Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA. **Rect@**
Volumen 22(2021). Páginas 77 a 97
DOI: 10.24309/recta.2021.22.2.02

ANÁLISIS MULTIOBJETIVO Y MODELOS DE REGRESIÓN. UNA APLICACIÓN PARA ANALIZAR EL BIENESTAR DE LOS ESTUDIANTES ESPAÑOLES

SANDRA GONZÁLEZ-GALLARDO
sandragg@uma.es
Departamento de Economía Aplicada (Matemáticas), Universidad de Málaga
Calle Ejido, 6, 29013, Málaga

Recibido (31/03/2021)
Revisado (20/10/2021)
Aceptado (20/11/2021)

RESUMEN: En este trabajo se propone un nuevo entorno metodológico en el que se combinan técnicas econométricas y de optimización multiobjetivo con el propósito de resolver problemas socioeconómicos. Dicha metodología consta de dos etapas, la primera consiste en desarrollar un modelo de regresión a partir del cual, en la segunda etapa, se construye un problema de optimización multiobjetivo. Una ventaja de esta combinación es la posibilidad de introducir preferencias en la resolución del problema. En concreto, se aplicará la metodología propuesta para analizar el bienestar de los estudiantes españoles usando cuatro índices (positividad, motivación, sentido de pertenencia y acoso escolar). El análisis del bienestar está tomando cada vez mayor relevancia por su relación con el rendimiento académico de los jóvenes. Para ello, se realizan cuatro regresiones (una por índice) en función de un conjunto de variables explicativas, con las que se construye un problema multiobjetivo para estudiar el conflicto entre dichos índices. Los resultados obtenidos mediante programación multiobjetivo intervalar no solo proporcionan información sobre cómo afecta la mejora de un índice del bienestar al resto de los mismos, sino que también nos permite conocer el perfil del estudiante que alcanza unos niveles óptimos y equilibrados entre los distintos índices.

Palabras Clave: Programación multiobjetivo intervalar; Análisis econométrico; Bienestar de los estudiantes; Economía de la educación.

ABSTRACT: In this work, a novel approach is proposed in which econometric and multiobjective optimization techniques are combined with the aim of analysing socio-economic problems. This approach consists of two stages. Firstly, a regression model is carried out, from which, in the second stage, a multiobjective optimization problem is defined. An advantage of this combination is the possibility of introducing preferences of decision makers -desired values- for solving the problem. Particularly, we apply this approach to analyse the well-being of Spanish students through four indexes (positive feelings, motivation, sense of belonging, and bullying). In recent years, studying the students' well-being has become very relevant because of its relation with their academic performance. Thus, four regressions are obtained (one per index) with respect to a set of explanatory variables, from which the multiobjective optimization problem is built. The results obtained using interval multi-objective programming provide us information both about how the improvement of one index can affect the values of the remaining ones, and also about the student's profile who achieves an optimum balance among the well-being indexes.

Keywords: Multiobjective interval programming; Econometric analysis; Students' well-being; Economics of education

1. Introducción

La existencia de numerosas técnicas matemáticas permite la resolución de problemas de distinta naturaleza desde diferentes enfoques. Particularmente, la econometría tiene la finalidad de proveer de contenido empírico a los modelos matemáticos o económicos, con el fin de predecir cómo le afecta a una variable los cambios de otras. En la práctica, se basa en estimaciones para cuantificar y verificar las predicciones de teorías matemáticas y económicas mediante un análisis de regresión, que estudia la relación existente entre las variables de respuesta y las variables explicativas. Con este análisis podemos encontrar si existe cierto grado de conflicto entre las variables respuesta, es decir, que la mejora de una implique el empeoramiento de la otra. Sin embargo, este tipo de análisis no puede aportar mucho más que la naturaleza conflictiva de las variables, de forma que no puede aportar información de cómo alcanzar simultáneamente valores óptimos para todas las variables de respuesta. Por lo tanto, la aplicación de técnicas de optimización multiobjetivo tras realizar un análisis econométrico puede permitirnos obtener información más allá de la que nos aporta el análisis estadístico.

La optimización multiobjetivo resuelve problemas con varios objetivos en conflicto dentro de un conjunto de posibilidades (conjunto factible). En estos problemas, debido al conflicto entre las funciones objetivo, es imposible obtener una solución que alcance el óptimo de cada objetivo de forma simultánea. Así, en las soluciones obtenidas a través de la optimización ningún objetivo puede mejorar sin empeorar, al menos, uno de los restantes.

En este artículo proponemos un nuevo entorno de trabajo en el que se combinan metodologías econométricas y multiobjetivo para ayudar a los investigadores (o decisores si hablamos de un problema real) a tener una visión más exhaustiva del problema sujeto a análisis. El entorno propuesto está formado por dos etapas. La primera de ellas consiste en un análisis econométrico con el objetivo de obtener estimaciones de la relación existente entre las variables explicativas y las variables de decisión consideradas. La relación de dependencia entre las variables puede obtenerse siguiendo varios modelos, como son lineal, polinómico, exponencial, etc. La segunda parte del entorno metodológico pretende definir un problema de optimización multiobjetivo a partir del análisis econométrico previo y las correlaciones entre las variables explicativas. De forma más específica, el análisis de la primera etapa se usa para definir las funciones objetivo y delimitar el conjunto de posibles valores de las variables explicativas mediante restricciones. Seguidamente, se estudia el problema mediante técnicas de optimización multiobjetivo desde un punto de vista de la decisión multicriterio.

Por lo tanto, con este enfoque permitimos definir un marco teórico en el que los decisores tienen información sobre los valores que tienen que tomar las variables explicativas para que las variables respuesta alcancen el valor óptimo simultáneamente. Los resultados de nuestra propuesta pueden sugerir estrategias para alcanzar ciertos valores en las variables explicativas mejorando así las variables respuesta tanto como sea posible.

El análisis de optimización multiobjetivo del modelo econométrico propuesto puede ser de interés en diferentes situaciones. Por ejemplo, cuando los investigadores quieren saber qué combinación de valores de las variables explicativas haría que las variables de decisión alcancen los mejores valores posibles simultáneamente. Además, el entorno propuesto nos puede arrojar información sobre cómo de sensible son los valores óptimos a posibles variaciones de las variables explicativas. Por otro lado, también podemos medir el impacto que tiene dar más importancia a alguna de las variables respuesta (objetivos del problema), teniendo la información de cuánto empeora el resto de las variables respuesta en dicha situación.

En este sentido, nosotros proponemos un enfoque metodológico aplicado a economía de la educación en el que se combinan técnicas econométricas y de optimización multiobjetivo, particularmente, de programación multiobjetivo intervalar para así conocer cómo llegar a ciertos niveles outputs relacionados con el bienestar de los estudiante mediante variables explicativas del día a día relacionadas con el colegio como son, el uso de internet y dispositivos digitales, horas dedicadas a estudiar otro idioma, además de otras variables sociodemográficas.

En los últimos años, el bienestar de los estudiantes está captando más atención por su relación con el rendimiento académico de los jóvenes (Korhonen *et al.*, 2014; Jogi *et al.*, 2015; Parhiala *et al.*, 2018). Hay muchos autores que afirman que la educación contribuye al crecimiento económico, y consecuentemente existe una fuerte relación entre la educación y el desarrollo de un país, ya que si el bienestar no es adecuado en la adolescencia, puede afectar a la productividad posterior como trabajadores (Warr, 1999). Además, existen evidencias afirmando que el rendimiento académico está influenciado por factores emocionales relacionados con el ámbito educativo tales como la motivación, el apoyo del profesorado, el acoso escolar entre otros (Scheerens y Hendriks, 2014; Baker *et al.*, 2004). Diferentes estudios muestran que aquellos programas educativos enfocados a mejorar el bienestar de los estudiantes mejoran los resultados académicos de los estudiantes. De esta forma, las políticas educativas deben centrarse en el estado de bienestar de los estudiantes con el objetivo de alcanzar mejor eficiencia y calidad en el proceso de enseñanza-aprendizaje. Por ejemplo, los sistemas educativos de las regiones con mayor rendimiento académico, como son los países del norte de Europa (Finlandia y Suecia) (OECD, 2019), hacen mucho hincapié en el desarrollo general de los estudiantes y en el aprendizaje socio-emocional de estos. Lo que sugiere que para obtener buen rendimiento académico no solo hay que tener en cuenta la materia impartida sino también otros aspectos de la vida de los estudiantes.

En este contexto, los últimos informes de PISA [1] reportan datos sobre el bienestar de los estudiantes y lo definen como un concepto multidimensional (OECD, 2017, 2019), en el que involucran cuatro dimensiones: psicológica, social, cognitiva y fisica. En este sentido, consideraremos los siguientes indicadores para medir el bienestar: índice de positividad, de motivación, de sentido de pertenencia al centro y de acoso escolar. Existen trabajos que respaldan el análisis del bienestar mediante estos indicadores. En cuanto al sentido de pertenencia, existen evidencias que encuentran relación entre la necesidad de sentirse integrado en el colegio con el comportamiento, la salud y el bienestar de los estudiantes (Ryan y Deci, 2018; Chinneck y Ramadan, 2000). Por otro lado, diferentes estudios observan el impacto de la motivación y el bienestar en el rendimiento académico de los estudiantes Parhiala *et al.* (2018); Jogi *et al.* (2015). Por último, la lacra del acoso escolar entre adolescente es un problema que concierne a nivel mundial y tiene graves consecuencias en la vida de los jóvenes, y también en su rendimiento escolar Strøm *et al.* (2013).

Teniendo en cuenta la naturaleza multidisciplinar del concepto de bienestar estudiantil, es adecuado utilizar técnicas de optimización multiobjetivo. El escenario ideal sería alcanzar los mejores valores para cada indicador, teniendo en cuenta que los índices de sentimientos positivos, motivación y sentido de pertenencia deben alcanzar los valores más altos posibles mientras que el índice de acoso escolar debe de alcanzar el menor valor posible. Sin embargo, debido al grado de conflicto existente entre los índices, es imposible alcanzar el mejor valor para cada índice en una misma solución. Por lo tanto el resultado dará el balance óptimo entre los cuatro indicadores, además del perfil de estudiante que alcanza dicho balance.

