define research intensity as the key indicator of the quality of universities (Mägi and Beerkens, 2016). On the other hand, Spanish policies and the changes in the universities' regulatory framework strive to achieve greater knowledge transfer results (Berbegal-Mirabent *et al.*, 2013).

However, despite the Spanish Government attempts to adapt to change in times of crisis and to align with "the Modernization Agenda for Universities" introduced by the European Commission (2006), a lot still needs to be done. In the 2019 edition of *Academic Ranking of World Universities (ARWU)*[9], only 13 Spanish universities are positioned among the top 500 institutions in the world, and all of them are below the 150th place.

Furthermore, despite being considered among the largest *HE* systems in Europe, the Spanish *HE* system has been seriously affected by the economic reces-

[8] https://www.boe.es/eli/es/lo/2007/04/12/4
[9] http://www.shanghairanking.com/es/

sion, since 2010. The investment in knowledge in Spain is not comparable to the gross investment of other countries, in fact, Spanish R&D investment continues to lose ground. According to European Commission (2018), among other European countries, Spain has exhibited disappointing R&D investment-intensity records, assuming a reduction of approximately 20.32% from 2009 to 2014 in the Government Budget Appropriations or Outlays for R&D (*GBOARD*) of Spain (Heitor *et al.*, 2016). However, in terms of integration in the *ERA* and beyond, Spain has managed to raise its scientific publications significantly compared to 2000, due to the higher research requirements set by the academic authorities for staff hiring and promotion purposes. Taking into account the world share of top 10% highly cited scientific publications, the Spanish share has increased in almost 0.9% from 2000 to 2016 (European Commission, 2018). Furthermore, due to the importance of fostering the production of new knowledge worldwide and stimulating positive impacts in scientific performance, Spain has increased its international scientific co-publications by 16% over the period 2000-2011 (European Commission, 2014). Nevertheless, the level of Spain's international co-publication (29.1%) is still below other comparable European countries (France 35.2% or Portugal 41%).

In Spain, public and private universities fulfil the different tasks required from any university nowadays, which are the construction of knowledge (research), the diffusion of knowledge (teaching), and the use of knowledge (technology transfer) (Dehon *et al.*, 2009). However, private universities usually present low levels of research activity and of the technology transfer for its use by society (Casani *et al.*, 2014).

Furthermore, the Spanish public and private universities are regulated by the National Government, which exercises the competencies that ensure the consistency and uniformity of the *HE* system, and also by the Regional Governments, which have competencies for the creation, modification and elimination of university programs, and also for the core funding of public universities (Martínez-Campillo and Fernandez-Santos, 2019). On the other hand, private universities enjoy greater flexibility which enables them to respond more quickly and efficiently to market changes, while public universities are coming under increasing pressure and greater rigidity, due to the strict regulation of their activities and public accounts. The Spanish public universities include in their decision-making processes criteria related with the social benefits. In general, the Spanish public universities are integrated by university schools, faculties, departments, university institutes for research, doctoral colleges and by other necessary schools or structures for the development of their functions. This PhD thesis is focused on the 47 Spanish public on-site universities, where the data comes from two sources: the Spanish University Rectors' Conference[10] (*CRUE*) and the Observatory of Spanish University Research[11] (*IUNE*).

In this time of economic uncertainty, the evaluation and the pursuit of productivity and efficiency in the allocation of resources to the Spanish public universities is becoming increasingly important, due to the limitation of public funding following the economic recession in 2008 (Martínez-Campillo and Fernandez-Santos, 2019). This issue has put greater pressure on the Spanish public universities to achieve excellence and improve competitiveness. There

[10] http://www.crue.org/Publicaciones/Paginas/UEC.aspx?Mobile=0
[11] http://www.iune.es

is no doubt that the Spanish public *HE* institutions are key drivers of growth performance, prosperity and competitiveness, at the national and international levels. In this context, *HE* rankings, an outcome of the competitiveness boosted by globalization, have put pressure on the Spanish public universities to adapt to market forces without compromising quality (Heitor and Horta, 2013). However, the literature notes that most university rankings produce single scores, making it "difficult for users to distinguish among institutions based on the characteristics they find most important" (Bonaccorsi and Cicero, 2016). However, it is necessary to bear in mind that the Spanish public *HE* system is quite complex, diverse and it includes different tasks that are interconnected with each other. It can be asserted that the Spanish public university system is basically a multidimensional concept, which makes it difficult to condense the diversified work going on within universities into a single number or ranking. In this direction, in today's increasingly globalised and complex environment, the Spanish public *HE* system should have a powerful tool to detect improvement areas, rather than just ranking the universities. Consequently, the evaluation of the Spanish *HE* system requires the use of techniques that are able to manage its complexity (Attardi *et al.*, 2017).

1.2. Composite indicators as decision making tools

In the current era of complex environment that is flooded with data, the measurement of multidimensional frameworks, such as the *HE* field, has received special attention in the worldwide scientific community. In this line, the need to summarize such complex and multidimensional concepts, in view of supporting decision-makers and attracting public interest, has been brought the concept of Composite Indicators (*CI*s) into the focus of public and policy interest. Due to their aim to provide simple comparisons that can be used to illustrate complex and sometimes elusive issues in wide-ranging fields, such as performance of *HE* institutions, *CI*s are increasingly recognized as useful tools in policy analysis and public communication. There is no doubt that the use of *CI*s has become indispensable for managing huge amounts of information.

In the field of *HE*, few initiatives have tackled the evaluation of university performance without using *CI*s, such as the *U-Multirank* and its Spanish version, *CYD Ranking*, which consider the scores of universities on individual indicators and place these in five performance groups ("very good" through to "weak"). Opposedly, within the performance of the *HE* field, a general widespread trend towards the construction of *CI*s has been established. In this line, most of the rankings existing nowadays have developed and used *CI*s to provide rankings of *HE* institutions or countries (the *ARWU*; the Times Higher Education World University Ranking (*THE*); the QS World University Ranking (*QS*); the *U-Ranking* project of the BBVA Foundation and the IVIE[12]; Aguillo *et al.*, 2008; Alasehir *et al.*, 2014; Giannoulis and Ishizaka, 2010; Torres-Salinas *et al.*, 2011). Other papers adopt *CI*s in order to analyze the quality of universities, such as Murias *et al.* (2008), who estimate a *CI* for quality assessment in the Spanish public *HE* system.

The literature on *CI*s is vast and provides a wide range of methodological approaches (the common practice in constructing *CI*s is well synthesized in Nardo

[12]The Valencian Institute of Economic Research

et al., 2008; Gana *et al.*, 2017; El Gibari *et al.*, 2019). Within the construction of *CI*s, the use of Multiple Criteria Decision Making (MCDM) approaches is highly suitable in multidimensional frameworks (Freudenberg, 2003; Jacobs *et al.*, 2004; Ibañez-Forés *et al.*, 2014). Multicriteria decision making is a set of techniques that can be used to support the process of decision making in a flexible manner when designing and analyzing systems by considering various indicators and objectives (Hernández Santoyo *et al.*, 2010; Cinelli *et al.*, 2014). According to Teodorovic and Edara (2007), one way of dealing with MCDM problems is to convert them into a single criterion problem. To this end, decision makers could use some of the compensatory methods (that allow advantages of one indicator to be traded for disadvantages of another), or non-compensatory methods (that do not allow advantages of one indicator to be traded for disadvantages of another). In this direction, according to Munda (2008), depending on the aggregation approach chosen, the compensatory character between indicators can be: full (additive methods and compensatory multicriteria approaches), partial (geometric aggregations and other partially-compensatory methods), or zero (non-compensatory techniques).

Anyway, the selection of an appropriate methodology is central to any exercise attempting to capture and summarize the interactions among the individual indicators included in a *CI* or ranking system (Saisana and D'Hombres, 2008). In this direction, according to Mazziotta and Pareto (2017), the complex and multidimensional nature of the *HE* field requires the definition of intermediate objectives whose achievement can be observed and measured by individual indicators.

1.3. MCDM methods used to construct composite indicators: compensability and use of reference levels

In recent years, MCDM methods have attracted attention of decision makers in a wide scope of fields. MCDM is a well known branch of decision making, consisting of analysing (ranking, classifying, choosing) a series of possible alternatives, taking into account different criteria simultaneously (Contreras *et al.*, 2011). Decision-making is a term that describes the process of how people choose to solve problems that are presented, and the trigger of decision-making is the existence of a problem. Nowadays, MCDM analysis is considered a valuable tool in solving problems characterized with multiple stakeholders, indicators, and objectives, as is the case of the *HE* field (Kumar *et al.*, 2017).

