



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS
INDUSTRIALES

TRABAJO FIN DE GRADO

**Análisis Cuantitativo de la Relación entre el Aumento de la
Edad de Jubilación y el Desempleo Juvenil**

Autor: Paula González-Ferrer Redondo

Director: David Roch Dupré

Madrid

Julio de 2023

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

Análisis Cuantitativo de la Relación entre el Aumento de la Edad de Jubilación y el Desempleo Juvenil

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el curso académico 2022-2023 es de mi autoría, original e inédito y no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente, y la información que ha sido tomada de otros documentos está debidamente referenciada.



Fdo.: Paula González-Ferrer Redondo, Fecha: 10/07/2023

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO



Fdo.: David Roch Dupré, Fecha: 10/07/2023

ANÁLISIS CUANTITATIVO DE LA RELACIÓN ENTRE EL AUMENTO DE LA EDAD DE JUBILACIÓN Y EL DESEMPLEO JUVENIL

Autor: González-Ferrer Redondo, Paula

Director: Roch Dupré, David

Entidad colaboradora: ICAI – Universidad Pontificia de Comillas

RESUMEN

El envejecimiento de la población y su impacto laboral son desafíos cada vez más importantes para la sociedad actual. Este presente Trabajo de Fin de Grado se centra en estudiar la relación cuantitativa existente entre la edad de jubilación y el desempleo juvenil, con el objetivo de analizar si retrasar la edad de jubilación puede ser una solución viable para abordar los desafíos que supone una sociedad cada vez más longeva.

Palabras clave: edad de jubilación, desempleo juvenil, trabajadores seniors, jóvenes, efectos fijos, efectos aleatorios, GDP, inflación, educación, gasto público, panel data.

INTRODUCCIÓN

El envejecimiento de la población es una consecuencia natural que surge debido a un aumento de la esperanza de vida y una reducción de la tasa de natalidad. Como resultado, se espera que el porcentaje de personas mayores de 65 años continúe aumentando, lo que gradualmente invertirá la estructura de la pirámide poblacional.

Esto conlleva consecuencias demográficas, económicas, sanitarias, políticas... Centrándonos en las económicas, destaca el problema de las pensiones, ya que las cotizaciones de ahora no serán suficientes para abastecer la demanda futura. Es por ello por lo que se plantean diversas soluciones para abordar este problema, como puede ser el retrasar la edad de jubilación. Entre los argumentos que ofrecen los expertos que respaldan esta medida, destaca el hecho de que la edad de jubilación española no se ha adaptado a la realidad social, ya que tener 65 años ahora no es lo mismo que tener 65 años hace 50 años. Además, varios estudios han demostrado que esta medida podría suponer un impulso para la producción del país. Sin embargo, surgen muchas preocupaciones. Por ejemplo, se cree que en el mercado laboral no hay espacio para

todos, y que, si los mayores no se retiran, no hay espacio para los jóvenes. Por tanto, mucha gente cree que retrasar la edad de jubilación fomenta el desempleo entre jóvenes.

DEFINICIÓN DEL PROYECTO

En este presente Trabajo de Fin de Grado se estudia cuantitativamente la relación existente entre la edad de jubilación y la tasa de desempleo juvenil. Para ello, se estudia las variables que podrían definir la tasa de desempleo juvenil (según un estudio de la literatura previo), y se analiza la representatividad de la variable edad de jubilación. De esta manera, se puede comprobar hasta qué punto es real la creencia de que, si no se jubilan los mayores, no hay espacio en el mercado laboral para que puedan entrar los jóvenes.

DESCRIPCIÓN DEL MODELO

Para estudiar la relación previamente explicada, se utiliza un modelo de regresión lineal con series temporales. Para ello, se han recopilado datos de 17 países de la Eurozona, durante una franja temporal de 16 años (2005-2020).

Previamente se ha realizado un estudio académico para determinar qué variables serán utilizadas en el modelo: el PIB (GPD), la inflación (CPI), distintos niveles de educación (básica, obligatoria y terciaria), gasto público en educación y en políticas de empleo, y edad de jubilación. Por otro lado, se estudia la inclusión de efectos fijos y/o aleatorios en el modelo. Esta consideración surge al analizar la variabilidad de la tasa de desempleo juvenil a lo largo de los años y entre distintos países. Se sugiere utilizar efectos aleatorios para tener en cuenta las variaciones inusuales en la tasa de desempleo juvenil durante esos años, así como la variabilidad existente entre países. Por otra parte, se propone utilizar efectos fijos para considerar la presencia de factores estructurales constantes que hacen que la tasa de desempleo mantenga el mismo valor durante un conjunto de años.

Por lo que, siendo el modelo del que se parte el siguiente: $YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed$, se propone el uso de factores fijos $Country, Year$ (), o aleatorios $((1|Country), (1|Year))$, lo que resulta en un total de 7 modelos distintos.

RESULTADOS

El mejor modelo de las diferentes opciones propuestas es aquel que incluye “Country” y “Year” como efectos aleatorios. En este modelo, las únicas variables independientes con p-valor inferior a 0,05 son el GDP y la inflación, siendo las únicas variables representativas con un nivel de confianza del 95%. La edad de jubilación resulta no representativa, y, por tanto, no existe evidencia de que el desempleo juvenil y la edad de jubilación estén directamente relacionados. El coeficiente de las variables representativas es negativo, lo que confirma una asociación negativa entre variable dependiente e independiente y concuerda con la literatura previa. En cuanto a los efectos aleatorios de Country” y “Year”, se afirma que dichos efectos capturan las diferencias para cada país y cada año que no se explican por las variables independientes incluidas por el modelo.

La siguiente etapa del proyecto consiste en determinar qué variables son las que influyen en el desempleo juvenil a través de un enfoque incremental en el modelo. Se ha seguido un enfoque en el que se agregan variables hasta que se observa que ya no mejoran el modelo, ya sea porque tienen un p-valor no significativo o porque no mejoran el coeficiente BIC (Coeficiente de Información de Bayes), que ha sido el criterio utilizado. En este enfoque incremental se ha determinado que las variables que mejor explican el desempleo juvenil son el GDP y el CPI, lo que concuerda con la Ley de Okun y la Curva de Phillips.

CONCLUSIONES

Los resultados indican que la variable de edad de jubilación no es representativa en la mayoría de los modelos analizados, lo que indica que no existe evidencia suficiente para afirmar una relación entre la edad de jubilación y el desempleo juvenil. En los modelos donde sí es representativa, su coeficiente es negativo. Esto indica que la relación entre ellas es negativa, es decir, que, a mayor edad de jubilación, menor tasa de desempleo juvenil.

Por tanto, estos resultados sugieren que prolongar la vida laboral de los seniors no perjudica necesariamente las oportunidades de empleo de los jóvenes. Por tanto, al retrasar la edad de jubilación, nos encontraríamos frente a una manera para sacar provecho de una sociedad cada vez más longeva.

ABSTRACT

The aging population and its impact on the labour market are increasingly important challenges for today's society. This present bachelor's thesis focuses on studying the quantitative relationship between retirement age and youth unemployment, with the aim of analysing whether delaying the retirement age can be a viable solution to address the challenges posed by an increasingly aging society.

Keywords: retirement age, youth unemployment, senior workers, youth, fixed effects, random effects, GDP, inflation, education, public spending, panel data.

INTRODUCTION

Population aging is a natural consequence resulting from increased life expectancy and reduced birth rates. As a result, the percentage of people over 65 is expected to continue rising, gradually reversing the population pyramid structure.

This has demographic, economic, health, and political consequences. From an economic perspective, the issue of pensions stands out, as current contributions will not be sufficient to meet future demand. Various solutions have been proposed to address this problem, such as delaying the retirement age. One of the arguments supporting this measure is that the Spanish retirement age has not adapted to the social reality, as being 65 years old now is not the same as being 65 years old 50 years ago. Furthermore, several studies have shown that this measure could boost a country's productivity. However, many concerns arise. For example, it is believed that there is no room for everyone in the labour market, and if older workers do not retire, there is no space for young people. Therefore, many people believe that delaying the retirement age exacerbates youth unemployment.

PROJECT DEFINITION

This present bachelor's thesis quantitatively studies the relationship between retirement age and the youth unemployment rate. To do so, it examines the variables that could define the youth unemployment rate (based on a previous literature review) and analyses the representativeness of the retirement age variable. This way, it is possible to determine the extent to which the belief that not retiring older workers leaves no room in the labour market for young people is true.

MODEL DESCRIPTION

To study the relationship, a linear regression model with time series is used. Data from 17 Eurozone countries over a 16-year period (2005-2020) have been collected.

An academic study was conducted beforehand to determine which variables would be used in the model: GDP, inflation (CPI), different levels of education (basic, compulsory, and tertiary), public spending on education and employment policies, and retirement age. Additionally, the inclusion of fixed and/or random effects in the model is examined. This consideration arises from analysing the variability of the youth unemployment rate over the years and across different countries. It is suggested to use random effects to account for unusual variations in the youth unemployment rate during those years, as well as the existing variability between countries. Furthermore, fixed effects are proposed to consider the presence of constant structural factors that keep the unemployment rate at the same level over a set of years.

Thus, starting with the model: $YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed$, the use of fixed factors ($Country, Year$) or random factors ($(1|Country), (1|Year)$) is proposed, resulting in a total of 7 different models.

RESULTS

The best model among the proposed options is the one that includes "Country" and "Year" as random effects. In this model, the only independent variables with a p-value below 0.05 are GDP and inflation, representing the only statistically significant variables with a 95% confidence level. The retirement age variable is not representative, and therefore, there is no evidence of a direct relationship between youth unemployment and retirement age. The coefficients of the significant variables are negative, confirming a negative association between the dependent and independent variables and aligning with previous literature. As for the random effects of "Country" and "Year," it is stated that these effects capture the differences for each country and year that are not explained by the independent variables included in the model.

The next stage of the project is to determine which variables influence youth unemployment through an incremental approach in the model. An approach was

followed where variables are added until it is observed that they no longer improve the model, either because they have a non-significant p-value or because they do not improve the Bayesian Information Criterion (BIC), which was the criterion used. In this incremental approach, it was determined that the variables that best explain youth unemployment are GDP and CPI, consistent with Okun's Law and the Phillips Curve.

CONCLUSIONS

The results indicate that the retirement age variable is not representative in most of the analysed models, suggesting that there is not enough evidence to support a relationship between retirement age and youth unemployment. In the models where it is representative, its coefficient is negative. This indicates a negative relationship between them, meaning that a higher retirement age corresponds to a lower youth unemployment rate.

Therefore, these results suggest that prolonging the working life of seniors does not necessarily harm the employment opportunities of young people. Thus, delaying the retirement age could be a way to benefit from an increasingly aging society.