Por otro lado, es necesario identificar variables significativas en el proceso enseñanza-aprendizaje que permitirán alcanzar los valores deseados para cada indicador. Las variables usadas también las proporciona PISA y están relacionadas con las características socio-demográficas de los alumnos, el uso de los dispositivos digitales e internet, el apoyo por parte del profesorado, entre otras.

---

[1] *The Programme for Internantional Students Assessment (PISA)*, es un estudio en el que se mide el rendimiento de los estudiantes de 15 años en los países de la OCDE.

Tras esta introducción, el resto del documento se estructura en 6 secciones. La primera sección recoge la notación y conceptos básicos de la econometría y la optimización multiobjetivo (Sección 2). La segunda sección, Sección 3, detalla la metodología propuesta en este artículo para analizar problemas socioeconómicos. A continuación, en la Sección 4 se desarrolla la forma de resolver un problema multiobjtivo intervalar. En la siguiente parte del artículo, Sección 5, se muestran los resultados obtenido. Por último, se presentan las conclusiones más relevantes.

2. Notación y conceptos básicos

En esta sección daremos los conceptos básicos de las dos técnicas que forman el entorno metodológico propuesto.

2.1. Econometría

Supongamos que queremos estudiar una situación socioeconómica y queremos encontrar la relación de dependencia de un conjunto de variables que la representan (variables respuesta), denotadas por y_j para $j = 1, \ldots, k$, con respecto a un conjunto de variables que la explican (variables explicativas), denotadas por x_i con $i = 1, \ldots, n$. En este sentido queremos establecer una correspondencia entre los dos conjuntos de variables de la siguiente forma:

$$(x_1(r), \ldots, x_n(r)) \to (y_1(r), \ldots, y_k(r)), \text{para } r = 1, \ldots, l, \tag{1}$$

siendo l el número de observaciones. Para encontrar las dependencias entre los conjuntos de variables, se realiza un análisis econométrico mediante la estimación de un modelo regresión en el que se regresan las variables respuesta con las variables explicativas. Vamos a denotar un modelo de regresión de la siguiente forma:

$$y_j = \hat{y}_j(x_1, \ldots, x_n) + \epsilon_j, \text{para } j = 1, \ldots, k, \tag{2}$$

donde $\hat{y}_j$ es la predicción de la variable respuesta y_j. Esta se predice mediante una función de las variables explicativas $x_1, \ldots, x_n$ más el error de predicción denotado por ϵ_j. Para simplificar la notación, vamos a suprimir el parámetro r que denota el número de observación. De esta forma, las variables explicativas y respuesta quedan denotadas como $x_1, \ldots, x_n$ y $y_1, \ldots, y_k$. Las funciones $\hat{y}_j$, para $j = 1, \ldots, k$, se denominan predictores de las variables respuesta, y pueden ser funciones de diferentes tipos: lineales, polinómicas, exponenciales, entre otras; dependiendo de la naturaleza de los datos y el modelo considerado. Para el caso lineal, la formulación es la siguiente:

$$\hat{y}_j = \hat{\alpha}^j + \hat{\beta}_1^j x_1 + \cdots + \hat{\beta}_n^j x_n + \epsilon_j, \ j = 1, \ldots, k, \tag{3}$$

donde $\hat{\beta}^{\mathbf{j}} = (\hat{\beta}_1^j, \ldots, \hat{\beta}_n^j)^T$ es el vector de coeficientes de regresión y $\hat{\alpha}^j$ es la constante de estimación. Los coeficientes de regresión pueden estimarse mediante diferentes metodologías. En nuestro caso, los estimaremos mediante mínimos cuadrados ordinarios (Ordinary Least Squares - OLS).

2.2. Optimización multiobjetivo

Un problema de optimización multiobjetivo (MOP) se puede definir de la siguiente forma (Miettinen, 1999) [2]:

$$\begin{array}{ll} \max & \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\} \\ \text{sujeto a} & \mathbf{x} \in X, \end{array} \tag{4}$$

[2]Sin pérdida de generalidad, definimos el problema multiobjetivo en sentido de máximo, pero si una o más funciones objetivo quieren ser minimizadas, se multiplican por -1.

donde $f_j : X \to \mathbb{R}$ con $j = 1, \ldots, k$, son las k funciones objetivo en conflicto que queremos optimizar simultáneamente y $X \in \mathbb{R}^n$ es la región o conjunto factible formado por los vectores de decisión factibles denotados por $\mathbf{x} = (x_1, \ldots, x_n)^T$. La imagen de cada vector de decisión factible se denomina vector objetivo y se denotan por $\mathbf{f}(\mathbf{x}) = (f(x_1), \ldots, f(x_n))^T$. Además, todos los vectores objetivo forman la región objetivo factible $F = \mathbf{f}(\mathbf{X}) \subseteq \mathbb{R}^m$. A partir de ahora, suponemos que X es un conjunto no vacio y compacto de $\mathbb{R}^k$.

Desafortunadamente, debido al conflicto existente entre los objetivos es casi imposible encontrar una solución que optimice todos los objetivos al mismo tiempo. En este sentido, se definen las soluciones eficientes o Pareto óptimas. Se dice que un vector $\mathbf{x}^0 \in X$ es eficiente o Pareto óptimo del problema (4) si no existe otro vector $\mathbf{x} \in X$ de forma que $f_j(\mathbf{x}^0) \leq f_j(\mathbf{x})$ para cada $j = 1, \ldots, k$ y $f_i(\mathbf{x}^0) < f_i(\mathbf{x})$ para al menos un índice i. Su correspondiente vector objetivo $\mathbf{f}(\mathbf{x^0})$ se denomina vector objetivo Pareto óptimo. El conjunto formado por todas las soluciones eficientes o Pareto óptimas del espacio de decisión se denomina Pareto óptimo (E), mientras que su imagen en el espacio objetivo, se llama frontera Pareto óptima (PF) y la denotamos por $f(E)$.

Por la definición de solución Pareto óptima, un problema multiobjetivo, normalmente, tiene más de una solución de este tipo. Por lo tanto, es útil saber cuál es el rango de variación de las funciones objetivo en la frontera Pareto óptima. En primer lugar, las cotas inferiores constituyen el vector nadir $\mathbf{z}^{nad} = (z_1^{nad}, \ldots, z_k^{nad})^T$, donde para cada $j = 1, \ldots, k$, se define como $z_j^{nad} = \text{mín}_{\mathbf{x} \in E} f_i(\mathbf{x})$. Por otro lado, el vector ideal $\mathbf{z}^\star = (z_1^\star, \ldots, z_k^\star)$ define las cotas superiores, y se calcula $z_j^* = \text{máx}_{\mathbf{x} \in E} f_i(\mathbf{x})$, para cada $j = 1, \ldots, k$. A veces, para facilitar la resolución de los problemas multiobjetivo es útil utilizar un vector que domine estrictamente a todas las soluciones Pareto óptimas. Para ello definimos el vector utopía, que se denota por $z_i^{\star\star} = z_i^\star + \epsilon$, donde $\epsilon > 0$ es un valor real pequeño, para cada $j = 1, \ldots, k$.

Mientras que el vector ideal se puede calcular en cualquier caso, el vector nadir no es fácil de obtener y este puede ser estimado, por ejemplo, usando la matriz de pagos del problema, aunque esta aproximación no tiene por qué ser buena (para más detalles, ver Ehrgott y Tenfelde-Podehl (2003); Miettinen (1999)). En Deb y Miettinen (2010) proponen diferentes enfoques para estimar el vector nadir. En la práctica, estos vectores se utilizan con frecuencia para normalizar las funciones objetivo, dividiendo cada valor de la función objetivo por la diferencia entre los valores ideal y nadir.

3. Metodología

Analizando un modelo de regresión, podemos identificar la relación de dependencia entre las diferentes variables explicativas y respuesta. Además, a partir de estas dependencias, podemos extraer conclusiones sobre el problema a estudiar. Sin embargo, a veces es interesante estudiar cómo se pueden alcanzar los mejores valores para las variables respuesta, de forma simultánea, considerando todos los valores que las variables explicativas pueden tomar. Esta información permite apoyar las malas o buenas decisiones tomadas en relación con el problema estudiado mediante la identificación de los factores claves referentes a las variables explicativas que permitirían alcanzar simultáneamente dichos valores óptimos de las variables respuesta.

Teniendo en cuenta esta idea, si los datos analizados muestran cierto grado de conflicto entre las variables respuesta, podemos definir un problema de optimización multiobjetivo basándonos en ciertas estimaciones econométricas previas. Una vez construido el problema multiobjetivo, que explicaremos de aquí en adelante, se podrán extraer interesantes conclusiones del problema estudiado. En este sentido, los decisores-investigadores, directivos, políticos, o la persona encargada del estudio- pueden expresar sus preferencias sobre el problema determinando, así como información sobre qué solución Pareto óptima se ajusta más a las expectativas definidas a través de las preferencias. En general, la principal novedad de esta combinación de técnicas es que la información que arrojan los resultados no se puede obtener solo haciendo el análisis econométrico de los datos.

En lo que sigue, vamos a describir en detalle cómo podemos hacer la combinación de dichas técnicas con datos reales referentes a economía de la educación.

3.1. Análsis Econométrico

En esta sección, se describen la base de datos que usamos, las características de la misma, las variables de decisión y explicativas, así como las estimaciones realizadas para obtener las funciones objetivo y las restricciones del problema multiobjetivo que vamos a construir.

3.1.1. Datos

Los datos utilizados en este artículo son de la base de datos de PISA 2018 (OECD, 2019), estos datos ofrecen una amplia información personal y académica de los estudiantes. PISA 2018 (OECD, 2019) evalúa los conocimientos en ciencias, matemáticas y lectura en 80 países y está dirigido aproximadamente a 612,000 estudiantes de 15 años.

