



---

# Universidad de Valladolid

## Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

Grado en Economía

### Influencia de los factores socioeconómicos en los resultados para los países del informe PISA: Análisis a través de un modelo econométrico

Presentado por:

***Micaela Casas Pérez***

Tutelado por:

***Pilar Zarzosa Espina***

Valladolid, 21 de Septiembre de 2020



## **RESUMEN**

El objetivo principal en este trabajo es el análisis de los factores socioculturales que afectan en el rendimiento académico de los adolescentes. Primero, se ha delimitado los conceptos necesarios para este estudio y se ha explicado la elección del Informe PISA como base de datos. Tras esto, se han definido las variables que se utilizarán cuando se aplique la metodología econométrica. Además, se ha desempeñado una observación de las diferencias entre los países seleccionados.

Para realizar el estudio, se han extraído los datos del Informe PISA. Después se ha introducido el modelo a estimar y se han efectuado las pruebas pertinentes con el fin de validar el modelo. Por último, han sido expuestas los resultados del trabajo junto a las conclusiones.

**Palabras clave:** estudio sociocultural, Informe PISA, econometría.

**Jel Classification:** I2: Education and Research Institutions, C51: Model Construction and Estimation, O57: Comparative Studies of Countries.

## **ABSTRACT**

The main objective in this project is analyzing the sociocultural factors that affects the adolescent academic performance. First, the necessary concepts for this project have been delimited and the choice of the PISA Report as the database has been explained. After this, the variables that will be used when applying the econometric methodology have been defined. Moreover, an observation of the differences between selected countries has also been carried out.

In order to carry out the study, the data from PISA Report have been extracted. Afterwards, the model to be estimated has been introduced and the pertinent tests have been effectuated for the purpose to validate the model. Finally, the results of the work have been presented together with the conclusions.

**Key words:** sociocultural studies, PISA Report, econometrics.

**Jel Classification:** I2: Education and Research Institutions, C51: Model Construction and Estimation, O57: Comparative Studies of Countries.



## ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN.....	1
2. CONSIDERACIONES PREVIAS DEL INFORME PISA.....	3
3. OBJETIVOS PRINCIPALES Y PARTICULARES .....	4
4. VARIABLES ELEGIDAS .....	5
4.1 - Puntuaciones medias en matemáticas y ciencias .....	5
4.2 - Índice de miedo al fracaso en las mujeres ( $x_1$ ).....	6
4.3 - Porcentaje de alumnos inmigrantes que esperan acabar sus estudios ( $x_2$ ) .....	6
4.4 - Porcentaje de colegios públicos ( $x_3$ ).....	7
4.5 - ISEC: índice socioeconómico y cultural ( $x_4$ ).....	8
4.6 - Porcentaje de <i>bullying</i> ( $x_5$ ).....	8
4.7 - Ratio de profesor-alumno ( $x_6$ ) .....	9
4.8 - Gasto público en educación en porcentaje del PIB ( $x_7$ ).....	9
4.9 - Índice de Gini ( $x_8$ ) .....	9
5. ANÁLISIS DE LAS DIFERENCIAS SEGÚN GÉNERO, NACIONALIDAD Y CENTRO .....	10
6. DISCUSION ADICIONAL .....	15
7. APLICACIÓN DE LA METODOLOGÍA ECONOMETRICA Y RESULTADOS .....	20
6.1 – Estimación del modelo .....	20
6.2 - Análisis de normalidad .....	24
6.3 - Análisis de heteroscedasticidad .....	26
7.4 – Multicolinealidad .....	29
8. RESULTADOS .....	34
9. CONCLUSIONES.....	34

<b>REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>41</b>
<b>ANEXO.....</b>	<b>44</b>



## ÍNDICE DE TABLAS, FIGURAS Y GRÁFICOS

GRÁFICO 5.1. Diferencias en las puntuaciones según el género del estudiante.....	10
GRÁFICO 5.2. Diferencias en las puntuaciones según la nacionalidad del estudiante.....	12
GRÁFICO 5.3. Diferencias en las puntuaciones según la titularidad del centro educativo.....	14
GRÁFICO 6.1. Desempleo joven en la Unión Europea en porcentaje.....	16
GRÁFICO 6.2. Porcentaje de jóvenes infracualificados.....	17
TABLA 7.1.1. Salida de la estimación.....	22
GRÁFICO 7.2.1. Salida Test de Jarque-Bera.....	22
GRÁFICO 7.2.2. Análisis Q-Q.....	26
GRÁFICO 7.3.1. Representación homocedasticidad y heteroscedasticidad.....	27
TABLA 7.3.1. Contraste de White.....	28
TABLA 7.4.1. Matriz de correlaciones.....	30
TABLA 7.4.2. Salida FIV.....	31
TABLA 7.4.3. Regresión de $x_{8i}$ respecto al resto de variables independientes...	31
TABLA 7.4.4. Regresión de $x_{6i}$ respecto al resto de variables independientes...	32
TABLA 7.4.5. Regresión de $x_{4i}$ respecto al resto de variables independientes...	33



# 1. INTRODUCCIÓN

El ámbito de la educación afecta en muchos aspectos a la economía, tanto desde un enfoque macroeconómico como desde uno microeconómico. A pesar de la existente relación, en los modelos contemporáneos muchas veces se olvidan los aspectos educativos o se reducen a una variable como, por ejemplo, el capital humano; la mayoría de las veces relegándola a un lugar inferior, ocupando el nivel superior todas aquellas variables puramente económicas.

Hoy en día, cabe cuestionarse: ¿se puede explicar mejor el desarrollo de un país mediante el grado de avance de su sistema educativo nacional o deberíamos fijar criterios puramente económicos? ¿Se podrían aplicar criterios mixtos, en los que los criterios económicos se completen con los criterios educativos? ¿Existen contradicciones entre ambos criterios y, de ser así, podemos paliarlos mediante algún otro instrumento?

Todas esas dudas se han convertido en debates que han enriquecido la literatura económica y que nos permiten, a día de hoy, seguir trabajando en modelos que mejoren la educación y, con ello, la realidad económica del país.

Para empezar, se considera que el inicio de los estudios sobre cómo influye el sistema educativo en la sociedad data de 1911, cuando Émile Durkheim escribe un artículo llamado Educación y Sociología. A partir de ahí, se empieza a desarrollar lo que conocemos como Sociología de la Educación, cuyo interés central es la Educación como instrumento analítico para comprender las relaciones sociales.

Durkheim formula en Educación y Sociología la siguiente definición de la educación: “La educación es la acción ejercida por las generaciones adultas sobre aquellas que no han alcanzado todavía el grado de madurez necesario para la vida social. Tiene por objeto el suscitar y desarrollar en el niño un cierto número de estados físicos, intelectuales y morales que exigen de él tanto la sociedad política en su conjunto como el medio ambiente específico al que está especialmente destinado.” (Émile Durkheim, 1911, pp. 4).

En cambio, la RAE define educación, en una de sus acepciones, como: “crianza, enseñanza y doctrina que se da a los niños y a los jóvenes” (Real Academia Española, s.f., definición 2). Con esta acepción podemos entender que la educación no enseña a los adolescentes a desarrollarse, a comprender y poner en práctica lo enseñado, sino a estudiar aquellos conocimientos que han sido heredados generación tras generación.

A primera vista, esto puede ser un asunto baladí, pero si nos quedamos con esta acepción de la RAE se crea un paradigma que definieron los lingüistas Benjamin Whorf y Edward Sapir y se conoce como la hipótesis de Sapir-Whorf (Hojer, 1954).

Esta hipótesis supone que el lenguaje determina la forma de pensar y, por tanto, la realidad que percibimos, es decir, cómo definimos lo que nos rodea influye en la percepción que tenemos de ello.

Dependiendo si se toma la acepción que nos proporciona la RAE o la dada por Durkheim, los estudios sobre el sistema educativo deberán reflejar, en el primer caso, el alcance de unos conocimientos y, en el segundo caso, la adquisición de unas competencias a través de los conocimientos impartidos.

Siguiendo la acepción que nos proporciona la RAE, la educación se basaría en memorizar para adquirir conocimiento que nos ha sido enseñado; además este dilema se acrecienta con el aumento de la tecnología que nos acerca el conocimiento de forma gratuita, directa e inmediata.

En cambio, siguiendo la definición de Durkheim, se entiende educación como el motor clave para el desarrollo de un niño. Además de ser más completa esta definición, es más completa a la hora de estudiar las problemáticas de los sistemas educativos en la actualidad.

Optando por esta última definición, utilizaré el informe PISA para valorar la educación ya que, como explico en el siguiente punto, en el cual definiré su estructura y metodología, no mide la adquisición de conocimiento, sino la capacidad de aplicar dicho conocimiento a los complejos lazos e interacciones de la actualidad.

## 2. CONSIDERACIONES PREVIAS DEL INFORME PISA

El Informe del Programa internacional para la Evaluación de Estudiantes o Informe PISA, por sus siglas en inglés: *Programme for International Student Assessment*, desde ahora en adelante PISA, es un proyecto de la Organización para la Cooperación y el Desarrollo Económico (OCDE) con el fin de ofrecer numerosa y detallada información, así en el futuro poder adoptar aquellas políticas públicas más adecuadas y necesarias para mejorar la educación.

Se elabora mediante cuestionarios realizados a los estudiantes de 15 años, indiferentemente de si el alumno ha repetido o no, es decir, el alumno puede hallarse cursando tanto cuarto de la ESO (educación obligatoria en el sistema educativo español) como en segundo de la ESO, no habiendo superado algún curso en este último caso; de esta forma, pueden medir las diferencias entre los primeros y los segundos.

Tanto para los alumnos de escuelas privadas como para los alumnos de escuelas públicas el intervalo se define entre 4.500 y 10.000 estudiantes por país. Sin embargo, como es el caso de España, si un estado lo desea puede pedir aumentar la muestra para poder hacer análisis regionales.

Un error frecuente en los análisis sobre el Informe PISA es pensar que éste trata de medir si el estudiantado ha asimilado o no contenidos de ciertas asignaturas o áreas de estudio. El objetivo de la prueba trata de medir qué son capaces de hacer los estudiantes con lo aprendido durante la etapa anterior a iniciar los estudios superiores o su inserción laboral. El diseño de PISA no busca ni evaluar el aprendizaje de los contenidos específicos ni evaluar el desempeño de los docentes.

Los alumnos hacen tres pruebas con resultados independientes: competencia lectora, matemática y científica. De estas competencias, el modelo sólo refleja las competencias científica y matemática. Esta decisión ha venido motivada porque el Informe PISA cada año se centra en una competencia, en el 2018 fue la comprensión lectora, con énfasis en el entorno digital. La competencia en la que se centra cada año tiene más preguntas y es más detallada en la evaluación; lo cual se traduce en que las otras dos competencias,

este año la matemática y la científica, tienen similar esquema y un número de preguntas similares, y, además, la OCDE publica los datos de la competencia central después que los de las otras dos competencias.

Los datos recogidos en el modelo no están para todos los países que realizan los cuestionarios. El informe PISA diferencia dos grupos de países a la hora de catalogar y analizar los datos: los miembros de la OCDE, países que forman parte de la organización; y los *partners*, aquellos países que no forman parte de la organización, pero mantienen relaciones con ella y realizan la evaluación, como puede ser China o Bulgaria. De este último grupo, hay una carencia de datos para muchas variables, que explicaré en el próximo apartado.

### **3. OBJETIVOS PRINCIPALES Y PARTICULARES**

El objetivo principal de este trabajo es estudiar las implicaciones del entorno socioeconómico y cultural en las puntuaciones de los estudiantes utilizando el Informe Pisa como base de datos.

Para la consecución del objetivo principal, voy a seguir una serie de pasos con objetivos particulares:

1. La elaboración de un modelo econométrico que relacione el entorno socioeconómico y el rendimiento académico. Este modelo tiene como objetivo explicar la relación entre las variables endógena y exógenas, expuestas en el punto 4. Para poder plantear el modelo correctamente es necesario hacer una breve y concisa introducción a las hipótesis clásicas que restringirán al modelo y sus estimaciones.
2. Tras la especificación del modelo, el siguiente paso será la constatación y verificación de las hipótesis clásicas. El estudio econométrico se hará mediante el software Econometrics Views (Eviews). Con la realización de los cálculos pertinentes se buscarán aquellos estimadores que cumplan la mayor cantidad de propiedades.
3. También se enunciarán las posibles soluciones a aquellos problemas que hayan surgido durante la aplicación de la metodología econométrica.
4. El último paso será exponer los resultados hallados con el estudio del modelo y, en caso de que hubiere algún error, corregirlo.

Este camino hacia el estudio del entorno de los estudiantes terminará con las conclusiones extraídas de todo el proceso.

#### **4. VARIABLES ELEGIDAS**

El modelo econométrico que se va a plantear pretende analizar cómo influye la realidad cultural y económica en el rendimiento académico.

En este apartado voy a definir las variables, tanto la variable dependiente como las independientes, además de exponer las razones de la elección de dichas variables junto a sus alternativas.

Primero, voy a introducir las variables:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_7 x_{7i} + \beta_8 x_{8i} + \varepsilon_i$$

$$\text{Para } i = 1 \dots 54$$

Siendo el subíndice  $i$  la referencia al país:

$y_i$  = puntuaciones medias en matemáticas y ciencias

$x_{1i}$  = índice de miedo al fracaso en las mujeres

$x_{2i}$  = porcentaje de alumnos inmigrantes que esperan acabar sus estudios

$x_{3i}$  = porcentaje de colegios públicos

$x_{4i}$  = índice socioeconómico y cultural

$x_{5i}$  = porcentaje de bullying

$x_{6i}$  = ratio de profesor-alumno

$x_{7i}$  = gasto público en porcentaje del PIB

$x_{8i}$  = índice de Gini

En forma matricial, el modelo quedaría de forma tal que:

$$Y = X\beta + \varepsilon$$

##### **4.1 - Puntuaciones medias en matemáticas y ciencias**

Como medida del rendimiento académico, que constituye la variable endógena del modelo, he utilizado las medias de las puntuaciones en matemáticas y ciencias

Dicha variable va a venir explicada por las variables que voy a exponer a continuación.