ÍNDICE

1. INTRODUCCIÓN AL PROYECTO	9
2. ESTADO DE LA CUESTIÓN	11
3. MOTIVACIÓN.....	12
4. OBJETIVOS DEL PROYECTO	13
5. OBJETIVOS DE DESARROLLO SOSTENIBLE (ODS)	13
6. SELECCIÓN Y PREPROCESADO DE LAS VARIABLES DEL MODELO.....	14
1) Selección de variables y recopilación de datos	15
2) Matriz de correlaciones entre indicadores	19
3) Panel Data.....	22
7. CONSTRUCCIÓN DEL MODELO DE REGRESIÓN LINEAL.....	23
Análisis previo de la variabilidad de los datos.....	23
Modelo de efectos fijos y modelo de efectos aleatorios	25
i. Modelo simple (modelo 1)	26
ii. Modelo con Fixed Effects (FEM).....	26
a. Fixed Effect Country (modelo 2).....	27
b. Fixed Effect Year (modelo 3).....	27
c. Fixed Effect Country y Year (modelo 4).....	27
iii. Modelo con Random Effects (REM)	28
d. Random Effect Country (modelo 5)	28
f. Random Effect Country y Year (modelo 7)	29
8. ANÁLISIS DE LOS RESULTADOS PARA CADA MODELO	29
1) Tendencias comunes entre modelos	36
2) Elección del mejor modelo	38
3) Tendencias comunes entre los mejores modelos	39
9. CONSTRUCCIÓN DEL MODELO INCREMENTAL.....	41

1)	Modelo incremental mediante BIC.....	42
2)	Modelo incremental mediante AIC.....	44
10.	ANÁLISIS DE LOS RESULTADOS DEL MODELO INCREMENTAL.....	45
11.	CONCLUSIONES.....	46
12.	BIBLIOGRAFÍA.....	48
13.	ANEXO	50

1. INTRODUCCIÓN AL PROYECTO

El envejecimiento de la población resulta una consecuencia natural derivada de un aumento de la esperanza de vida y de una reducción de la tasa de natalidad (número de nacimientos por cada 1000 habitantes en 1 año). Se trata de un creciente e importante reto, ya que afecta a todos los países del mundo con una mayor o menor intensidad.

Enfocándonos en España, en 1960, las personas mayores de 65 años representaban un 8,2% de la población total. Según las previsiones más recientes del INE, se prevé que, para 2050, esta cifra llegue hasta valores del 30,4%, aumentando cada vez más (INE, 2022).

Por otro lado, la tasa de natalidad varía de manera inversa al porcentaje anterior. Además, se introduce un nuevo concepto, el del índice de fecundación (número medio de hijos por mujer en 1 año). En los años 60, dicho índice era de 2,88 con una tasa de natalidad del 21,7%. Por el contrario, en 2021, la tasa de natalidad tomaba valores del 7,10%, junto con un índice de fecundación de 1,19 (Expansión, 2021). Se prevé que, para el año 2050, dicho índice oscile entre valores de 0,87 (escenario bajo) y 1,74 (escenario alto), mientras que la tasa de natalidad siga reduciéndose (INE, 2022). Para Mestres Domènech (2019), la reducción de la tasa de natalidad podría deberse a un aumento del coste de vida, del número de años de formación educativa, y una mayor disponibilidad y accesibilidad de distintos métodos anticonceptivos.

Ambos puntos mencionados previamente explican por qué la pirámide poblacional se está invirtiendo. Esto no sólo tiene consecuencias demográficas, si no también económicas, sanitarias, políticas... Centrándonos en las macroeconómicas, destaca el problema que surge con el pago de las pensiones. Como la esperanza de vida crece cada vez más, las cotizaciones de ahora no serán suficientes para satisfacer la demanda futura, lo cual supondrá un problema grave que urge solucionar cuánto antes.

Para Hernández de Cos, Jimeno y Ramos (2017), el gasto público en pensiones (expresado en términos del PIB) depende de tres factores. El primero es demográfico, y se basa en que el gasto en pensiones es mayor cuando aumenta la tasa de dependencia (índice que establece cuántas personas dependen de la población activa). Por tanto, cuantas más personas de la tercera edad haya, mayor gasto en pensiones. El segundo factor está relacionado con el mercado de trabajo: si hay menos personas en edad de

trabajar con empleo, el PIB será menor, y, por tanto, el gasto en pensiones con relación al PIB será más alto. El tercer y último factor se enfoca en la relación entre la pensión media y la productividad económica. Esta relación está determinada por la tasa de sustitución de las pensiones y el peso de los salarios en el PIB. Por tanto, cuanto mayor sea la tasa de sustitución de las pensiones, mayor será su gasto público.

Por consiguiente, para contrarrestar el efecto de las pensiones, se estudia la posibilidad de retrasar la edad de jubilación. González González (2021) defiende claramente que es una medida que urge tomar. Entre sus razones destaca que “la edad de jubilación no se ha adaptado a la realidad social española”, ya que “en 1919, una persona que llegaba a los 65 años de edad podía esperar vivir cinco años más; hoy son veinte años”. Otros estudios del Reino Unido comparten dicha postura, y defienden que “retrasar la edad de jubilación 1 año hace que aumente el PIB en un 1%” (Scott, 2021a). De esta manera, “al obligar a los mayores a jubilarse más tarde, las reformas de las pensiones que aumentan la edad de jubilación incrementarían la oferta de mano de obra y el empleo de los mayores” (Bertoni & Brunello, 2020).

Sin embargo, surgen muchas preocupaciones en cuanto a esta última declaración. En primer lugar, se cree que el nivel de producción de los seniors puede ser un obstáculo o impedimento, debido a que las personas de la tercera edad no rinden tanto como las personas más jóvenes. También preocupa el estado de salud que puedan presentar, ya que no se sabe si “existe capacidad sanitaria a esas edades para soportar trabajar más tiempo” (Scott, 2021a). Por otra parte, en cuanto a lo que es el mercado laboral como tal, existe una creencia popular de que no hay espacio para todos. Por tanto, si los mayores siguen trabajando, se cree que no dejarán hueco en el mercado para que entren los trabajadores más jóvenes, fomentando el desempleo entre dicha parte de la población.

El presente Trabajo de Fin de Grado pretende contrastar esta última afirmación. Se formulará un modelo de regresión lineal en el que, con suficientes datos y evidencia, se estudiará la relación existente entre la edad de jubilación y el empleo juvenil. Se partirá de la hipótesis de que retrasar la edad de jubilación no fomenta el desempleo entre los jóvenes. En caso de que la hipótesis resultase correcta, nos encontraríamos frente a una posible solución para sacar el mayor provecho a una sociedad cada vez más longeva.

2. ESTADO DE LA CUESTIÓN

Tal y como se ha presentado en la introducción, para poder sacar más provecho a una sociedad longeva, habría realizar una serie de cambios a nivel macroeconómico. Esto se fundamenta en la principal diferencia entre una sociedad envejecida y una longeva, ya que la primera se centra en “los cambios en la estructura de edad de la población, mientras que una sociedad longeva busca aprovechar las ventajas de una vida más larga mediante cambios en la forma de envejecer” (Scott, 2021b).

Volviendo a la creencia previamente expuesta (si los mayores no se jubilan, no dejan que los jóvenes entren al mercado laboral, causando que aumente el desempleo que sufre dicha generación). Varios economistas han desarrollado diferentes estudios para intentar demostrar cuál es el verdadero origen de dicho desempleo juvenil, para así profundizar en la afirmación previa.

Como establece Moreno Mínguez en 2015, existe debate sobre las causas del desempleo juvenil en España, ya que son numerosas y a menudo contradictorias. Un estudio del BBVA realizado por García (2011) establece que las causas del desempleo juvenil en España hay que buscarlas en deficiencias tanto del sistema educativo como del mercado de trabajo. Entre las primeras destacan la magnitud del abandono escolar temprano y la polarización de la educación. Por otro lado, entre las segundas sobresalen la segmentación del mercado laboral y la ineficacia de las políticas activas de empleo (pág. 3).

Lo llamativo de las causas previamente expuestas, es que en ninguna se menciona que retrasar la edad de jubilación fomente el desempleo juvenil. Entonces, ¿de dónde sale esta creencia, por qué la gente lo piensa?

Esto es un ejemplo de la falacia del mercado laboral. Dicha falacia afirma que, si la producción es fija, el fomento del empleo entre los trabajadores de más edad provocará la falta de puestos de trabajo para los jóvenes. Sin embargo, dado que el mantenimiento de los trabajadores de más edad en empleo implica mayores ingresos, la idea de que haya una cantidad invariable de producción parece improbable. Al igual que el aumento del empleo femenino en el Reino Unido, que pasó de 7 a 16 millones de personas entre 1951 y 2019, no provocó un descenso del empleo masculino, tampoco un aumento de los trabajadores de más edad debería provocar un descenso del número de

puestos de trabajo disponibles para los trabajadores más jóvenes (Scott, 2021a, pág. e831).

Böheim y Nice (2019), dos economistas de Austria y Alemania respectivamente establecen que, a medida que las economías se vuelven más complejas y crecen debido a la especialización, se crean nuevos puestos de trabajo adicionales. El progreso tecnológico, aunque hace que algunas ocupaciones sean obsoletas, crea constantemente una demanda de nuevos puestos. Las investigaciones demuestran que el hecho de que las personas mayores sean activas y productivas beneficia a todas las edades y estimula la creación de más puestos de trabajo.

Por otra parte, se cree que los trabajadores jóvenes son complementarios a los seniors, y no sustitutivos. Para que fuesen sustitutivos, ambos grupos poblacionales deberían de aportar las mismas habilidades al mercado, y no es así, ya que un grupo consta con la experiencia (por ejemplo), y otro no. En realidad, “los datos sugieren que incluso los trabajadores con cualificaciones similares que difieren en tan sólo cinco años de edad ya son sustitutos imperfectos de edad” (Böheim & Nice, 2019).

Por lo tanto, “se debe buscar maximizar la aportación de valor de generaciones diversas” (Sánchez-Cabezudo Bayón & del Olmo García, 2021), fomentando el empleo entre seniors y jóvenes para que puedan aprender mutuamente.

3. MOTIVACIÓN

Como se ha mencionado previamente, nos dirigimos hacia una sociedad cada vez más longeva. “El impacto de este cambio demográfico no se circunscribe solo al sistema de bienestar, sino que afecta a toda la sociedad en general y a la economía en particular” (Mestres Domènech, 2019). Es por ello por lo que se trata de un tema actual y de interés general, ya que afectará a todos los ciudadanos en un momento u otro.

Si no se aprovechan las ventajas que este reto puede ofrecer, y nos dejamos llevar como sociedad por los distintos prejuicios y estereotipos que supone (más gente mayor equivaldría a menos productividad, menor capacidad de aprendizaje, menos curiosidad...), no conseguiremos prosperar como país, afectando tanto a nuestra economía como a la sociedad. Este Trabajo de Fin de Grado pretende averiguar si existe alguna relación entre retrasar la edad de jubilación y el desempleo juvenil, como tan

comúnmente se cree. Para ello, se formulará un modelo cuantitativo basado en modelos de regresión que permita estudiar dicha relación. En caso de que no exista ninguna relación, retrasar la edad de jubilación podría ser una posible solución al problema de las pensiones, además de permitir el impulso de la economía del país.

Vinculado a esto, Scott (2021a) postula que “una vida más larga y saludable debería ser positiva para la economía, y adaptarse a la longevidad ofrece la oportunidad de contrarrestar las implicaciones negativas del envejecimiento de la sociedad” (pág. e834). Por otra parte, un estudio realizado por Oxford Economics y la Universidad de Salamanca (2021) establece que “es esencial analizar el alcance del impacto de la economía de la longevidad para aprovechar sus ventajas y fomentar la prosperidad y el desarrollo económico en España”.

Por todos estos motivos, se considera que resulta de gran interés relacionar el empleo juvenil con el senior para ver cuál es el efecto real que tiene uno sobre otro.

4. OBJETIVOS DEL PROYECTO

El objetivo principal de este Trabajo de Fin de Grado es estudiar desde un punto de vista cuantitativo la relación existente entre los dos mercados laborales (el de los seniors y el de los jóvenes). Este estudio nos permitirá medir el verdadero impacto de la edad de jubilación en el desempleo juvenil a nivel nacional e internacional. De esta manera, se podrá comprobar hasta qué punto es real la creencia de que, si no se jubilan los mayores, no hay espacio en el mercado laboral para que puedan entrar los jóvenes.