La evaluación académica se realiza mediante diferentes test, compuestos por preguntas de múltiple respuesta y preguntas en las que los estudiantes tienen que construir la respuesta, incluyendo cuestiones de las tres competencias anteriormente citadas (ciencias, matemáticas y lectura). PISA también incluye medidas de competencias transversales como la resolución de problemas colaborativos. Especialmente, PISA está diseñado para enfatizar las habilidades funcionales que los estudiantes han adquirido a medida que se acercan al final de la enseñanza secundaria obligatoria. Concretamente, los estudiantes de 15 años rellenan un cuestionario general que nos da información sobre ellos, sus casas y su experiencia en el proceso de aprendizaje. En algunos países como es el caso de España, los alumnos responden dos cuestionarios adicionales. El primero de ellos relacionado con las Tecnologías de la Información y Comunicación (TICs), y el segundo relacionado con el bienestar de los estudiantes en algunos aspectos escolares y de la vida. Basándose en este cuestionario, PISA reporta información del bienestar a través de los siguientes índices: índice de positividad, índice de motivación, índice de sentido de pertenencia e índice de acoso escolar. Así, estos índices se construyen mediante análisis factorial haciendo uso de las siguientes preguntas del cuestionario:

- Índice de positividad: ¿Con qué frencuencia ("nunca", "rara vez", "a veces", "siempre") te sientes feliz, alegre o animado?
- Índice de motivación: ¿En qué medida está usted de acuerdo o en desacuerdo ("muy en desacuerdo", "en desacuerdo", "de acuerdo", "muy de acuerdo") con las siguientes afirmaciones?:"Me satisface trabajar lo máximo posible"; "Una vez que empiezo una tarea, persisto hasta que la termino"; "Si no soy bueno en algo, prefiero seguir trabajando hasta dominarlo en lugar de hacer otra cosa en lo que soy bueno".
- Índice de sentido de pertenencia: ¿En qué medida está usted de acuerdo o en desacuerdo ("muy en desacuerdo", "en desacuerdo", "de acuerdo", "muy de acuerdo") con las siguientes afirmaciones?: "Me siento como un extraño (o excluido de las cosas) en el colegio; Hago amigos fácilmente en el colegio"; "Siento que pertenezco al colegio"; "Me siento incómodo y fuera de lugar en mi colegio"; "Parece que les caigo bien a otros compañeros"; "Me siento solo en el colegio".
- Índice de acoso escolar: ¿Con qué frecuencia ("nunca o casi nunca", "algunas veces al año", "algunas veces al mes", "una vez a la semana o más") has tenido las siguientes experiencias en la escuela?: "Otros estudiantes me dejaron fuera de las cosas a propósito"; "Otros estudiantes se burlaron de mí"; "Fui amenazado por otros estudiantes"; "Otros

estudiantes me quitaron o destruyeron mis cosas"; "Fui golpeado o empujado por otros estudiantes".

El sentido de los índices de positividad, motivación y sentido de pertenencia es el siguiente: un valor más alto significa que el alumno reporta mayor positividad, motivación y sentido de pertenencia. Sin embargo, un valor alto en el índice de acoso escolar indica que el estudiante estuvo más expuesto a acoso, mientras que un valor bajo significa que el estudiante ha sufrido menos acoso escolar.

Específicamente, 35.943 estudiantes de 1089 colegios españoles cumplimentaron dicha evaluación. De este análisis excluimos a los estudiantes matriculados en colegios privados, de esta forma, solo consideramos alumnos de colegios públicos y concertados. Además, consideramos solo a estudiante no repetidores. La muestra también se verá reducida debido a los datos *missing* de las variables consideradas. Por lo tanto, las regresiones finales utilizan 12,169 observaciones en total.

La Tabla 1 y 2 reportan la descripción de las variables explicativas usadas en el modelo y los correspondientes estadísticos descriptivos de todas las variables, respectivamente. Después de analizar estos datos, podemos afirmar que en la muestra seleccionada el 29.5 % de los colegios son concertados mientras que el 70.5 % son públicos. En cuanto a género, existe aproximadamente el mismo número de chicos que de chicas. En media, los estudiantes españoles empiezan a utilizar dispositivos digitales cuando tienen 7 años. Las variables apoyo del profesorado y ambiente escolar son variables obtenidas mediante análisis factorial de una serie de preguntas del cuestionario de PISA. En este punto, es importante notar, que cuanto menor sea el valor de la variable apoyo del profesorado significa que se ha recibido más apoyo. Sin embargo, cuanto menor sea el valor de la variable ambiente escolar, peor es el ambiente escolar percibido por los estudiantes. Por último, los estudiantes españoles emplean 4.25 horas de media en aprender otro idioma.

Tabla 1. Descripción de las variables explicativas

Nombre	Notación	Variable	Tipo
escs	x_1	Nivel socioeconómico	Continua
concertado	x_2	Tipo de colegio (Concertado 1 - público 0)	Binaria
género	x_3	Chico (0) o chica (1)	Binaria
horasInternet	x_4	Horas de internet fuera del colegio	Continua
dispositivosDigitales	x_5	Edad de inicio de uso de dispositivos digitales	Continua
apoyoProfesorado	x_6	Apoyo del profesorado	Continua
ambienteEscolar	x_7	Ambiente escolar	Continua
idiomas	x_8	Horas dedicadas a aprender otro idioma	Continua

Tabla 2. Estadísticos descriptivos de las variables

Nombre	Notación	Media	Desviación típica	Mínimo	Máximo
Índice de positividad	y_1	0.0370	0.8795	-4.0655	0.9016
Índice de motivación	y_2	0.0633	0.7944	-3.3026	1.2486
Índice de sentido de pertenencia	y_3	-0.0150	0.8740	-2.8899	1.2096
Índice de acoso escolar	y_4	-0.0801	0.7831	-0.4818	5.4689
Nivel socioeconómico	x_1	0.0338	0.9466	-4.316	3.717
Tipo de colegio	x_2	0.2955	0.4563	0.0000	1.0000
Género	x_3	0.5445	0.4980	0.0000	1.0000
Horas de Internet fuera del colegio	x_4	3.6811	2.0533	0.0000	7.0000
Edad de uso de dispositivos digitales	x_5	7.5441	2.7196	4.5000	14.5000
Apoyo del profesorado	x_6	0.0194	0.9231	-0.9720	2.4555
Ambiente escolar	x_7	0.0349	0.9653	-2.1020	1.4652
Horas dedicadas a aprender otro idioma	x_8	4.2554	1.3111	3.0000	7.0000

3.1.2. Estimaciones econométricas

En esta sección vamos a describir el modelo de regresión que hemos utilizado para posteriormente construir el problema multiobjetivo. Se consideran como variables explicadas o respuesta los cuatro índices descritos previamente (sentimientos positivos, motivación, sentido de pertenencia y acoso escolar). Cada uno de estos índices se estima mediante un conjunto de ocho variables explicativas, descritas en la Tabla 1, utilizando mínimos cuadrados ordinarios. La idea que sustenta esta técnica es minimizar el ruido estadístico, denotado por ϵ, lo máximo posible. Además, esta metodología asume que dicho error está normalmente distribuido.

Suponiendo que los estudiantes están denotados por r, y cada uno de los índices se denota por j, el modelo de regresión se puede representar como sigue:

$$W^j(r) = \alpha^j + \beta_1^j escs(r) + \beta_2^j concertado(r) + \cdots + \beta_8^j idiomas(r) + \epsilon^j(r) \tag{5}$$

donde $j = 1, \ldots, 4$ (uno para cada índice), $r = 1, \ldots, 12{,}169$ (estudiantes en la muestra), $W^j(r)$ es la predicción de cada índice del bienestar j para cada estudiante r; $escs(r)$, $colegio(r), \ldots, idiomas(r)$ son el conjunto de variables explicativas, $\epsilon^j(s)$ es el término de error, β^j es el vector de coeficientes y α^j la constante de la estimación.

En la Tabla 3 se muestran las cotas superiores e inferiores de los intervalos de confianza de los coeficientes de cada estimación. Cabe notar que se muestran los coeficientes de las variables explicativas que son significativas al 10 %, 5 % y 1 %, aquellas variables que no son significativas tienen como coeficiente 0 . Las cotas superiores e inferiores de los coeficientes corresponden con los intervalos de confianza al 95 % obtenidos de las estimaciones.

Tabla 3. Cotas inferiores y superiores de los coeficientes de estimación

		Positividad		**Motivación**		**Sentido de pertenencia**		**Acoso escolar**	
Variables		**Inferior**	**Superior**	**Inferior**	**Superior**	**Inferior**	**Superior**	**Inferior**	**Superior**
Nivel socioeconómico	x_1	0,0006	0,0554	0,0087	0,0603	0,0159	0,0661	0	0
Tipo de colegio	x_2	0	0	0	0	0,0395	0,1508	0	0
Género	x_3	-0,0980	-0,0138	0,1274	0,2164	-0,1006	-0,0202	-0,1707	-0,0781
Horas de Internet fuera del colegio	x_4	0	0	-0,0413	-0,0209	0,0117	0,0341	-0,0009	0,0250
Edad de uso de dispositivos digitales	x_5	0,0013	0,0160	0,0083	0,0218	0	0	0	0
Apoyo del profesorado	x_6	-0,1131	-0,0614	-0,1622	-0,1111	-0,1408	-0,0803	0,0370	0,0843
Ambiente escolar	x_7	0,0284	0,0835	0,0084	0,0500	0,0367	0,0876	-0,1278	-0,0767
Horas dedicadas a aprender otro idioma	x_8	0	0	0,0087	0,0429	0	0	0	0
Constante	α	-0,0903	0,1106	-0,2298	-0,0383	-0,2944	-0,0391	-0,1050	0,1258
Observaciones	N	12169	12169	12169	12169	12169	12169	12169	12169
R-cuadrado	R^2	0,0181	0,0181	0,0555	0,0555	0,0319	0,0319	0,0329	0,0329

Empezando con el índice de positividad, la variable explicativa índice socioeconómico tiene un efecto positivo sobre este índice, lo que significa que los estudiantes con mayor nivel socioeconómico reportarán un mayor índice de positividad. En cuanto a género, las estimaciones sugieren que los chicos muestran mayor positividad que las chicas. Las variables apoyo del profesorado y ambiente escolar también tienen un efecto positivo en el índice de positividad. Analizando la edad con la que empiezan a utilizar dispositivos digitales, se intuye que cuanto más mayores empiecen más alto será el índice de positividad de los estudiantes.