As mentioned in Section 1.2, MCDM methods are highly suitable in multidimensional frameworks, when aggregating individual indicators into a composite one, since this process involves making choices when combining criteria of different natures, and it requires a number of steps in which decisions must be made (Freudenberg, 2003; Jacobs *et al.*, 2004). In this direction, many works have attempted to construct *CI*s by falling back on MCDM methods. A large number of papers that use MCDM methods to construct *CI*s have been published, and this trend seems set to continue.

In the literature, one classification of the MCDM approaches used to construct *CI*s is focused on the aggregation of the individual indicators, regardless the weighting procedure. This way, El Gibari *et al.* (2019) identify five different categories: the elementary methods, the value and utility based methods, the outranking relation approach, the DEA based methods and the distance

functions based methods. Furthermore, one of the most extended classifications in the literature of *CI*s differentiates between compensatory and non-compensatory techniques (Asadzadeh *et al.*, 2017). The former refers to the existence of trade-offs among the different criteria, while the latter limits the compensatory character among the individual indicators.

Following this classification, depending on the MCDM methods used to construct *CI*s, the compensatory character between indicators varies. First, some MCDM approaches allow for full compensation among the criteria, such as the Simple Additive Weighting (SAW), the Utility Theory Additive (UTA), the Simple Multi-Attribute Rating Technique ($SMART$), the Data Envelopment Analysis (DEA) and The Technique for Order Preferences by Similarity to Ideal Solutions ($TOPSIS$). Second, other MCDM methods limit the compensation degree, such as the Multi-Utility Theory ($MAUT$), the Multi-Attribute Theory ($MAVT$) or the Weighted Product (WP) technique, which adopt multiplicative instead of additive functions and they only allow compensation between indicators within certain limitations. Third, some MCDM techniques used to construct *CI*s limit or completely prevent compensation, such as outranking methods: Elimination and Choice Expressing Reality ($ELECTRE$) and Preference Ranking Organization Method for Enrichment Evaluations ($PROMETHEE$). Some of these techniques have been applied in the *HE* field, such as Murias *et al.* (2008) who present a *CI* for quality assessment in the Spanish public university system using DEA, or Ding and Zeng (2015) who use TOPSIS combined with information entropy weight to analyse the overall performance of 68 Chinese universities. Giannoulis and Ishizaka (2010) adopt ELECTRE III to evaluate the performance of British universities, reflecting personal preferences, while Jati and Durai Dominic (2017) propose a complete ranking of university websites using the PROMETHEE II method.

Furthermore, some MCDM techniques allow for different compensation degrees depending on the aggregation scenario chosen. In this line, Blancas *et al.* (2010) use a goal programming ($GPCI$) approach to develop two *CI*s, depending on the degree of compensation allowed between the criteria. Ruiz *et al.* (2020) generalize the original reference point scheme proposed by Wierzbicki (1980) and Wierzbicki *et al.* (2000), which was adapted later on to build *CI*s based on two reference levels (Ruiz *et al.*, 2011). This *MRP-WSCI* method allows to construct different *CI*s depending on the compensation degree among the indicators. Two different *CI*s are provided: the weak indicator (WCI), allowing for full compensation among the individual indicators, and the strong indicator (SCI), not allowing for any compensation.

When constructing *CI*s, a natural and comfortable way to provide preferential information for the decision makers consists of giving reference levels (Miettinen, 1999). In this regard, some MCDM methods based on the minimization of distance functions allow the assessment of reference levels by the decision maker. These reference levels can naturally define performance levels for the individual indicators and the corresponding distance function measures the position of each unit with respect to these levels. The reference levels can be established, in absolute or relative terms, according to the target of the analysis. First, they can be given by one or a group of decision makers, if they have enough knowledge about the problem and they wish to do so. In this case, the *CI* gives us an absolute measure of performance, with respect to these values, and these levels could be named as "absolute levels". Second, they can be set

statistically, given a data set. In this case, the *CI* measures the relative position of the units with respect to those belonging to the data set, and these levels could be named as "relative levels".

Within the distance functions based techniques family, the most popular MCDM techniques are TOPSIS (Hwang and Yoon, 1981), the compromise programming (Yu, 1973), the goal programming (Ijiri, 1965) and the reference point approach (Wierzbicki, 1980). However, not all of them allow the construction of *CI*s for different compensation degrees. Within this context, TOPSIS and the compromise programming method allow for a full compensation among the criteria, while only the *GPCI* approach, the double reference point method and its generalization, the *MRP-WSCI* approach, allow different compensation degrees. Specifically, the *GPCI* allows establishing just one reference level, while the double reference point method allows establishing two reference levels. This can be a limitation in some cases, since the use of only one or two reference levels may not be enough for the users. Instead, the *MRP-WSCI* approach allows the decision maker to establish, for each indicator, a desired number of reference levels, and the final outcome can be interpreted in terms of the position with respect to these levels.

1.4. Efficiency assessment: Data Envelopment Analysis

When considering the role of universities in contributing to the social, economic and cultural development, their level of the operational efficiency should be taken into account (Agasisti *et al.*, 2019). In this regard, governments around the world have been faced with increasing pressure on their finances, giving rise to the need to operate universities with a higher degree of efficiency (Abbott and Doucouliagos, 2003). Specifically, the general trend in reducing public funds for Spanish universities has led to define the efficiency in the Spanish *HE* sector as a priority (de la Torre *et al.*, 2017b). According to Salas Velasco (2020), performance evaluation of *HE* institutions is vital for judging the degree to which resources made available to the *HE* field are utilized efficiently in the process of obtaining desired outcomes.

In this direction, since the pioneering work by Charnes *et al.* (1978) and Charnes *et al.* (1981), the comparative analysis of the efficiency in education has been brought into the focus of public and policy interest (De Witte and López-Torres, 2017). There is no doubt that this topic will continue to dominate headlines in the field of *HE*, where a university is considered as efficient if it makes the best possible use of available inputs.

In practice, Data Envelopment Analysis (DEA) and its variations, have been the most used methodologies for measuring the efficiency in the context of *HE*. Specifically, in the last years, empirical research on the *HE* efficiency using DEA models is becoming increasingly important at the international level (e.g. Abbott and Doucouliagos, 2003; Johnes, 2006; Agasisti *et al.*, 2019). While in the Spanish panorama, according to Martínez-Campillo and Fernandez-Santos (2019), there has been limited research on the technical efficiency of Spanish public *HE* system (Berbegal-Mirabent *et al.*, 2013; de la Torre *et al.*, 2017a,b; Berbegal-Mirabent, 2018; de Jorge Moreno *et al.*, 2019; Salas Velasco, 2019; Martínez-Campillo and Fernandez-Santos, 2019; Salas Velasco, 2020). In this direction, the assessment of the Spanish public *HE* system has been tackled from different perspectives. For instance, some papers use universities, facul-

ties or university departments, among others, as units of evaluation (de Jorge Moreno *et al.*, 2019), while other papers analyse the relative effectiveness of the Spanish *HE* institutions in terms of "general-model", "teaching-only" (Fuentes *et al.*, 2016), "research-only" (Expósito-García and Velasco-Morente, 2018) and "knowledge transfer-only" model (de la Torre *et al.*, 2017b). In this thesis, we will deal with the comparison among all the 47 Spanish public universities as a unit of analysis.

DEA is a non-parametric technique, originally introduced by Farrell (1957) and extended by Charnes *et al.* (1978) and Charnes *et al.* (1981). In a DEA model, technical efficiency is defined as the relative ability of each decision making unit (DMU) in producing outputs given a certain set of inputs (Yang *et al.*, 2017). An important aspect to emphasize in the DEA model is the choice of a set of weights which combines several outputs and several inputs. This issue is left to DEA through a linear programming technique which chooses the best set of weights for each DMU to maximize the efficiency ratio (outputs/inputs).

Furthermore, both DEA mathematical formulation (variable returns, VRS, and constant returns to scale, CRS) have been used to evaluate the Spanish universities. For instance, de la Torre *et al.* (2017b) and Berbegal-Mirabent (2018) applied the VRS formulation, where each Spanish public university is compared with any other homogeneous university, that is, to another of the same "relative" size. Salas Velasco (2019) used, first, the CRS model, where the single university's dimension has no importance in defining efficiency performance, that is, Spanish universities face the same efficiency frontier, independent of their relative size. Then, the results are compared to the VRS model.

On the other hand, given that, in general, the Spanish public universities have little decision-making power on the inputs used, since a great part of them are given by the public *HE* funding, the output oriented of a DEA model specification is the most widely used in the Spanish *HE* field. In this direction, Berbegal-Mirabent *et al.* (2013); Fuentes *et al.* (2016); de la Torre *et al.* (2017b) and Martínez-Campillo and Fernandez-Santos (2019) used the output oriented model, where the level of inputs remains at the same level while the outputs are maximized. It should be noted that CRS results are invariant to the choice of input or output orientation, whereas this is not the case for VRS results.