#### **4.2 - Índice de miedo al fracaso en las mujeres ( $x_1$ )**

La primera variable considerada,  $x_{1i}$ , es el índice de miedo al fracaso femenino donde valores más altos del índice indican mayor miedo al fracaso. Esta variable trata de explicar cómo las mujeres prevén su entrada en el mercado laboral tras la finalización de sus estudios (véase gráfica 1 del anexo).

En este caso, se puede observar cómo los países más desarrollados económicamente y más igualitarios tienen mejores puntuaciones; sin embargo, hay ciertos casos como Arabia Saudí que podemos observar como el índice es menor mientras que los EE. UU. tiene una puntuación mucho más alta.

Estas disparidades se deben a la socialización de las adolescentes, sobre todo en la edad escolar. Por ello, este índice hay que entenderlo, cuidadosamente, desde una perspectiva de género.

Con este índice se busca modelizar el comportamiento femenino y estudiar si influye en las puntuaciones. Como observamos en las salidas del *Eviews*, se cometería un error de omisión de variable en el caso de eliminar  $x_{1t}$ .

#### **4.3 - Porcentaje de alumnos inmigrantes que esperan acabar sus estudios ( $x_2$ )**

El reto de modelizar la inmigración para ver cómo influye en las puntuaciones reside en la carencia de datos de muchos países sobre los inmigrantes, en especial entre los países subdesarrollados. Esto se debe principalmente a un abandono temprano del sistema educativo y la falta de recursos de los países menos desarrollados.

Para poder representar la inmigración en el modelo, se planteó primero el “porcentaje de alumnos inmigrantes” como variable y no resultó significativa. Después, se introdujo la variable “porcentaje de alumnos inmigrantes que esperan acabar sus estudios” (véase gráfica 2 del anexo) que, a pesar de seguir

sin ser significativa, ahora mostraba un coeficiente positivo, siendo negativo el de la primera variable.

Según un estudio de Portes, Aparicio, Haller y Vickstrom (2011) y Portes, Vickstrom, Haller y Aparicio (2013) son destacables las aspiraciones de los adolescentes como variable relevante a la hora de explicar el rendimiento académico.

Aparicio y Portes (2014) exponen que las familias inmigrantes que se centran en el futuro de sus hijos e hijas les trasladan unas altas expectativas. Sin embargo, el trabajo de Suarez-Orozco (2003) defiende que las expectativas vienen dadas por el nivel de clase social.

#### **4.4 - Porcentaje de colegios públicos ( $x_3$ )**

Como he mencionado al inicio, la proporción entre alumnos de escuelas públicas y privadas es paritaria. Sin embargo, el porcentaje de escuelas públicas y privadas o concertadas no tiene por qué serlo.

Según Mancebón y Pérez-Ximénez (2007) las diferencias (véase gráfico 3 en el anexo) dadas entre las puntuaciones de los colegios públicos y privados se deben a la selección de alumnos que hacen los centros privados.

También según Mancebón y Pérez-Ximénez (*op. cit.*), los alumnos de las escuelas privados y concertadas tienen mayores niveles de autoconfianza, ambición académica y respaldo económico familiar.

La selección de alumnos y las expectativas o nivel de autoconfianza del propio alumno, como se ha mencionado en el punto anterior, están relacionadas con el porcentaje de colegios públicos, debido a que una selección de alumnos suele dar lugar alumnado con mayor nivel de confianza. En este caso es más relevante las ambiciones y expectativas que otras variables, como sucede en el caso de la inmigración.

Estudios más antiguos (véanse Jimenez *et al.*, 1991, Williams and Carpenter, 1991, Witte, 1992, o Figlio y Stone, 1997) también siguen la línea de que la mejora del rendimiento académico en los centros privados respecto a los públicos se debe a una segregación socioeconómica en los primeros.

Por todas estas razones, se considera relevante la variable porcentaje de escuelas públicas en el modelo.

#### **4.5- ISEC: índice socioeconómico y cultural ( $x_4$ )**

El Índice socioeconómico y cultural, de ahora en adelante, ISEC (en inglés ESCS, *Economic, Social and Cultural Status*) es una variable que mide el contexto social y familiar del alumnado y se elabora a partir de los siguientes datos:

- El nivel de estudios de la madre y el padre: se toma como valor sólo el más alto de los dos.
- La profesión de la madre y del padre: se asigna el valor de mayor nivel de ambos.
- Nivel de recursos doméstico: por ejemplo, disponer de ordenador, escritorio o mesa para uso personal, libros propios, habitación propia, conexión a Internet (véase gráfica 4 del anexo), etc.

Es una variable continua tipificada para cada país, con media 0 y desviación típica 1. Para su obtención, se lleva a cabo la media ponderada de los ISEC de los alumnos del país en cuestión.

#### **4.6 - Porcentaje de *bullying* ( $x_5$ )**

Vamos a comenzar con la definición que da la RAE del acoso escolar (o en inglés, *bullying*): “en centros de enseñanza, acoso que uno o varios alumnos ejercen sobre otro con el fin de denigrarlo y vejearlo ante los demás”.

Quizás, *a priori*, la relación entre el *bullying* y el rendimiento académico no es evidente; por ello, en este apartado vamos a ver cómo afecta el acoso escolar y la diferencia entre aquellos países con mayores porcentajes y con menores porcentajes (véase gráfica 4 del anexo).

Según un artículo (Florencia Gerenni y Leila Fridman, 2015, pp. 71-82): “Esta relación entre *bullying* y rendimiento académico se debe a que el primero obstruye la capacidad de respuesta de los adolescentes que son víctima a las metas educativas propuestas, entre otras cosas por el ausentismo que comienzan a manifestar por temor al maltrato”.

Otro estudio (Navarro, Miguel y Germes, 2002) constata que el acoso escolar influye en la percepción de uno mismo, variable que viene recogida en la base de datos del Informe Pisa. Esto se debe a la socialización en el centro escolar en una época muy importante a la hora de forjar la autoestima y la personalidad. Esto lleva a peor percepción de uno mismo, menores ambiciones, rechazo al colegio y absentismo escolar. Esta misma idea es la que viene recogida en el estudio de Vickstrom, Haller y Aparicio (2013).

Hay que tener en cuenta también que los países con más recursos frente al acoso escolar (como el modelo KiVa, en Finlandia) tienen porcentajes menores de *bullying* y mayores puntuaciones en el Informe PISA. Los países que dedican menos recursos tienen un nivel de desarrollo menor y también menores puntuaciones.

#### **4.7 - Ratio de profesor-alumno ( $x_6$ )**

La ratio profesor-alumnos es una variable muy utilizada a la hora de estudiar las diferencias en el rendimiento académico entre países. Sin embargo, cuando introducimos la variable se observa que no es significativa y si la eliminamos no caemos en un error de omisión de variable.

#### **4.8 - Gasto público en educación en porcentaje del PIB ( $x_7$ )**

Esta variable, al contrario que el resto, es la única que proviene directamente de la base de datos del Banco Mundial, ya que a base de datos de la OCDE carecía de la mayoría de los datos para los países *partners*. Sin embargo, no todos los países tienen datos coincidentes en año con PISA, por lo que para aquellos que no coincidían, he escogido el dato más cercano.

El gasto público en educación es una variable muy interesante a la hora de representar los recursos que un país dedica a la educación, tanto pública como privada.

#### **4.9 - Índice de Gini ( $x_8$ )**

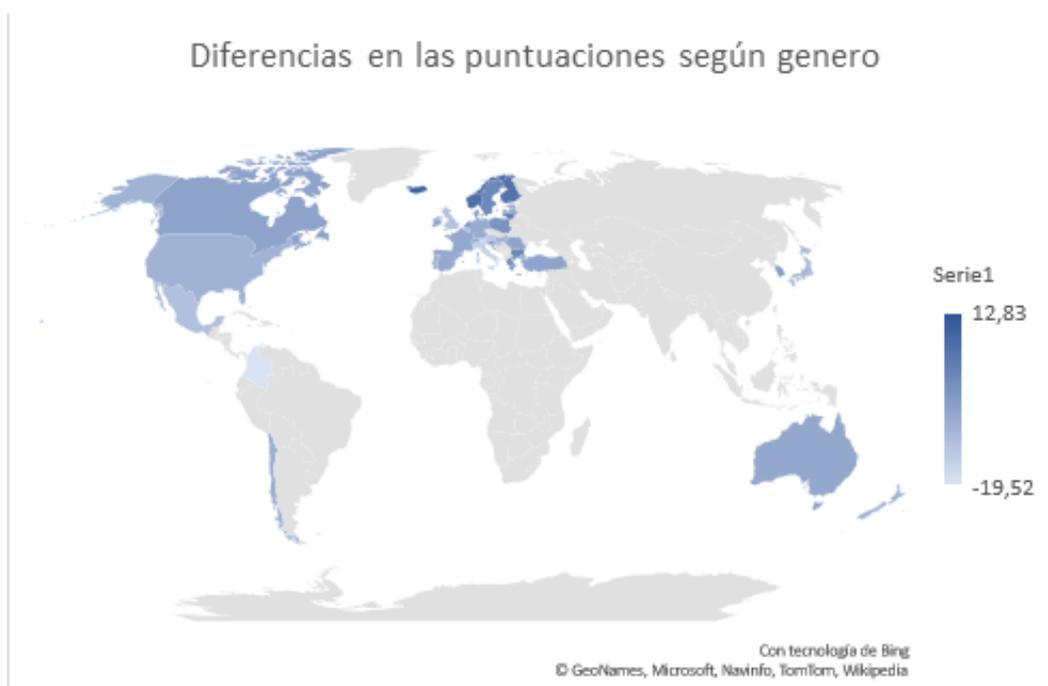
El Índice de Gini es una variable que mide la desigualdad de un país, que está entre 0 y 1, siendo 0 la perfecta igualdad y 1 la perfecta desigualdad. Al igual que el gasto público en educación es una variable muy generalizada.

Normalmente es una variable utilizada para medir la desigualdad en los ingresos, sin embargo, en el informe PISA se utiliza como medida de desigualdad en general. Cabe destacar que una mayor desigualdad, tanto en los niveles de ingresos como en general, produce mayores diferencias a la hora de acceder al sistema educativo y a los recursos que las propias familias invierten en sus hijos e hijas.

## 5. ANÁLISIS DE LAS DIFERENCIAS SEGÚN GÉNERO, NACIONALIDAD Y CENTRO

En este apartado, examino las diferencias en las puntuaciones según el género y la nacionalidad del estudiante y la titularidad del centro de una selección de países. La específica selección de países se debe a cuestiones de recolección y publicación de los datos, ya que los países restantes que sí se utilizan en el modelo no tienen publicados estos datos.

Gráfico 5.1. *Diferencias en las puntuaciones según el género del estudiante.*



Fuente: Elaboración propia a través de datos de la OCDE.

Para empezar a hablar sobre las diferencias según el género, existen diversos estudios que muestran que el género tiene incidencias en el rendimiento en matemáticas del alumnado (Rodríguez, Delgado y Bakieva, 2011; Rodríguez-Mantilla, Fernández-Díaz y Jover, 2018). En cuanto a ciencias,

las diferencias no son tan significativas (Cercadillo, Lis, 2015). A pesar de que en este trabajo no se mencionan las puntuaciones en lectura, cabe destacar que las mujeres logran un mejor rendimiento en comprensión lectora como norma.

En el Gráfico 5.1. *Diferencias en las puntuaciones según el género del estudiante* los niveles positivos de las diferencias son aquellos países donde las mujeres alcanzan mejores puntuaciones mientras que cuando la diferencia toma puntos negativos los hombres son los que obtienen mejores resultados.

Estas diferencias vienen explicadas por las diferentes pautas de socialización entre hombres y mujeres. Estos patrones tienen la excepción de los países del norte de Europa, donde se presentan mayores niveles de equidad. Igualmente, hay que mencionar que estas diferencias no son tan considerables como las diferencias según la nacionalidad del estudiante y el tipo de centro, que estudiaremos a continuación.

Uno de los estudios más reveladores de los últimos años (Balart, P., Oosterveen, M., 2019) revela que la duración de las pruebas no afecta al rendimiento de los estudiantes. Barlat defiende que las mujeres mantienen de una forma más adecuada su rendimiento durante la prueba, independientemente de la ventaja o desventaja que puedan tener en un principio en la materia que se evalúa. Por ejemplo, el estudio señala que en la mayoría de los países donde las alumnas tenían una desventaja inicial en matemáticas y ciencias, esta desventaja disminuyó a la mitad después de llevar dos horas de examen.

Otra de las explicaciones sobre la diferencia de las puntuaciones es que, a los chicos, desde pequeños se les socializa en juegos de lógica mientras que a las chicas se les socializa en juegos de cuidados, y estas experiencias permiten a los niños crear una serie de estrategias mentales de cara a la resolución de problemas. Estos estereotipos se transmiten sutilmente mediante la socialización de hombres y mujeres.

En el estudio de Barnat y Oosterveen (2019), no se ha proporcionado una explicación definitiva sobre la causa de estos sesgos de género a la hora de mantener el rendimiento, a pesar de haber descartado algunas posibilidades como, por ejemplo, que chicos y chicas sigan distintas estrategias para resolver el examen o que distribuyan de forma distinta el tiempo disponible.

Sin embargo, el factor más significativo a la hora de explicar las diferencias se remite a la falta de confianza en las chicas respecto a asignaturas como matemáticas, incluso entre las chicas que obtienen mejores puntuaciones (OCDE, 2015).

En este punto del trabajo cabría considerar el porqué de que las chicas de 15 años puedan tener estos sesgos en la percepción de ellas mismas y falta de confianza. La literatura ofrece muchas respuestas y muy variadas, sin embargo, los autores coinciden en que dichos sesgos suelen estar vinculados a determinados patrones culturales y valores familiares que existen en la sociedad actual y que los datos del Informe Pisa hacen explícitos.

Gráfico 5.2. *Diferencias en las puntuaciones según la nacionalidad del estudiante.*



Fuente: Elaboración propia a través de datos de la OCDE.