Siguiendo con el índice de motivación, parece claro que los estudiantes aventajados socioeconómicamente muestran mayor motivación. De acuerdo con Yu *et al.* (2020), el análisis econométrico sugiere que las niñas reportan mayor motivación que los niños. El uso de internet fuera del colegio afecta de forma negativa en la motivación estudiantil, tal y como se muestra en Demir y Kutlu (2018). Además, el análisis parece mostrar que cuanto mejor sea el ambiente escolar mayor será la motivación de los estudiantes, esto va en concordancia con los resultados de Fan y Williams (2018). Por último, y como muestra Yıldırım (2012), aquellos estudiantes que tienen más apoyo por parte del profesorado son propensos a tener más motivación.

En cuanto al índice de sentido de pertenencia, como en otros estudios (OECD, 2019) los chicos tienen mayor sentido de pertenencia al colegio que las chicas, y también aquellos que tienen mayor nivel socioeconómico. Por otro lado, es posible percibir que el apoyo del profesorado y el ambiente escolar afectan positivamente al sentido de pertenencia de los estudiantes al colegio (Allen *et al.*, 2018; OECD, 2017).

En lo que respecta al índice de acoso escolar, de acuerdo con Veenstra *et al.* (2005), las chicas están menos implicadas en casos de acoso que los chicos. Además, los resultados sugieren que aquellos estudiantes que usan más horas Internet fuera del colegio, reportan mayor índice de acoso. Estos resultados están en concordancia con Athanasiou *et al.* (2018), donde demuestran que cuanto más tiempo usando internet y redes sociales mayor es el riesgo de sufrir acoso escolar. Por otro lado, y coincidiendo con Mischel y Kitsantas (2020), los resultados afirman que los estudiantes que tienen mayor apoyo de los profesores tienen menos riesgo de ser acosados. Además, el ambiente escolar afecta negativamente al índice de acoso, lo que significa que un buen ambiente escolar reduce el índice de acoso en los colegios (Mischel y Kitsantas, 2020).

Por lo tanto, los resultados indican que el modelo estimado tiene cohereencia ya que se apoyan otros trabajos previamente prublicados (Yu *et al.*, 2020; Yıldırım, 2012; Mischel y Kitsantas, 2020; Demir y Kutlu, 2018). Por ello, el problema multiobjetivo que se propone en la Sección 3.2 recoge la realidad presente en los colegios.

3.2. Problema de optimización multiobjetivo intervalar

En esta sección, basándonos en el análisis econométrico del apartado anterior (primera etapa), vamos a construir un problema de optimización multiobjetivo. Para ello, definimos las variables de decisión, las restricciones y las funciones objetivo.

3.2.1. Variables de decisión

Las variables explicativas, $x_1, \ldots, x_k$, consideradas como variables clave en el modelo econométrico, consituyen las variables de decisión del problema de optimización multiobjetivo. Dependiendo del problema socioeconómico estudiado, las variables explicativas pueden ser continuas, enteras o binarias, por lo tanto, las variables de decisión serán de la misma naturaleza.

En nuestro caso, no tendremos variables binarias categóricas, pero sí variables binarias y continuas. Las variables binarias son las variables referidas al tipo de centro (público o concertado) y al género (chico o chica). El resto son variables continuas: nivel socioeconómico, horas de uso de internet fuera del colegio, edad con la que empieza a utilizar dispositivos digitales, apoyo del profesorado, ambiente escolar y horas dedicadas a aprender otro idioma.

3.2.2. Restricciones

En lo que sigue, vamos a definir el conjunto factible del problema de optimización multiobjetivo a través de restricciones, teniendo en cuenta el significado de las variables de decisión (explicativas) y la dependencia existente entre ellas.

En el caso que haya variables binarias categóricas, se definen una serie de restricciones técnicas para controlar los diferentes valores que estas pueden tomar. En primer lugar, debemos asegurar que en el problema de optimización formulado, las variables binarias categóricas referidas al mismo aspecto no tomen el valor 1 simultáneamente. Así, en el problema de optimización multiobjetivo se incluye una restricción de este tipo $\sum_{i\in H} x_i = 1$, donde $H \subset \{1,\ldots,k\}$ es el conjunto de subíndices de las variables que no pueden tomar el valor 1 a la vez. En segundo lugar, cuando no podemos considerar la variable de referencia como variable de decisión, la restricción anterior pasa a ser una restricción de desigualdad, es decir, $\sum_{i\in H} x_i \leq 1$. En nuestro caso, al no tener variables binarias categóricas, el problema multiobjetivo no tendrá restricciones de este tipo.

Por otro lado, para definir de forma realista el conjunto factible, se considera un conjunto de restricciones basadas en las dependencias existentes entre las variables de decisión. Es posible que existan variables explicativas que muestren fuertes dependencias, por lo tanto, deben controlarse mediante restricciones. Estas se pueden construir a partir de un análisis de regresión mediante intervalos de confianza. Normalmente, se consideran los intervalos de confianza al 95 %.

En el problema de optimización multiobjetivo que estamos estudiando se considerarán restricciones de este tipo. En este caso, se estudia la dependencia del nivel socioeconómico (x_1) y el género (x_3) con las variables: horas de uso de internet fuera del colegio (x_4), edad con la que empiezan a usar dispositivos digitales (x_5) y horas dedicadas a aprender otro idioma (x_8). Por ejemplo, consideremos la dependencia entre x_1 y x_3 con x_4. En este caso, x_4 se puede aproximar con x_1 y x_3 mediante una regresión por mínimos cuadrados ordinarios como sigue:

$$x_4 = \hat{a}_1 x_1 + \hat{a}_2 x_3 + \hat{b} \tag{6}$$

donde $\hat{a}_1 \in [\hat{a}_1^l, \hat{a}_1^u]$, $\hat{a}_2 \in [\hat{a}_2^l, \hat{a}_2^u]$ y $\hat{b} \in [\hat{b}^l, \hat{b}^u]$ son los intervalos de confianza obtenidos de la regresión. De esta forma, siguiendo (6), la variable x_4 puede tomar valores entre las siguientes cotas:

$$\hat{a}_1^l x_1 + \hat{a}_2^l x_3 + \hat{b}^l \leq x_4 \leq \hat{a}_1^u x_1 + \hat{a}_2^u x_3 + \hat{b}^u$$

Transformando esta desigualdad obtenemos dos restricciones para el problema de optimización multiobjetivo que estamos construyendo.

$$\begin{gathered} x_4 - (\hat{a}_1^l x_1 + \hat{a}_2^l x_3 + \hat{b}^l) \geq 0 \\ x_4 - (\hat{a}_1^u x_1 + \hat{a}_2^u x_3 + \hat{b}^u) \leq 0 \end{gathered}$$

En la Tabla 4 se muestran los valores de los intervalos de confianza de las restricciones consideradas para el problema.

Además, el mínimo y el máximo de cada una de las variables de decisión alcanza en la base de datos se usan para definir restricciones en forma de cotas. El mínimo se asocia como cota inferior de la variable y el máximo como cota superior de la variable. De esta forma, para cada $i = 1,\ldots,8$, tendremos dos restricciones adicionales, $x_i \geq l_i$ y $x_i \leq u_i$. En la Tabla 5 se presentan dichos valores, que formarán el resto de restricciones del problema.

De esta forma, las restricciones que delimitan el conjunto factible del problema son las siguientes:

$$\begin{array}{cl} -0{,}3070 \cdot x_1 + 0{,}1443 \cdot x_3 - x_4 \leq -3{,}4583 & \text{(C1)} \\ 0{,}1906 \cdot x_1 - 0{,}3677 \cdot x_3 + x_4 \leq 3{,}6419 & \text{(C2)} \\ -0{,}3852 \cdot x_1 + 0{,}3944 \cdot x_3 - x_5 \leq -7{,}1713 & \text{(C3)} \\ 0{,}2479 \cdot x_1 - 0{,}6532 \cdot x_3 + x_5 \leq 7{,}3679 & \text{(C4)} \\ 0{,}0601 \cdot x_1 + 0{,}0327 \cdot x_3 - x_8 \leq -4{,}1200 & \text{(C5)} \\ -0{,}1724 \cdot x_1 - 0{,}1742 \cdot x_3 + x_8 \leq 4{,}2703 & \text{(C6)} \\ l_i \leq x_i \leq u_i \text{ for all } i = 1,\ldots,8 & \text{(C7-C14)} \end{array} \tag{7}$$

Tabla 4. Valores de los coeficientes para las restricciones del problema

Restricciones		
Horas de Internet fuera del colegio (x_4)	$a^l_{x_i}$	$b^u_{x_i}$
Nivel socioeconómico (x_1)	-0.3070	-0.1906
Género (x_3)	0.1443	0.3677
Constante (α)	3.4583	3.6419
Edad de uso de dispositivos digitales (x_5)	$a^l_{x_i}$	$b^u_{x_i}$
Nivel socioeconómico (x_1)	-0.38520	-0.24790
Género (x_3)	0.39436	0.65323
Constante (α)	7.17126	7.36786
Horas dedicadas a aprender otro idioma (x_8)	$a^l_{x_i}$	$b^u_{x_i}$
Nivel socioeconómico (x_1)	0.06013	0.17244
Género (x_3)	0.03272	0.17420
Constante (α)	4.11997	4.27029

Tabla 5. Cota inferior y superior de las variables de decisión

Variables		**Cotas $[l_i, u_i]$**
Nivel socioeconómico	x_1	$[-4{,}316, 3{,}717]$
Tipo de colegio	x_2	0 or 1
Género	x_3	0 or 1
Horas de Internet fuera del colegio	x_4	[0,7]
Edad de uso de dispositivos digitales	x_5	[4.5,14.5]
Apoyo del profesorado	x_6	[-0.972,2.455]
Ambiente escolar	x_7	[-2.102,1.465]
Horas dedicadas a aprender otro idioma	x_8	[3,7]