Nevertheless, according to Palomares-Montero and García-Aracil (2011), the evaluation of *HE* institutions' activities is not easy because they are multi-product firms with complex production processes, and the perspective adopted depends on the aims of the assessment. Furthermore, most rankings existing nowadays, operate under the concept of "the higher the value for all the indicators, the higher the performance achieved" (González-Garay *et al.*, 2019). Hence, they can assign worse (or better) performances to universities regardless of whether they make an efficient (or inefficient) use of their resources.

1.5. Purposes and structure of the PhD thesis

In conclusion, the important role that *HE* plays towards the global competitiveness of countries confirms universities as key players of socio-economic development for national economies. There is no doubt that universities are key players as sources of knowledge and innovation: they build human capital, and enhance the social and cultural fabric that ensures that innovation can thrive (OECD, 2007). To this end, according to Dehon *et al.* (2009) and

Sánchez-Barrioluengo (2012), *HE* institutions fulfill different functions, such as the construction of knowledge (research), the diffusion of knowledge (teaching), and the valorization and use of knowledge (service to society).

In the course of the last years, university rankings have become one of the most favorite issues in the assessment of *HE* institutions (Sanz-Casado and De-Filippo, 2011; Glanzel and Debackere, 2009). In fact, since the publication of the first edition of the *ARWU Ranking* in 2003 and the successive international lists, such as the *Times Higher Education World University Ranking*[13], the *QS World University Ranking*[14], the *CWTS Leiden Ranking*[15] and the *Webometrics Ranking of World Universities*[16], the comparative evaluation of the quality of *HE* institutions has been brought into the focus of public and policy interest. In the Spanish panorama, the *U-Ranking* project[17] of the BBVA foundation and the IVIE institute, and the *CyD Ranking*[18] of the Foundation "Conocimiento y Desarrollo" are highlighted.

According to Marope *et al.* (2013), the existence of the well-known world university rankings shows that universities are really living "in an age of assessments and comparisons", which has led to define the rankings as "a global strategic instrument", instead of a simple "consumer product" (Hazelkorn, 2013). This is mainly due to the simple and synthetic way in which they offer the information, which facilitates comparisons among the different universities while simplifying them. Unfortunately, according to the latest edition (2019) of the *ARWU Ranking*, only 15% of the total of Spanish universities are among the top 500 *HE* institutions in the world, and only one of them, which is the university of Barcelona, is among the top 200. This data is frequently mentioned as proof of the limited quality and the low involvement of the Spanish university system in the international projection.

However, despite their popularity, rankings have come under some criticism for their purposes and for using poor research methods (Climent *et al.*, 2013). Meanwhile, Rauhvargers (2011) highlights that the criticism received by university rankings is mainly due to their limited coverage and opacity. In this direction, most of rankings, especially the commercial ones, use simple aggregation techniques. This implies a compensatory character among the criteria, producing single scores, which makes it difficult for users to detect the differences among universities according to their performance on the different indicators. In this way, the value given to the rankings is linked to how they are made as well as the metrics used for it (Rauhvargers, 2011).

On these premises, the present doctoral thesis aims to go beyond simply providing a ranking of the Spanish public university system. To this end, in order to evaluate the performance of the Spanish public universities, first, we propose a novel methodology based on the multicriteria reference point scheme (*MRP-WSCI*; Ruiz *et al.*, 2020). Then, a combination of the *MRP-WSCI* approach with the DEA analysis is proposed.

As mentioned in Section 1.4, there is no doubt that the literature provides a wide range of papers that use the non-parametric DEA method to evaluate *HE*

[13]https://www.timeshighereducation.com/world-university-rankings
[14]http://www.topuniversities.com/university-rankings
[15]http://www.leidenranking.com/
[16]http://www.webometrics.info/es
[17]http://www.u-ranking.es/index2.php
[18]http://www.rankingcyd.org/

institutions. In fact, since the doctoral thesis of Rhodes (1978), where the DEA method was proposed and used, *HE* has been one of the focus of the efficiency evaluation through the DEA approach.

However, the number of outputs considered in most published papers that use the DEA method ranges from two to four, being the number of graduate students and the number of publications the most frequently used. However, according to González-Garay *et al.* (2019), the academic quality in *HE* is inherently multi-faceted, and therefore, difficult to assess using just a few indicators. In this direction, the evaluation of such complex phenomena, as is the case of the *HE* field, requires to take into account two important aspects. First, the *HE* system is quite complex, diverse and it includes different tasks interconnected with each other. Second, in today's globalized and highly competitive environment, the evaluation of the university system requires the use of techniques that are able to manage its complexity (Attardi *et al.*, 2017). Thus, given the complex and multidimensional nature of the *HE* field, the purpose of this thesis consists of considering all the aspects regarded as relevant to assess university performance. To this end, the performance of the Spanish public universities is assessed using the *MRP-WSCI* approach. Then, these results are used as outputs in the DEA analysis.

Besides, the assessment of universities using the *MRP-WSCI* approach has two main advantages. First, reference levels are used for each indicator and thus, the scores are easily interpreted as the current position of the university with respect to these levels. Also, when measuring multidimensional frameworks, such as the *HE* field, the use of *CI*s is highly suitable. According to Saisana and Tarantola (2002), the *CI* should ideally measure multidimensional concepts which cannot be captured by a single indicator, providing simple comparisons that can be used to illustrate complex and sometimes elusive issues in wide-ranging fields. Thus, a *CI* is a mathematical combination of all individual indicators, with the main advantage of offering a simplified view of a much more complex problem. In this direction, the use of the *MRP-WSCI* method allows to construct *CI*s for different compensation degrees in such a way that the scores, apart from giving an overall performance measure of the universities, also provide warning signals that let the user detect improvement areas.

In pursuit of our aims, this PhD thesis is presented as a compendium of articles and it is organized in the following way:

- **Chapter 2** comprises the three publications that constitute the compendium:
  - **Subchapter 2.1** is dedicated to the theoretical and methodological framework:
    - Since the literature today provides a wide range of methodologies for the construction of university performance *CI*s, the first article, El Gibari *et al.* (2019), aims to identify the different MCDM methods used to construct *CI*s. Besides, in this paper, these methods have been classified in five categories: the elementary methods, the value and utility based methods, the outranking relation approach, the data envelopment analysis based methods and the distance functions based methods. In general, our review has shown a clear tendency towards an increasing number of papers that use MCDM methods to construct *CI*s since 2014.

However, this review has also allowed us to identify a shortage in the construction of *CI*s using MCDM methods in the field of *HE*. This article is published in a *Scopus* database with an impact index of 0.378 (*Q2*). Its Citation Index is 33 in *Google Scholar* and its Field-Weighted Citation Impact is 8.36 in a *Scopus* database (date of reference: 31 May 2020).

- The second paper, Ruiz *et al.* (2020), comprises a generalization of the double method reference point method (Ruiz *et al.*, 2011), called *MRP-WSCI*. The reference point method was originally proposed by Wierzbicki (1980), with the idea of solving multi-objective programming problems, by generating efficient solutions that were closer to certain reference levels (desired) for the stated goals. Wierzbicki *et al.* (2000) extended this methodology to a double reference point scheme, suggesting its use for the construction of objective rankings, where a reservation level (minimum level regarded as admissible) and an aspiration level (level regarded as desirable) can be specified for each objective. Later on, Ruiz *et al.* (2011) and Cabello *et al.* (2014) further adapted and developed this approach for the construction of composite indicators with different compensation degrees and based on the weights given by the user. However, despite the advantages of the aforementioned double reference point method, it has two possible drawbacks. First, in some cases, the use of only two reference levels may not be enough for the users. Second, the results are given in a fixed scale (from -1 to 2). This can be a limitation when combining this methodology with others, such as traditional DEA models. In this direction, the decision makers may wish to use a scale with different ranges. Thus, the *MRP-WSCI* method can be applied in a more general context, where the decision makers can establish any number of reference levels, which somehow define the performance levels (e.g., very poor, poor, fair, good, very good,...), and any scale can be used. This paper is published in the *OMEGA: The International Journal of Management Science*, which is indexed in the *JCR* with a noteworthy impact factor of 5.341 (*Q1*). Specifically, *OMEGA* is the second best journal within its category "*Operations Research & Management Science*". The Citation Index of this publication is 3 in Google Scholar and its Field-Weighted Citation Impact is 0.52 in a *Scopus* database (date of reference: 31 May 2020).