Para continuar, en el Gráfico 5.2. *Diferencias en las puntuaciones según la nacionalidad* las diferencias positivas vienen dadas cuando los nativos obtienen mejores puntuaciones, mientras que las diferencias negativas ocurren en el caso de que los nativos obtengan menores puntuaciones. En este gráfico cabe destacar el caso de Corea del Sur con una diferencia de 150,640121 puntos a favor de los nativos, mientras que la última diferencia negativa estaría en Australia con sólo un nivel de -18,138153 puntos (véase tabla 1 del anexo).

La dificultad que existe en torno a las investigaciones sobre inmigración reside en las disparidades entre los inmigrantes de los diferentes países estudiados en este punto. La lógica nos lleva a pensar que no existen muchos puntos en común entre la inmigración en países asiáticos como Japón o Corea del Sur y países como Canadá, pudiéndose considerar ambos países desarrollados en términos educativos.

Incluso en países en los que podríamos vislumbrar cierta cercanía, como en los europeos, la realidad de la inmigración no es tan cercana a la hora de estudiarla. Igualmente, en los países del norte de Europa, conocidos por sus notables niveles de igualdad en todos los aspectos, se puede observar destacables diferencias con puntuaciones alrededor de los 60.

Continuando en Europa, en los países del mediterráneo es donde se puede dar una mayor semejanza en los niveles de diferencia. Hay varios razonamientos válidos para este acontecimiento, entre ellos, la similitud entre el nivel económico de las familias inmigrantes, la procedencia del inmigrante y el motivo de migración.

Por ello, para poder acercarse a un enfoque más idóneo, se debería hacer un estudio sobre la procedencia de la inmigración para poder realizar una primera comparación entre aquellos países en los que los inmigrantes proceden de países menos desarrollados; y, una segunda, entre aquellos países en los que los inmigrantes provienen de países igual o más desarrollados.

Otra circunstancia que puede afectar a las diferencias es el porcentaje de estudiantes inmigrantes en las escuelas. Pasando del 1% de Corea del Sur y Turquía hasta un 12% en Noruega o un 20% en Suecia (véase tabla 1 del anexo).

De hecho, si lo comparamos con las diferencias según género, la varianza es mucho menor en este caso, es decir, las diferencias no son tan notables según el género.

Estas particularidades de la inmigración conllevan que las investigaciones y hallazgos se enfoquen sólo en determinados países o regiones. A la hora de

estudiar los antecedentes en inmigración, hay que tener una mirada muy completa sobre estas implicaciones.

Gráfico 5.3. *Diferencias en las puntuaciones según la titularidad del centro educativo.*



Fuente: Elaboración propia a través de datos de la OCDE.

Por último, en el Gráfico 5.3. *Diferencias en las puntuaciones según la titularidad del centro*, las diferencias positivas se dan cuando los resultados en las escuelas públicas están por encima de los resultados en las escuelas privadas.

A pesar de que en la mayoría de los países se observan mejores puntuaciones en los colegios privados o concertados, cuando existe igualdad de condiciones de renta, los colegios públicos obtienen mejores puntuaciones respecto a los privados. El rendimiento en las escuelas públicas mejora respecto a las escuelas privadas cuando el efecto renta sobre los resultados académicos se elimina.

Esto se puede observar, como hemos visto en el punto 4.4, en el estudio de Mancebón y Pérez-Ximénez (2007), ya que las diferencias dadas entre las puntuaciones de los colegios públicos y privados se deben a la selección de alumnos que hacen los centros privados.

## 6. DISCUSION ADICIONAL

Un aspecto a tener en cuenta es que el Informe Pisa, a pesar de ser bastante representativo para entender ciertas cuestiones educativas muy interesantes, olvida otros aspectos muy importantes a la hora de desarrollar los sistemas educativos nacionales.

De hecho, hay autores en España que consideran que es muy difícil entender las disparidades y peculiaridades geográficas españolas desde dicho informe. Una de estas particulares es la densidad de población española frente a la del resto de Europa.

España es un caso muy particular, con una densidad de población relativamente baja, 93 habitantes por kilómetro cuadrado, existen zonas totalmente deshabitado o con muy poca población, pero también existen otras zonas con una alta densidad de población, 215 habitantes por kilómetro cuadrado. Por ello, cuando medimos la densidad de población en zonas habitadas, la densidad aumenta drásticamente.

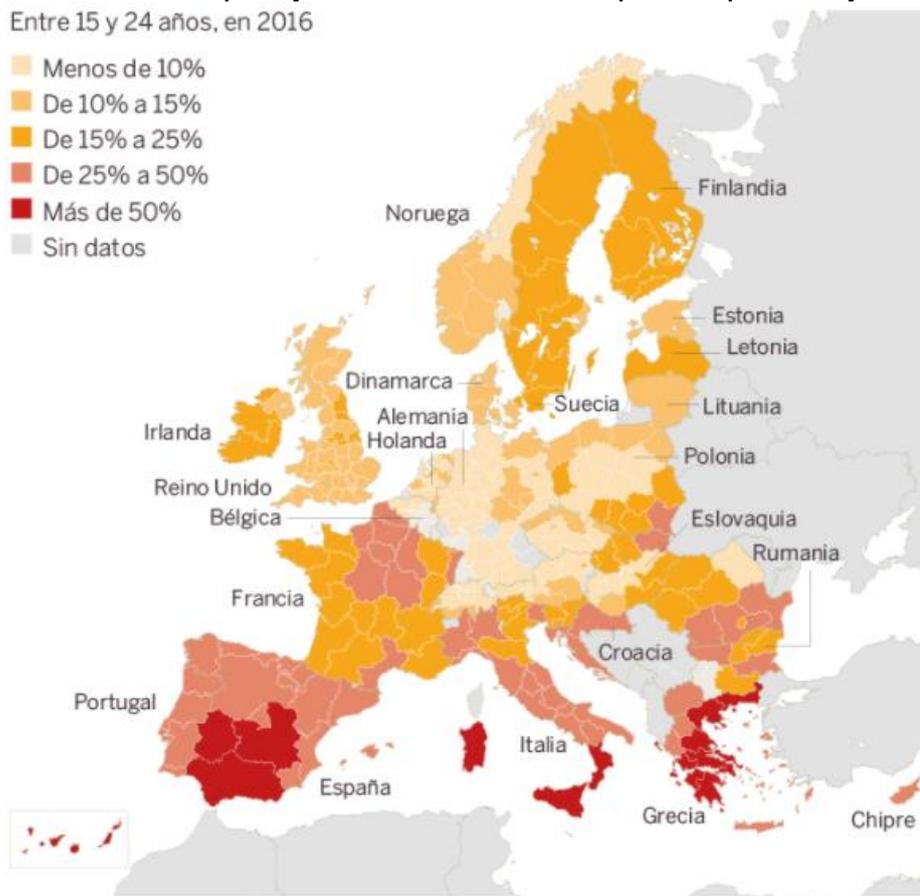
Este análisis es muy importante en el momento de desarrollar o reformar un sistema educativo, ya que no es lo mismo construir escuelas en París que en Copenhague. De hecho, las escuelas de Castilla y León, Navarra o Galicia se puedan parecer más a las escuelas de Noruega que a las del resto de Comunidades Autónomas de España.

Las puntuaciones en Castilla y León y Galicia, zonas mayormente rurales, (véase Gráfica 6 en el anexo) se acercan más a los países nórdicos que a otras zonas de España. Entre las explicaciones que se dan están el nivel de implicación de los padres en las escuelas o la toma de medidas rápidas a problemas concretos en consenso.

Otra particularidad es el modelo productivo de las naciones. Cabe cuestionarse si en la actualidad el modelo productivo actual demanda obra de mano cualificada.

Gráfico 6.1. Desempleo joven en la Unión Europea en porcentaje.

Entre 15 y 24 años, en 2016



Fuente: Eurostat.

Como se puede observar en el gráfico, los países del sur de Europa, con sistemas productivos asociados al turismo, tienen índices de desempleo más altos. Este sector de la economía no precisa de mano de obra cualificada, mientras que los países del norte de Europa tienen porcentajes menores.

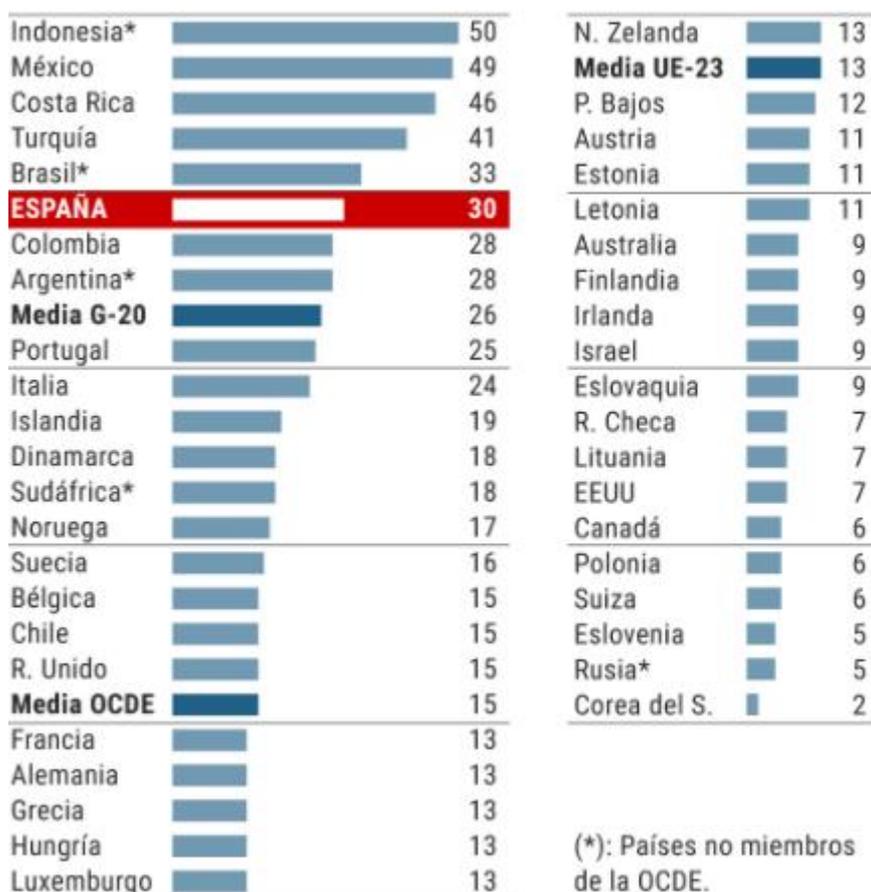
Esto indica un grave problema sobre el sistema productivo, pero también sobre el sistema educativo. Aquí existen dos objetos claves:

1. Los jóvenes infracualificados: por un lado, el 30% de los jóvenes entre 25 y 34 años en España no tienen título de bachillerato o estudios equivalente, véase el Gráfico 6.2. Un sistema educativo fracasa cuando no es capaz de formar a las personas para su entrada en el mercado laboral.

Gráfico 6.2. Porcentaje de jóvenes infracualificados.

### ESPAÑA, LÍDER EN JÓVENES INFRACUALIFICADOS

% de población de entre 25 y 35 años de edad que no tiene estudios de Bachillerato o equivalente de FP



(\*): Países no miembros de la OCDE.

FUENTE: OCDE.

A.H. | EL MUNDO GRÁFICOS

Fuente: OCDE a través de El Mundo.

- Los jóvenes sobrecualificados: por otro lado, el empleo precario y la falta de trabajo provoca que muchos jóvenes trabajen en empleos que necesitan menos formación. A este problema se le añade el fenómeno conocido como “fuga de cerebros”, jóvenes muy cualificados que emigran a países con mejores condiciones de vida y mejores sueldos porque en su país de origen no encuentran trabajo estable acorde a su formación. Este problema es bastante grave, ya que mientras el Estado se encarga de invertir en la formación de estos jóvenes, otros países recogen los beneficios del talento.

Ahora bien, una de las consecuencias de estas dos realidades es la existencia de un grupo social llamado *ninis*, haciendo referencia a los jóvenes que ni estudian ni trabajan. Son aquellos profesionales que no estudian porque ya alcanzaron un nivel alto de formación y no encuentran trabajo o aquellos que se han visto expulsados por el sistema educativo y no son capaces de entrar en el mercado laboral por falta de competencias

En cuanto a la movilidad social, hoy en día es punto clave a la hora de analizar los sistemas educativos. Se entiende como ascensor social la facilidad o dificultad para ascender de clase social perteneciendo a una clase social más baja. Dentro de la movilidad social, los inmigrantes tienen mayores dificultades (Garreta J.,1994),

Mientras que Becker y Tomes (1986) sostienen que “casi todas las ventajas o desventajas de riquezas de los antepasados son eliminadas en tres generaciones”. En cambio, en 2016, Barone y Mocetti analizaron las familias más ricas de Florencia en 1427 y en 2011 y concluyeron que los apellidos de las familias más ricas lo seguían siendo 600 años después.

Además, el ascensor social no sólo no funciona hacia arriba, sino que además no funciona hacia abajo (Requena, 2016). Los hijos de clases altas es muy difícil que bajen de escala social, aunque tengan menores capacidades.

En Europa, la Ilustración supuso el desarrollo de los sistemas educativos con ideas de gratuidad, obligatoriedad y universalización. La expansión de la educación sirvió como herramienta para que las clases más bajas pudiesen acceder a mejores condiciones de vida. Hoy en día, esa cualidad de la educación se ha perdido, por lo que cabe reflexionar si la actual educación debe tomar valores diferentes.

La universalización de la educación, en la actualidad, no se puede considerar que haya concluido (véase gráfica 23 en el anexo). Se estima que existen casi un siglo de diferencias educativas entre las zonas más desarrolladas y las menos desarrolladas.

Uno de los últimos aspectos que me gustaría destacar es la importancia de la escuela pública en la atenuación de las desigualdades económicas. El informe

PISA, como hemos visto, muestra que las escuelas privadas obtienen mejores resultados que las escuelas públicas.

En los estudios posteriores, cuando se ha corregido el efecto renta de estas puntuaciones, se demostraba que los colegios públicos estaban por encima de aquellas. Es decir, la segregación social de las escuelas privadas tiene como consecuencia una mejora en las puntuaciones.