Entre los objetivos específicos destaca identificar y analizar las variables que influyen en el desempleo juvenil, considerando factores como la educación y las políticas de empleo, además de otros indicadores relevantes.

5. OBJETIVOS DE DESARROLLO SOSTENIBLE (ODS)

Los ODS se elaboraron para la agenda 2030 y tienen como objetivo garantizar un futuro mejor para todos. No son obligatorios, pero cada país asume la responsabilidad de trabajar por su cumplimiento. El presente trabajo está alineado con el cumplimiento de varios de ellos.

El ODS número 8 busca el trabajo decente y el crecimiento económico. Si resultase que el empleo juvenil y el senior son complementarios en vez de sustitutivos, nos encontraríamos frente a una posible solución para maximizar los beneficios de una sociedad cada vez más longeva. Al impulsar el trabajo de los seniors, aumentaría la producción del país, lo cual provocaría un crecimiento económico. Sin embargo, para poder fomentar el trabajo de los seniors, habría que reestructurar el mercado laboral fomentando políticas de gestión de la edad. De esta manera, se conseguiría que trabajar fuese más atractivo para los seniors ya que se estarían ofreciendo trabajos más decentes.

Por otra parte, el ODS número 10 busca una reducción de las desigualdades en y entre países. Este trabajo se centra en desigualdades dentro de países, en concreto, en desigualdades que existen en el mercado laboral entre los mayores y los jóvenes. Este tipo de discriminación basada en la edad se llama edadismo, y es “junto con el sexismo y el racismo, una de las tres grandes formas de discriminación en la sociedad” (Sánchez-Cabezudo Bayón & del Olmo García, 2021, pág. 28).

En este trabajo se pretende mostrar los beneficios que supondría aprovecharse de las ventajas laborales que ofrecen los trabajadores de la tercera edad. De tal manera, se fomentaría el trabajo decente, el crecimiento económico, y se reduciría la discriminación entre grupos poblacionales, cumpliendo 2 de los 17 ODS que hay.

6. SELECCIÓN Y PREPROCESADO DE LAS VARIABLES DEL MODELO

Como se ha explicado en el apartado anterior, en el presente Trabajo Final de Grado se estudia la relación existente entre el desempleo juvenil y la edad de jubilación. Para ello, se formula un modelo de regresión lineal con series temporales, en el que se intenta explicar el desempleo juvenil por una serie de variables. Se ha utilizado el software Matlab para realizar los cálculos estadísticos, cuyo código completo está incluido en el Anexo.

A continuación, se detalla la metodología seguida para la construcción del modelo.

1) Selección de variables y recopilación de datos

El primer paso consiste en un estudio bibliográfico para entender, desde un punto de vista académico, qué variables influyen (de manera significativa o marginal) en el desempleo de los trabajadores jóvenes.

Según Martín Román (2002), la ley de Okun establece que, por cada tres puntos de crecimiento real porcentual del Producto Interior Bruto (GPD) por encima de la tasa normal de un año (o el periodo considerado), la tasa de desempleo se reduce porcentualmente en un punto. Por lo que, esta ley sostiene la existencia de una relación negativa entre la variación de la producción y la evolución de la tasa de desempleo en una economía.

Por otro lado, un artículo del BBVA (2022) menciona que la curva de Phillips, propuesta por William Phillips en 1958, establece una relación inversa entre las variables desempleo e inflación (CPI). Dicha teoría sostiene que un aumento en el desempleo reduce la inflación, mientras que una disminución en el desempleo provoca un aumento en la inflación.

Otro punto de vista, como es el de Mitrakos, Tsakloglou y Cholezas (2010), defiende que el desempleo juvenil es resultado de la falta de educación, ya que un bajo nivel de educación aumenta la duración del periodo en el que un joven permanece desempleado. A diferencia de estos autores, en este Trabajo de Fin de Grado se tiene en cuenta la distinción entre niveles de educación, los cuales se comentarán más adelante.

Otra variable que puede afectar al desempleo juvenil es el gasto público, ya que “se muestra que el gasto público es no neutral y que en el periodo que se ejerce incrementa los niveles de empleo y producción” (Velázquez Orihuela, 2015). Dado que uno de los objetivos del presente Trabajo de Fin de Grado es explicar el desempleo juvenil, resulta de relevancia centrarse en el gasto público destinado a la educación y el empleo (**GPed y GPempl**).

Para finalizar, se incluye la variable senior edad de jubilación (Retirement Age) para poder estudiar si, como se ha planteado en apartados anteriores, un aumento en la edad de jubilación de los trabajadores seniors perjudicaría al desempleo de la población joven.

Por lo tanto, siendo Y la variable dependiente, y haciendo referencia al desempleo juvenil (Youth Unemployment), el primer modelo propuesto se formula de la siguiente manera: $Y \sim 1 + GDP + CPI + RetirementAge + ObEd + SecEd + TerEd + GPempl + GPed$.

A continuación, se explica qué mide cada una de las variables escogidas, y en qué unidades se expresa.

GDP: proporciona el valor agregado creado por la producción de bienes y servicios en un país en un periodo de tiempo determinado (normalmente 1 año). Se expresa en millones de dólares americanos.

CPI: expresa la variación en los precios de la cesta de bienes y servicios de los hogares. La medida se expresa en forma de tasa de crecimiento anual, y se utiliza un índice con el año base establecido en 2015.

Retirement Age: indica la edad promedio de jubilación para cada país, expresada en años. Los datos utilizados se han obtenido del informe "*Pensions at a Glance*" de la OECD. Adicionalmente, se llevó a cabo una búsqueda de información para obtener los datos de los años más recientes. Los datos de esta variable se presentaban desglosados por género, por lo que se calculó la media ponderada teniendo en cuenta la proporción de hombres y mujeres en cada país. Este enfoque se implementó con el fin de obtener resultados lo más precisos posible.

Niveles de educación (ObEd, SecEd, TerEd): representa el porcentaje de la población que ha completado como máximo dicho nivel de educación. Eurostat lleva a cabo la siguiente clasificación de los distintos niveles de educación que hay:

- Educación primaria (ObEd): el último nivel de educación completado corresponde a Educación Primaria y Educación Secundaria Obligatoria (ESO).
- Educación secundaria (SecEd): el último nivel de educación completado corresponde al Bachillerato, y a programas orientados hacia conocimientos específicos, como pueden ser las formaciones profesionales de duración anual (FP), o programas de capacitación laboral o preparación para el empleo.
- Educación terciaria (TerEd): el último nivel de educación completado corresponde a las formaciones profesionales de 2 años de duración, carreras universitarias, másteres, doctorados...

Por ejemplo, el indicador SecEd expresa el porcentaje de la población cuyo nivel educativo máximo alcanzado es la educación secundaria (incluyendo la educación primaria).

GPempl: expresa la fracción del gasto público que va destinada a políticas del mercado laboral (en inglés LPM). Se centra en intervenciones públicas dirigidas a personas con dificultades en el mercado laboral, y se divide en tres secciones:

- Servicios de LPM (1): incluye los gastos relacionados con el servicio público de empleo, así como cualquier otro servicio de apoyo financiado públicamente destinado a los solicitantes de empleo.
- Medidas de LMP (2-7): incluye los gastos relacionados con medidas de activación dirigidas a desempleados y otros grupos (programas de capacitación, rotación y trabajo compartido, creación directa de empleo, etc.)
- Apoysos de LPM (8-9): incluye los gastos de mantenimiento de ingresos y apoyo durante el desempleo, junto con los beneficios de jubilación anticipada.

Se ha escogido el filtro “Total LPM (categories 2-7)”, ya que es la única variable que incluye financiación de medidas de fomento del empleo. En cuanto a su unidad de medida, se expresa como porcentaje del GDP.

GPed: indica el porcentaje del gasto público que el gobierno dedica a la educación. Este indicador no se distingue entre niveles de educación que hay. Se expresa como porcentaje del GDP.

A continuación, en la Tabla 1 se muestra un resumen de la relación esperada entre cada variable independiente (GDP, CPI, Retirement Age, ObEd, SecEd, TerEd, GPempl, GPed) y la dependiente (Youth Unemployment). El signo “+” corresponde a una relación directa (si X aumenta, Y lo hace también), mientras que el signo “-” corresponde a una relación inversa (si X aumenta, Y disminuye). Este signo viene dado en base a la literatura previamente explicada.

	GDP	CPI	Ret. Age	ObEd	SecEd	TerEd	GPempl	GPed
Youth Unemployment	-	-	-	-	-	-	-	-

Tabla 1: Relación esperada entre variables independientes y dependiente

Por otro lado, en la Tabla 2 se muestra un resumen de las variables del modelo, junto con el enlace de la base de datos utilizada y los filtros empleados para obtener los datos de cada una de las mismas.

Variable	Ud de medida	Fuente	Filtro utilizado
Desempleo juvenil (Y)	%	OECD	–
https://data.oecd.org/unemp/youth-unemployment-rate.htm			
PIB (GDP)	M\$	OECD	–
https://data.oecd.org/gdp/gross-domestic-product-gdp.htm			
Inflación (CPI)	% sobre 2015	OECD	–
https://data.oecd.org/price/inflation-cpi.htm			
Edad de jubilación (Retirement Age)	Años	OECD	–
<i>“Pensions at a Glance” by OECD</i>			
Tipo de educación (ObEd, SecEd, TerEd)	%	EuroStat	Age class:15-64
https://ec.europa.eu/eurostat/databrowser/view/EDAT_LFS_9911__custom_5333517/default/table?lang=en			
Gasto público empleo (GPempl)	% del GDP	EuroStat	Total LMP
https://webgate.ec.europa.eu/empl/redisstat/databrowser/view/LMP_EXPSUMM\$TPS00076/default/table?lang=en			
Gasto público educación (GPed)	% del GDP	OECD	Sección educación
https://data.oecd.org/gga/general-government-spending.htm			

Tabla 2: Tabla resumen de las variables del modelo

Una vez definidos los indicadores, se procede a la recogida de datos. Para ello, se utilizan diversas bases de datos, destacando la OECD y EuroStat. Se ha recopilado datos de 17 países de la Eurozona, durante una franja temporal de 16 años (desde 2005 hasta 2020). No se ha incluido los valores de las variables correspondientes a los años 2021 y 2022 debido a la numerosa falta de información de varios países.

Los 17 países de la Eurozona incluidos en el presente TFG son los siguientes: Alemania, Austria, Bélgica, España, Eslovaquia, Eslovenia, Estonia, Finlandia, Francia, Grecia, Irlanda, Italia, Letonia, Lituania, Luxemburgo, Países Bajos y Portugal.

2) Matriz de correlaciones entre indicadores

El siguiente paso es evaluar la relación lineal existente entre todas las variables (dependiente-independiente, y dependiente-dependiente). Para ello, se calcula el coeficiente de correlación de Pearson (R^2) entre cada par de variables, lo que permite evaluar la fuerza y dirección de la relación entre ellas. Además, este coeficiente permite seleccionar las variables más importantes para el modelo, e identificar posibles problemas de multicolinealidad entre las variables independientes.

Inicialmente se ha planteado el problema de la siguiente manera: se ha calculado, para cada país, la matriz de correlación entre pares de variables. El problema que se ha encontrado es que los resultados varían mucho entre países, lo que impide llegar a una buena conclusión común.

A modo de ejemplo, en la Imagen 1 y en la Imagen 2 se observa la matriz de correlación para los datos de Italia y Países Bajos respectivamente. Como se puede observar, no existe ningún patrón para los valores obtenidos. Estas diferencias son notables entre todas las matrices, por lo que se concluye que no se puede llegar a ninguna buena conclusión utilizando este planteamiento.