3.2.3. Funciones objetivo

La idea principal de un problema de programación multiobjetivo es optimizar simultáneamente varias funciones objetivo (dos o más). Como el análisis econométrico nos ha permitido expresar los índices relacionados con el bienestar (variables respuesta) en función de un conjunto de variables significativas y se percibe cierto conflicto en las regresiones, ya que existen diferencias en los signos de los coeficientes, es adecuado definir un problema multiobjetivo. Las funciones objetivo serán denotadas por $f_j(x_1, \ldots, x_8)$ con $j = 1, \ldots, 4$, y los coeficientes se definen a partir de los intervalos de confianza obtenidos por las estimaciones y mostrados en la Tabla 3. Por lo tanto, las funciones objetivo se expresan como sigue:

$$\begin{aligned}
f_1(x_1, \ldots, x_8) &= [0{,}0006, 0{,}0554] \cdot x_1 + [-0{,}0980, -0{,}0138] \cdot x_3 + \cdots + [-0{,}0903, 0{,}1106] \\
f_2(x_1, \ldots, x_8) &= [0{,}0087, 0{,}0603] \cdot x_1 + [0{,}1274, 0{,}2164] \cdot x_3 + \cdots + [-0{,}2298, -0{,}0383] \\
f_3(x_1, \ldots, x_8) &= [0{,}0159, 0{,}0661] \cdot x_1 + [0{,}0395, 0{,}1508] \cdot x_2 + \cdots + [-0{,}29443, -0{,}0391] \\
f_4(x_1, \ldots, x_8) &= [-0{,}1707, -0{,}0781] \cdot x_3 + [-0{,}1707, -0{,}0781] \cdot x_4 + \cdots + [-0{,}1050, 0{,}1258]
\end{aligned} \tag{8}$$

Así, el problema multiobjetivo que hemos diseñado está formado por cuatro funciones objetivo, tres de ellas a maximizar (índice de positividad, motivación y sentido de pertenencia) y la última a minimizar (índice de acoso escolar). Además, el conjunto factible está delimitado por catorce restricciones, de las cuales ocho son cotas. Así, el problema tiene la siguiente expresión:

$$\begin{array}{lll}
\max & f_1 = [0{,}0006, 0{,}0554] \cdot x_1 + \cdots + [-0{,}0903, 0{,}1106] & \\
\max & f_2 = [0{,}0087, 0{,}0603] \cdot x_1 + \cdots + [-0{,}2298, -0{,}0383] & \\
\max & f_3 = [0{,}0159, 0{,}0661] \cdot x_1 + \cdots + [-0{,}29443, -0{,}0391] & \\
\min & f_4 = [-0{,}1707, -0{,}0781] \cdot x_3 + \cdots + [-0{,}1050, 0{,}1258] & \\
\text{sujeto a} & -0{,}3070 \cdot x_1 + 0{,}1443 \cdot x_3 - x_4 \leq -3{,}4583 & \text{(C1)} \\
& 0{,}1906 \cdot x_1 - 0{,}3677 \cdot x_3 + x_4 \leq 3{,}6419 & \text{(C2)} \\
& -0{,}3852 \cdot x_1 + 0{,}3944 \cdot x_3 - x_5 \leq -7{,}1713 & \text{(C3)} \\
& 0{,}2479 \cdot x_1 - 0{,}6532 \cdot x_3 + x_5 \leq 7{,}3679 & \text{(C4)} \\
& 0{,}0601 \cdot x_1 + 0{,}0327 \cdot x_3 - x_8 \leq -4{,}1200 & \text{(C5)} \\
& -0{,}1724 \cdot x_1 - 0{,}1742 \cdot x_3 + x_8 \leq 4{,}2703 & \text{(C6)} \\
& l_i \leq x_i \leq u_i \text{ for all } i = 1, \ldots, 8 & \text{(C7-C14)}
\end{array} \tag{9}$$

4. Resolución del problema de optimización multiobjetivo

Los problemas de programación multiobjetivo deben tener en cuenta explícitamente el tratamiento de la incertidumbre inherente a los coeficientes del modelo [3]. La programación intervalar o programación por intervalos es uno de los enfoques utilizados para tratar la incertidumbre en los problemas de programación multiobjetivo. A diferencia de la programación estocástica o de la programación difusa, que parten de la especificación o la asunción de distribuciones probabilísticas y distribuciones posibilistas, respectivamente, la programación por intervalos solo requiere información sobre el rango de variación de algunos (o todos) los coeficientes del problema (Oliveira y Antunes, 2007). En esta sección, se presenta un nuevo enfoque para resolver problemas de programación multiobjetivo por intervalos en los que los coeficientes de las funciones objetivo son intervalos, como el problema que hemos definido en la sección anterior. El nuevo enfoque permite generar las denominadas soluciones posible y necesariamente eficientes, si existen, sin requerir un esfuerzo computacional considerable y utilizando técnicas de optimización de un solo objetivo.

4.1. Conceptos y notación

Sin perdida de generalidad, vamos a considerar el siguiente problema de programación multiobjetivo con intervalos en los coeficientes de las funciones objetivo:

$$\begin{array}{lll}
\text{maximizar} & F_k(\mathbf{x}) = \sum_{j=1}^{n} [c_{kj}^L, c_{kj}^U] x_j, & k = 1, \ldots, p, \\
\text{sujeto a} & \sum_{j=1}^{n} a_{ji} x_j \leq b_i & i = 1, \ldots, m, \\
& x_j \geq 0 & j = 1, \ldots, n.
\end{array} \tag{10}$$

donde $\mathbf{x} = (x_1, \ldots, x_n)^T$ es el vector de variables de decisión y $X = \{\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n | \sum_{j=1}^{n} a_{ji} x_j \leq b_i, i = 1, \ldots, m, x_j \geq 0, j = 1, \ldots, n\}$ es la región factible del problema.

Es evidente que para cada uno de los coeficientes $\bar{c_k j} \in [c_{kj}^L, c_{kj}^U]$ es un valor real, entonces el problema (10) es un problema de optimización multiobjetivo como el siguiente:

$$\begin{array}{lll}
\text{maximizar} & \bar{F}_k(\mathbf{x}) = \sum_{j=1}^{n} \bar{c}_k x_j, & k = 1, \ldots, p, \\
\text{sujeto a} & \sum_{j=1}^{n} a_{ji} x_j \leq b_i & i = 1, \ldots, m, \\
& x_j \geq 0 & j = 1, \ldots, n.
\end{array} \tag{11}$$

donde $\bar{c}_{kj} \in [c_{kj}^L, c_{kj}^U]$ para todo $k = 1, \ldots, p, j = 1, \ldots, n$.

El concepto de valor ideal en programación multiobjetivo, pasa a ser intervalo ideal en programación intervalar. De forma que, la cota inferior del intervalo ideal corresponde con el valor

[3]Nos referimos con coeficientes del modelo, a los coeficientes de las funciones objetivo y de las restricciones.

ideal cuando cosideramos la cota inferior de los intervalos en las funciones objetivos, y, De forma análoga, la cota superior del intervalo ideal está formada por el valor ideal cuando consideramos la cota superior de los intervalos en las funciones objetivo. Así, para cada función objetivo $F_k(x)$, se resuelven los siguientes problemas lineales (Chinneck y Ramadan, 2000):

$$\begin{array}{lll} \text{maximizar} & F_k^U(\mathbf{x}) = \sum_{j=1}^n c_{kj}^U x_j, & \\ \text{sujeto a} & \sum_{j=1}^n a_{ji} x_j \leq b_i & i = 1, \ldots, m, \\ & x_j \geq 0 & j = 1, \ldots, n. \end{array} \tag{12}$$

y

$$\begin{array}{lll} \text{maximizar} & F_k^L(\mathbf{x}) = \sum_{j=1}^n c_{kj}^L x_j, & \\ \text{sujeto a} & \sum_{j=1}^n a_{ji} x_j \leq b_i & i = 1, \ldots, m, \\ & x_j \geq 0 & j = 1, \ldots, n. \end{array} \tag{13}$$

denotando a las soluciones óptimas de (12) and (13) como $\mathbf{x}_k^U$ y $\mathbf{x}_k^L$, respectivamente, y a los correspondientes valores objetivo como:

$$F_k^{U^*} = F_k^U(\mathbf{x}_k^U),\ k = 1, \ldots, p, \tag{14}$$

$$F_k^{L^*} = F_k^L(\mathbf{x}_k^L),\ k = 1, \ldots, p. \tag{15}$$

De esta forma, los intervalos ideales son $[F_k^{L^*}, F_k^{U^*}]$ para $k = 1, \ldots, p$ y representan las cotas de la solución ideal intervalar (Oliveira y Antunes, 2007).

Además, el concepto de eficiencia en programación multiobjetivo (Miettinen, 1999), definición 1, también se adapta la programación multiobjetivo intervalar (Bitran, 1980) en las siguientes definiciones (2) y (3).

Definición 1 *Una solución $x' \in X$ es eficiente para el problema (11) sí y solo sí no existe otra solución $x \in X$ de forma que $\bar{F}_k(\mathbf{x}') \leq \bar{F}_k(\mathbf{x})$ para cada objetivo $k = 1, \ldots, p$ cumpliendo, al menos, una desigualdad estricta.*

Definición 2 *Una solución $\mathbf{x}' \in X$ es necesariamente eficiente para el problema (10) sí y solo sí es eficiente para el problema (11) para cualquier $\bar{cj}_k \in [c_{kj}^L, c_{kj}^U]$ con $k = 1, \ldots, p, j = 1, \ldots, n.$. El conjunto de soluciones necesariamente eficientes se donota por N_E.*

Definición 3 *Una solución $\mathbf{x}' \in X$ es posiblemente eficiente para el problema (10) si es eficiente para (11) para al menos un $\bar{c}_{kj} \in [c_{kj}^L, c_{kj}^U]$ con $k = 1, \ldots, p, j = 1, \ldots, n$. El conjunto de soluciones posiblemente eficientes se denota como P_E.*

Con estas definiciones, es trivial que $N_E \subseteq P_E$. Cualquier solución adecuada para un problema multiobjectivo intervalar podría escogerse de los conjuntos N_E o P_E. Sin embargo, debido a que estos conjuntos pueden contener infinitos elementos, se requiere un método para obtener soluciones.