La mercantilización educativa en los últimos años se ha ido cuestionando y poniendo el foco en los países nórdicos como ejemplo para paliar este tipo de desigualdad. La baja cantidad de colegios públicos en los países escandinavos se le considera una de las causas de las puntuaciones en el Informe Pisa,

Para terminar, me gustaría destacar la cuestión de si la educación está realmente reconocida económica y socialmente. Este sector tan particular y esencial en la actualidad ha venido modificándose con nuevas teorías en busca de acercarse a la realidad político-social.

Tras la crisis económica de 2008, además de aumentar los niveles de desigualdad, se ha perdido una importante fe en el sector educativo. Con la pérdida de la movilidad social, las diferencias entre colegios de alta clase y de baja clase y las escasas inversiones en educación, la influencia de la educación ha ido disminuyendo poco a poco.

En conclusión, la educación no cumple las funciones sociales que le han sido atribuidas durante mucho tiempo y debería servir para mejorar las condiciones de vida y disminuir las desigualdades. Los Informes como el Pisa nos dan una ligera panorámica de cómo están funcionando el sistema educativo; sin embargo, hay que sopesar aquellos factores que inciden más allá de los meramente educativos,

Una nación debe ser capaz de paliar estos factores mediante un sistema educativo para conseguir que el elemento clave del rendimiento de los alumnos sean el esfuerzo individual y los recursos estatales.

## 7. APLICACIÓN DE LA METODOLOGÍA ECONOMÉTRICA Y RESULTADOS

En este punto, procederé a aplicar la metodología mediante EViews en el modelo previamente especificado en el punto 4 de este trabajo. El Modelo de Regresión Lineal Normal Clásico (MRLNC) trata de evaluar la relación causa-efecto que se produce en la variable dependiente a través de unas variables independientes. Esto también va a permitir llevar a cabo predicciones.

Hipótesis básicas:

1. Hipótesis de normalidad: Para cada uno de los países, la perturbación aleatoria, y también la variable dependiente, tiene una distribución estadística normal.

$$\varepsilon \sim N(0, \sigma_{\varepsilon}^2 I)$$

2. Cómo se recoge en la expresión anterior, la perturbación aleatoria tiene media nula.
3. Hipótesis de homoscedasticidad: La varianza de las perturbaciones es constante en todas las observaciones (países).
4. Hipótesis de incorrelación entre las perturbaciones.
5. Regresores no estocásticos: Si pudiéramos repetir el mismo experimento, los valores de las variables explicativas serían los mismos.
6.  $Rg(X)=k+1 < N$ , siendo  $k$  es el número de variables explicativas: si no hay multicolinealidad perfecta, no hay relaciones lineales entre los regresores exactas.
7. Linealidad en los parámetros: El modelo está correctamente especificado mediante su forma lineal, recogida en el apartado.

### 7.1 – Estimación del modelo

La estimación mediante Mínimos Cuadrados Ordinarios (MCO) es un método que consiste en buscar el valor de  $\beta$  que minimiza la suma de los cuadrados de los residuos  $e$ . Los residuos son las diferencias entre el valor observado y el esperado.

Matemáticamente, partiendo del modelo  $Y = X + \beta$  la función objetivo a minimizar:

$$\sum e_i^2 = e'e = (Y - X\hat{\beta})'(Y - X\hat{\beta})$$

Para resolver este problema, aplicaremos las condiciones de mínimo. Primero, derivaremos la función objetivo respecto al estimador que pretendemos hallar y lo igualaremos a 0.

Después comprobaremos la condición de segundo orden para mínimo, es decir que la segunda derivada es una matriz definida positiva<sup>1</sup>. De esta manera se obtiene el estimador MCO.

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Propiedades del estimador  $\hat{\beta}$  obtenido por MCO:

1. Lineal en Y: cada estimador es una función lineal de las variables aleatorias dependientes  $Y_1, Y_2, \dots, Y_n$ .
2. Insesgado: la media del estimador es igual al parámetro.
3. Óptimo: es el estimador de mínima varianza, entre todos los estimadores lineales en Y e insesgados
4. Eficiente: es el estimador de mínima varianza entre todos los estimadores insesgados (lineales y no lineales).
5. Consistente: converge en probabilidad hacia el verdadero valor del parámetro.

Por eso, decimos que el estimado por MCO es ELIO (Estimador Lineal, Insesgado y Óptimo).

---

<sup>1</sup> Condición de primer orden:

$$\frac{\partial e'e}{\partial \beta} = 0$$

Condición de segundo orden:

$$\frac{\partial^2 e'e}{\partial \beta' \partial \beta} \text{ matriz definida positiva}$$

Una vez especificado el modelo (ecuación e hipótesis clásicas) y recogidas las propiedades que, bajo esas hipótesis, tendrían los estimadores, procedo a estimar el modelo del apartado 4, obteniendo el siguiente resultado.

Tabla 7.1.1. *Salida de la estimación.*

Dependent Variable: Y  
 Method: Least Squares  
 Date: 06/29/20 Time: 21:03  
 Sample (adjusted): 1 53  
 Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	575.7894	36.89799	15.60490	0.0000
X1	60.82218	18.16584	3.348163	0.0017
X2	0.077856	0.279973	0.278086	0.7823
X3	0.466697	0.201299	2.318429	0.0252
X4	20.21510	8.578695	2.356430	0.0231
X5	-2.311093	0.454305	-5.087100	0.0000
X6	1.062140	0.957587	1.109184	0.2735
X7	-9.057034	2.471481	-3.664618	0.0007
X8	-2.110942	0.714589	-2.954064	0.0051

R-squared	0.807526	Mean dependent var	467.8586
Adjusted R-squared	0.771716	S.D. dependent var	48.63739
S.E. of regression	23.23847	Akaike info criterion	9.285606
Sum squared resid	23221.15	Schwarz criterion	9.623321
Log likelihood	-232.4258	Hannan-Quinn criter.	9.415078
F-statistic	22.55079	Durbin-Watson stat	1.685705
Prob(F-statistic)	0.000000		

Fuente: Elaboración propia a través de datos de la OCDE.

El modelo entonces quedaría como

$$\hat{y}_i = 575,7894 + 60,82218x_{1i} + 0,077856x_{2i} + 0,466697x_{3i} + 20,21510x_{4i} - 2,311093x_{5i} + 1,062140x_{6i} - 9,057034x_{7i} - 2,110942x_{8i}$$

Siendo el subíndice  $i$  la referencia al país, recordamos las variables:

*Para  $i = 1..54$*

$y_i$  = puntuaciones medias en matemáticas y ciencias

$x_{1i}$  = índice de miedo al fracaso en las mujeres

$x_{2i}$  = porcentaje de alumnos inmigrantes que esperan acabar sus estudios

$x_{3i}$  = porcentaje de colegios públicos

$x_{4i}$  = índice socioeconómico y cultural

$x_{5i}$  = porcentaje de bullying

$x_{6i}$  = ratio de profesor-alumno

$x_{7i}$  = gasto público en educación en porcentaje del PIB

$x_{8i}$  = índice de Gini

Como se observa en la Tabla 7.1.1. *Salida de la estimación* la variable  $x_{2i}$  no es significativa ya que tiene un p-valor muy alto. Por ello, voy a repetir la regresión eliminando la variable en cuestión para observar si con este paso se incurriría en un error por omisión de variable relevante.

En la Tabla 2 del anexo, *salida de la estimación sin  $x_{2i}$  (porcentaje de inmigrantes que esperan acabar sus estudios)*, se comprueba que si eliminamos la variable  $x_{2i}$  no se produce el error de omisión de variable. Sin embargo, hay que tener en cuenta que la variable es muy interesante cuando se tengan que analizar las conclusiones.

Como se indicó en el apartado 4.3, una de las variables que se consideró para modelizar la inmigración fue el porcentaje de alumnos inmigrantes. En esta salida,  $x_2$  es el porcentaje de inmigrantes.

En la tabla 3 del anexo, vemos que si quitamos esta variable el  $R^2$  disminuye levemente. Esta bajada del  $R^2$  sumado a el coeficiente negativo de la variable  $x_2$  ponía en cuestión la introducción de dicha variable (véase ecuación 2 del anexo).

Como el factor de la nacionalidad de los estudiantes es importante a la hora de analizar los resultados, busque una variable sustitutiva que representase la inmigración en el modelo.

Otra variable cuyo p-valor es bastante alto es  $x_{6i}$ , la ratio profesor-alumnos, así que también realice la estimación sin dicha variable para ver si es significativa o no.

Como se observa, al igual que sucede en la Tabla 4 de anexo, *salida de la estimación sin  $x_{6i}$  ratio profesor-alumnos*, se produce una pequeña reducción del  $R^2$  (véase ecuación 3 en el anexo)

Esta variable es muy relevante a la hora de comparar los países ya que existen diversos estudios, como veremos a la hora de analizar las conclusiones, que indican fuertes relaciones entre la ratio profesor-alumnos y los resultados de los alumnos. Por estas razones, la ratio profesor-alumno se mantendrá en el modelo.

## 7.2 - Análisis de normalidad

La hipótesis clásica de normalidad, como se ha mencionado, supone que los errores siguen una distribución normal.

Consecuencias del no cumplimiento de la normalidad:

- Los estimadores MCO cumplirán todas las propiedades, excepto la eficiencia, es decir, ya no serán de mínima varianza entre todos los insesgados, pero siguen siendo óptimos, o sea de mínima varianza entre los lineales e insesgados.
- Debido al anterior punto, se desconoce qué distribución exacta tienen los estimadores, pero su distribución asintótica sigue siendo Normal. Además, la inferencia basada en las distribuciones habituales, como la Normal, la t, o la F, deja de tener validez exacta; pero sí mantiene validez asintótica.

La prueba más generalizada para comprobar la normalidad es la prueba de Jarque-Bera, se basa en la búsqueda de las desviaciones de la curtosis y de la asimetría respecto a una distribución normal con la misma media y varianza.

### Test de normalidad de Jarque-Bera

Contraste de hipótesis:

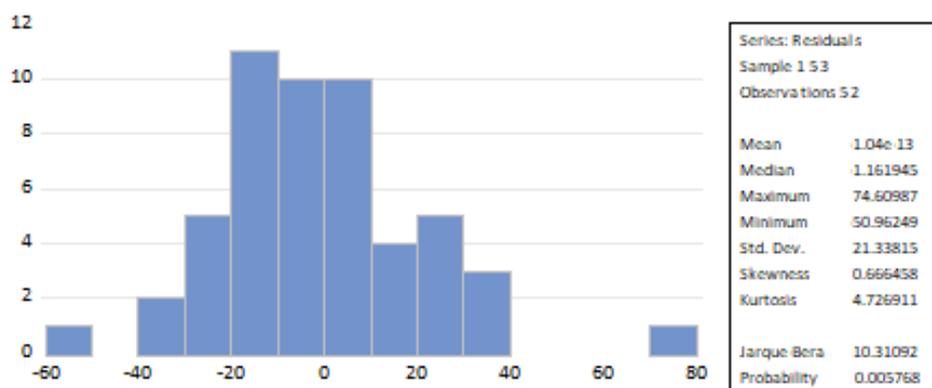
$H_0: \text{normalidad}$

$H_1: \text{no normalidad}$

Estadístico de Jarque-Bera:

$$d_{JB} = \frac{N - k - 1}{6} (g_1^2 + \frac{1}{4} (g_2 - 3)^2) \sim \chi_2^2$$

Gráfico 7.2.1. Salida Test de Jarque-Bera.



Fuente: Elaboración propia a través de datos de la OCDE.

Como tenemos un p-valor de 0,005768, se rechaza la hipótesis nula. Como tenemos una distribución no normal, analizaremos los valores de curtosis y de simetría.

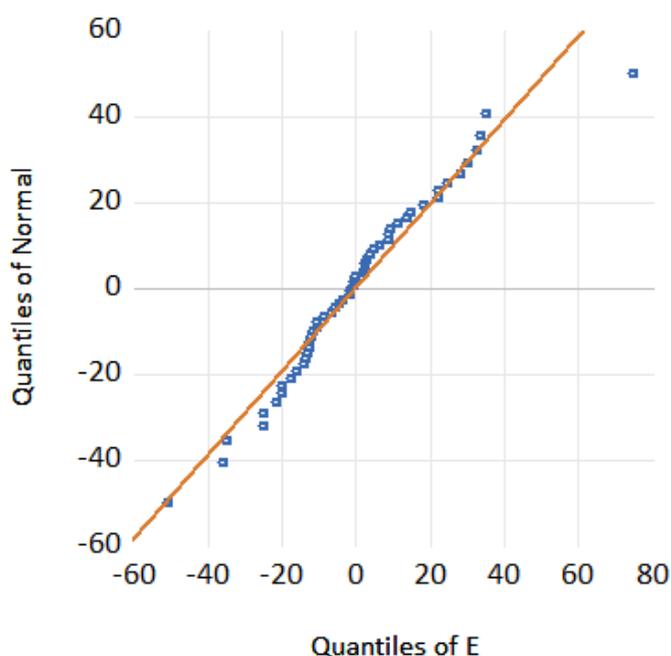
La asimetría, en inglés *skewness*, mide el grado de simetría de la distribución, es decir, si es totalmente simétrica. El coeficiente se calcula a partir del tercer momento respecto a la media, en este caso, es igual a 0,666458. Como es mayor que 0 nos encontramos ante asimetría por la derecha.

En cuanto a la curtosis, en inglés *kurtosis*, mide el grado de apuntamiento de la curva y se calcula a partir del cuarto momento respecto de la media. Con un valor de 4,726911 decimos que tiene una forma platicúrtica porque es mayor que 3.

También podemos observar la normalidad mediante la gráfica cuantil-cuantil (Q-Q) que permite comparar cómo es la distribución normal con la forma de la distribución de los residuos.

En el siguiente gráfico, podemos observar que, aunque se asemeje mucho a una normal, tiene una asimetría hacia la derecha, como nos había indicado el contraste de Jarque-Bera. Los puntos alejados son un indicio de datos extremos en el modelo. Por último, se puede observar una ligera curtosis ya que, cruza la línea de la normal varias veces.

Gráfico 7.2.2. *Análisis Q-Q.*



Fuente: Elaboración propia a través de datos de la OCDE.