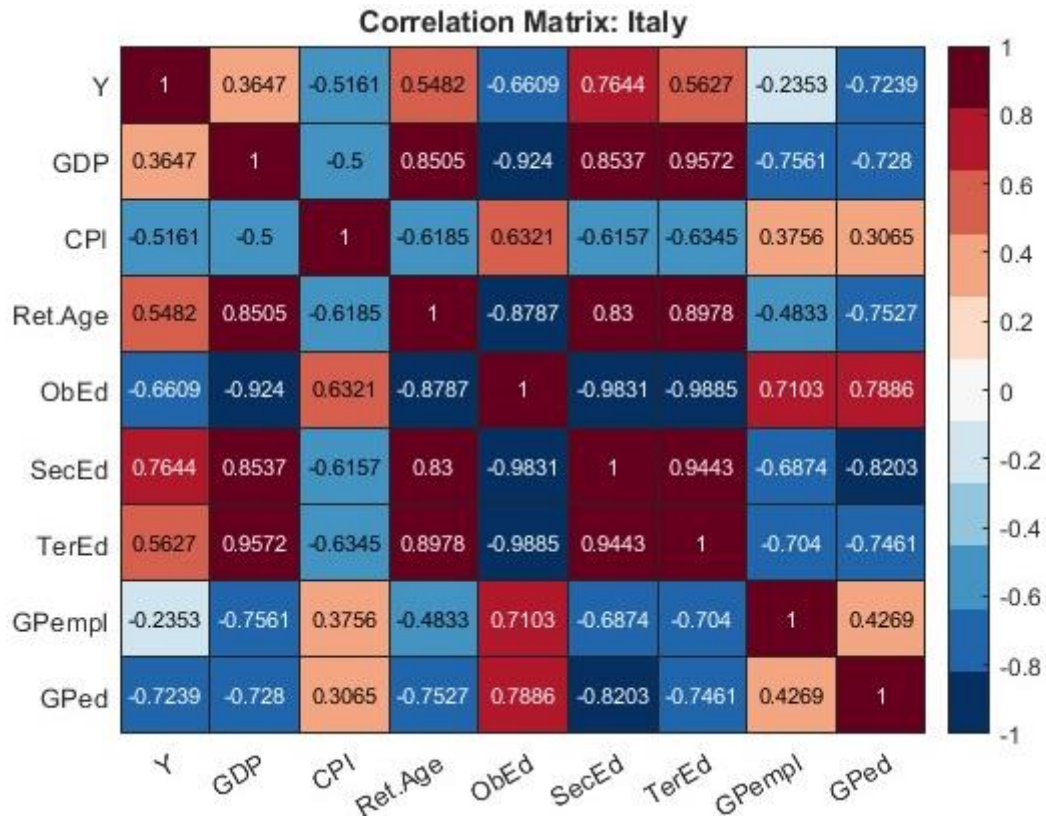


Imagen 1: Matriz de correlación para Italia

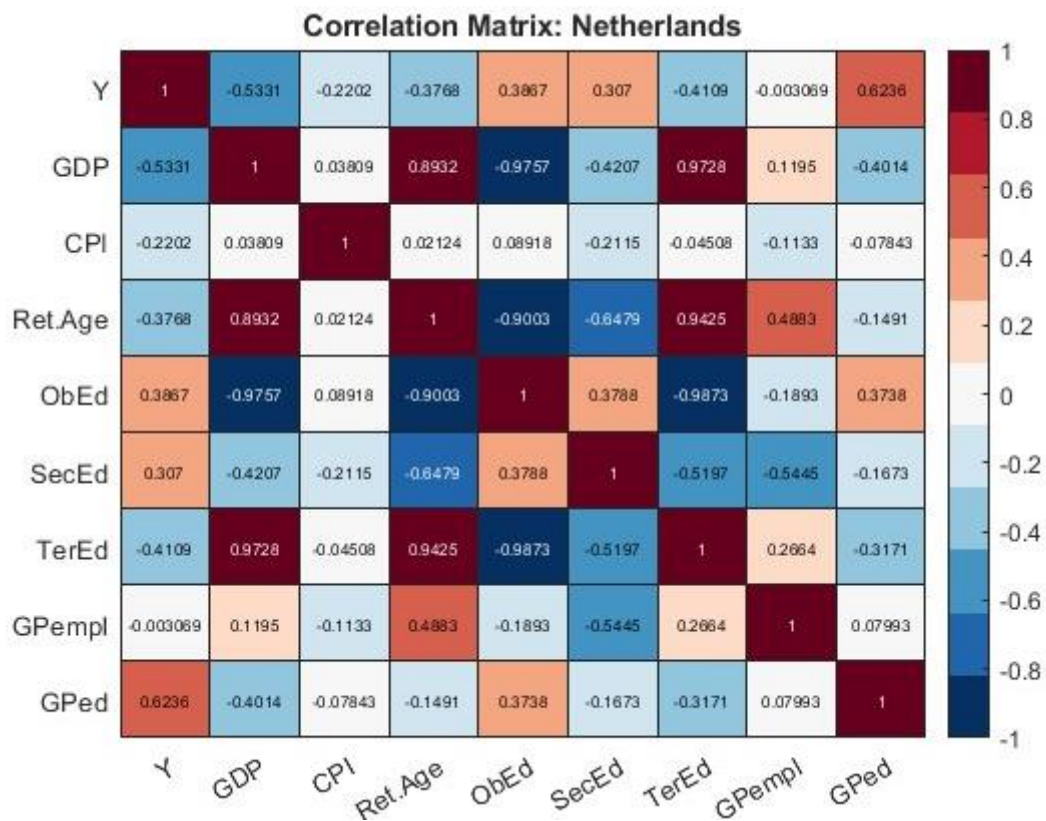


Imagen 2: Matriz de correlación para Países Bajos

Por lo que, ante la variabilidad observada en los resultados de la matriz de correlación para cada país, se combinan los datos de los 17 países en una única serie temporal ampliada, que representa un “macro país” con un total de 272 observaciones, como si fuesen 272 años. De esta manera, se busca generar una muestra más robusta y amplia que permita obtener conclusiones más significativas y generalizables. Dicha matriz se puede ver en la Imagen 3.

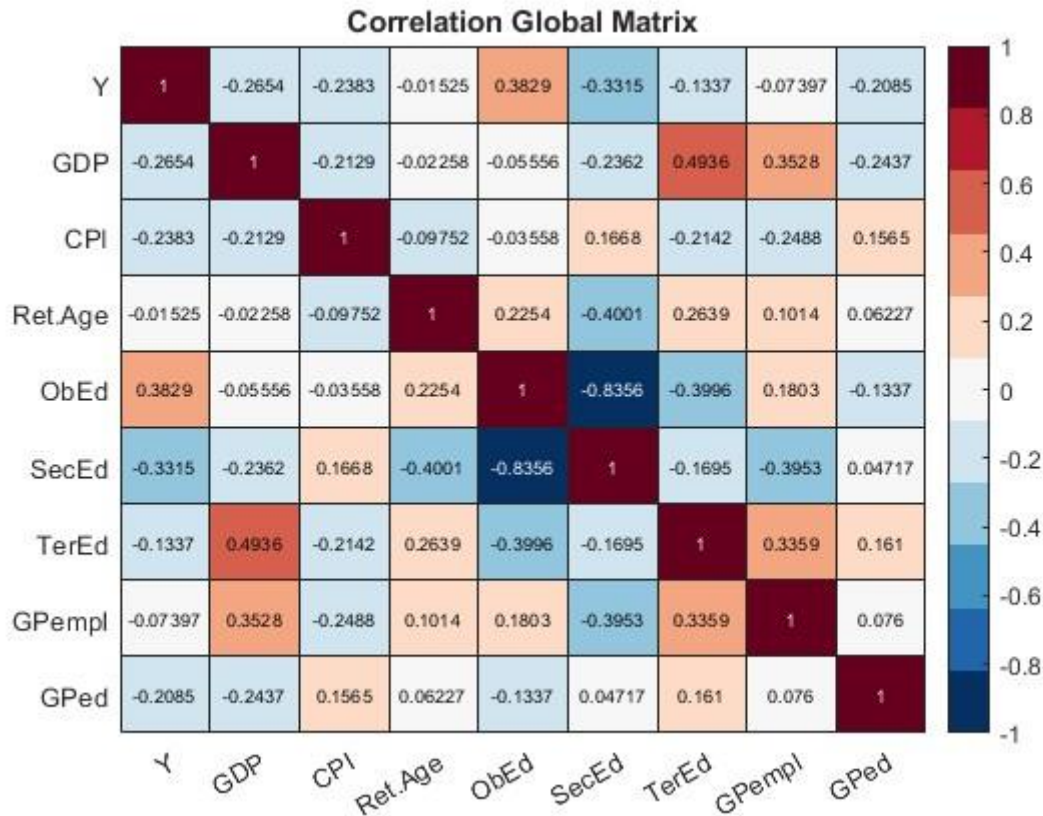


Imagen 3: Matriz de correlación global

Mukaka (2012) establece que se considera una correlación alta (de intensidad media, alta o muy alta) entre dos variables cuando su valor supera $\pm 0,5$. Por tanto, según este criterio, la única correlación entre variables que puede generar problemas de multicolinealidad es la existente entre la Educación Básica y la Secundaria. Esto resulta lógico, debido a la definición de las propias variables que se dio anteriormente.

Si consideramos que la definición de cada variable es el "nivel máximo de educación completada", podemos observar que la variable "Educación Secundaria" (SecEd) engloba tanto la educación obligatoria (Educación Primaria y Educación Secundaria Obligatoria) como el bachillerato y formación profesional de 1 año de

duración. Aunque el bachillerato no es obligatorio en España, cada vez más jóvenes eligen cursarlo antes de salir al mercado laboral. Y si no, en caso de no hacerlo, es común realizar una formación profesional. Es por ello por lo que se entiende que la educación secundaria abarca la finalización de la educación básica y, en consecuencia, ambas variables expresan la misma información. Esta situación en la que existe una alta correlación entre dos variables independientes se llama multicolinealidad, y puede dificultar la interpretación de los coeficientes y generar problemas en la estimación y fiabilidad de los resultados. Es por ello por lo que, para solucionar este problema de multicolinealidad, se escoge eliminar la variable “Educación Básica”. Es importante destacar que esta correlación no se presenta en el caso de la "Educación Terciaria", ya que obtener un título universitario no es un requisito obligatorio para ingresar al mercado laboral.

Una vez resuelta la multicolinealidad, pasamos a analizar la correlación entre la variable dependiente Y (Youth Unemployment) y la variable independiente senior (Retirement Age). El valor del coeficiente R^2 obtenido es de -0,01525. Esto indica que existe una correlación muy débil entre ambas variables, lo que podría sugerir que la edad de jubilación no es crucial a la hora de explicar las variaciones en el desempleo juvenil. Sin embargo, para obtener una comprensión más completa, y poder llegar a conclusiones válidas, se considera necesario analizar los resultados del modelo completo propuesto.

3) Panel Data

Por tanto, una vez hecho en análisis de correlación entre variables, el modelo final es el siguiente:

$$\text{Youth Unemployment} \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd \\ + GPempl + GPed$$

Dado que se está trabajando con datos de serie temporal (ya que, como se ha explicado anteriormente, queremos evaluar el efecto del desempleo juvenil en 17 países a lo largo de 16 años), la mejor manera de trabajar con ellos es ordenarlos en formato de Panel Data. Al ordenar las observaciones de esta manera, se obtiene una visión más completa de cómo varían los efectos del desempleo juvenil en cada país a lo largo del tiempo.