Existen diferentes formas de introducir preferencias en la resolución de un problema multiobjetivo. Una de ellas es mediante un punto de referencia $\mathbf{q} = (q_1, \ldots, q_p)^T$, y consiste en dar valores de referencia o valores deseados para las funciones objetivo.

Una de las funciones comúnmente usadas para incorporar preferencias es la función de logro propuesta por (Wierzbicki, 1980):

$$s(\mathbf{q}, \mathbf{f}(\mathbf{x}), \mu) = \max_{k=1,\ldots,p} \{\ \mu_k(f_k(\mathbf{x}) - q_k)\ \} + \rho \sum_{k=1}^{p} \mu_k(f_k(\mathbf{x}) - q_k), \tag{16}$$

donde $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_p(\mathbf{x}))$ es el vector de funciones objetivo, $\mu = (\mu_1, \ldots, \mu_p)^T$ es el vector de pesos, $\mathbf{q} = (q_1, \ldots, q_p)^T$ es el punto de referencia y $\rho > 0$ se denomina coeficiente de aumento. Esta función debe ser minimizada en la la región factible:

$$\begin{aligned} &\min \max_{k=1,\ldots,p} \{\, \mu_k(f_k(\mathbf{x}) - q_k) \,\} + \rho \sum_{k=1}^{p} \mu_i(f_k(\mathbf{x}) - q_k), \\ &\text{sujeto a :}\ \mathbf{x} \in X \end{aligned} \tag{17}$$

Para este tipo de problemas escalarizados, es necesario considerar un problema multiobjetivo determinista como (11). En este caso, $f_k(x) = \bar{F}_k(\mathbf{x}) = \sum_{j=1}^{n} \bar{c}_{kj} x_j$ para todo $k = 1 \ldots, p$, $j = 1, \ldots, n$.

Se demuestra que cada solución de (17) es una solución Pareto óptima de (11) y, además, cualquier solución Pareto óptima se obtiene resolviendo el problema (17) utilizando como punto de referencia el punto ideal (o cualquier vector objetivo que domine a este, como el punto utopía), y variando el vector de pesos en todo el espacio de vectores de pesos (Wierzbicki, 1980; Miettinen, 1999). Por otro lado, se demuestra que cualquier solución Pareto óptima se puede obtener resolviendo el problema (17) fijando un vector de peso y variando el punto de referencia (Miettinen, 1999).

En lo que sigue, vamos a describir brevemente, algunos problemas escalarizados para obtener soluciones posiblemente eficientes de (10) propuestas en (Henriques *et al.*, 2019). Vamos a considerar que queremos minimizar la peor distancia posible entre el intervalo de la función objetivo con el intervalo deseado. Si consideramos el intervalo de las funciones objetivo:

$$F_k(\mathbf{x}) = \left[F_k^L(\mathbf{x}) = \sum_{j}^{n} c_{kj}^L x_j, F_k^U(\mathbf{x}) = \sum_{j}^{n} c_{kj}^U x_j \right]$$

y queremos minimizar la distancia a un intervalo deseado $T_k = [t_k^L, t_k^U]$, podemos definir el siguiente problema

$$\begin{array}{lll} \text{minimizar} & S(e_1^L, \ldots, e_p^L, e_1^U, \ldots, e_p^U) & \\ \text{sujeto a} & \sum_{j=1}^{n} c_{kj}^L x_j + e_k^L = t_k^L, & k = 1, \ldots, p, \\ & \sum_{j=1}^{n} c_{kj}^U x_j + e_k^U = t_k^U, & k = 1, \ldots, p, \\ & \sum_{j=1}^{n} a_{ji} x_j \leq b_i & i = 1, \ldots, m, \\ & x_j \geq 0 & j = 1, \ldots, n. \end{array} \tag{18}$$

donde las variables son x_j, e_k^L y e_k^U con $j = 1, \ldots, n$, $k = 1, \ldots, p$. Estas últimas (e_k^L y e_k^U) representan las desviaciones de los intervalos y $S(e_1^L, \ldots, e_p^L, e_1^U, \ldots, e_p^U)$ es una función de logro de dichas desviaciones.

Si consideramos que T_k está definido por las cotas de la solución intervalar ideal, es decir, $T_k = [F_k^{L*}, F_k^{U*}]$, tenemos el siguiente problema:

$$\begin{array}{lll} \text{minimizar} & S(e_1^L, \ldots, e_p^L, e_1^U, \ldots, e_p^U) & \\ \text{sujeto a} & \sum_{j=1}^{n} c_{kj}^L x_j + e_k^L = F_k^{L*}, & k = 1, \ldots, p, \\ & \sum_{j=1}^{n} c_{kj}^U x_j + e_k^U = F_k^{U*}, & k = 1, \ldots, p, \\ & \sum_{j=1}^{n} a_{ji} x_j \leq b_i & i = 1, \ldots, m, \\ & x_j \geq 0 & j = 1, \ldots, n. \end{array} \tag{19}$$

donde las variables de desviación e_k^L,e_k^U, son positivas o cero.

El problema (19) puede resolverse considerando un procedimiento optimista (el límite inferior de la desviación necesaria) o un procedimiento pesimista (el límite superior de la desviación necesaria). El segundo enfoque permite obtener la menor distancia con respecto el objetivo considerado (Oliveira y Antunes, 2007; Inuiguchi y Kume, 1991).

Si consideramos la siguiente función de logro:

$$S(e_1^L, \dots, e_p^L, e_1^U, \dots, e_p^U) = \max_{k=1,\dots,p}\{\mu_k^L e_k^L + \mu_k^U e_k^U\} \tag{20}$$

el problema (19) viene dado por:

$$\begin{array}{lll} \text{minimizar} & max_{k=1,\dots,p}\{\mu_k^L e_k^L + \mu_k^U e_k^U\} & \\ \text{sujeto a} & \sum_{j=1}^n c_{kj}^L x_j + e_k^L = F_k^{L*}, & k = 1, \dots, p, \\ & \sum_{j=1}^n c_{kj}^U x_j + e_k^U = F_k^{U*}, & k = 1, \dots, p, \\ & \sum_{j=1}^n a_{ji} x_j \leq b_i & i = 1, \dots, m, \\ & x_j \geq 0 & j = 1, \dots, n. \end{array} \tag{21}$$

y se intenta alcanzar las cotas de la solución ideal intervalar mediante un enfoque minimax usando un término aditivo para cada intervalo.

Por un lado, está demostrado en Henriques *et al.* (2019) que si la solución del problema

$$\begin{array}{lll} \text{minimizar} & v & \\ \text{sujeto a} & \mu_k^L(F_k^{L*} - \sum_{j=1}^n c_{kj}^L x_j) + \mu_k^U(F_k^{U*} - \sum_{j=1}^n c_{kj}^U x_j) \leq v, & k = 1, \dots, p, \\ & \sum_{j=1}^n a_{ji} x_j \leq b_i & i = 1, \dots, m, \\ & x_j \geq 0 & j = 1, \dots, n. \end{array} \tag{22}$$

es única, entonces es una solución posiblemente eficiente de (10).

Se puede demostrar que cada solución de (21) es también una solución de (22). Así, tenemos que, si la solución de (21) es única, entonces es una solución posiblemente eficiente de (10). Consideremos $v^* = \max_{k=1,\dots,p}\{\mu_k^L e_k^{*L} + \mu_k^U e_k^{*U}\}$. Teniendo en cuenta las restricciones del problema (21):

$$\begin{aligned} e_k^{*L} &= F_k^{L*} - \textstyle\sum_{j=1}^n c_{kj}^L x_j^*, \; k = 1, \dots, p, \\ e_k^{*U} &= F_k^{U*} - \textstyle\sum_{j=1}^n c_{kj}^U x_j^*, \; k = 1, \dots, p, \end{aligned}$$

Vamos a probar que $\mathbf{x}^* = (x_1^*, \dots, x_n^*)^T$ y v^* es también una solución óptima de (22). Supongamos que $\mathbf{x}^* = (x_1^*, \dots, x_n^*)^T$ y v^* no es una solución óptima de (22). Sean $\bar{\mathbf{x}} = (\bar{x}_1, \dots, \bar{x}_n)^T$ y $\bar{v}$ una solución óptima de (22). Por lo tanto, tenemos que $\bar{v} < v^*$ y en consecuencia:

$$\begin{aligned} \bar{v} =& \max_{k=1,\dots,p}(\mu_k^L(F_k^{L*} - \sum_{j=1}^n c_{kj}^L \bar{x}_j) + \mu_k^U(F_k^{U*} - \sum_{j=1}^n c_{kj}^U \bar{x}_j)) < \\ &< \max_{k=1,\dots,p}(\mu_k^L(F_k^{L*} - \sum_{j=1}^n c_{kj}^L x_j^*) + \mu_k^U(F_k^{U*} - \sum_{j=1}^n c_{kj}^U x_j^*)) = v^* \end{aligned} \tag{23}$$

Considerando de nuevo las restricciones del problema (21):

$$\begin{aligned} \bar{e}_k^L &= F_k^{L*} - \textstyle\sum_{j=1}^n c_{kj}^L \bar{x}_j, \; k = 1, \dots, p, \\ \bar{e}_k^U &= F_k^{U*} - \textstyle\sum_{j=1}^n c_{kj}^U \bar{x}_j, \; k = 1, \dots, p, \end{aligned}$$

y sustituyendo esto en lo anterior (23), se obtiene:

$$\max_{k=1,\dots,p}\{\mu_k^L \bar{e}_k^L + \mu_k^U \bar{e}_k^U\} < \max_{k=1,\dots,p}\{\mu_k^L e_k^* + \mu_k^U e_k^*\}$$

Esto es una contradicción con que $\mathbf{x}^* = (x_1^*, \dots, x_n^*)^T$, e_k^{*U}, e_k^{*L}, con $k = 1, \dots, p$ es solucion óptima de (21).