### **7.3 - Análisis de heteroscedasticidad**

En econometría la heteroscedasticidad ocurre en un modelo de regresión lineal cuando las varianzas de las perturbaciones no coinciden. Esta ausencia de homocedasticidad supone el incumplimiento de una de las hipótesis básicas del modelo de regresión lineal.

La heteroscedasticidad suele ocurrir en dos casos:

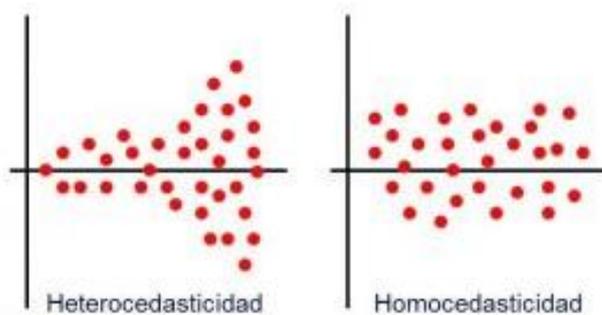
1. En datos de corte transversal, como es el caso de este trabajo, se presentan con mayor frecuencia problemas de heteroscedasticidad; por el contrario, las series de tiempo es menos probable que presenten heteroscedasticidad. La heteroscedasticidad en los modelos de corte transversal se debe a que los diferentes elementos, en este modelo los países, incluyen datos con distinta dispersión.
2. En datos agrupados como sumas o medias de grupos. Se puede demostrar que la varianza de cada perturbación depende del tamaño del grupo al que corresponde.
3. Errores en la especificación del modelo.

Consecuencias de la heteroscedasticidad:

- Con heteroscedasticidad, los estimadores son insesgados y consistentes, pero no de varianza mínima.
- La varianza de los estimadores no es la habitual y por ellos los contrastes habituales no son válidos.

Hay dos formas de examinar la heteroscedasticidad, mediante análisis gráficos y mediante contrastes de hipótesis. A continuación, podemos observar las diferencias gráficas entre la homocedasticidad (a la derecha del gráfico) y la heteroscedasticidad (a la izquierda del gráfico). En el eje de ordenadas se coloca la variable dependiente y en el eje de abscisas la variable explicativa.

Gráfico 7.3.1. *Representación homocedasticidad y heteroscedasticidad.*



Fuente: Google Imágenes.

Podemos analizar la heteroscedasticidad de las siguientes formas:

1. Análisis gráficos de los residuos frente a las variables independientes.

Existen dos diagramas de dispersión que nos permiten observar la heteroscedasticidad:

1. Diagrama de dispersión de las variables explicativas en el eje de abscisas frente a los residuos en el eje de ordenadas (véase gráficos del 7 al 14 del anexo).
2. Diagrama de dispersión de las variables explicativas en el eje de abscisas frente a los residuos al cuadrado en el eje de ordenadas (véase gráficos del 15 al 22 del anexo).

Como podemos observar en estos gráficos, en alguna variable no se observan claramente estructuras de homocedasticidad. Por ello, acudiremos a las pruebas estadísticas como forma de apoyo para constatar si hay o no heteroscedasticidad.

## 2. Contrastes de hipótesis

### **Contraste de White**

El contraste de White es el más general de todos los contrastes debido a que detecta la heteroscedasticidad en cualquiera de sus formas, es decir, no se precisa imponer restricciones sobre qué variable es la causa de la heteroscedasticidad.

Tras la estimación principal del modelo por MCO, se calcula una regresión de los residuos al cuadrado en función del término constante, de todas las variables independientes, sus cuadrados y los productos cruzados, excepto aquellos productos que se repitan.

Por último, se empleará como estadístico  $N$  multiplicado por el  $R$  cuadrado de la regresión auxiliar, que asintóticamente se distribuye como una Chi cuadrado con un número de grados de libertad igual al número de regresores de la regresión auxiliar. Se rechazará la hipótesis nula cuando los valores del estadístico sean grandes.

Este proceso, lo hace directamente EViews sin tener que realizar la regresión. Además, para los p-valor asociados al test de White que sean pequeños se rechazará la hipótesis nula.

$H_0$ : homocedasticidad

$H_1$ : heterocedasticidad

Tabla 7.3.1. *Contraste de White.*

Heteroskedasticity Test: White  
Null hypothesis: Homoskedasticity

F-statistic	1.890985	Prob. F(44,7)	0.1921
Obs*R-squared	47.96467	Prob. Chi-Square(44)	0.3152

Fuente: Elaboración propia a través de datos de la OCDE.

Como para valores pequeños del p-valor se rechaza la hipótesis nula, en nuestro caso, con un 0,3152 no rechazamos la hipótesis nula.

En el caso de haber hallado heteroscedasticidad, ésta se corregiría mediante:

- Depende del tipo de datos se pueden utilizar técnicas para arreglar la heteroscedasticidad, como puede ser tomar logaritmos en todas las variables del modelo original.
- Otra opción es estimar mediante Mínimos Cuadrados Ponderados (MCP).
- Por último, se puede estimar por MCO, pero estimando las varianzas mediante la propuesta de White.

#### **7.4– Multicolinealidad**

La multicolinealidad hace referencia a la existencia de relaciones lineales entre las variables explicativas del modelo. Hay dos tipos de multicolinealidad: multicolinealidad perfecta, relación total; y, multicolinealidad imperfecta, menos fuerte que la primera.

Para entender las consecuencias de la multicolinealidad, se debe separar entre multicolinealidad perfecta e imperfecta:

- Cuando existe multicolinealidad perfecta, no se cumple que el rango de la matriz  $X$  sea pleno, es decir,  $rg(X) < k + 1$ ; por esta razón, no se puede estimar por MCO.
- Cuando existe multicolinealidad imperfecta,
  - Aumenta la sensibilidad de los coeficientes estimados, lo cual dificulta la precisión e interpretación de los estimadores y sus resultados.
  - Los contrastes de significación individual no rechazarán la hipótesis nula, mientras que no haya problema al realizar contrastes conjuntos ya que, la multicolinealidad no afecta a  $R^2$ .

El primer paso para buscar la multicolinealidad es hallar la matriz de correlaciones en la que se muestra el grado de relación lineal entre los

regresores. Los valores altos de la matriz se traducen en una fuerte correlación, mientras que los valores cercanos a 0 indican que la correlación es débil.

Tabla 7.4.1. *Matriz de correlaciones.*

	x1	x2	x3	x4	x5	x6	x7	x8
x1	1	0,36844	-0,17939	0,33743	-0,03034	-0,00119	0,02344	-0,15977
x2	0,36844	1	-0,28235	0,16500	-0,24816	-0,01724	-0,05602	0,03662
x3	-0,17939	-0,28235	1	-0,12958	0,16705	-0,20085	0,07517	-0,05412
x4	0,33743	0,16500	-0,12958	1	-0,49587	-0,48053	-0,17780	-0,66478
x5	-0,03034	-0,24816	0,16705	-0,49587	1	0,33344	0,04356	0,43586
x6	-0,00119	-0,01724	-0,20085	-0,48053	0,33344	1	0,04007	0,57161
x7	0,02344	-0,05602	0,07517	-0,17780	0,04356	0,04007	1	0,10871
x8	-0,15977	0,03662	-0,05412	-0,66478	0,43586	0,57161	0,10871	1

Fuente: Elaboración propia a partir de datos de la OCDE.

Como se observa en la Tabla 6.4.1. *Matriz de correlaciones*, las correlaciones son bastante bajas en todas las variables excepto entre la  $x_6$  y la  $x_8$ , con un valor de 0,57161 y la  $x_4$  y la  $x_8$  con un valor de 0,66478. Para profundizar en el análisis de la multicolinealidad, hallaremos el Factor de Inflación de la Varianza (FIV) para todas las variables.

El FIV muestra la relación entre la varianza que se observa para un estimador y la que se obtendría en caso de que el correspondiente regresor fuese ortogonal al resto, es decir, estuviese totalmente incorrelacionado con el resto de los regresores. La fórmula es:

$$FIV_i = \frac{1}{1 - R_i^2}$$

Para calcularlo es preciso realizar una regresión con la variable que puede causar la multicolinealidad como variable dependiente frente al resto de variables independientes. Se toma el  $R^2$  de esta regresión y se halla el FIV. Cuanto más grande sea el FIV, mayor será la proporción de la varianza debida a la multicolinealidad.

EViews hace directamente todos los Factores de Inflación de la Varianza:

Tabla 7.4.2. *Salida FIV.*

Regresores	FIV
x1	1,43270
x2	1,38010
x3	1,22610
x4	2,45449
x5	1,58275
x6	1,70302
x7	1,06131
x8	2,27438

Fuente: Elaboración propia a partir de datos de la OCDE.

A continuación, indicamos la obtención de los 3 valores de FIV que hemos destacado en el análisis de la correlación para  $x_4$ ,  $x_6$  y  $x_8$ , que, en coherencia con dicho análisis, son los que poseen unos coeficientes con mayor FIV.

Tabla 7.4.3. *Regresión de  $x_{8i}$  respecto al resto de variables independientes.*

Dependent Variable: X8  
 Method: Least Squares  
 Date: 08/27/20 Time: 12:06  
 Sample (adjusted): 1 53  
 Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	20.33580	7.155184	2.842107	0.0068
X1	-2.446438	3.814628	-0.641331	0.5246
X2	0.086600	0.057605	1.503349	0.1399
X3	-0.018001	0.042381	-0.424747	0.6731
X4	-5.155458	1.634451	-3.154244	0.0029
X5	0.124492	0.093988	1.324543	0.1922
X6	0.457156	0.189901	2.407336	0.0203
X7	0.128663	0.521043	0.246933	0.8061
R-squared	0.560320	Mean dependent var		35.08077
Adjusted R-squared	0.490371	S.D. dependent var		6.867486
S.E. of regression	4.902581	Akaike info criterion		6.158039
Sum squared resid	1057.553	Schwarz criterion		6.458230
Log likelihood	-152.1090	Hannan-Quinn criter.		6.273125
F-statistic	8.010407	Durbin-Watson stat		1.594709
Prob(F-statistic)	0.000003			

Fuente: Elaboración propia a partir de datos de la OCDE.

Cálculo del FIV:

$$FIV = \frac{1}{1 - 0,560320} = 2,274381368$$

La varianza sería 2,2743 veces la que tendría en caso de ortogonalidad. A partir de valores por encima de 10 se puede considerar la presencia de multicolinealidad; por ello, podemos decir que  $x_{8i}$  no produce multicolinealidad.

Repetimos los pasos para  $x_{6i}$  como variable dependiente y el resto de los regresores como independientes.

Tabla 7.4.4. *Regresión de  $x_{6i}$  respecto al resto de variables independientes.*

Dependent Variable: X6  
Method: Least Squares  
Date: 08/27/20 Time: 12:43  
Sample (adjusted): 1 53  
Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	9.239177	5.639494	1.638299	0.1085
X1	3.090714	2.821688	1.095342	0.2793
X2	-0.030845	0.043831	-0.703725	0.4853
X3	-0.054209	0.030619	-1.770433	0.0836
X4	-1.975804	1.317312	-1.499876	0.1408
X5	0.024490	0.071427	0.342873	0.7333
X7	-0.145173	0.388476	-0.373700	0.7104
X8	0.254578	0.105751	2.407336	0.0203
R-squared	0.412808	Mean dependent var		12.72191
Adjusted R-squared	0.319391	S.D. dependent var		4.434594
S.E. of regression	3.658498	Akaike info criterion		5.572621
Sum squared resid	588.9228	Schwarz criterion		5.872812
Log likelihood	-136.8881	Hannan-Quinn criter.		5.687707
F-statistic	4.418978	Durbin-Watson stat		1.570319
Prob(F-statistic)	0.000866			

Fuente: Elaboración propia a partir de datos de la OCDE.

Cálculo del FIV de  $x_{6i}$ :

$$FIV = \frac{1}{1 - 0,412808} = 1,703020477$$

En este caso, la varianza es 1,703 veces la que se habría obtenido si hubiera ortogonalidad. Al igual que en el anterior caso, no se puede considerar la presencia de multicolinealidad.

Por último, repetimos el procedimiento para  $x_4$ :

Tabla 7.4.5. Regresión de  $x_{4i}$  respecto al resto de variables independientes.

Dependent Variable: X4  
 Method: Least Squares  
 Date: 09/18/20 Time: 10:27  
 Sample (adjusted): 1 53  
 Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.075110	0.567962	3.653608	0.0007
X1	0.697214	0.301433	2.312998	0.0255
X2	-0.000278	0.004920	-0.056430	0.9553
X3	-0.003288	0.003503	-0.938619	0.3531
X5	-0.014961	0.007658	-1.953529	0.0571
X6	-0.024618	0.016414	-1.499876	0.1408
X7	-0.050131	0.042769	-1.172133	0.2475
X8	-0.035772	0.011341	-3.154244	0.0029

R-squared	0.592583	Mean dependent var	-0.262722
Adjusted R-squared	0.527767	S.D. dependent var	0.594267
S.E. of regression	0.408376	Akaike info criterion	1.187381
Sum squared resid	7.337917	Schwarz criterion	1.487572
Log likelihood	-22.87190	Hannan-Quinn criter.	1.302467
F-statistic	9.142510	Durbin-Watson stat	1.385497
Prob(F-statistic)	0.000001		

Fuente: Elaboración propia a partir de datos de la OCDE.

Cálculo del FIV de  $x_{4i}$ :

$$FIV = \frac{1}{1 - 0,592583} = 2,45448766$$

La varianza es 2,4545 veces la que se habría obtenido si hubiera ortogonalidad. Como en los anteriores casos, no se puede considerar la presencia de multicolinealidad.

Cuando exista multicolinealidad, puede deberse a problemas de la recogida de la muestra. Se podría paliar la multicolinealidad de las siguientes formas:

- Eliminar las variables que causan la multicolinealidad. Si eliminamos una variable que no se ve fuertemente afectada, podemos sesgar el modelo y cometer un error de omisión de variable.
- Aumentar la muestra, introduciendo observaciones adicionales. Esto no es siempre posible debido a la carencia de datos que pudiera existir en la fuente que a la que se ha acudido para la recolección de estos.