A continuación, en la Tabla 3, se proporciona una muestra, a modo de ejemplo, de la estructura final de los datos al ordenarlos en forma de panel. No se incluyen todas las observaciones debido a su cantidad de 272.

Country	Year	Y	GDP	CPI	Retirement Age	SecEd	TerEd	GPempl	GPed
Austria	2005	10,98	35024,56	2,30	58,24	59,60	14,80	0,441	4,80
Austria	2006	9,75	37659,84	1,44	58,93	59,40	14,50	0,521	4,80
Austria	2007	9,41	39436,42	2,17	58,04	59,30	14,30	0,49	4,68
Austria	2008	8,49	41316,02	3,22	57,72	59,70	14,80	0,495	4,84
Austria	2009	10,65	40929,34	0,51	57,98	59,70	15,80	0,637	5,11
Austria	2010	9,44	42020,55	1,81	58,74	60,10	16,10	0,638	5,11
Austria	2011	8,91	44469,21	3,29	59,20	60,30	16,20	0,551	4,98
Austria	2012	9,39	46477,66	2,49	60,44	60,50	16,70	0,551	5,00
Austria	2013	9,66	47936,68	2,00	60,96	60,00	17,40	0,587	5,02
Austria	2014	10,27	48813,53	1,61	61,19	52,40	26,80	0,617	4,93
Austria	2015	10,57	49942,06	0,90	61,04	52,40	27,40	0,576	4,93
Austria	2016	11,23	52665,09	0,89	61,29	51,70	28,20	0,586	4,89
Austria	2017	9,78	54188,36	2,08	62,03	51,20	28,90	0,596	4,84
Austria	2018	9,36	56956,11	2,00	62,11	51,10	29,40	0,581	4,78
Austria	2019	8,51	59719,33	1,53	62,50	50,40	30,40	0,525	4,79
Austria	2020	10,44	57253,30	1,38	62,46	50,20	30,50	0,561	5,05
Belgium	2005	21,47	33176,68	2,78	58,15	35,40	26,40	0,476	5,69
Belgium	2006	20,50	35253,92	1,79	58,46	35,50	27,00	0,459	5,63
Belgium	2007	18,83	36794,23	1,82	59,27	36,20	27,40	0,472	5,52
Belgium	2008	17,98	37883,23	4,49	59,60	37,10	27,80	0,513	5,72

Tabla 3: Datos ordenados en formato panel

7. CONSTRUCCIÓN DEL MODELO DE REGRESIÓN LINEAL

Análisis previo de la variabilidad de los datos

Antes de proponer los modelos de regresión lineal, se ha considerado importante realizar un análisis de la variabilidad en la tasa de desempleo juvenil a lo largo de los años y entre países. Para ello, se han utilizado diagramas de cajas. Dichos diagramas permiten identificar posibles patrones, tendencias y diferencias significativas en la tasa de desempleo juvenil a nivel temporal y geográfico. De tal manera, se obtiene una

buena representación visual de la variabilidad de los datos recopilados, lo que facilita la interpretación de los resultados obtenidos.

Diagrama de cajas por año

En el siguiente Diagrama 1, cada caja representa un año, y dentro de cada caja están los datos correspondientes a todos los países. Este diagrama permite detectar la variabilidad de la tasa de desempleo juvenil en cada año, y posibles diferencias entre países para cada año específico.

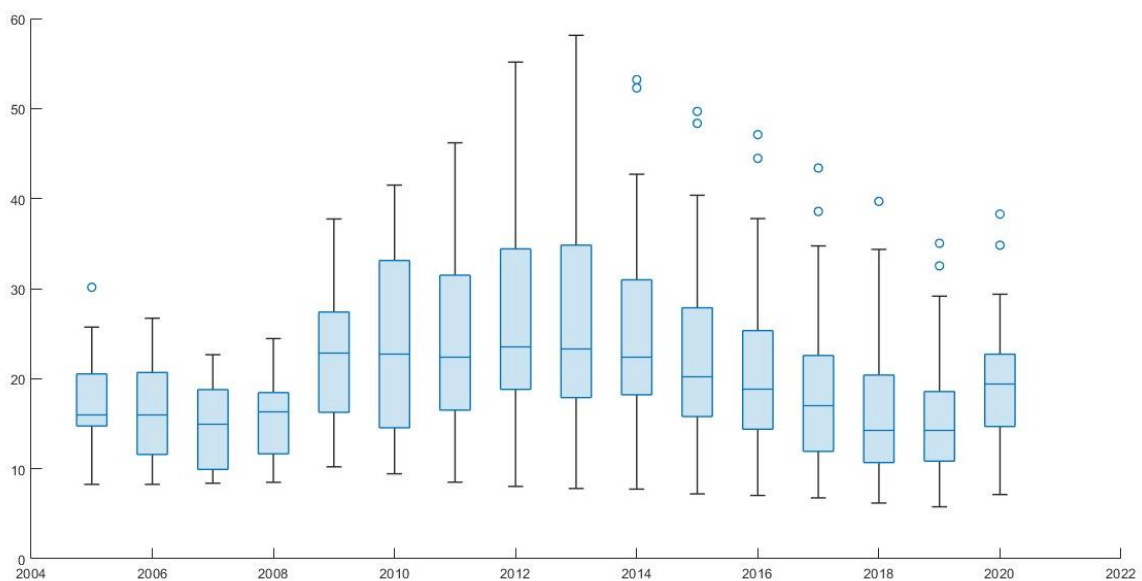


Diagrama 1: Diagrama de cajas por año

Al analizar el diagrama de cajas, se observan varios valores atípicos durante los años 2014 y 2020. Esto podría indicar situaciones excepcionales y fluctuaciones en la tasa de desempleo juvenil entre países durante ese periodo.

Por otro lado, se observa una similitud en los valores de la mediana desde 2010 hasta 2014, lo que sugiere estabilidad en la tasa de desempleo juvenil durante ese intervalo. Adicionalmente, la dispersión de las cajas en esos años también muestra una variabilidad similar, lo que indicaría que los factores estructurales que influyen en la tasa de desempleo juvenil se mantuvieron relativamente constantes durante esos años.

Además, los valores extremos no considerados como *outliers* varían para cada año, lo que refleja diferencias individuales entre los países en términos de la tasa de desempleo juvenil.

Diagrama de cajas por país

En el siguiente Diagrama 2, cada caja representa un país, y dentro de cada caja están los datos correspondientes a todos los años para cada país. Este diagrama permite detectar la variabilidad de la tasa de desempleo juvenil en cada país, y posibles diferencias entre años para cada país específico.

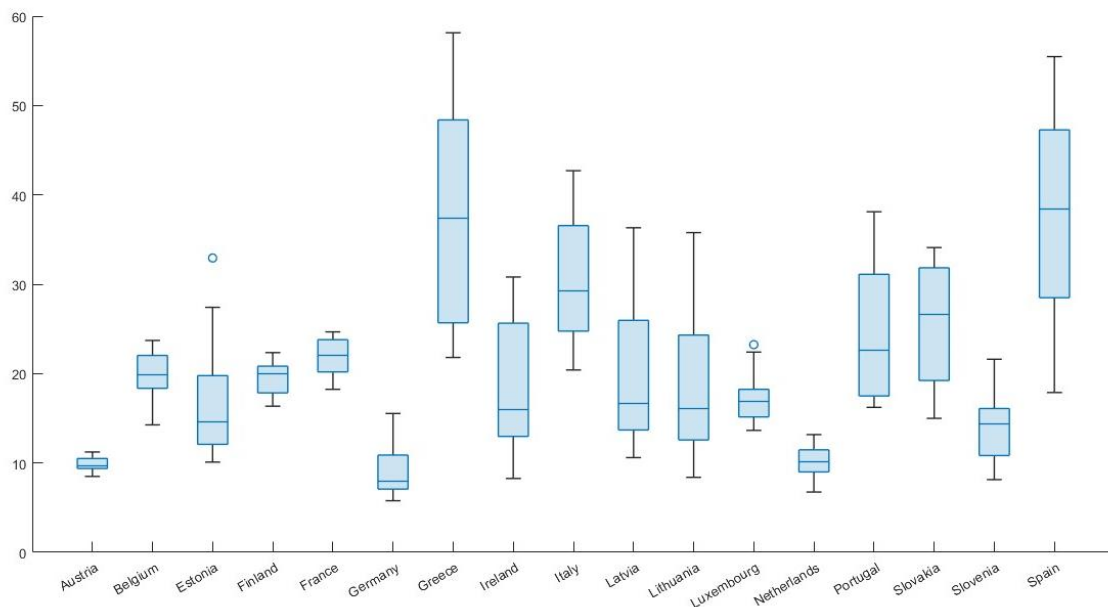


Diagrama 2: Diagrama de cajas por países

En el diagrama de cajas por países se observa una variabilidad en las dispersiones de las cajas, sin seguir un patrón específico. Algunos países, como Grecia, muestran una alta dispersión, mientras que otros, como Austria, presentan una dispersión más leve. Por otro lado, los intervalos de confianza no se superponen, lo que indicaría el uso de algún factor de agrupamiento en el modelo de regresión.

Modelo de efectos fijos y modelo de efectos aleatorios

En un análisis de datos de tipo panel, resulta de especial relevancia considerar la inclusión de efectos fijos y efectos aleatorios en el modelo. Bollen y Brand (2010)

establecen que la principal ventaja de esto es que se controlan las variables omitidas invariantes en el tiempo. No obstante, surgen algunos problemas, como la incertidumbre de decidir entre aplicar efectos fijos (FEM), aleatorios (REM), o ambos.

Por esta razón, se exploran distintas opciones de modelos para analizar la variabilidad temporal en la tasa de desempleo juvenil. Entre ellos está el modelo simple (sin FEM o REM), modelos que utilizan exclusivamente FEM (efectos fijos) y modelos que emplean únicamente REM (efectos aleatorios). Los resultados previamente explicados en el análisis de los diagramas de cajas sugieren la existencia de diferencias significativas en la tasa de desempleo juvenil entre países y a lo largo de los años. Los *outliers* observados durante 2014-2020 sugieren el uso de efectos aleatorios para considerar dichas variaciones atípicas. Por otro lado, la similitud entre medianas y dispersión entre los años 2010 y 2014 sugiere la presencia de factores estructurales constantes durante esos años, lo que podría sugerir la incorporación de efectos fijos en el modelo. Por otro lado, la variabilidad en las dispersiones de las cajas del diagrama de cajas entre países indica la existencia de diferencias individuales entre países, lo que podría sugerir la inclusión de efectos aleatorios en el modelo.

Sin embargo, es importante realizar una comparación entre los diferentes modelos para determinar cuál de ellos describe mejor la variabilidad temporal en la tasa de desempleo juvenil. En el momento de interpretar resultados a nivel internacional, nos centramos únicamente en España, con el fin de resumir las principales conclusiones.

i. Modelo simple (modelo 1)

En este modelo, no se incluye ningún efecto fijo ni variable.

$$\text{YouthUnemployment} \sim 1 + \text{GDP} + \text{CPI} + \text{RetirementAge} + \text{SecEd} \\ + \text{TerEd} + \text{GPempl} + \text{GPed}$$

ii. Modelo con Fixed Effects (FEM)

“La diferencia más obvia entre el FEM y el REM es la ausencia de las variables observadas invariantes en el tiempo y sus coeficientes. El FEM habitual no incluye explícitamente estas variables, sino que las integra en las variables latentes invariantes con el tiempo” (Bollen & Brand, 2010). Adicionalmente, estos autores establecen que el

uso de FEM permite que dichas variables invariantes se correlacionen con las variables que varían con el tiempo, como puede ser el desempleo.