- **Subchapter 2.2** is devoted to the empirical application:
  - The third publication, El Gibari *et al.* (2018), comprises an application of multicriteria analysis techniques, based on the double reference point method (Ruiz *et al.*, 2011), to the Andalusian public university system. In order to get a more accurate vision of each university mission, the three basic missions (research, teaching and technology transfer) are analyzed separately. Within the scopes of multicriteria and *HE* field, the results obtained are particularly useful because they serve not only to measure the relative position of each university, but also

to offer a warning system to assist in strategic decision making. Besides, aggregations for different compensation degrees are provided for each Andalusian university. This paper is published in the *Journal of Informetrics*, which is indexed in the *JCR* with an impact factor of 3.879 (*Q1*), in the category "Computer Science & Interdisciplinary Applications". Its Citation Index is 7 in Google Scholar and its Field-Weighted Citation Impact is 0.90 in a *Scopus* database (date of reference: 31 May 2020).

- **Chapter 3** presents some additional research questions related to the aforementioned publications:
  - **Subchapter 3.1** proposes an extension of the theoretical and methodological framework:
    - In order to make decisions in our dynamic environment, the use of *CI*s has become indispensable, since they are powerful tools for summarizing, focusing and condensing the complexity of such environment. However, the information provided by the *CI* should be maximal, allowing to detect improvement areas, rather than just ranking the units. In this direction, the literature provides few approaches that allow the construction of *CI*s for different compensation degrees. On these premises, we illustrate the behavior of some methods that build *CI*s allowing different compensation degrees. Specifically, we analyze the additional information provided by each of these methods when the compensatory and non-compensatory scenarios are jointly considered. An extended version of this research work has been submitted to the "*International Journal of Information Technology and Decision Making*" and is presently under review. This journal is indexed in the *JCR* with an impact factor of 2.862 (*Q2*), in the categories "*Operations Research & Management Science*" and "Computer Science & Information Systems".
  - **Subchapter 3.2** proposes an expansion of the multicriteria analysis to the Spanish public university system:
    - Finally, we evaluate the Spanish public universities performance by combining the *MRP-WSCI* approach and the DEA analysis. We discuss how this combination can add value to traditional *CI*s, not only in terms of securing high performance of the Spanish public universities, but also in how to make an appropriate use of resources. Our analysis highlights the need to assess the university performance by using both compensatory and non-compensatory scenarios. Furthermore, combining the use of *CI*s with the DEA analysis gives us a more complete picture of the institutions under assessment, allowing universities to detect their weaknesses and strengths. The combination is illustrated using data on research, teaching and technology transfer for 47 Spanish public universities for the most recent academic year available, 2016-2017. It should be noted that an extended version of this research work is currently under development, which we plan to

submit shortly to a journal indexed in the *JCR* database and of wide international dissemination.

- **Chapter 4** summarizes the most significant results of this PhD thesis. Finally, the work ends with some conclusions and future research lines.

2. The compendium of publications

The flexibility of using *CI*s to manage today's complex environment has led to a considerable increase in their number, as well as the existence of numerous methods for their construction. Thus, this Chapter presents the compendium of publications, where the main relevant aspects related to the theoretical and methodological framework in the construction of *CI*s are tackled in Subchapter 2.1. First, we identify the different MCDM methods used for aggregating single indicators into composite ones. Then, we present an innovative methodology based on the multicriteria reference point scheme. Finally, in Subchapter 2.2, we show and analyze the empirical application of the double reference point method to the Andalusian public university system.

2.1. The theoretical and methodological framework

2.1.1. Building composite indicators using multicriteria methods: a review

El Gibari, S., Gómez, T. & Ruiz, F. (2019). Building composite indicators using multicriteria methods: a review. *Journal of Business Economics*, 89, 1-24.

DOI: https://doi.org/10.1007/s11573-018-0902-z

This article is published in the *Scopus* database with an impact index of 0.378 (*Q2*). Its Citation Index is 33 in *Google Scholar* and its Field-Weighted Citation Impact is 8.36 in a *Scopus* database (date of reference: 31 May 2020).

Abstract

Composite indicators are increasingly recognized as a useful tool in policy analysis and public communication. They provide simple comparisons of units that can be used to illustrate the complexity of our dynamic environment in wide-ranging fields, such as competitiveness, governance, environment, press, development, peacefulness, tourism, economy, universities, etc. Their construction has been dealt with from several angles. Some authors claim that MCDM techniques are highly suitable in multidimensional frameworks when aggregating single indicators into a composite one, since this process involves making choices when combining criteria of different natures, and it requires a number of steps in which decisions must be made. In this paper, we conduct a literature review of papers published after 2002 in leading international journals indexed in a recognized database (*JCR*), in order to identify the different MCDM methods used for aggregating single indicators into composite ones. They have been classified in five categories: the elementary methods, the value and utility based methods, the outranking relation approach, the data envelopment analysis based methods and the distance functions based methods. In general, our review has shown a clear tendency towards an increasing number of papers that use MCDM methods to construct composite indicators since 2014.

2.1.2. *MRP-WSCI*: Multiple reference point based weak and strong composite indicators

Ruiz, F., El Gibari, S., Cabello, J.M. & Gómez, T. (2020). *MRP-WSCI*: Multiple reference point based weak and strong composite indicators. *Omega*, 95, 102060.

DOI: https://doi.org/10.1016/j.omega.2019.04.003

This article is published in the *JCR* database with an impact factor of 5.341 (*Q1*). Its Citation Index is 3 in *Google Scholar* and its Field-Weighted Citation Impact is 0.52 in a *Scopus* database (date of reference: 31 May 2020).

Abstract

The construction of composite indicators is becoming an increasingly important way of managing great amounts of information obtained from multiple single indicators. Obviously, the aggregation of several indicators in a single synthetic measure always implies some loss of information on the way. For this reason, this aggregation must be done in the most informative possible way, so that the results obtained can be easily interpreted, and in an efficient way, so that the information contained in the composite indicator is maximal. In this paper, we propose a novel methodology based on the multicriteria reference point scheme. The decision maker can establish any number of reference levels for each indicator, and the final outcome can be interpreted in terms of the position with respect to these levels. Besides, two different aggregations are proposed: the weak indicator, allowing for full compensation among the single indicators, and the strong indicator, not allowing for any compensation. The joint visualization of both composite indicators provides valuable information that may be unnoticed using other existing approaches. The approach is illustrated using an example based on the EU-Regional Social Progress Index (EU-SPI).

2.2. The empirical application. Performance measurement

2.2.1. Evaluating university performance using reference point based composite indicators

El Gibari, S., Gómez, T. & Ruiz, F. (2018). Evaluating university performance using reference point based composite indicators. *Journal of Informetrics*, 12 (4), 1235-1250.

DOI: https://doi.org/10.1016/j.joi.2018.10.003

This article is published in the *JCR* database with an impact factor of 3.789 (*Q1*). Its Citation Index is 7 in *Google Scholar* and its Field-Weighted Citation Impact is 0.90 in a *Scopus* database (date of reference: 31 May 2020).

Abstract

In this paper, we propose the application of a novel methodology to build composite indicators, in order to evaluate university performance. We analyze separately the three basic dimensions of our university system (research, teaching and technology transfer), because we are interested in getting a more accurate vision of each of them. In order to build the composite indicators, we use a multicriteria analysis technique, based on the double reference point method. One advantage of this technique is the possibility to use reference levels, in such a way that the results obtained are easily interpreted in terms of the performance of the university with respect to these levels. Besides, aggregations for different compensation degrees are provided. In order to illustrate the advantages of this method, it has been applied to evaluate the performance of the public universities of the Spanish region of Andalucía, for year 2008. The results show that the performance of the Andalusian public universities in the teaching block is better than in the research and technology transfer blocks. The application lets us conclude that the methodology offers a warning system to assist in strategic decision making, and the values of the indicators allow us to find fields of improvement in all areas.

3. Additional research questions

3.1. Comparisons of the *MRP-WSCI* approach with other schemes for building compensatory and non-compensatory composite indicators

According to Saaty and Ergu (2015), MCDM research has become a hot research topic because many complex practical decision problems involve multiple criteria as well as multiple objectives. In this line, one way of dealing with MCDM problems is to convert them into a single criterion problem (Teodorovic and Edara, 2007). To this end, decision makers could use some of the compensatory methods (that allow advantages of one indicator to be traded for disadvantages of another), or non-compensatory methods (that do not allow advantages of one indicator to be traded for disadvantages of another).

Specifically, depending on the aggregation method chosen, the compensatory character between indicators varies (Munda, 2008). This compensation can be: full (additive methods and compensatory multicriteria approaches), partial (geometric aggregations and other partially-compensatory methods), or zero (non-compensatory techniques).

In this direction, the compensatory *CI*s are designed in order to provide an overall measure of the performance of each unit, while the non-compensatory *CI*s are useful tools to point out bad behaviors. Of course, when evaluating complex phenomena, it is necessary to seek information on the different dimensions which contribute to characterize a multidimensional phenomenon (Mazziotta and Pareto, 2017). In this line, the *CI* reduces the dimensions with an evident loss of information, but it manages this complexity, allowing a single measure that is more communicative. In general, the joint consideration of both compensatory and non-compensatory *CI*s should ideally provide a richer amount of information about the performance of the units in a simple and easy to understand way.