- Utilizar información extramuestral: para ello, se pueden emplear estimadores procedentes de otros modelos o establecer ciertos valores para los parámetros, obtenidos de la teoría económica. Cabe destacar que tanto las restricciones como los estimadores prestados deben dotarse de un sentido económico acorde con el modelo en cuestión.
- Otras soluciones.

En nuestro caso, el análisis de la multicolinealidad concluye que no es un problema importante en nuestro modelo.

## 8. RESULTADOS

Para empezar, es necesario realizar un pequeño comentario sobre el significado de los coeficientes. El coeficiente de cada regresor mide el efecto separado del regresor sobre la variable dependiente:

- $\hat{\beta}_1 = 60,82218$ : son los puntos que aumentan las calificaciones debido a la variación positiva del índice de miedo al fracaso laboral en las mujeres. A pesar de la general tendencia del superior rendimiento femenino en el ámbito académico (OCDE, 2018), cuando se enfrentan al mundo laboral las expectativas de las mujeres empeoran y aumenta el miedo al fracaso laboral. Esto se debe a aspectos sociales y culturas, al tener consciencia de este trato desigual en el mercado laboral, se crea una necesidad de hacer mayores esfuerzos en su formación y aumenta el miedo al fracaso.
- $\hat{\beta}_2 = 0,077856$ : el efecto del porcentaje de alumnos inmigrantes que esperan acabar sus estudios es casi nulo, esto puede ser explicado porque su p-valor es alto. Como se ha comentado anteriormente, el coeficiente negativo de la variable “porcentaje de alumnos inmigrantes” condujo a la selección de la variable “porcentaje de alumnos inmigrantes que esperan acabar sus estudios”, cuyo coeficiente era positivo.
- $\hat{\beta}_3 = 0,466697$ : el porcentaje de colegios públicos tampoco tiene un efecto cuantitativamente muy elevado sobre la variable endógena, es decir, en las puntuaciones medias en ciencias y matemáticas.

- $\hat{\beta}_4 = 20,21510$ : el índice socioeconómico y cultural tiene un peso notable sobre las puntuaciones medias en ciencias y matemáticas.
- $\hat{\beta}_5 = -2,311093$ : las variaciones del porcentaje de *bullying* tienen un efecto indirecto en las puntuaciones medias, esto es, cuando se reduce el acoso escolar, aumentan las notas levemente.
- $\hat{\beta}_6 = 1,062140$ : Cuando disminuye la ratio profesor-alumno, aumentan las puntuaciones, aunque no es un efecto muy notable.
- $\hat{\beta}_7 = -9,057034$ : en este caso, existe una contradicción en el signo y la lógica económica, puesto que debería obtenerse un coeficiente positivo. Cabe destacar que la variable  $x_{7i}$ , gasto público en educación en porcentaje del PIB, es muy importante, ya que es capaz de representar los recursos que el Estado pone a disposición de la educación, tanto pública como privada y su recolección ha sido complicada debido a la carencia de datos para varios países. Teniendo un signo negativo la variable  $x_{7i}$ , los resultados académicos indicarían que los países con mayor gasto en educación tienen peores resultados académicos, lo cual no es coherente. Esto se puede explicar porque existe la posibilidad de que esta variable sea una variable endógena o por los problemas inherentes a la base de datos en lo que a esta variable se refiere. En el caso de que fuese variable endógena, habría que buscar otro método de estimación. Por otro lado, no importa tanto la inversión sino como se materializa, ya que existen algunos países con un gasto público en educación más bajo y mejores resultados; sin embargo, esto no es la norma general.
- $\hat{\beta}_8 = -2,110942$ : el índice de Gini cuando se acerca a 0 muestra más igualdad. Por ello, cuando aumenta este índice, disminuyen las puntuaciones. El efecto cuantitativo es similar al efecto del porcentaje de bullying.

Tras haber analizado los resultados de los coeficientes, se puede confirmar que el rendimiento académico está influenciado fundamentalmente por factores socioculturales y por las expectativas laborales del colectivo femenino. En el siguiente punto, se desarrollarán el resto de las conclusiones.

## 9. CONCLUSIONES

A continuación de esto, en este apartado voy a comentar las conclusiones a las que he llegado durante el trabajo. En primer lugar, resumiré las conclusiones obtenidas durante el estudio teórico y posteriormente las obtenidas durante el análisis aplicado.

Este modelo nos viene a demostrar que las puntuaciones no están basadas en el esfuerzo, esto no quiere decir que no influya, sino que no es concluyente. Son los factores del ambiente sociocultural aquellos verdaderos determinantes a la hora de comprender el éxito de un sistema educativo y las expectativas y motivaciones respecto al futuro laboral que este sistema es capaz de fomentar, al menos en el colectivo que obtiene mejores calificaciones, que son las mujeres.

Primero, respecto a las conclusiones del estudio teórico podemos confirmar:

- Los patrones sociales y culturales que desarrollan los hombres y las mujeres debido a los estereotipos de género influyen en el rendimiento académico, las elecciones de carrera profesional y en su futura adaptación al mercado laboral.
- Las ambiciones, la confianza en uno mismo y las expectativas de los progenitores en sus hijos son variables relevantes cuando se estudian los resultados académicos.
- Existe una gran dificultad a la hora de estudiar la inmigración en un modelo ya que se pueden encontrar notables diferencias entre los propios inmigrantes, sobre todo en relación con el país de procedencia y con el nivel socioeconómico.
- Las escuelas privadas y concertadas efectúan divisiones respecto a la clase social que producen leves mejoras en los resultados frente a las escuelas públicas.
- Existen factores cualitativos como la geografía o la cultura que no vienen reflejados en las evaluaciones internacionales como el Informe Pisa y que son destacables en el examen y comparación de los distintos sistemas educativos.
- En la actualidad, existen dificultades cuando el tejido productivo de una nación es incapaz de absorber la mano de obra cualificada

nacional. La globalización aumenta este efecto debido a la facilidad de deslocalización de una empresa a zonas con mano de obra más barata.

En la parte práctica, se ha estimado un modelo econométrico llegando a las siguientes conclusiones:

- Los factores que mayor efecto positivo tienen en las puntuaciones medias en matemáticas y ciencias son el miedo de las mujeres al fracaso laboral y el índice socioeconómico y cultural.
- Los datos nos muestran que las mujeres obtienen mejores notas, sin embargo, tienen mayores problemas a la hora de enfrentarse al mercado laboral. La posibilidad de tener hijos en edad laboral y la dificultad de la conciliación entre la vida laboral y familiar y el techo de cristal explica mayores índices de miedo al fracaso laboral.
- Acorde con los estudios ya mencionados de Portes, Aparicio, Haller y Vickstrom (2011) y Aparicio y Portes (2014), las ambiciones influyen en la inmigración. Esto explicaría el cambio de signo entre la variable “porcentaje de inmigrantes” y la variable “porcentaje de inmigrantes que esperan acabar sus estudios”. En el caso de un aumento de la variable “porcentaje de inmigrantes”, con un coeficiente negativo, explicaría que al aumentar el porcentaje de inmigración empeorasen los resultados académicos. Esto se puede deber al aprendizaje o no del idioma en el que se imparten las lecciones, la falta de integración o la presencia de mayor acoso escolar. Mientras que los autores anteriormente citados, confirmarían que las ambiciones en los adolescentes son destacables en la explicación del rendimiento académico, es decir, la ambición de acabar los estudios daría lugar al coeficiente positivo.
- El informe PISA, como hemos visto, muestra que las escuelas privadas obtienen mejores resultados que las escuelas públicas. Sin embargo, de acuerdo con Mancebón y Pérez-Ximénez (2007) cuando se ha corregido el efecto renta de las puntuaciones de las escuelas privadas, se demostraba que los colegios públicos estaban por encima de los privados; esto demuestra que un incremento en el

porcentaje de escuelas públicas produciría una mejora del rendimiento académico. En otras palabras, la discriminación que hacen los colegios privados hacia las clases con menos recursos tiene como consecuencia una mejora en las puntuaciones.

- En cuanto a la variable  $x_{6i}$ , ratio profesor-alumno, hemos comprobado que no es relevante a la hora de explicar el rendimiento del alumnado; esta variable puede ser interesante al hacer un análisis más global de todo el sistema educativo, ya que existen correlaciones con los resultados académicos, pero no existe causalidad. La variable ratio de alumnos por clase, puede ser interesante para estudiar este tipo de relaciones y poder sustituir a la ratio profesor-alumno para ver si es significativa en futuras investigaciones. Esta posición podría constituirse como una línea futura de investigación.
- Se comprueba que el porcentaje de *bullying*, de acuerdo con Florencia Gerenni y Leila Fridman (2015, pp. 71-82), afecta al rendimiento académico. Además, existe una relación entre los países que disponen de más recursos para hacer frente al acoso escolar y los países con mayores puntuaciones en el Informe PISA.
- Como comentamos en el punto 4.8, la variable  $x_{7i}$ , es decir, el gasto público en educación en porcentaje del PIB, es relevante; sin embargo, es una variable que ha creado problemas ya que fue la única variable sacada de una base de datos diferente a la del resto de variables, con desigualdades en los años seleccionados debido a falta de información. Otro punto clave de esta variable es la posibilidad de que actúe como variable endógena, esto podría deberse la existencia de correlaciones con otros aspectos no cuantificables recogidos en las perturbaciones.
- Con perspectivas a líneas abiertas de investigación, en primer lugar, se debería depurar la base de datos, especialmente respecto a los datos de la variable  $x_{7i}$ , el gasto público en educación en porcentaje del PIB. Para lograr una base de datos depurada, se debe plantear la introducción de una variable ficticia con el fin de incluir las diferencias metodológicas entre los países incluidos en este trabajo.

Tras esto, habría que estudiar la consideración de  $x_{7i}$  como posible variable, es decir, su posible correlación con las perturbaciones. En este caso, cabría plantearse la obtención del estimador de Mínimos Cuadrados en dos Etapas, como mejor estimador de Variables Instrumentales.

Este Trabajo Fin de Grado aspira a constituir una aportación al conocimiento de los agentes que influyen en la educación, especialmente en el rendimiento académico. La educación adopta un papel clave para el desarrollo de los individuos en sociedad y también es esencial para aumentar la calidad de vida y reducir las desigualdades de los ciudadanos.



## REFERENCIAS BIBLIOGRÁFICAS

- [1] Eviews (Versión 11). (2020). [Paquete econométrico].  
<http://www.eviews.com/download/university11/>
- [2] Beltrán, J. Y Hernández, F.J. (2013). *Sociología de la educación*. Madrid: McGraw Hill.
- [3] Besozzi, E. (1993). *Elementi di sociologia dell'educazione*. Roma: La Nuova Italia Scientifica.
- [4] Rajimon, John (2010). La economía y la función de producción en educación. Revista Científica "Visión de Futuro", 13(1).  
<https://www.redalyc.org/articulo.oa?id=3579/357935475006>
- [5] Fuentes de Frutos, Silvia; and Renobell Santarén, Víctor (2020). La influencia del género en el aprendizaje matemático en España. Evidencias desde PISA. *Revista de Sociología de la Educación-RASE*, 13(1), 63-80.  
<http://dx.doi.org/10.7203/RASE.13.1.16042>.
- [6] OECD (2015), *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*, PISA, OECD Publishing, Paris.  
<https://doi.org/10.1787/9789264229945-en>.
- [7] Balart, P., Oosterveen, M. (2019). Females show more sustained performance during test-taking than males. *Nature Communications*, 10(3798).  
<https://doi.org/10.1038/s41467-019-11691-y>
- [7] Bausela Herreras, Esperanza (2016). Análisis Comparativo de la Probabilidad de Tener Bajo Rendimiento en Función del Sexo en las Competencias PISA 2012. *Revista Iberoamericana de Diagnóstico y Evalua*

ción - e Avaliação Psicológica, 1 (41), 58-65. [fecha de consulta: 22 de Agosto de 2020].

<https://www.redalyc.org/articulo.oa?id=4596/459646901006>

[8] Fernandez-Cano, Antonio (2016). Una crítica metodológica de las evaluaciones PISA. *Revista Electrónica de Investigación y Evaluación Educativa-RELIEVE*, 22(1), art. M15. <http://dx.doi.org/10.7203/relieve.22.1.8806>

[9] Calero, Jorge, & Oriol Escardíbul, J. (2016). Proceso educativo y resultados del alumnado nativo y de origen inmigrante en España. Un análisis basado en PISA-2012. *Estudios de Economía Aplicada*, 34(2), 413-438. [fecha de consulta 22 de Agosto de 2020]. <https://www.redalyc.org/articulo.oa?id=301/30146038006>

[10] Julio Carabaña (2008). Las diferencias entre países y regiones en las pruebas PISA. Universidad Complutense de Madrid.

[11] Entorf, Horst and Minoiu, Nicoleta, PISA Results: What a Difference Immigration Law Makes (2004). <https://ssrn.com/abstract=508544>

[12] Jarque J.M. and Bera A.K.(1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 55(2). [http://academicos.fciencias.unam.mx/wp-content/uploads/sites/91/2015/04/jarque\\_bera\\_87.pdf](http://academicos.fciencias.unam.mx/wp-content/uploads/sites/91/2015/04/jarque_bera_87.pdf)

[13] Gujarati, Damodar N. (2010). *Econometría*. Madrid: McGraw Hill.

[14] Cavero Álvarez, J y Zarzosa Espina, P. "Material docente de la asignatura. Econometría. Curso 2015-2016". *Universidad de Valladolid. Facultad de Ciencias Económicas y Empresariales*.

[15] Organización para la Cooperación y el Desarrollo Económico (2018): Informe PISA, Metodología 2020. Base de datos. Disponible en: <https://www.oecd.org/pisa/data/2018database/>

[16] Banco Mundial (2018): Índice de Gini. Disponible en: <https://datos.bancomundial.org/indicador/SI.POV.GINI>

[17] Esping-Andersen G. y Cimentada J. (2020, julio). ¿Qué influye más en la posición social de una persona, sus habilidades o su origen familiar? *Observatorio Social de la Caixa*.