Las variables latentes invariantes con el tiempo es lo que definimos como efectos fijos. En nuestro modelo, usaremos como efectos fijos la variable país (“Country”) y año (“Year”), previamente definida en el código como variable categórica, para que el modelo no la reconozca como variable numérica.

Al incluir la variable “Country” como efecto fijo, se registran las variaciones sistemáticas entre los países que pueden influir en la variable dependiente. Por tanto, se controlaría el impacto que cada país tiene sobre el desempleo juvenil, considerando las diferencias no observadas que podrían afectar los resultados. Esto es coherente con los resultados obtenidos al calcular la matriz de correlación de cada país, en el que se vio que cada país tenía su propio comportamiento independiente, ya que dichas matrices carecían de una tendencia en común. Esto también es coherente con lo observado en el diagrama de cajas del desempleo juvenil por países, en el que se vio que las dispersiones de las cajas no seguían un patrón específico.

Por otra parte, al incluir la variable “Year” como efecto fijo, se registran las variaciones sistemáticas que ocurren de un año a otro y que podrían influir en la variable dependiente. De esta manera, se controlan las variaciones temporales comunes a todos los países, lo que permite analizar cómo cambia el desempleo juvenil a lo largo de la franja de tiempo seleccionada.

a. Fixed Effect Country (modelo 2)

$$YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed + Country$$

b. Fixed Effect Year (modelo 3)

$$YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed + Year$$

c. Fixed Effect Country y Year (modelo 4)

$$YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed + Country + Year$$

iii. Modelo con Random Effects (REM)

Un modelo de efectos aleatorios (REM) considera la variabilidad aleatoria de ciertos factores que influyen en el resultado. Según Torres-Reyna (2007), se justifica este modelo en base a la variación entre los datos, la cual se asume como aleatoria y no correlacionada con las variables independientes incluidas en el modelo. Estos “efectos aleatorios permiten generalizar las inferencias más allá de la muestra utilizada en el modelo” (Torres-Reyna, 2007).

Con relación a los efectos aleatorios, Bates (2005) postula que están asociados a un factor de agrupamiento (variable categórica utilizada para agrupar las observaciones). Adicionalmente, explica que los términos de efectos aleatorios se especifican mediante un término lineal separado por “|” del factor de agrupamiento, y se interpretan como “dado” o “condicional a”. En nuestro modelo, las variables de agrupamiento serán país (“Country”) y año (“Year”), siendo los efectos aleatorios (1|Country) y (1|Year). Esto significa que, para cada observación dentro de un país específico, se asignará un efecto aleatorio que varía entre las observaciones dentro de ese país. Del mismo modo, para cada observación en un año particular, se asignará un efecto aleatorio que varía entre las observaciones dentro de ese año.

Por tanto, al incluir estas variables como factores de agrupamiento, se está controlando el efecto constante de cada país y año en la variable independiente (desempleo juvenil), teniendo en cuenta las diferencias no observadas entre los países, y las variaciones temporales comunes a todos los casos dentro de cada año. De esta manera, se considera y se ajusta los aspectos individuales y específicos de cada país y año que podrían influir en el desempleo juvenil, pero que no están siendo capturados por las variables del modelo.

d. Random Effect Country (modelo 5)

$$YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed + (1|Country)$$

e. Random Effect Year (modelo 6)

$$YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed + (1|Year)$$

f. Random Effect Country y Year (modelo 7)

$$YouthUnemployment \sim 1 + GDP + CPI + RetirementAge + SecEd + TerEd + GPempl + GPed + (1|Country) + (1|Year)$$

8. ANÁLISIS DE LOS RESULTADOS PARA CADA MODELO

Modelo 1

Model fit statistics:

AIC	BIC	LogLikelihood	Deviance
1913.3	1945.7	-947.63	1895.3

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
{'(Intercept)'} }	157.02	17.321	9.0655	264	2.8915e-17	122.92	191.12
{'GDP' }	-0.00029746	3.3035e-05	-9.0045	264	4.4276e-17	-0.00036251	-0.00023242
{'CPI' }	-1.3633	0.25957	-5.252	264	3.0995e-07	-1.8743	-0.85217
{'RetirementAge' }	-1.3981	0.26068	-5.3632	264	1.7867e-07	-1.9114	-0.88481
{'SecEd' }	-0.48443	0.04765	-10.166	264	1.052e-20	-0.57825	-0.39061
{'TerEd' }	0.23281	0.086391	2.6949	264	0.0074936	0.062709	0.40292
{'GPempl' }	-7.3263	2.3072	-3.1754	264	0.0016737	-11.869	-2.7835
{'GPed' }	-3.2878	0.62052	-5.2985	264	2.4648e-07	-4.5096	-2.066

Random effects covariance parameters (95% CIs):

Group: Error

Name	Estimate	Lower	Upper
{'Res Std'}	7.8852	7.2497	8.5765

En este modelo, todas las variables independientes tienen un p-valor < 0,05. Esto significa que todas son representativas con un nivel de confianza del 95%. Destaca que, aunque la variable Retirement Age sea significativa, su coeficiente es negativo. Esto establece una relación negativa entre la variable dependiente e independiente. Por lo tanto, sería “a mayor edad de jubilación, menor desempleo juvenil”.

En cuanto al signo de los coeficientes, todos son negativos menos TerEd. Esto significaría que, a mayor porcentaje de personas cuyo máximo nivel de educación completado es la educación terciaria, mayor tasa de paro juvenil. Esta afirmación parece contradictoria con lo popularmente entendido, y se comentará en profundidad más adelante. El signo del resto de los coeficientes del modelo confirma la relación esperada entre variables mencionada al principio del Trabajo de Fin de Grado.

Modelo 2

Model fit statistics:

AIC	BIC	LogLikelihood	Deviance
1732.4	1822.6	-841.22	1682.4

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
{'(Intercept)'} }	-48.442	25.25	-1.9185	248	0.056191	-98.174	1.2888
{'Country_Belgium' }	11.959	4.1199	2.9027	248	0.0040327	3.8444	20.073
{'Country_Estonia' }	-4.5716	3.7601	-1.2158	248	0.22521	-11.977	2.8342
{'Country_Finland' }	3.6636	3.3102	1.1068	248	0.26946	-2.856	10.183

{'Country_France' }	13.505	3.3703	4.0072	248	8.1275e-05	6.8675	20.143
{'Country_Germany' }	1.0279	2.2304	0.46085	248	0.64531	-3.365	5.4207
{'Country_Greece' }	38.908	4.1795	9.3093	248	7.137e-18	30.676	47.14
{'Country_Ireland' }	23.345	4.1311	5.6509	248	4.3689e-08	15.208	31.481
{'Country_Italy' }	40.442	3.9965	10.119	248	2.2541e-20	32.571	48.313
{'Country_Latvia' }	-1.8511	3.3393	-0.55433	248	0.57985	-8.4281	4.726
{'Country_Lithuania' }	-1.65	3.3411	-0.49386	248	0.62184	-8.2306	4.9305
{'Country_Luxembourg' }	33.004	4.6524	7.0938	248	1.36e-11	23.84	42.167
{'Country_Netherlands' }	6.9743	3.301	2.1128	248	0.03562	0.47269	13.476
{'Country_Portugal' }	39.182	6.7772	5.7815	248	2.2173e-08	25.834	52.53
{'Country_Slovakia' }	8.7242	2.8747	3.0348	248	0.0026631	3.0622	14.386
{'Country_Slovenia' }	-7.0579	2.6267	-2.687	248	0.0076974	-12.231	-1.8844
{'Country_Spain' }	49.529	6.6341	7.4658	248	1.4032e-12	36.462	62.595
{'GDP' }	-0.00030582	7.5437e-05	-4.0541	248	6.7423e-05	-0.0004544	-0.00015725
{'CPI' }	-0.86968	0.20351	-4.2735	248	2.7458e-05	-1.2705	-0.46886
{'RetirementAge' }	-0.20646	0.36288	-0.56894	248	0.56991	-0.92119	0.50827
{'SecEd' }	0.81732	0.17619	4.6388	248	5.6763e-06	0.4703	1.1643
{'TerEd' }	0.68227	0.15485	4.4059	248	1.5683e-05	0.37727	0.98726
{'GPempl' }	1.7281	2.2251	0.77662	248	0.43812	-2.6545	6.1107
{'GPed' }	5.1478	0.98275	5.2381	248	3.465e-07	3.2122	7.0833

Random effects covariance parameters (95% CIs):

Group: Error

Name	Estimate	Lower	Upper
{'Res Std'}	5.3323	4.9026	5.7998

En este modelo, todas las variables independientes tienen un p-valor $< 0,05$ menos Retirement Age y GPempl. Esto significa que todas son representativas con un nivel de confianza del 95%, excepto la edad de jubilación y el gasto público en empleo.

El p-valor de la variable Retirement Age es de 0,56991. Esto implica que no hay evidencia suficiente para afirmar que la edad de jubilación tiene una relación significativa con el desempleo juvenil.

En cuanto al signo del coeficiente de las variables significativas, las únicas que poseen una asociación inversa con la variable Desempleo Juvenil son GDP y CPI, lo cual ya se esperaba tras la justificación literaria de las variables del modelo. Destaca el signo del coeficiente asociado a la variable GPed, ya que estaría indicando que, a mayor porcentaje de gasto público destinado a la educación, mayor tasa de desempleo juvenil. Lo mismo pasa con la variable TerEd. Se profundizará en ello más adelante.

En este modelo se incluye la variable “Country” como efecto fijo. Los coeficientes (*estimates*) de cada país representan el impacto específico de cada país en el modelo, además de que proporciona información sobre cómo difiere la tasa de desempleo juvenil de cada país con respecto al país de referencia (Austria, al ser el primer país de nuestro panel data). Por otro lado, si nos fijamos en el p-valor asociado al coeficiente de cada país, un p-valor significativo (con un nivel de confianza del 95%) indica que el efecto del país correspondiente en el desempleo juvenil es significativamente diferente del efecto del país de referencia. Es decir, existen diferencias significativas en el desempleo juvenil entre países.

Enfocándonos en los resultados de España, se obtiene un coeficiente positivo de valor 49,529 que, sumado al valor del coeficiente de referencia, da 1,087. La interpretación de este resultado indica que, en comparación con Austria, se espera un incremento del 1,087% en la tasa de desempleo juvenil española, considerando todos los demás factores constantes. En cuanto al p-valor asociado, su valor es de aproximadamente 0. Esto quiere decir que existen diferencias estadísticamente significativas entre España y Austria en términos de su efecto individual en el desempleo juvenil, y dichas diferencias no se deben al azar.