Con el objetivo de no requerir la unicidad de la solución, se puede añadir un termino de aumento a la función de logro de las desviaciones de la siguiente forma:

$$\begin{aligned} &S(e_1^L,\ldots,e_p^L,e_1^U,\ldots,e_p^U)\\ &=\max_{k=1,\ldots,p}\{\mu_k^L e_k^L+\mu_k^U e_k^U\}+\rho\sum_{k=1}^{p}(\mu_k^L e_k^L+\mu_k^U e_k^U) \end{aligned} \tag{24}$$

donde $\rho > 0$ se denomina término de aumento. Por lo tanto, el problema a minimizar viene dado ahora por:

$$\begin{array}{lll} \text{minimizar} & max_{k=1,\ldots,p}\{\mu_k^L e_k^L+\mu_k^U e_k^U\}+\rho\sum_{k=1}^{p}(\mu_k^L e_k^L+\mu_k^U e_k^U) & \\ \text{sujeto a} & \sum_{j=1}^{n} c_{kj}^L x_j + e_k^L = F_k^{L*}, & k=1,\ldots,p,\\ & \sum_{j=1}^{n} c_{kj}^U x_j + e_k^U = F_k^{U*}, & k=1,\ldots,p,\\ & \sum_{j=1}^{n} a_{ji}x_j \le b_i & i=1,\ldots,m,\\ & x_j, e_k^L, e_k^U \ge 0 & j=1,\ldots,n. \end{array} \tag{25}$$

Está demostrado en Henriques *et al.* (2019) que la solución del problema:

$$\begin{array}{lll} \text{minimize} & v+\rho\sum_{k=1}^{p}(\mu_k^L(F_k^{L*}-\sum_{j=1}^{n}c_{kj}^L x_j)+\mu_k^U(F_k^{U*}-\sum_{j=1}^{n}c_{kj}^U x_j)) & \\ \text{sujeto a} & \mu_k^L(F_k^{L*}-\sum_{j=1}^{n}c_{kj}^L x_j)+\mu_k^U(F_k^{U*}-\sum_{j=1}^{n}c_{kj}^U x_j)\le v, & k=1,\ldots,p,\\ & \sum_{j=1}^{n} a_{ji}x_j \le b_i & i=1,\ldots,m,\\ & x_j \ge 0 & j=1,\ldots,n. \end{array} \tag{26}$$

es siempre una solución posiblemente eficiente de (10).

Por otro lado, se puede demostrar que si el valor óptimo v del problema (21) es cero, la solución x^* es también una solución necesariamente eficiente. Es decir, sea $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)^T$, e_k^{*L},e_k^{*U} con $k = 1, \ldots, p$ la solución óptima de (21), donde μ_k^L, $\mu_k^U > 0$ y $\mu_k^L + \mu_k^U = 1$ para todo $k = 1, \ldots, p$.

Supongamos que x^* no es una solución necesariamente eficiente del problema (10), por lo tanto $x^* \notin N_E$, y consideramos $x^1 \in N_E$ una solución necesariamente eficiente. De esta forma, para cualquier coeficiente de las funciones objetivo, no existe otra solución que domine a x^1. Lo que significa que si consideramos, por ejemplo, los coeficientes c_{kj}^L para $k = 1, \ldots, p$, $j = 1, \ldots, n$ y teniendo en cuenta que x^* es una solución factible, existe al menos un k_0 que cumple lo siguiente:

$$\sum_{j=1}^{n} c_{k_0 j}^L x_j^* < \sum_{j=1}^{n} c_{k_0 j}^L x_j^1 \tag{27}$$

Si tenemos en cuenta las definiciones de F_k^{U*} y F_k^{L*} en (14) y (15) respectivamente, para cada $\mathbf{x} = (x_1, \ldots, x_n)^T \in X$:

$$F_k^{L^*} - \sum_{j=1}^{n} c_k^L x_j \ge 0 \text{ y } F_k^{U^*} - \sum_{j=1}^{n} c_k^U x_j \ge 0 \text{ para todo } k = 1, \ldots, p.$$

Como tenemos que $v^* = 0$ para $\mathbf{x}^*$ y μ_k^L, $\mu_k^U > 0$ para todo $k = 1, \ldots, p$, es verifica:

$$F_k^{L^*} - \sum_{j=1}^{n} c_{kj}^L x_j^* = 0 \text{ y } F_k^{U^*} - \sum_{j=1}^{n} c_{kj}^U x_j^* = 0 \text{ para todo } k = 1, \ldots, p.$$

Entonces, para cada $\mathbf{x} \in X$:

$$\sum_{j=1}^{n} c_{kj}^L x_j \le \sum_{j=1}^{n} c_{kj}^L x_j^* \text{ and } \sum_{j=1}^{n} c_{kj}^U x_j \le \sum_{j=1}^{n} c_{kj}^U x_j^* \text{ for all } k = 1, \ldots, p.$$

Paricularmente para $\mathbf{x}^1$:

$$\sum_{j=1}^{n} c_{kj}^L x_j^1 \le \sum_{j=1}^{n} c_{kj}^L x_j^* \text{ and } \sum_{j=1}^{n} c_{kj}^U x_j^1 \le \sum_{j=1}^{n} c_{kj}^U x_j^* \text{ for all } k = 1, \ldots, p.$$

Pero esto se contradice con (27), así obtenemos que $\mathbf{x}^* \in N_E$.

Esta nueva propuesta de problema escalarizado (26) nos permitirá generar las denominadas soluciones posiblemente eficientes (incluso necesariamente eficientes si $v^* = 0$) teniendo en cuenta la distancia ponderada a los valores ideales inferior y superior. De esta forma, el problema de optimización multiobjetivo intervalar definido en la sección anterior (Sección 3), será resuelto utilizando la formulación del problema (26) y teniendo en cuenta los resultados teóricos descritos en esta sección.

5. Resultados

En esta sección, vamos a resolver el problema anteriormente definido mediante un análisis econométrico (9):

$$
\begin{array}{lll}
\max & f_1 = [0{,}0006, 0{,}0554] \cdot x_1 + \cdots + [-0{,}0903, 0{,}1106] & \\
\max & f_2 = [0{,}0087, 0{,}0603] \cdot x_1 + \cdots + [-0{,}2298, -0{,}0383] & \\
\max & f_3 = [0{,}0159, 0{,}0661] \cdot x_1 + \cdots + [-0{,}29443, -0{,}0391] & \\
\min & f_4 = [-0{,}1707, -0{,}0781] \cdot x_3 + \cdots + [-0{,}1050, 0{,}1258] & \\
\text{sujeto a} & -0{,}3070 \cdot x_1 + 0{,}1443 \cdot x_3 - x_4 \leq -3{,}4583 & \text{(C1)} \\
& 0{,}1906 \cdot x_1 - 0{,}3677 \cdot x_3 + x_4 \leq 3{,}6419 & \text{(C2)} \\
& -0{,}3852 \cdot x_1 + 0{,}3944 \cdot x_3 - x_5 \leq -7{,}1713 & \text{(C3)} \\
& 0{,}2479 \cdot x_1 - 0{,}6532 \cdot x_3 + x_5 \leq 7{,}3679 & \text{(C4)} \\
& 0{,}0601 \cdot x_1 + 0{,}0327 \cdot x_3 - x_8 \leq -4{,}1200 & \text{(C5)} \\
& -0{,}1724 \cdot x_1 - 0{,}1742 \cdot x_3 + x_8 \leq 4{,}2703 & \text{(C6)} \\
& l_i \leq x_i \leq u_i \text{ for all } i = 1, \ldots, 8 & \text{(C7-C14)}
\end{array}
$$

En este problema pretendemos optimizar simultáneamente los índices de positividad, motivación, sentido de pertenencia y acoso escolar. Para ello vamos a seguir la formulación (26) propuesta en la sección anterior (Sección 4). Así, en primer lugar, tenemos que obtener los valores óptimos de cada una de las funciones objetivos usando la cota superior e inferior de los coeficientes. Esto se corresponde con la resolución de los problemas, (12) y (13), respectivamente. Los valores obtenidos se muestran en la Tabla 6.

Tabla 6. Valores ideales de cada objetivo para las cotas inferiores y superiores

Función objetivo	**Valor ideal inferior**	**Valor ideal superior**
Índice de positividad	0,0717	0,6020
Índice de motivación	0,1184	0,9872
Índice de sentido de pertenencia	0,0292	0,6637
Índice de acoso escolar	-0,7406	-0,0838

Un vez obtenidos los valores ideales para las cotas superiores e inferiores, estos se utilizan en la resolución del problema aplicando la formulación del problema escalarizado (26).

Como adelantábamos en la secciones previas, la resolución de este tipo de problemas, da información sobre el perfil de estudiante que alcanza el equilibrio óptimo entre los objetivos (Tabla 7) y el valor de cada función objetivo en el óptimo en ambos casos, cotas superiores y cotas inferiores (Tabla 8).

Analizando los valores obtenidos por las variables de decisión, podemos obtener varias conclusiones del perfil de estudiante que alcanzaría el balance óptimo entre los índices del bienestar. En primer lugar, el estudiante debería alcanzar un nivel socioeconómico de $3,7171$. Además, este balance lo alcanzan las chicas matriculadas en centros concertados. Siguiendo con las variables de decisión, estos estudiantes deberían emplear 2,4616 horas de uso de Internet fuera del colegio

Tabla 7. Solución del problema multiobjetivo de programación intervalar

Variables		**Valor en la solución**
Nivel socioeconómico	x_1	3,7171
Tipo de colegio	x_2	1
Género	x_3	1
Horas de Internet fuera del colegio	x_4	2,4616
Edad de uso de dispositivos digitales	x_5	7,0996
Apoyo del profesorado	x_6	-0,9723
Ambiente escolar	x_7	1,4652
Horas dedicadas a aprender otro idioma	x_8	5,0855

Tabla 8. Valor de las funciones objetivo en la solución

Funciones objetivo	**Peor escenario**	**Mejor escenario**
Índice de positividad	-0,0256	0,5987
Índice de motivación	0,1016	0,9050
Índice de sentido de pertenencia	-0,0769	0,6275
Índice de acoso escolar	-0,0829	-0,6737

y empezar a utilizar dispositivos digitales a los 7 años. El apoyo por parte del profesorado y el ambiente escolar deberían alcanzar su mejor valor. Por último, estos estudiantes deberían dedicar aproximadamente 5 horas aprendiendo otro idioma.