According to Ciommi *et al.* (2017), the simplest and most common aggregation methods are based on the arithmetic average and the geometric average of individual indicators. In general, the former allows for full compensation, while the geometric average aggregation limits the compensatory character among the criteria. However, in addition to the *MRP-WSCI* approach (Ruiz *et al.*, 2020) proposed in this PhD thesis, there are other procedures to build *CI*s for different compensation degrees in the literature, like in Tarabusi and Guarini (2013) building a function that somehow incorporates the two extreme cases of compensation: the full compensation represented by the weighted arithmetic mean and the zero compensation by the minimum function. Furthermore, within the MCDM framework, Blancas *et al.* (2010) use a goal programming (*GPCI*) approach to develop two *CI*s, depending on the degree of compensation allowed between the strengths and weaknesses of the units.

Thus, the aim of this Subchapter is to propose comparisons of the *MRP-WSCI* approach with other methods that specifically build *CI*s with different compensation degrees among indicators. Namely, comparisons will be made between the joint use of the arithmetical and geometrical averages, the *Mean-Min* Function (Tarabusi and Guarini, 2013), a *GPCI* approach (Blancas *et al.*, 2010) and the *MRP-WSCI* method (Ruiz *et al.*, 2020). To this end, we test the information provided by both compensatory and non-compensatory scenarios by applying them to a real case, through the construction of the childhood vulnerability index, in the African context. We will compare the actual meaning of the results obtained by each of the methods compared, as well as the amount of added information provided by the non-compensatory scheme.

3.2. Evaluating the Spanish public universities performance and efficiency by combining the *MRP-WSCI* approach with Data Envelopment Analysis

The literature notes that most university rankings existing nowadays, operate under the concept of "the higher the value for all the indicators, the higher the performance achieved" (González-Garay *et al.*, 2019). Hence, they can assign worse (or better) performances to universities regardless of whether they make an efficient (or inefficient) use of their resources.

Furthermore, governments around the world have been faced with increasing pressure on their finances, giving rise to the need to operate universities with a higher degree of efficiency (Abbott and Doucouliagos, 2003). In this direction, since the pioneering work by Charnes *et al.* (1978) and Charnes *et al.* (1981), the comparative analysis of the efficiency in education has been brought into the focus of public and policy interest (De Witte and López-Torres, 2017). Specifically, following the economic recession in 2008, the general trend in reducing public funds for Spanish universities has led to define the allocation of resources in an efficient way in the Spanish public higher education sector as a priority (de la Torre *et al.*, 2017b; Martínez-Campillo and Fernandez-Santos, 2019).

There is no doubt that DEA is an excellent tool to measure the relative performance of a set of units with multiple inputs and outputs (Akbari *et al.*, 2020). In the literature of DEA, some authors, such as Dyson *et al.* (2001), claim that the number of units has to be at least twice the number of inputs and outputs to be used, while others suggest that the number of units has to be three times higher. In this direction, problems of discrimination between efficient and inefficient units often arise when there is a relatively large number of variables when compared to the number of units (Charles *et al.*, 2019).

However, the academic quality in higher education is inherently multi-faceted, and therefore, difficult to assess using just a few indicators (González-Garay *et al.*, 2019). In this direction, the evaluation of such complex phenomena, as is the case of the higher education field, requires to take into account two important aspects. First, the higher education system is quite complex, diverse and it includes different tasks interconnected with each other. It can be asserted that the university system is basically a multidimensional concept, which makes it difficult to condense the diversified work going on within universities into a single number or ranking. Second, in today's globalized and highly competitive environment, the higher education system should have a powerful tool to detect improvement areas, rather than just ranking the universities. Consequently, the evaluation of the university system requires the use of techniques that are able to manage its complexity (Attardi *et al.*, 2017).

Therefore, in this additional research question, we discuss how the combination of the *MRP-WSCI* approach proposed in this PhD thesis with the well-known DEA method can add value to traditional composite indicators, not only in terms of securing high performance but also in how to make an appropriate use of resources. First, we analyze the added information provided by the joint consideration of the *MRP-WSCI* composite indicators and the DEA analysis. Second, we test the robustness of the DEA efficiency scores for the Spanish public universities, with respect to changes of the weights when constructing the *MRP-WSCI* composite indicators. To this end, we apply the suggested methodology to the case of the Spanish public universities.

4. Results, conclusions and future research lines

4.1. Results

In the last years, the assessment of the quality and the excellence of *HE* institutions has been brought into the focus of public and policy interest. In this direction, most university rankings have developed and used *CI*s to provide rankings of universities or countries, which is not surprising, due to their aims of summarizing, focusing and condensing the complexity of our dynamic environment. There is no doubt that *CI*s are becoming an increasingly important way of bench-marking universities' performance.

In general, the construction of *CI*s involves making choices, and combining criteria of different natures. Thus, many works have attempted to construct them by falling back on MCDM methods. In particular, our results have shown a clear tendency towards an increasing number of papers that use MCDM methods to build *CI*s since 2014. More specifically, our review has shown that, since 2012, there has been an increasing tendency to adopt the distance functions based methods. Furthermore, although the MCDM methods have been applied to construct *CI*s in a wide variety of fields, such as sustainability (29%), environment (12%), business (8%), energy (6%), tourism (5%), human development (5%), water service (4%) and investment (4%), this review has allowed us to identify a shortage in the field of *HE*. Moreover, most of the papers reviewed have been published in journals belonging to the category of "Environmental Sciences" (48%), which seems logical since sustainability and environment are the most assessed fields using MCDM methods to construct *CI*s. Other papers have chosen journals focusing on the category of "Operations Research & Management Science" (27%). However, there are just a few articles published in the category of "Education & Educational Research".

Besides, when aggregating individual indicators into a composite one, an important aspect to emphasize is the compensation degree among the different criteria. This way, in the field of MCDM methods used to construct *CI*s, some techniques allow for full compensation (e.g. the Simple Additive Weighting, SAW, the Utility Theory Additive, UTA, the Simple Multi-Attribute Rating Technique, SMART, the DEA and TOPSIS methods), partial compensation (e.g. the Multi-Utility Theory, MAUT, the Multi-Attribute Theory, MAVT, or the Weighted Product method, WP), or no compensation (e.g. ELECTRE and PROMETHEE).

Other MCDM methods allow for different compensation degrees, such as Blancas *et al.* (2010), who use a goal programming (*GPCI*) approach to develop two *CI*s, depending on the degree of compensation allowed between the strengths and weaknesses of the units, or the *MRP-WSCI* method proposed in this PhD. In particular, both approaches carry out a distance based normalization. However, the *GPCI* approach does not force the normalized indicators within a predetermined range. In this line, the minimum and maximum values of the normalized indicators, and therefore, those of the corresponding *CI*s, are not bounded a priori. The *GPCI* approach allows to identify the strengths and weaknesses of the units, taking into account the position of each unit with respect to one reference level. However, the *MRP-WSCI* method normalizes the indicators to have an identical and bounded range, allowing for multiple reference levels. This way, the *MRP-WSCI* approach can be applied in more general

contexts, where the decision makers may wish to establish more performance levels, or use different measurement units. Therefore, the values of the *CI*s proposed can be easily interpreted as the global position of the corresponding university with respect to hypothetical global reference levels. Thus, the final score is not just a number, but a more informative measure. Besides, in this PhD tesis, we have shown that the *MRP-WSCI* composite indicators satisfy a series of properties which are regarded in the scientific literature as highly desirable (existence and uniqueness, symmetry, transitivity, monotonicity, invariance and exhaustivity properties).

Ideally, the information provided by any aggregation method should be a powerful tool for decision makers, allowing to detect improvement areas, rather than just ranking the units. This way, we have proved in our research that the joint consideration of both compensatory and non-compensatory scenarios provides richer information about the performance of the units in a simple and easy to understand way. In fact, due to the complexity of our dynamic environment, the information provided should be as maximal as possible. In this direction, in order to test the amount of added information provided by the joint consideration of the compensatory and non-compensatory scenarios, we have proposed a comparison among several methods that allow for different compensation degrees, depending on the aggregation scenario chosen: the *Mean-Min* Function, the *GPCI* method and the *MRP-WSCI* approach. Besides, since the arithmetic and geometric averages are the most common aggregation methods in the literature, both averages have been also used.