Modelo 3

Model fit statistics:

AIC	BIC	LogLikelihood	Deviance
1840.4	1926.9	-896.19	1792.4

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
{'Intercept' }	183.75	15.761	11.659	249	2.3617e-25	152.71	214.79
{'Year_2006' }	-1.0175	2.2422	-0.45378	249	0.65038	-5.4336	3.3987
{'Year_2007' }	-2.5934	2.2608	-1.1471	249	0.25245	-7.0462	1.8595
{'Year_2008' }	0.55179	2.372	0.23263	249	0.81624	-4.1199	5.2234
{'Year_2009' }	7.8451	2.3664	3.3152	249	0.0010522	3.1844	12.506
{'Year_2010' }	10.668	2.2915	4.6553	249	5.2643e-06	6.1544	15.181
{'Year_2011' }	11.32	2.2681	4.9909	249	1.1297e-06	6.8527	15.787
{'Year_2012' }	13.386	2.2878	5.8513	249	1.5294e-08	8.8805	17.892
{'Year_2013' }	13.286	2.3632	5.6221	249	5.05e-08	8.6315	17.94
{'Year_2014' }	11.567	2.4645	4.6935	249	4.4373e-06	6.7132	16.421
{'Year_2015' }	9.1254	2.5233	3.6165	249	0.00036144	4.1557	14.095
{'Year_2016' }	8.3468	2.5181	3.3148	249	0.0010538	3.3874	13.306
{'Year_2017' }	7.0107	2.4351	2.879	249	0.0043362	2.2147	11.807
{'Year_2018' }	5.8633	2.4812	2.3631	249	0.018893	0.97648	10.75
{'Year_2019' }	5.4179	2.5215	2.1487	249	0.032624	0.45171	10.384
{'Year_2020' }	13.24	2.6493	4.9975	249	1.095e-06	8.0221	18.458
{'Year_2020' }	13.24	2.6493	4.9975	249	1.095e-06	8.0221	18.458
{'GDP' }	-0.00031179	2.7861e-05	-11.191	249	8.0971e-24	-0.00036666	-0.00025691
{'CPI' }	-0.28827	0.31169	-0.92486	249	0.35593	-0.90215	0.32561
{'RetirementAge' }	-1.8034	0.23807	-7.5754	249	7.0182e-13	-2.2723	-1.3346
{'SecEd' }	-0.56035	0.041625	-13.462	249	2.0917e-31	-0.64233	-0.47837
{'TerEd' }	0.11974	0.078863	1.5184	249	0.13019	-0.03558	0.27507
{'GPempl' }	-7.1175	2.0277	-3.5101	249	0.00053158	-11.111	-3.1238
{'GPed' }	-4.0679	0.54017	-7.5308	249	9.2707e-13	-5.1318	-3.0041

Random effects covariance parameters (95% CIs):

Name	Estimate	Lower	Upper
{'Res Std'}	6.5265	6.0005	7.0986

En este modelo, todas las variables independientes tienen un p-valor < 0,05 excepto CPI y TerEd. Esto significa que todas son representativas con un nivel de confianza del 95%, excepto la inflación y la educación terciaria, que no serían representativas.

En este caso, ocurre igual que con el modelo 1. La variable Retirement Age sí es representativa, pero su coeficiente es negativo, lo cual expresaría una relación inversa entre la edad de jubilación y el desempleo juvenil.

En cuanto al signo de los coeficientes, todos son negativos menos el asociado a la variable TerEd, que, al no ser representativa, no resulta de importancia. El resto de las variables con coeficiente negativo respaldan lo establecido en la parte de justificación literaria de las variables escogidas en el modelo.

En este modelo se incluye la variable “Year” como efecto fijo, previamente definida como variable categórica. Los coeficientes de cada año representan el impacto específico de cada año en el modelo, proporcionando información sobre cómo difiere la tasa de desempleo juvenil para cada año en comparación con el año de referencia (2005).

Centrándonos en los resultados obtenidos para 2020 (el año más actual disponible), se obtiene un valor de coeficiente estimado de 13,24. Este valor se interpreta como el cambio promedio en la tasa de desempleo juvenil en el año 2020 en comparación con el año 2005, para todos los países considerados. Su p-valor asociado es de valor aproximadamente 0, lo cual quiere decir que el efecto del año 2020 en la tasa de desempleo juvenil en comparación con el del año 2005 es estadísticamente significativo.

Modelo 4

Model fit statistics:

AIC	BIC	LogLikelihood	Deviance
1686.7	1831	-803.37	1606.7

Fixed effects coefficients (95% CIs):

Name	Estimate	SE	tStat	DF	pValue	Lower	Upper
{'Intercept' }	-31.244	30.681	-1.0184	233	0.30956	-91.691	29.203
{'Country_Belgium' }	12.04	3.6072	3.3376	233	0.00098351	4.9326	19.147
{'Country_Estonia' }	0.81851	3.5832	0.22843	233	0.81951	-6.241	7.8781
{'Country_Finland' }	6.0825	3.1849	1.9098	233	0.057392	-0.19248	12.357
{'Country_France' }	12.709	2.9976	4.2399	233	3.2258e-05	6.8035	18.615
{'Country_Germany' }	-0.49294	2.053	-0.2401	233	0.81046	-4.5378	3.5519
{'Country_Greece' }	33.381	4.3894	7.605	233	6.9694e-13	24.733	42.029
{'Country_Ireland' }	16.644	3.9629	4.2001	233	3.8007e-05	8.8367	24.452
{'Country_Italy' }	31.186	5.0433	6.1838	233	2.7803e-09	21.25	41.123
{'Country_Latvia' }	3.1343	3.0379	1.0317	233	0.30327	-2.851	9.1196
{'Country_Lithuania' }	2.6453	3.1202	0.84779	233	0.39742	-3.5021	8.7927
{'Country_Luxembourg' }	25.292	4.4165	5.7267	233	3.1423e-08	16.59	33.993
{'Country_Netherlands' }	4.0469	2.923	1.3845	233	0.16753	-1.712	9.8058
{'Country_Portugal' }	27.33	7.2265	3.7819	233	0.00019783	13.093	41.568
{'Country_Slovakia' }	10.736	2.6297	4.0828	233	6.1186e-05	5.5554	15.917
{'Country_Slovenia' }	-1.7525	2.4278	-0.72187	233	0.4711	-6.5357	3.0306
{'Country_Spain' }	39.032	6.5479	5.9609	233	9.2149e-09	26.131	51.932
{'Year_2006' }	-0.38232	1.607	-0.23791	233	0.81216	-3.5484	2.7838
{'Year_2007' }	-1.0989	1.6564	-0.66342	233	0.50771	-4.3622	2.1645
{'Year_2008' }	0.88455	1.8023	0.49078	233	0.62404	-2.6664	4.4355
{'Year_2009' }	1.8202	2.0219	0.90026	233	0.36891	-2.1633	5.8038
{'Year_2010' }	4.8703	2.0129	2.4195	233	0.016309	0.90439	8.8362
{'Year_2011' }	6.7553	2.0835	3.2423	233	0.0013593	2.6504	10.86
{'Year_2012' }	8.3051	2.2222	3.7373	233	0.00023409	3.9268	12.683
{'Year_2013' }	7.0738	2.4355	2.9045	233	0.0040324	2.2754	11.872
{'Year_2014' }	5.5137	2.6178	2.1062	233	0.036256	0.35605	10.671
{'Year_2015' }	3.7819	2.7108	1.3952	233	0.1643	-1.5588	9.1227
{'Year_2016' }	3.0719	2.8438	1.0802	233	0.28116	-2.5309	8.6747
{'Year_2017' }	2.4639	2.9052	0.8481	233	0.39725	-3.2599	8.1878
{'Year_2018' }	0.35273	3.152	0.11191	233	0.91099	-5.8574	6.5628
{'Year_2019' }	-0.30344	3.3732	-0.089957	233	0.9284	-6.9493	6.3424
{'Year_2020' }	0.096935	3.788	0.02559	233	0.97961	-7.3661	7.56

{'GDP' }	-0.00023104	7.4445e-05	-3.1035	233	0.002149	-0.00037771	-8.4368e-05
{'CPI' }	-0.8097	0.24122	-3.3567	233	0.00092096	-1.2849	-0.33445
{'RetirementAge' }	0.051287	0.36462	0.14066	233	0.88826	-0.66709	0.76967
{'SecEd' }	0.47538	0.19177	2.4789	233	0.013887	0.097557	0.8532
{'TerEd' }	0.3694	0.21953	1.6827	233	0.093769	-0.063109	0.80191
{'GPempl' }	2.7293	2.0928	1.3041	233	0.19348	-1.394	6.8526
{'GPed' }	2.3522	1.1014	2.1356	233	0.033751	0.18223	4.5221

Random effects covariance parameters (95% CIs):

Group: Error			
Name	Estimate	Lower	Upper
{'Res Std'}	4.6395	4.2656	5.0463

En este modelo, todas las variables independientes tienen un p-valor $< 0,05$ menos Retirement Age, TerEd y GPempl. Esto significa que todas son representativas con un nivel de confianza del 95%, excepto las mencionadas anteriormente, que no son representativas. Esto quiere decir que no existe evidencia suficiente para afirmar que dichas variables tienen una relación significativa con el desempleo juvenil.

Las únicas variables que poseen una asociación inversa con la variable Desempleo Juvenil son GDP y CPI, lo cual ya se esperaba tras la búsqueda literaria. Destaca el signo positivo del coeficiente asociado a GPed y SecEd, lo que resulta contradictorio con lo popularmente entendido. Como se ha mencionado anteriormente, se profundizará en ello más adelante.

En este modelo se incluye la variable “Country” y la variable “Year” (categórica) como efectos fijos. Centrándonos en los resultados del coeficiente asociado a España, se obtiene un valor de coeficiente de 39,032 que, sumado al valor del coeficiente de referencia, da 7,79. La interpretación de este resultado indica que, en comparación con Austria, se espera un incremento del 7,79% en la tasa de desempleo juvenil española, considerando todos los demás factores constantes. En cuanto al p-valor asociado, su valor es de aproximadamente 0. Esto quiere decir que existen diferencias estadísticamente significativas entre España y Austria en términos de su efecto individual en el desempleo juvenil, y dichas diferencias no se deben al azar. Estos resultados son muy parecidos a los que se obtuvieron con el modelo 2, que únicamente incluía como efecto fijo la variable “Country”.

Por otro lado, los resultados obtenidos del coeficiente asociado al año 2020 son un valor de coeficiente de 0,0996, y un p-valor asociado de valor 0,979. Como su p-valor es aproximadamente 1, el efecto del año 2020 en la tasa de desempleo juvenil en comparación con el del año 2005 no es estadísticamente significativo, y, por tanto, dicha relación carece de importancia estadística.

Modelo 5

```

Model fit statistics:
  AIC      BIC      LogLikelihood  Deviance
  1803.5   1839.5   -891.74       1783.5

Fixed effects coefficients (95% CIs):
  Name      Estimate      SE      tStat      DF      pValue      Lower      Upper
{'(Intercept)'} }      -6.0274      24.561     -0.2454     264      0.80634     -54.388     42.334
{'GDP'      }      -0.00029409      7.4019e-05     -3.9731     264      9.1578e-05     -0.00043983     -0.00014834
{'CPI'      }      -0.97627      0.20986     -4.6521     264      5.2026e-06     -1.3895     -0.56306
{'RetirementAge'} }      -0.25901      0.36844     -0.703      264      0.48267     -0.98447     0.46644
{'SecEd'    }      0.44964      0.15563     2.8893      264      0.0041818     0.14322     0.75607
{'TerEd'    }      0.57685      0.15446     3.7346      264      0.00023037     0.27272     0.88098
{'GPempl'   }      1.7051      2.3043      0.73995     264      0.45999     -2.8321     6.2422
{'GPed'     }      4.0991      0.98599     4.1573      264      4.3562e-05     2.1577     6.0405

Random effects covariance parameters (95% CIs):
Group: Country (17 Levels)
  Name1      Name2      Type      Estimate      Lower      Upper
{'(Intercept)'} }      {'(Intercept)'} }      {'std' }      13.912      8.9641     21.591

Group: Error
  Name      Estimate      Lower      Upper
{'Res Std' }      5.5578      5.086     6.0735

```

En este modelo, todas las variables independientes tienen un p-valor $< 0,05$ menos Retirement Age y GPempl. Esto significa que todas son representativas con un nivel de confianza del 95%, excepto la edad de jubilación y el gasto público en empleo.