Los valores alcanzados por los índices se dan tanto para el peor y mejor escenario (utilizando la cota inferior (respectivamente superior) del intervalo en el caso que el índice sea a minimizar y superior (respectivamente inferior) en el caso que el índice sea a maximizar. Para el peor escenario, los índices de positividad y sentido de pertenencia empeoran sus correspondientes medias en la muestra. Sin embargo, los índices de motivación y acoso escolar mejoran sus medias, este último obteniendo un valor muy próximo a ella. En El mejor escenario, todos los índices mejoran su media.

Tabla 9. Distancia de la solución al ideal

Funciones objetivo	$d(F_k^*, F_k(x))$
Índice de positividad	0,0973
Índice de motivación	0,1082
Índice de sentido de pertenencia	0,1061
Índice de acoso escolar	0,0669

Con el objetivo de medir cómo de cerca están las soluciones obtenidas del intervalo ideal (Tabla 6), se calculan las siguientes distancias (Tabla 9):

$$d(F_k^*, F_k(x)) = max_{k\in 1,2,3,4}(|F_k^{L^*} - F_k^L(x)|, |F_k^{U^*} - F_k^U(x)|). \tag{28}$$

Por lo tanto, cuando la distancia está cerca de cero, el intervalo de soluciones está más cera de su correspondiente intervalo ideal.

Centrándonos en las funciones objetivo, podemos decir que los índices de positividad y acoso están más cercanos a su ideal que los índices de motivación y sentido de pertenencia. Eso lo podemos ver comparando las Tablas 6 y 8, además de observando las distancias. En términos de nuestro problema, podemos concluir que, si deseamos alcanzar simultáneamente valores para los cuatro indicadores lo más cerca posible de sus correspondientes niveles ideales óptimos, los mejores resultados se alcanzan cuando los niveles de los índices de positividad y acosos se acerca a su intervalo ideal, a expensas de un sacrificio en los índices de motivación y sentido de pertenencia.

6. Conclusiones

En este artículo se propone un entorno metodológico en el que se combinan técnicas econométricas y de optimización multiobjetivo que pueden resultar muy útiles para investigadores y decisores. Utilizando esta combinación, es posible identificar qué valores de la variables explicativas, ajustándose a los datos observados, alcanzan simultáneamente valores óptimos de las variables que consideramos como objetivo o variables respuesta.

La metodología propuesta es muy útil cuando existe cierto grado de conflicto entre las variables explicativas consideradas. En esta situación, se sugiere aplicar la combinación de técnicas propuestas en este artículo. Para ello, definimos un problema multiobjetivo a partir de las regresiones en las que se basa el análisis econométrico considerando incertidumbre. Posteriormente, este problema se resolverá utilizando técnicas de programación multiobjetivo intervalar pudiendo así sacar conclusiones sobre el tema que estemos estudiando.

Con el resultado obtenido de la resolución del problema, los decisores/investigadores pueden sugerir políticas con el objetivo de mejorar las variables explicativas tanto como sea posible para poder alcanzar, simultáneamente, valores deseados en los objetivos. Además, podemos prever como afectará la mejora de alguno de los objetivos en el resto de ellos. Este tipo de información es imposible obtenerla haciendo solo un análisis econométrico del problema, por lo que el enfoque combinado que aquí se propone constituye una aportación de investigación muy relevante.

Particularmente, en este trabajo se propone estudiar un problema referido al bienestar de los estudiantes españoles, ya que cada vez son más las evidencias encontradas en las que se relaciona el rendimiento académico, y por consecuente su rendimiento posterior como persona adulta, con el bienestar de los adolescentes en el colegio. Para ello, hemos diseñado un problema multiobjetivo en el que se optimizan, de forma simultánea, cuatro índices que componen el bienestar de los estudiantes (positividad, motivación, sentido de pertenencia y acoso escolar) en función de un conjunto de variables explicativas relacionadas con el proceso de enseñanza-aprendizaje. Los resultados obtenidos en este problema permiten arrojar luz al desarrollo de políticas educativas orientadas con el objetivo de alcanzar niveles óptimos del bienestar. Además, las conclusiones alcanzadas permitirían a los responsables políticos anticiparse, a efectos prácticos, de la mejora de todos estos indicadores (positividad, motivación, sentido de pertenencia y acoso) al mismo tiempo. Así, esta aplicación demuestra los beneficios de nuestra propuesta en la práctica. Creemos que las posibilidades de aplicar este marco entorno a otros contextos pertenecientes a diferentes campos de investigación (como, por ejemplo, la industria, arquitectura o la ingeniería) y merecen futuros trabajos de investigación.

Agradecimientos

Este trabajo está respaldado por el Ministerio de Economía y Competitividad (proyecto ECO2017-88883-R) y por la Consejería de Economía, Conocimiento y Bienestar de la Junta de Andalucía (grupo PAI SEJ-532 y UMA18-FEDERJA-024). Además, la autora del trabajo, Sandra González-Gallardo, es beneficiaria de un contrato de investigación dentro del "Sistema Nacional de Garantia Juvenil y del Programa Operativo de Empleo Juvenil 2014-2020 - Fondos FEDER".

Referencias

Allen, K., Kern, M. L., Vella-Brodrick, D., Hattie, J., y Waters, L. (2018). What schools need to know about fostering school belonging: a meta-analysis. *Educational Psychology Review*, 30:1–34.

Athanasiou, K., Melegkovits, E., Andrie, E., Magoulas, C., Tzavara, C. K., Richardson, C., Greydanus, D., Tsolia, M., y Tsitsika, A. (2018). Cross-national aspects of cyberbullying victimization among 14–17-year-old adolescents across seven european countries. *BMC Public Health*, 18.

Baker, D. P., Fabrega, R., Galindo, C., y Mishook, J. (2004). Instructional time and national achievement: Cross-national evidence. *PROSPECTS*, 34:311–334.

Bitran, G. R. (1980). Linear multiple objective problems with interval coefficients. *Management Science*, 26(7):694–706.

Chinneck, J. y Ramadan, K. (2000). Linear programming with interval coefficients. *Journal of the Operational Research Society*, 51.

Deb, K. y Miettinen, K. (2010). Nadir point estimation using evolutionary approaches: Better accuracy and computational speed through focused search. En Ehrgott, M., Naujoks, B., Stewart, T. J., y Wallenius, J., editores, *Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems*, pp. 339–354, Berlin, Heidelberg. Springer Berlin Heidelberg.

Demir, Y. y Kutlu, M. (2018). Relationships among internet addiction, academic motivation, academic procrastination and school attachment in adolescents. *International Online Journal of Educational Sciences*, (5):315–322.

Ehrgott, M. y Tenfelde-Podehl, D. (2003). Computation of ideal and nadir values and implications for their use in mcdm methods. *European Journal of Operational Research*, 151(1):119–139.

Fan, W. y Williams, C. (2018). The mediating role of student motivation in the linking of perceived school climate and achievement in reading and mathematics. *Frontiers in Education*, 3:50.

Henriques, C., Luque, M., Marcenaro-Gutierrez, O., y Lopez-Agudo, L. (2019). A multiobjective interval programming model to explore the trade-offs among different aspects of job satisfaction under different scenarios. *Socio-Economic Planning Sciences*, 66:35–46.

Inuiguchi, M. y Kume, Y. (1991). Goal programming problems with interval coefficients and target intervals. *European Journal of Operational Research*, 52(3):345 – 360.

Jogi, A.-L., Kikas, E., Lerkkanen, M.-K., y Mägi, K. (2015). Cross-lagged relations between math-related interest, performance goals and skills in groups of children with different general abilities. *Learning and Individual Differences*, 39:105–113.

Korhonen, T., Lavonen, J., Kukkonen, M., Sormunen, K., y Juuti, K. (2014). *The Innovative School as an Environment for the Design of Educational Innovations*, pp. 99–113. SensePublishers, Rotterdam.

Miettinen, K. (1999). *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston.

Mischel, J. y Kitsantas, A. (2020). Middle school students' perceptions of school climate, bullying prevalence, and social support and coping. *Social Psychology of Education*, 23:51–72.

OECD (2017). *PISA 2015 Results (Volume III)*.

OECD (2019). PISA 2018 results in focus. *OECD Publishing*.

Oliveira, C. H. y Antunes, C. H. (2007). Multiple objective linear programming models with interval coefficients – an illustrated overview. *European Journal of Operational Research*, 181(3):1434–1463.

Parhiala, P., Torppa, M., Vasalampi, K., Eklund, K., Poikkeus, A.-M., y Aro, T. (2018). Profiles of school motivation and emotional well-being among adolescents: Associations with math and reading performance. *Learning and Individual Differences*, 61:196–204.

Ryan, R. M. y Deci, E. L. (2018). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78.

Scheerens, J. y Hendriks, M. (2014). *State of the Art of Time Effectiveness*, pp. 7–29. Springer International Publishing, Cham.

Strøm, I. F., Thoresen, S., Wentzel-Larsen, T., y Dyb, G. (2013). Violence, bullying and academic achievement: A study of 15-year-old adolescents and their school environment. *Child Abuse & Neglect*, 37(4):243–251.

Veenstra, R., Lindenberg, S., Oldehinkel, A. J., De Winter, A. F., Verhulst, F. C., y Ormel, J. (2005). Bullying and victimization in elementary schools: A comparison of bullies, victims, bully/victims, and uninvolved preadolescents. *Developmental Psychology*, 41(4):672–682.

Warr, P. (1999). *Well-being and the Workplace*, pp. 392–412. New York: Russell Sage Foundation.

Wierzbicki, A. P. (1980). The use of reference objectives in multiobjective optimization. En Fandel, G. y Gal, T., editores, *Multiple Criteria Decision Making, Theory and Applications*, pp. 468–486. Springer.

Yu, J., McLellan, R., y Winter, L. (2020). Which boys and which girls are falling behind? Linking adolescents' gender role profiles to motivation, engagement, and achievement. *Journal of Youth and Adolescence*, 50:336–352.

Yıldırım, S. (2012). Teacher support, motivation, learning strategy use, and achievement: A multilevel mediation model. *The Journal of Experimental Education*, 80(2):150–172.