<https://observatoriosociallacaixa.org/es/-/posicion-social-habilidad-y-origen-familiar>

[18] Suárez-Grimalt, L. (2017). Migración y movilidad social: una aproximación desde las estrategias de acumulación de activos de la población latinoamericana en España. *Revista Española De Sociología*, 26(3). <https://doi.org/10.22325/fes/res.2017.27>

[19] Souto, X. M. (2018). La geografía escolar: deseos institucionales y vivencias de aula. *Boletín de la Asociación de Geógrafos Españoles*, 79, 2757, 1–31. <http://dx.doi.org/10.21138/bage.2757>



Gráfica 3: *Porcentaje de escuelas públicas.*

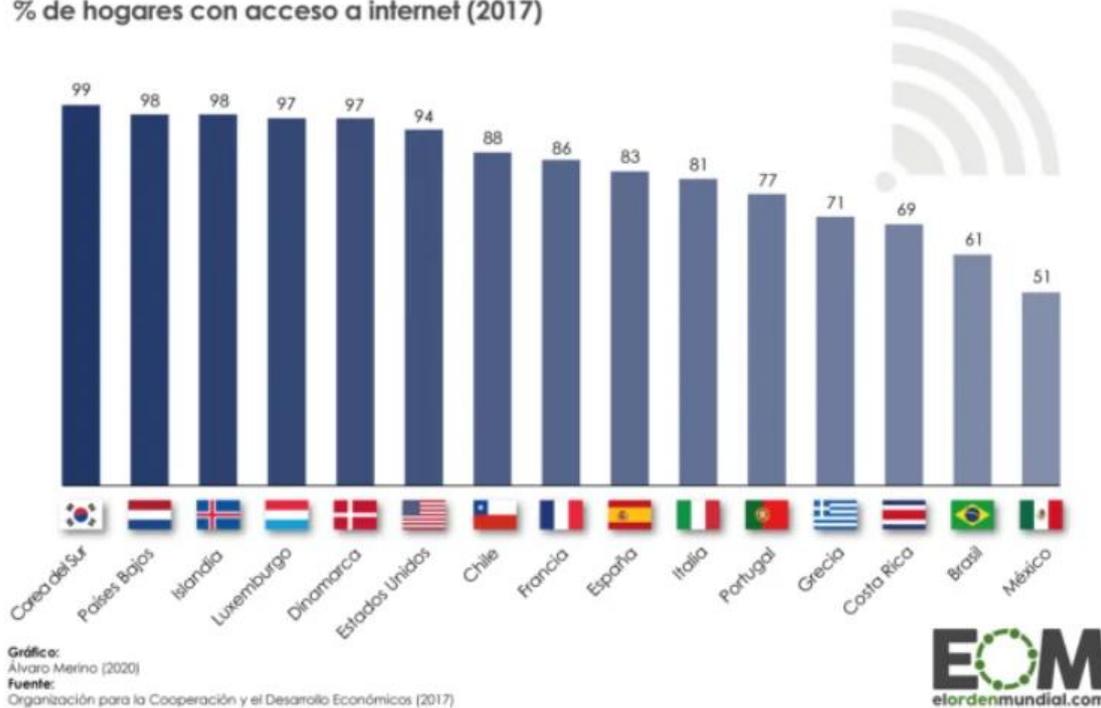


Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 4. *Porcentaje de hogares con acceso a internet.*

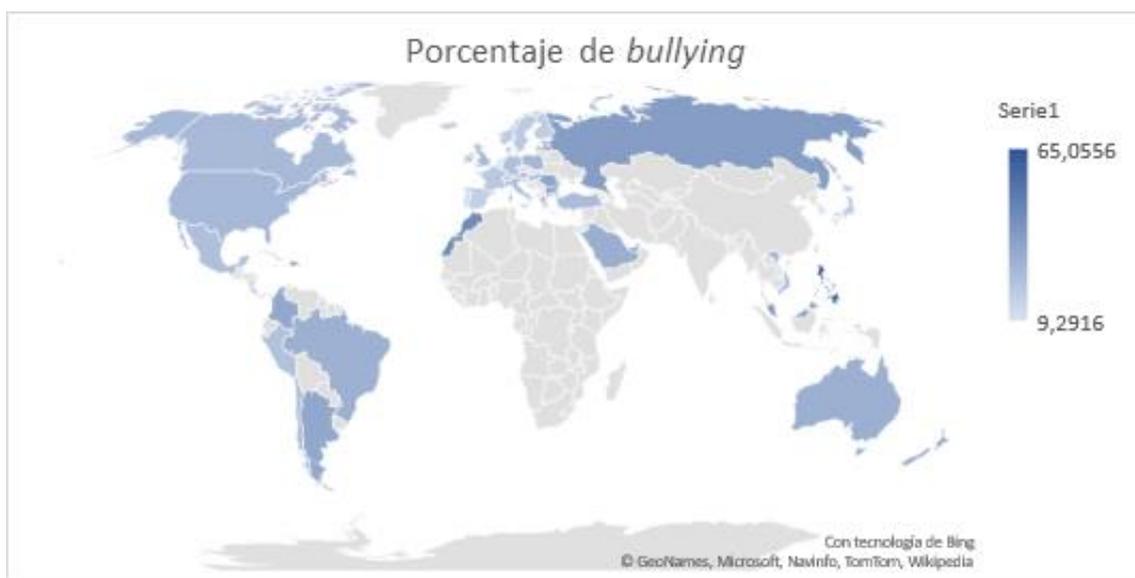
## Internet en la sociedad de la información

% de hogares con acceso a internet (2017)



Fuente: El Orden Mundial (EOM).

Gráfica 5: *Porcentaje de bullying en las escuelas.*



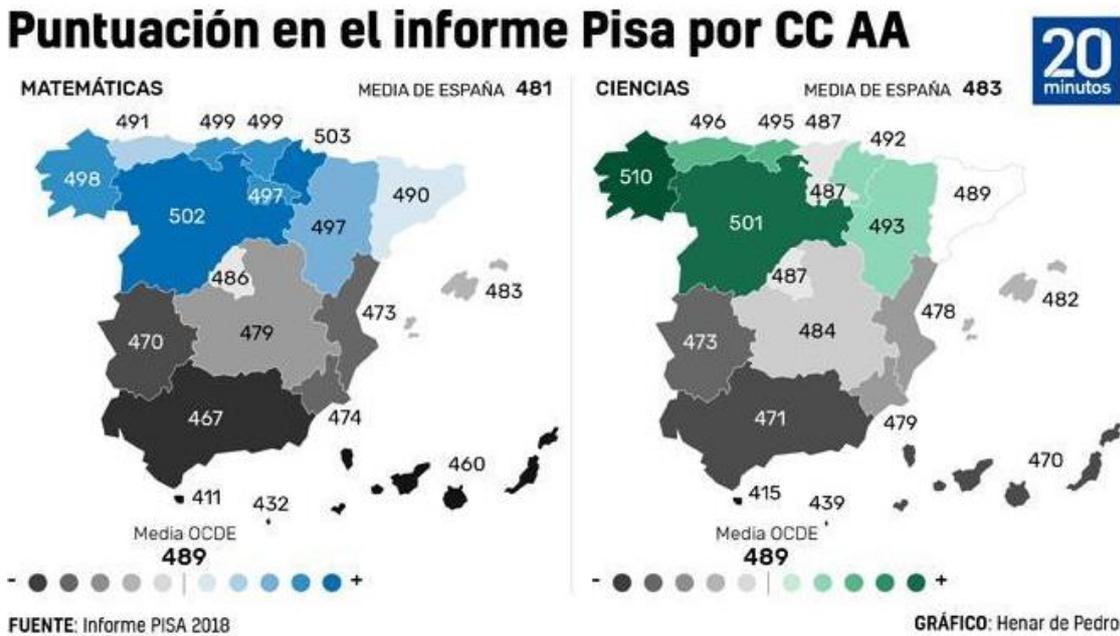
Fuente: Elaboración propia a partir de datos de la OCDE.

Tabla 1: *porcentaje de estudiantes inmigrantes.*

<b>Países</b>	<b>%</b>
Corea	0,09
Bulgaria	1
Colombia	1
Japón	1
Polonia	1
R. Eslovaca	1
Rumanía	1
Turquía	1
Lituania	2
México	2
Chile	3
Hungría	3
Letonia	4
R. Checa	4
Finlandia	6
Islandia	6
Portugal	7
Croacia	9
Eslovenia	9
Malta	9
Estonia	10
Italia	10
Dinamarca	11
España	12
Grecia	12
Noruega	12
Francia	14
Países Bajos	14
Israel	16
Bélgica	18
Irlanda	18
Reino Unido	20
Suecia	20
Alemania	22
Austria	23
Estados Unidos	23
Nueva Zelanda	26
Australia	28
Suiza	34
Canadá	35
Luxemburgo	55

Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 6. Puntuaciones por las Comunidades Autónomas.



Fuente: periódico 20 minutos.

Tabla 2. Salida de la estimación sin  $x_{2i}$  (porcentaje de alumnos inmigrantes que esperan acabar sus estudios).

Dependent Variable: Y  
 Method: Least Squares  
 Date: 08/27/20 Time: 13:58  
 Sample (adjusted): 1 53  
 Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	581.3764	30.62235	18.98536	0.0000
X1	62.70468	16.67972	3.759337	0.0005
X3	0.456634	0.195933	2.330569	0.0244
X4	20.19480	8.487965	2.379228	0.0218
X5	-2.345968	0.432049	-5.429860	0.0000
X6	1.034047	0.942207	1.097474	0.2784
X7	-9.109916	2.438181	-3.736358	0.0005
X8	-2.067019	0.689569	-2.997552	0.0045

R-squared	0.807179	Mean dependent var	467.8586
Adjusted R-squared	0.776503	S.D. dependent var	48.63739
S.E. of regression	22.99353	Akaike info criterion	9.248941
Sum squared resid	23262.91	Schwarz criterion	9.549132
Log likelihood	-232.4725	Hannan-Quinn criter.	9.364027
F-statistic	26.31307	Durbin-Watson stat	1.675617
Prob(F-statistic)	0.000000		

Fuente: Elaboración propia a través de datos de la OCDE.

En el modelo:

$$\hat{y}_i = 581,3764 + 62,70468x_{1i} + 0,456634x_{3i} + 20,1948x_{4i} - 2,345968x_{5i} + 1,034047x_{6i} - 9,109916x_{7i} - 2,067019x_{8i}$$

Tabla 3. *Salida de la estimación con x2 como porcentaje de alumnos inmigrantes.*

Dependent Variable: Y  
 Method: Least Squares  
 Date: 06/29/20 Time: 20:25  
 Sample (adjusted): 1 53  
 Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	583.6446	30.83963	18.92515	0.0000
X1	60.61334	16.91642	3.583107	0.0009
X2	-0.250104	0.296076	-0.844730	0.4029
X3	0.437471	0.197878	2.210810	0.0324
X4	23.57215	9.407589	2.505653	0.0161
X5	-2.293202	0.437940	-5.236333	0.0000
X6	0.975410	0.947834	1.029094	0.3092
X7	-9.429081	2.475163	-3.809478	0.0004
X8	-1.956812	0.704019	-2.779489	0.0080
R-squared	0.810327	Mean dependent var	467.8586	
Adjusted R-squared	0.775039	S.D. dependent var	48.63739	
S.E. of regression	23.06874	Akaike info criterion	9.270944	
Sum squared resid	22883.17	Schwarz criterion	9.608659	
Log likelihood	-232.0445	Hannan-Quinn criter.	9.400416	
F-statistic	22.96325	Durbin-Watson stat	1.637393	
Prob(F-statistic)	0.000000			

Fuente: Elaboración propia a través de datos de la OCDE.

En el modelo:

$$\hat{y}_i = 583,6446 + 60,61334x_{1i} - 0,250104x_{2i} + 0,437471x_{3i} + 23,57215x_{4i} - 2,293202x_{5i} + 0,97541x_{6i} - 9,429081x_{7i} - 1,956812x_{8i}$$

Tabla 4. Salida de la estimación sin  $x_{6i}$  ratio profesor-alumnos.

Dependent Variable: Y  
 Method: Least Squares  
 Date: 09/14/20 Time: 11:16  
 Sample (adjusted): 1 53  
 Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	585.6027	35.91521	16.30514	0.0000
X1	64.10496	17.96996	3.567339	0.0009
X2	0.045095	0.279138	0.161550	0.8724
X3	0.409119	0.194999	2.098060	0.0417
X4	18.11651	8.389324	2.159472	0.0363
X5	-2.285081	0.454885	-5.023429	0.0000
X7	-9.211228	2.474017	-3.723187	0.0006
X8	-1.840545	0.673476	-2.732902	0.0090

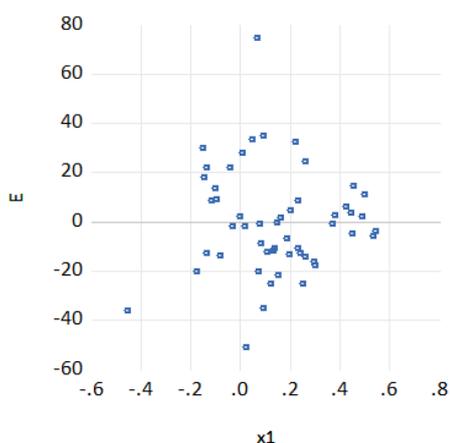
R-squared	0.802019	Mean dependent var	467.8586
Adjusted R-squared	0.770522	S.D. dependent var	48.63739
S.E. of regression	23.29921	Akaike info criterion	9.275354
Sum squared resid	23885.54	Schwarz criterion	9.575545
Log likelihood	-233.1592	Hannan-Quinn criter.	9.390440
F-statistic	25.46331	Durbin-Watson stat	1.791572
Prob(F-statistic)	0.000000		

Fuente: Elaboración propia a través de datos de la OCDE.

En el modelo:

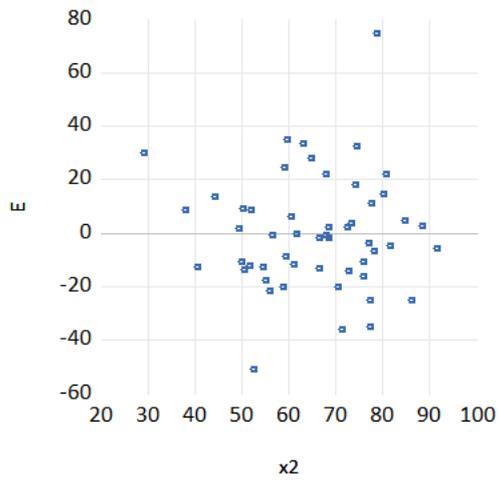
$$\hat{y}_i = 585,6027 + 64,10496x_{1i} + 0,045095x_{2i} + 0,409119x_{3i} + 18,11651x_{4i} - 2,211228x_{5i} - 9,21128x_{7i} - 1,84054x_{8i}$$

Gráfica 7. Análisis de los residuos frente a las variables  $x_{1i}$ .