Our results reveal that the non-compensatory scenarios of the classical average and the *GPCI* allow, in practice, for a partial compensation, while the non-compensatory *CI*s provided by the *Mean-Min* Function and the *MRP-WSCI* approach do not allow for any compensation among the indicators. Furthermore, the *Mean-Min* Function and the *MRP-WSCI* approach provide richer information when both compensatory and non-compensatory scenarios are considered. In order to prove this assertion, we have carried out a series of computational experiments, where all the methods have been applied to a number of problems that have been randomly generated. We can conclude that the results obtained are very homogeneous across all the simulation sizes considered. In fact, the *Mean-Min* Function and the *MRP-WSCI* method present the smallest correlation between the compensatory and the non-compensatory schemes, in almost a 50% of the simulations each, and only the *GPCI* method obtains the least correlated results in a residual number of cases.

On the other hand, the normalization procedure of the *Mean-Min* Function allows to identify the units with a performance above or below the average, while the *GPCI* and the *MRP-WSCI* approaches carry out a distance based normalization, taking into account the position of each unit with respect to a set of reference levels (one, in the case of the former, or more, in the latter). Moreover, the minimum and maximum values of the scale of the *CI*s are predefined for the classical averages and the *MRP-WSCI*, while in the case of the *Mean-Min* Function and a *GPCI* approach, there are no upper and lower levels, so the results can take any real value.

The complexity, diversity and multidimensional nature of the *HE* field make the use of MCDM methods to construct *CI*s highly suitable, since this process involves making choices when combining criteria of different natures, and it requires a number of steps in which decisions must be made. In fact, the appli-

cation of the double reference point method to the Andalusian public university system constitutes a contribution in both scopes, multicriteria and *HE*.

First, in opposition to the well-known university rankings, the main advantage of our research has been the possibility to analyze *HE* institutions from different perspectives simultaneously by developing and using *CI*s jointly with individual indicators of university performance. On the one hand, we minimize the amount of information lost in the aggregation steps, since the university performance has been assessed in such a way that reference levels are used for each indicator and thus, the scores are easily interpreted as the current position of the university with respect to these levels. Moreover, we have developed *CI*s, for different compensation degrees, for each of the three main missions of universities, in such a way that the scores, apart from giving an overall performance measure of the universities, also provide warning signals that may assist user in strategic decision making for policy purposes, by identifying the weaknesses and strengths of each university.

With respect to the Andalusian public university system, our results reveal that, in general, their performance in the teaching mission is better than in the research and technology transfer missions, although the values of the indicators suggest that all the universities have still much room for improvement. In fact, our results show that, in the teaching mission, the "international attractiveness" and "internships" are weaknesses for most of the Andalusian public universities. While, in the research missions, a large number of weaknesses takes place in the "publications" and in the "amounts of the National Plan projects per doctor". Thus, special attention has to be paid to these aspects, since the internationalization, the scientific research, and the public investment in R&D of *HE* institutions features among the sector's key transformations, specially in the European context.

In addition to providing valuable information about the performance of the universities through the joint consideration of the compensatory and the non-compensatory scenarios, rankings for different compensation degrees are also provided. In this case, we can affirm that the ranking for the research mission is more stable than the teaching and technology transfer missions. Anyway, no ranking is highly affected by slight changes in the compensation allowed. Even comparing extreme situations, there are only minor changes in the positions of the Andalusian public universities. It should be noted that the evaluation of the Andalusian public universities has served as an illustration of the double reference point method. Thus, the application is devoted to stating the potential advantages of the methodology as compared to the ones traditionally used. However, the assessment results would be more valuable if strategic decision makers in higher education had been involved in the research.

Finally, the use of techniques that are able to manage the present complex environment is highly suitable. Thus, in this doctoral thesis, a combination, in a robust environment, of the *MRP-WSCI* approach with the DEA analysis has been proposed, which can be used as a supporting tool so that decision makers can detect and improve important factors of the educational efficiency and effectiveness of the Spanish public universities. This combination allows us to consider as outputs all the aspects regarded as relevant to assess university performance. In this direction, in addition to providing an overall performance measure and detecting the actions needed to improve the performance of the Spanish universities through the application of the *MRP-WSCI*, these perfor-

mance indicators (compensatory and non-compensatory) are compared with the resources available by each Spanish public university. Thus, the DEA analysis gives us a more complete picture of the universities under assessment, allowing them to better understand their weaknesses and strengths.

Among the 47 Spanish public universities in our data set, our results show that 18 universities are efficient in the compensatory scenario, and 17 in the non-compensatory one, and 14 of them are efficient in both DEA. On the one hand, the results reveal that, on average, the scores among the Spanish public universities are quite similar in both analyses (1.11 and 1.12, respectively). This indicates that, keeping their inputs at the current levels, Spanish universities should expand, on average, their outputs by 11% in the compensatory scenario and 12% in the non-compensatory one, in order to operate efficiently.

In general, most universities tend to be less inefficient when the overall performance is analyzed (specifically, 21 out of 47). In fact, our results show that certain Spanish public universities are much more inefficient in the non-compensatory scenario, while in the compensatory one, their scores improve substantially. This indicates that these universities perform poorly for at least one criterion (worse performance in the non-compensatory scenario), but these unfavorable performances are compensated by good performances in others, resulting in a better overall performance (compensatory scenario). These findings highlight the need to assess the university performance by using scenarios, where the compensatory one gives an overall performance measure of the universities, while the non-compensatory one helps to detect the actions needed to improve the performances of the universities.

Furthermore, our analysis shows that, in addition to the Spanish public universities that are always present in the most prestigious university rankings, some other ones that occupy the lowest positions in some missions in the *MRP-WSCI* analysis are among the efficient ones in the DEA analysis. In fact, some Spanish public universities, even having quite "modest" resources, manage to make an efficient use of them, achieving "good" performances in some areas.

Summing up, the combination of the *MRP-WSCI* composite indicators and the DEA analysis provides a richer amount of information. First, since the overall performance measure is analyzed through the compensatory scenario, and the detection of the actions needed to improve the performances of the Spanish public universities is carried out through the non-compensatory one, the joint consideration of both scenarios is highly suitable. Second, combining both scenarios with the DEA analysis helps the decision maker(s) in deciding which aspects must be corrected, by better understanding the the weaknesses and strengths of each Spanish public university.

4.2. Final conclusions and future research lines

Nowadays, a better performance and commitment of the Spanish public university system to create a competitive advantage is crucial, since it is a clear evidence that, in the future, universities will have an even more vital contribution to our society. Taking into account the complexity, diversity and multidimensional nature of a university system, this doctoral thesis has attempted to analyze the performance and efficiency of the Spanish public *HE* system by building compensatory and non-compensatory *CI*s from a multicriteria approach.

In opposition to the well-known university rankings, the advantage of this thesis is threefold. First, the possibility to analyze the Spanish public universities from different perspectives simultaneously, by developing and using *CI*s, where the information about the performance of the individual indicators is not lost throughout all the procedure. Second, the interpretation of the *MRP-WSCI* composite indicators is more intuitive and easy for the decision maker, since they represent the global position of the corresponding university with respect to the given reference levels. This way, the final score is not just a number, but a more informative measure. Third, in order to evaluate the Spanish university system, as most of the rankings existing nowadays, we have developed and used *CI*s. However, aggregations for different compensation degrees for each of the three main missions of universities are provided, which, to our knowledge, has not yet been adopted to evaluate university performance. This way, the joint consideration of the compensatory and non-compensatory scenarios provides a richer amount of information for decision-making. Specifically, apart from providing an overall measure of the Spanish public universities, the *MRP-WSCI* methodology provides alert signs which let the decision maker detect bad performances in certain indicators that may remain unnoticed otherwise. Furthermore, the combination of the *MRP-WSCI* approach with the DEA analysis provides a fuller and more nuanced picture of each Spanish public university, allowing to detect the weaknesses and strengths of the *HE* institutions. In fact, this combination can be used as a supporting tool so that decision makers can improve important factors of the efficiency and effectiveness of the Spanish public *HE* system.

A relevant conclusion of this PhD thesis is a quite remarkable variety across results within the Spanish universities missions. Although the values of the indicators suggest that, in general, all the universities have still much room for improvement in all areas. In this line, our results reveal that, despite having a good performance in the teaching and research missions, some Spanish public universities perform poorly in the technology transfer mission, and consequently a special attention should be paid to this mission. Besides, in addition to the Spanish public universities that are always present in the most prestigious university rankings, there are others that, even having quite "modest" resources, they manage to make an efficient use of them, achieving "good" performances in some areas.

To conclude, the present doctoral thesis has focused on the Spanish public university system, analyzing a specific year. However, in order to study the evolution of the Spanish universities along a given period, it would also be interesting to extend this research to a dynamic scheme. Specifically, it is worth

examining whether the Spanish public universities manage to maintain their performance in each mission analyzed in the long term. To this end, the reference levels can be given in two different ways. First, in order to analyze the situation of each Spanish public university compared to a particular year, fixed reference levels for the whole period can be used. This allows analyzing the evolution of the universities compared to the situation they presented in the baseline year. Alternatively, these reference levels might change from one year to another according to new requirements established by the academic authorities. In this case, an absolute measure of the performance of the Spanish public universities is given.

Furthermore, it would also be interesting to make a wider comparison among the Spanish public and private universities, where the results will give an idea of the relative position of each Spanish university with respect to all the institutions of the Spanish *HE* system. Besides, in order to get the most out of the Bologna Process, a wider comparison among universities from different European countries could be useful for the Spanish Government purposes. It should be noted that this comparison can only be carried out whenever the quality and homogeneity of the data used for the different universities make it possible.

5. References

References

Abbott, M. and Doucouliagos, C. (2003). The efficiency of Australian universities: A data envelopment analysis. *Economics of Education Review*, **22(1)**, 89–97.

Agasisti, T., Barra, C., and Zotti, R. (2019). Research, knowledge transfer, and innovation: The effect of Italian universities' efficiency on local economic development 2006−2012. *Journal of Regional Science*, **59(5)**, 819–849.

Aguillo, I. F., Ortega, J. L., and Fernández, M. (2008). Webometric Ranking of World Universities: Introduction, Methodology, and Future Developments. *Higher Education in Europe*, **33(2-3)**, 233–244.

Akbari, N., Jones, D., and Treloar, R. (2020). A cross-European efficiency assessment of offshore wind farms: A DEA. *Renewable Energy*, **151**, 1186–1195.

Alasehir, O., Perit-Cakir, M., Acarturk, C., Baykal, N., and Akbulut, U. (2014). URAP-TR: a national ranking for Turkish universities based on academic performance. *Scientometrics*, **101**, 159–178.

Asadzadeh, A., Kötter, T., Salehi, P., and Birkmann, J. (2017). Operationalizing a concept: The systematic review of composite indicator building for measuring community disaster resilience. *International Journal of Disaster Risk Reduction*, **25**, 147–162.

Attardi, R., Cerreta, M., Sannicandro, V., and Torre, C. M. (2017). Non-compensatory composite indicators for the evaluation of urban planning policy: The Land-Use Policy Efficiency Index (LUPEI). *European Journal of Operational Research*, **264(2)**, 491–507.

Berbegal-Mirabent, J. (2018). The influence of regulatory frameworks on research and knowledge transfer outputs: An efficiency analysis of Spanish public universities. *Journal of Engineering and Technology Management*, **47**, 68–80.

Berbegal-Mirabent, J., Lafuente, E., and Solé, F. (2013). The pursuit of knowledge transfer activities: An efficiency analysis of Spanish universities. *Journal of Business Research*, **66**, 2051–2059.

Blancas, F. J., Caballero, R., González, M., Lozano-Oyola, M., and Pérez, F. (2010). Goal programming synthetic indicators: An application for sustainable tourism in Andalusian coastal countries. *Ecological Economics*, **69**, 2158–2172.

Bonaccorsi, A. and Cicero, T. (2016). Nondeterministic ranking of university departments. *Journal of Informetrics*, **10**, 224–237.

Cabello, J., Ruiz, F., Pérez-Gladish, B., and Méndez-Rodríguez, P. (2014). Synthetic indicators of mutual funds' environmental responsibility: An application of the reference point method. *European Journal of Operational Research*, **236**(1), 313–325.

Casani, F., De Filippo, C., García-Zorita, C., and Sanz-Casado, E. (2014). Public versus private universities: assessment of research performance; case study of the Spanish university system. *Research Evaluation*, **23(1)**, 48–61.

Cervo, A. L. and Bervian, P. A. (1980). *Metodología científica*. McGraw-Hill, Bogotá.

Charles, V., Aparicio, J., and Zhu, J. (2019). The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis. *European Journal of Operational Research*, **279(3)**, 929–940.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, **2**, 429–444.

Charnes, A., Cooper, W. W., and Rhodes, E. (1981). Evaluating program and managerial efficiency: An application of Data Envelopment Analysis to program Follow Through. *Management Science*, **27(6)**, 668–697.

Cinelli, M., Coles, S. R., and Kirwan, K. (2014). Analysis of the potentials of multicriteria decision analysis methods to conduct sustainability assessment. *Ecological Indicators*, **46**, 138–148.

Ciommi, M., Gigliarano, C., Emili, A., Taralli, S., and Chelli, F. M. (2017). A new class of composite indicators for measuring well-being at thelocal level: An application to the Equitable and Sustainable Well-being (BES) of the Italian Provinces. *Ecological Indicators*, **76**, 281–296.

Climent, V., Michavila, F., and Ripollés, M. (2013). *Los rankings universitarios, mitos y realidades*. Tecnos, Madrid.

Contreras, I., Hinojosa, M. A., and Mármol, A. M. (2011). Ranking Alternatives in Group Decision-Making with Partial Information: A Stable Approach. In E. Herrera Viedma, J. L. García Lapresta, J. Kacprzyk, M. Fedrizzi, H. Nurmi, and S. Zadrożny, editors, *Consensual Processess, STUDFUZZ 267*, pages 41–52. Springer-Verlag Berlin Heidelberg.

de Jorge Moreno, J., González Robles, A., Martinez, A., Minero Calvo, R., and Georgiana Miron, A. (2019). Assessing efficiency in the Spanish public universities through comparative non-radial and radial data envelopment analysis. *Tertiary Education and Management*, **25**, 195–210.

de la Torre, E. M., Gómez-Sancho, J. M., and Pérez-Esparrells, C. (2017a). Comparing university performance by legal status: a Malmquist-type index approach for the case of the Spanish higher education system. *Tertiary Education and Management*, **23(3)**, 206–221.

de la Torre, E. M., Agasisti, T., and Pérez-Esparrells, C. (2017b). The relevance of knowledge transfer for universities' efficiency scores: an empirical approximation on the Spanish public higher education system. *Research Evaluation*, **26(3)**, 211–229.

De Witte, K. and López-Torres, L. (2017). Efficiency in education: a review of literature and a way forward. *Journal of the Operational Research Society*, **68**, 339–363.

Dehon, C., Jacobs, D., and Vermandele, C. (2009). Ranking and research assessment in higher education: current and future challenges. In C. Dehon, D. Jacobs, and C. Vermandele, editors, *Ranking universities*, pages 1–10. Editions de l'Université de Bruxelles.

Ding, L. and Zeng, Y. (2015). Evaluation of Chinese higher education by TOPSIS and IEW — The case of 68 universities belonging to the Ministry of Education in China. *China Economic Review*, **36**, 341–358.

Dyson, R. G., Allen, R., Camanho, V. V., Sarrico, C. S., and Shale, E. A. (2001). Pitfall and protocols in DEA. *European Journal of Operational Research*, **132(2)**, 245–259.

El Gibari, S., Gómez, T., and Ruiz, F. (2018). Evaluating university performance using reference point based composite indicators. *Journal of Informetrics*, **12**, 1235–1250.

El Gibari, S., Gómez, T., and Ruiz, F. (2019). Building composite indicators using multicriteria methods: a review. *Journal of Business Economics*, **89**, 1–24.

European Commission (2006). Delivering on the modernisation agenda for universities: Education, research and innovation. Technical report, Communication from the Commission to the Council and the European Parliament, Brussels.

European Commission (2014). Research and Innovation performance in Spain: Country Profile 2014. Technical report, Publications Office of the European Union, Luxembourg.

European Commission (2018). Science, Research and Innovation Performance of the Eu 2018: Strengthening the foundations for Europe's future. Technical report, Publications Office of the European Union, Luxembourg.

Expósito-García, A. and Velasco-Morente, F. (2018). How efficient are universities at publishing research? A data envelopment analysis of Spanish state universities. *El profesional de la información*, **27(5)**, 1108.

Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society (Series A)*, **120**, 253–281.

Freudenberg, M. (2003). Composite Indicators of Country Performance: A Critical Assessment. http://dx.doi.org/10.1787/405566708255.

Fuentes, R., Fuster, B., and Lillo-Bañuls, A. (2016). A three-stage DEA model to evaluate learning-teaching technical efficiency: Key performance indicators and contextual variables. *Expert Systems With Applications*, **48**, 89–99.

Gana, X., Fernandez, I., Guo, J., Wilson, M., Zhao, Y., Zhou, B., and Wu, J. (2017). When to use what: Methods for weighting and aggregating sustainability indicators. *Ecological Indicators*, **81**, 491–502.

Giannoulis, C. and Ishizaka, A. (2010). A Web-based decision support system with ELECTRE III for a personalised ranking of British universities. *Decision Support Systems*, **48**, 488–497.

Glanzel, W. and Debackere, K. (2009). On the "multi-dimensionality" of ranking and the role of bibliometrics in university assessment. In C. Dehon, D. Jacobs, and C. Vermandele, editors, *Ranking universities*, pages 65–75. Editions de l'Université de Bruxelles.