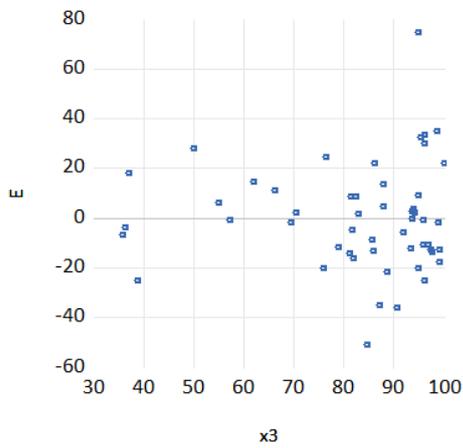
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfico 8. *Análisis de los residuos frente a las variables  $x_{2i}$ .*



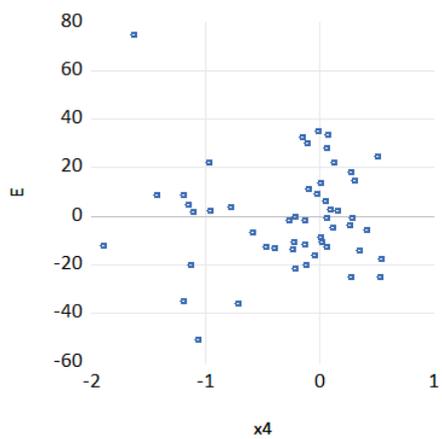
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfica 9. *Análisis de los residuos frente a las variables  $x_{3i}$ .*



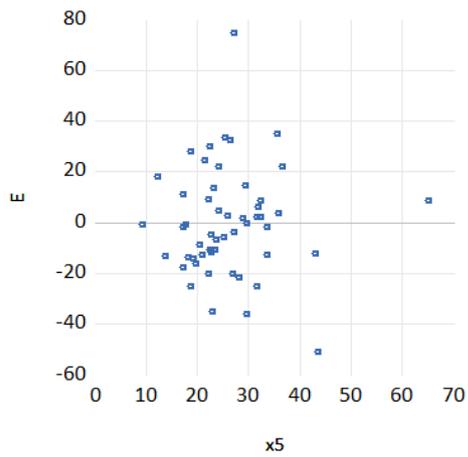
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfica 10. *Análisis de los residuos frente a las variables  $x_{4i}$ .*



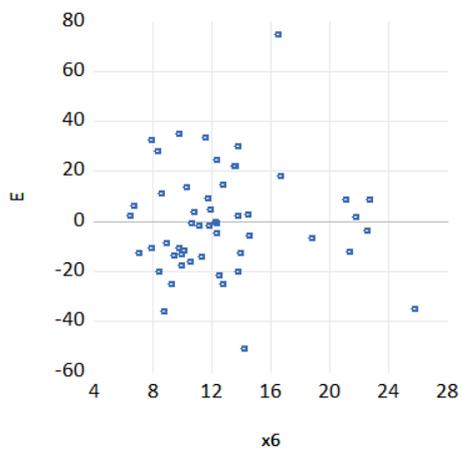
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfica 11. *Análisis de los residuos frente a las variables  $x_{5i}$ .*



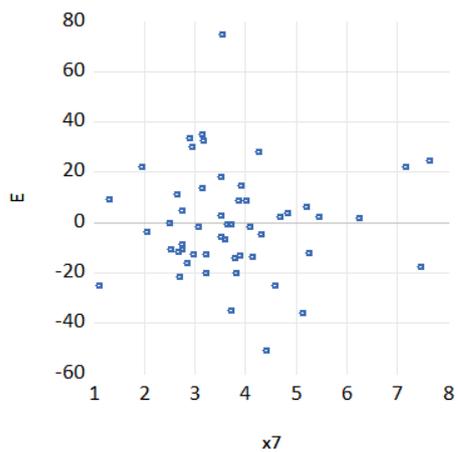
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfica 12. *Análisis de los residuos frente a las variables  $x_{6i}$ .*



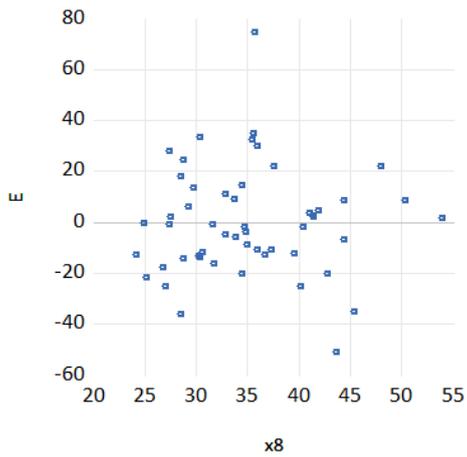
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfica 13. *Análisis de los residuos frente a las variables  $x_{7i}$ .*



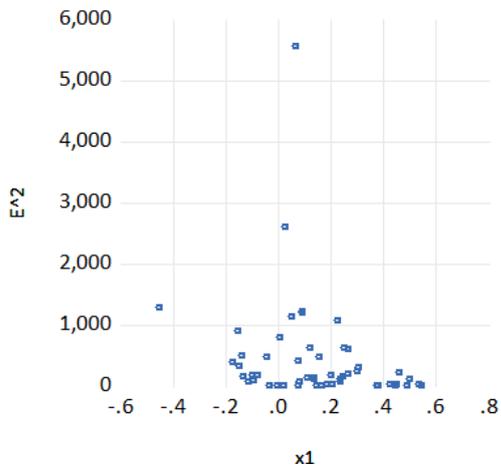
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfica 14. *Análisis de los residuos frente a las variables  $x_{8i}$ .*



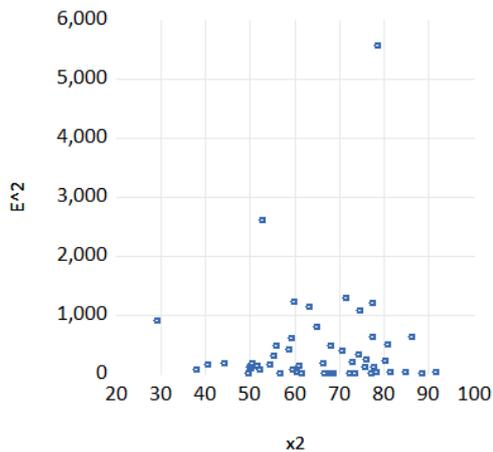
Fuente: Elaboración propia a través de datos de la OCDE.

Gráfica 15. *Análisis de los residuos al cuadrado frente a  $x_{1i}$ .*



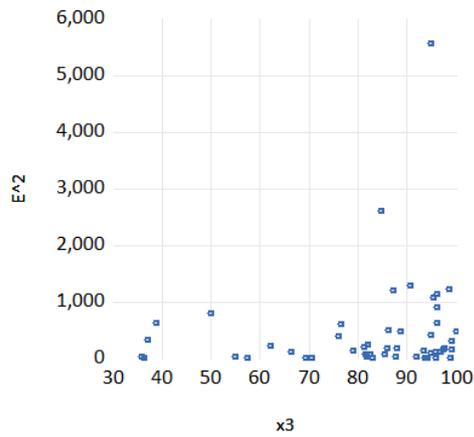
Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 16. *Análisis de los residuos al cuadrado frente a  $x_{2i}$ .*



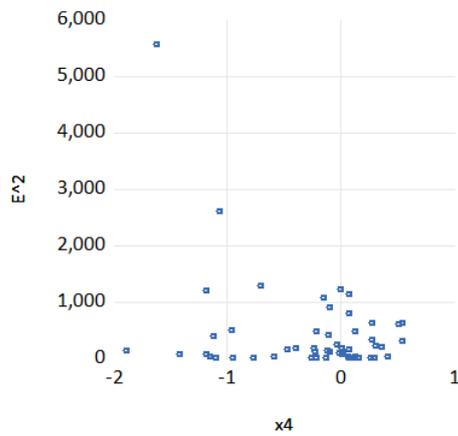
Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 17. *Análisis de los residuos al cuadrado frente a  $x_{3i}$ .*



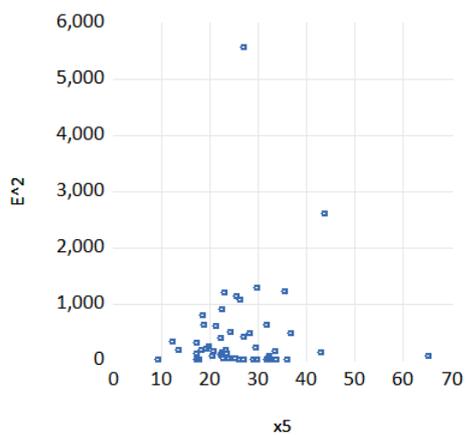
Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 18. *Análisis de los residuos al cuadrado frente a  $x_{4i}$ .*



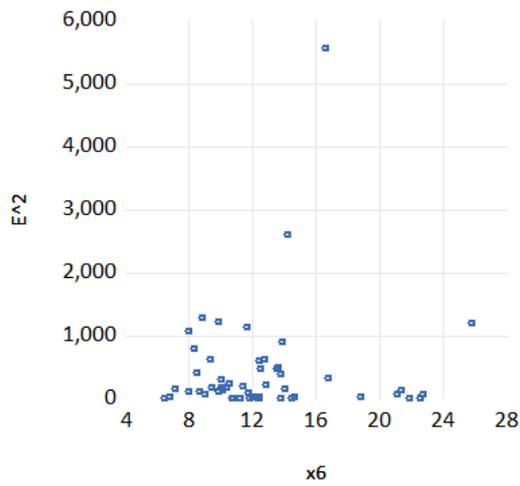
Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 19. *Análisis de los residuos al cuadrado frente a  $x_{5i}$ .*



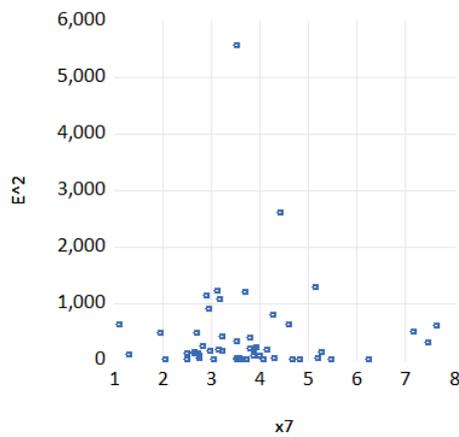
Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 20. *Análisis de los residuos al cuadrado frente a  $x_{6i}$ .*



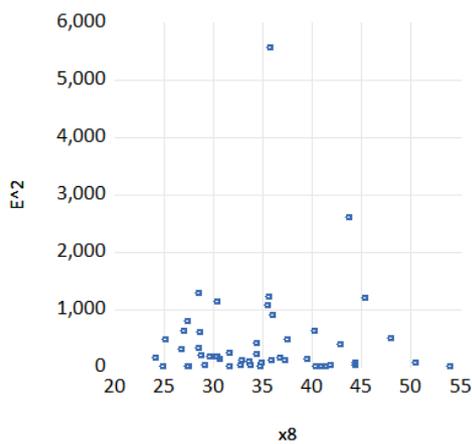
Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 21. *Análisis de los residuos al cuadrado frente a  $x_{7i}$ .*



Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 22. *Análisis de los residuos al cuadrado frente a  $x_{8i}$ .*

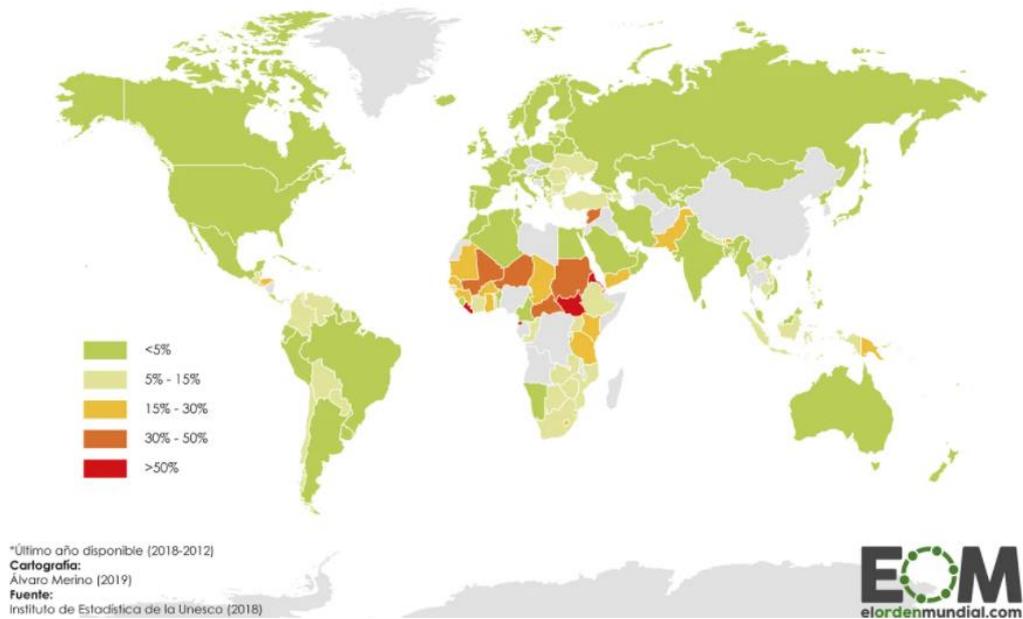


Fuente: Elaboración propia a partir de datos de la OCDE.

Gráfica 23. *Porcentaje de niños sin escolarizar.*

## Niños sin escolarizar en el mundo

% de niños en edad escolar primaria sin escolarizar\*



Fuente: El Orden Mundial (EOM).