González-Garay, A., Pozo, C., Galán-Martín, A., Brechtelsbauer, C., Chachuat, B., Chadha, D., Hale, C., Hellgardt, K., Kogelbauer, A., Matar, O. K., McDowell, N., Shah, N., and Guillén-Gosálbez, G. (2019). Assessing the performance of UK universities in the field of chemical engineering using data envelopment analysis. *Education for Chemical Engineers*, **29**, 29–41.

Hazelkorn, E. (2013). How Rankings are Reshaping Higher Education. In V. Climent, F. Michavila, and M. Ripollés, editors, *Los Rankings Universitarios: Mitos y Realidades*, pages 49–60. Tecnos (ANAYA), Madrid.

Heitor, M. V. and Horta, H. (2013). Democratizing Higher Education and Access to Science: The Portuguese Reform 2006-2010. *Higher Education Policy*, **27(2)**, 1–19.

Heitor, M. V., Horta, H., Dopazo, C., and Fueyo, N. (2016). Iberia thirty years after Saramago's Stone Raft: Opportunities for technical change and challenges for science and technology policy under increasing uncertainty. *Forecasting Social Change*, **113**, 129–145.

Hernández Santoyo, A., Casas Vilardell, M., León Sánchez, M. A., and Pérez León, V. E. (2010). Valoración económica de bienes y servicios ambientales en áreas protegidas: contribución al proceso de toma de decisiones. *Retos de la Dirección*, **4(2)**.

Hwang, C. L. and Yoon, K. (1981). *Multiple attribute decision making: Methods and applications*. Springer-Verlag, New York.

Ibañez-Forés, B., Bovea, M. D., and Pérez-Belis, V. (2014). A holistic review of applied methodologies for assessing and selecting the optimal technological alternative from a sustainability perspective. *Journal of Cleaner Production*, **70**, 259–281.

Ijiri, Y. (1965). *Management Goals and Accounting for Control*. Amsterdam: North-Holland.

Jacobs, R., Smith, P., and Goddard, M. (2004). Measuring performance: an examination of composite performance indicators. Technical report, The University of York and Centre for Health Economics, United Kingdom.

Jati, H. and Durai Dominic, D. (2017). A New Approach of Indonesian University Webometrics Ranking Using Entropy and PROMETHEE II. *Procedia Computer Science*, **124**, 444–451.

Johnes, J. (2006). Data envelopment analysis and its application to the measurement of efficiency in higher education. *Economics of Education Review*, **25**, 273–288.

Kumar, A., Sah, B., Singh, A., Deng, Y., He, X., Kumar, P., and Bansal, R. C. (2017). A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. *Renewable and Sustainable Energy Reviews*, **69**, 596–609.

Marope, P. T. M., Wells, P. J., and Hazelkorn, E. (2013). Rankings and accountability in higher education : uses and misuses. Technical report, UNESCO, Paris.

Martínez-Campillo, A. and Fernandez-Santos, Y. (2019). The impact of the economic crisis on the (in)efficiency of public Higher Education institutions in Southern europe: The case of Spanish universities. *Socio-Economic Planning Sciences*. https://doi.org/10.1016/j.seps.2019.100771.

Mazziotta, M. and Pareto, A. (2017). Chapter 7. Synthesis of Indicators: The Composite Indicators approach. In F. Maggino, editor, *Complexity in Society: From Indicators Construction to their Synthesis*, pages 159–191, Switzerland. Springer-Nature.

Mägi, E. and Beerkens, M. (2016). Linking research and teaching: Are research-active staff members different teachers. *Higher Education*, **72**, 241–258.

Miettinen, K. (1999). *Nonlinear multiobjective optimization*. Boston: Kluwer Academic Publishers.

Munda, G. (2008). *Social multi-criteria evaluation for a sustainable economy*. Springer-Verlag, New York.

Murias, P., de Miguel, J. C., and Rodríguez, D. (2008). A composite indicator for university quality assesment: The case of spanish higher education system. *Social Indicators Research*, **89**, 129–146.

Nardo, M., Saisana, M., Saltelli, A., and Tarantola, S. (2005). Tools for Composite Indicators Building. Technical report, European Commission, Ispra(Italy).

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Giovannini, E., and Hoffman, A. (2008). Handbook on Constructing Composite Indicators. Methodology and user guide. Technical report, OECD, Paris.

OECD (2007). Higher Education and Regions, Globally Competitive. Locally Engaged. Technical report, OECD, Paris.

OECD (2012). Assessment of Higher Education Learning Outcomes. Feasibility Study Report. Volume 1-Design and Implementation. Technical report, OECD, Paris.

Palomares-Montero, D. and García-Aracil, A. (2011). What are the key indicators for evaluating the activities of universities? *Research Evaluation*, **20(5)**, 353–363.

Rauhvargers, A. (2011). Global University Ranking and their impact: EUA Report on Rankings. Technical report, European University Association, Brussels.

Rhodes, E. (1978). *Data Envelopment Analysis and related approaches for measuring the efficiency of decision making units with an application to Program Follow Through.* Ph.D. thesis, Carnegie-Mellon University, School of Urban and Public Affairs.

Ruiz, F., Cabello, J. M., and Luque, M. (2011). An application of reference point techniques to the calculation of synthetic sustainability indicators. *Journal of the Operational Research Society*, **62**, 189–197.

Ruiz, F., El Gibari, S., Cabello, J. M., and Gómez, T. (2020). MRP-WSCI: Multiple reference point based weak and strong composite indicators. *Omega*, **95**, 102060. https://doi.org/10.1016/j.omega.2019.04.003.

Saaty, T. L. and Ergu, D. (2015). When is a Decision-Making Method Trustworthy? Criteria for Evaluating Multi-Criteria Decision-Making Methods. *International Journal of Information Technology & Decision Making*, **14(6)**, 1171–1187.

Sabino, C. (1992). *El proceso de investigación.* Panapo, Caracas.

Saisana, M. and D'Hombres, B. (2008). Higher Education Rankings: Robustness Issues and Critical Assessment. How much confidence can we have in Higher Education Rankings? Technical report, Joint Research Centre, European Commission., Ispra (Italy).

Saisana, M. and Tarantola, S. (2002). State-of-the-Art Report on Current Methodologies and Practices for Composite Indicator Development. Technical report, Joint Research Centre, European Commission, Ispra (Italy).

Salas Velasco, M. (2019). The technical efficiency performance of the higher education systems based on data envelopment analysis with an illustration for the Spanish case. *Educational Research for Policy and Practice.*

Salas Velasco, M. (2020). Measuring and explaining the production efficiency of Spanish universities using a non-parametric approach and a bootstrapped-truncated regression. *Scientometrics*, **122**, 825–846.

Sánchez-Barrioluengo, M. (2012). Cómo afronta la universidad el cumplimiento de sus misiones: El caso de las universidades públicas españolas. In E. P. Vieiral, editor, *Investigaciones de Economía de la Educación,* vol. 7, pages 543–560, Oporto. Asociación de Economía de la Educación.

Sanz-Casado, E. and De-Filippo, D. (2011). Observatorio IUNE: Una nueva herramiento para el seguimiento de la actividad investigadora del sistema universitario español. Bordón 63 (2), 2011, 101-115, ISSN: 0210-5934•101.

Tarabusi, E. C. and Guarini, G. (2013). An Unbalance Adjustment Method for Development Indicators. *Social Indicators Research*, **112**, 19–45.

Teodorovic, D. and Edara, P. (2007). A real-time road pricing system: The case of a two-link parallel network. *Computers & Operations Research*, **34**, 2–27.

Torres-Salinas, D., Moreno-Torres, J. G., Delgado-López-Cózar, E., and Herrera, F. (2011). A methodology for Institution-Field ranking based on a bidimensional analysis: the IFQ^{2}A index. *Scientometrics (2011) 88:771–786*, **88**, 771–786.

Wierzbicki, A. P. (1980). The use of reference objectives in multiobjective optimization. In G. Fandel and T. Gal, editors, *Lecture Notes in Economics and Mathematical Systems,* vol. 177, pages 468–486, Berlin. Springer-Verlag.

Wierzbicki, A. P., Makowski, M., and Wessels, J., editors (2000). *Model-Based Decision Support Methodology with Environmental Applications.* Kluwer Academic Publishers, Dordrecht.

Yang, G. L., Song, Y. Y., Xu, D. L., and Yang, J. B. (2017). Overall Efficiency and its Decomposition in a Two-Stage Network DEA Model. *Journal of Management Science and Engineering*, **2(3)**, 161–192.

Yu, P. L. (1973). A class of solutions for group decision problems. *Management Science*, **19**, 936–946.