Los únicos coeficientes positivos son los asociados al GDP y a la inflación. El resto son positivos, y, aunque parece que se contradice con lo que comúnmente se cree, se profundizará en la justificación de su signo más adelante.

En este modelo se está incluyendo como efecto aleatorio la variable “Country”. Esto permite capturar las diferencias entre los países en cuanto a su efecto en la variable dependiente Y. Por tanto, para interpretar los resultados de la parte aleatoria, resulta de interés fijarse en la varianza estimada asociada al efecto aleatorio de “Country”. Su valor es de 13,912 y su intervalo de confianza del 95% es [8,9641, 21,591]. Dado que este intervalo no incluye el valor 0, el término de efectos aleatorios es significativo. Esto quiere decir que el efecto aleatorio de “Country” captura las diferencias para cada país que no se explican por las variables independientes incluidas en el modelo

Modelo 6

```

Model fit statistics:
  AIC      BIC      LogLikelihood  Deviance
  1864.8   1900.8   -922.39       1844.8

Fixed effects coefficients (95% CIs):
  Name      Estimate      SE      tStat      DF      pValue      Lower      Upper
{'(Intercept)'} }      183.09      16.212     11.294      264      2.1627e-24     151.17     215.01
{'GDP'      }      -0.00030807      2.8679e-05     -10.742     264      1.4398e-22     -0.00036454     -0.0002516
{'CPI'      }      -0.56356      0.30081     -1.8735     264      0.062108     -1.1558     0.028734
{'RetirementAge'} }      -1.7029      0.24117     -7.061      264      1.4616e-11     -2.1778     -1.228
{'SecEd'    }      -0.54338      0.042491     -12.788     264      1.8184e-29     -0.62705     -0.45972
{'TerEd'    }      0.15028      0.079847     1.8822      264      0.060915     -0.0069334     0.3075

```

```

{'GPempl' } -7.3927 2.0739 -3.5646 264 0.00043247 -11.476 -3.3091
{'GPed' } -3.9712 0.554 -7.1682 264 7.6148e-12 -5.062 -2.8804

```

Random effects covariance parameters (95% CIs):

Group: Year (16 Levels)

```

Name1 Name2 Type Estimate Lower Upper
{'(Intercept)'} {'(Intercept)'} {'std'} 4.569 3.0184 6.9163

```

Group: Error

```

Name Estimate Lower Upper
{'Res Std'} 6.7409 6.1795 7.3532

```

En este modelo, todas las variables independientes tienen un p-valor $< 0,05$ menos CPI y TerEd. Esto significa que todas son representativas con un nivel de confianza del 95%, excepto la inflación y la educación terciaria.

A pesar de que la edad de jubilación sí sale representativa, su coeficiente es negativo, lo que establece una asociación inversa entre el desempleo juvenil y la edad de jubilación. El resto de los coeficientes de las variables representativas también son negativos, lo que se espera tras la justificación literaria de las variables del modelo.

Por otro lado, en este modelo se está incluyendo como efecto aleatorio la variable "Year". Esto permite capturar las diferencias entre los años en cuanto a su efecto en la variable dependiente Y. La varianza estimada asociada a este efecto aleatorio es de 4,569, y su intervalo de confianza del 95% es [3,0184, 6,9163]. Dado que este intervalo no incluye el valor 0, el término de efectos aleatorios es significativo. Esto quiere decir que el efecto aleatorio de "Year" captura las diferencias para cada año que no se explican por las variables independientes incluidas en el modelo

Modelo 7

```

AIC BIC LogLikelihood Deviance
1765.9 1805.5 -871.93 1743.9

```

Fixed effects coefficients (95% CIs):

```

Name Estimate SE tStat DF pValue Lower Upper
{'(Intercept)'} 29.79 24.698 1.2062 264 0.22883 -18.84 78.421
{'GDP'} -0.00025211 6.6086e-05 -3.8149 264 0.00016971 -0.00038223 -0.00012199
{'CPI'} -0.92236 0.23904 -3.8586 264 0.00014343 -1.393 -0.45169
{'RetirementAge'} -0.15936 0.3499 -0.45544 264 0.64916 -0.8483 0.52959
{'SecEd'} 0.037925 0.13075 0.29006 264 0.772 -0.21952 0.29537
{'TerEd'} 0.20472 0.16188 1.2647 264 0.2071 -0.11401 0.52346
{'GPempl'} 2.2861 2.1925 1.0427 264 0.29805 -2.0309 6.6032
{'GPed'} 0.93143 0.9869 0.94379 264 0.34614 -1.0118 2.8746

```

Random effects covariance parameters (95% CIs):

Group: Year (16 Levels)

```

Name1 Name2 Type Estimate Lower Upper
{'(Intercept)'} {'(Intercept)'} {'std'} 3.1526 2.0024 4.9633

```

Group: Country (17 Levels)

```

Name1 Name2 Type Estimate Lower Upper
{'(Intercept)'} {'(Intercept)'} {'std'} 8.6372 5.0831 14.676

```

Group: Error

```

Name Estimate Lower Upper
{'Res Std'} 4.9948 4.5578 5.4737

```

En este modelo, las únicas variables independientes que tienen un p-valor $< 0,05$ son GDP y CPI. Esto significa que únicamente el GDP y la inflación son representativas con un nivel de confianza del 95%.

En cuanto al signo de los coeficientes asociados a las variables representativas, es negativo. Esto afirma lo establecido en la justificación literaria de las variables del modelo de regresión.

En este modelo se incluye como efectos aleatorios las variables “Country” y “Year”. La varianza estimada asociada al efecto aleatorio “Country” es de valor 8,63, y su intervalo de confianza del 95% es [5,0831, 14,676]. Dado que este intervalo no incluye el valor 0, el término de efectos aleatorios “Country” es significativo. Por otro lado, la varianza asociada al efecto aleatorio “Year” es de valor 3,15 con un intervalo de confianza del 95% de [2,0024, 4,9633]. Dado que este intervalo no incluye al 0, el término de efectos aleatorios “Year” es significativo. Esto quiere decir que los efectos aleatorios de “Country” y “Year” capturan las diferencias para cada país y cada año que no se explican por las variables independientes incluidas en el modelo.

1) Tendencias comunes entre modelos

Estos modelos, a pesar de tener características distintas, comparten tendencias comunes entre ellos. A lo largo del análisis de los 7 modelos, se han identificado los siguientes patrones:

En primer lugar, la variable de Retirement Age no es significativa en los modelos 2, 4, 5 y 7, lo cual implica que no hay evidencia suficiente para afirmar que la edad de jubilación tiene una relación significativa con el desempleo juvenil. Por otro lado, aunque en los modelos 1, 3 y 6 sí sale representativa, el signo de su coeficiente es negativo. Esto significa que sí están relacionadas dichas variables, pero de manera inversa. Es decir, que, a mayor edad de jubilación, menor tasa de desempleo juvenil. Ambos resultados podrían responder a la cuestión planteada en la primera parte del Trabajo de Fin de Grado: en principio **no** existe ninguna relación entre el mercado juvenil y el senior, y, en caso de existir, es negativa.

La variable GDP es representativa en todos los modelos propuestos, y el signo de su coeficiente es negativo. La variable CPI tiene coeficiente negativo en aquellos

modelos en los que es representativa. Estas dos asociaciones concuerdan con lo establecido en el estudio literario.

En cuanto a las variables de educación (SecEd y TerEd), ninguna de ellas es significativa en todos los modelos. Destaca que el coeficiente de ambas variables es negativo en algunos modelos, lo cual indica una asociación indirecta (esperada tras el estudio bibliográfico de las variables). Sin embargo, en otros modelos se observa que dichos coeficientes son positivos. Dado que esta relación directa entre variables contradice las expectativas convencionales, se profundizará en ello en la sección siguiente.

Para finalizar, en cuanto a las variables de gasto público (GPempl y GPed), el gasto público en empleo es negativo en aquellos modelos en los que es una variable significativa. Esto quiere decir que, a mayor gasto público destinado a fomentar políticas de empleo, menor tasa de desempleo juvenil, lo cual coincide con la justificación literaria. Sin embargo, el signo de la variable GPed es distinto para cada modelo. Al igual que con las variables de educación, se profundizará en esto a continuación.

En la siguiente Tabla 4 se puede observar las similitudes entre todos los modelos en función de la significancia de las variables con un intervalo de confianza del 95%. Las variables representativas son aquellas cuya celda aparece coloreada en gris, y, para simplificar resultados, solo se pone el signo del coeficiente en aquellas variables que son representativas. El signo “+” o “-” hace referencia a la asociación de la correspondiente variable (+ = asociación directa, - = asociación indirecta).

	GDP	CPI	Ret. Age	SecEd	TerEd	GPempl	GPed
Modelo 1	-	-	-	-	-	-	-
Modelo 2	-	-		+	+		+
Modelo 3	-		-	-		-	-
Modelo 4	-	-		+			+
Modelo 5	-	-		+	+		+
Modelo 6	-		-	-		-	-
Modelo 7	-	-					

Tabla 4: Similitudes entre significancia de las variables para todos los modelos propuestos

2) Elección del mejor modelo

Una vez definidos las distintas opciones de modelos, se procede a escribir el correspondiente código en Matlab, para determinar cuál de ellos se ajusta mejor a los datos, y, por lo tanto, cuál es el más adecuado.

Existen diversos métodos para determinar cuál es el modelo óptimo, como puede ser el Akaike's Information Criterion (AIC), o el Bayesian's Information Criterion (BIC). Radosavljevic y Pituic (2023) postulan que “AIC se utiliza para predecir datos futuros como criterio de adecuación del modelo, mientras que BIC se utiliza para identificar modelos con las probabilidades más altas de ser modelos verdaderos para los datos”. De manera complementaria, Chakrabarti y Ghosh (2011) respaldan el uso del AIC como buen método cuando el objetivo principal es la predicción, al ser eficiente en muestras grandes al predecir un conjunto independiente de observaciones futuras basadas en los datos disponibles. A su vez, el BIC es más adecuado si se busca seleccionar el modelo correcto o verdadero, ya que elige el modelo que tiene la mayor probabilidad a posteriori.

Por otro lado, Vrieze (2012) establece que, según Shao (1997), el BIC es preferido cuando existe un modelo verdadero de dimensión fija y finita, mientras que el AIC es preferible cuando el modelo verdadero es demasiado complejo para estimar de manera paramétrica.

Es por ello por lo que se utilizará el coeficiente BIC para comparar modelos, siendo “el mejor modelo aquel que tenga un menor (...) BIC” (Mohammed, Naugler, & Far, 2015). Además, una ventaja del uso del coeficiente BIC es que “garantiza que el verdadero modelo será seleccionado a medida que aumenta el tamaño de la muestra, siempre y cuando el verdadero modelo esté entre los modelos candidatos que se están considerando (y se cumplan otras suposiciones)” (Vrieze, 2012).

A continuación, se presentan los valores de BIC obtenidos para cada modelo:

```
El modelo 1 tiene un BIC de 1945.710289, diferencia de BIC respecto al mejor modelo: 140.18
El modelo 2 tiene un BIC de 1822.588502, diferencia de BIC respecto al mejor modelo: 17.06
El modelo 3 tiene un BIC de 1926.913663, diferencia de BIC respecto al mejor modelo: 121.39
El modelo 4 tiene un BIC de 1830.965511, diferencia de BIC respecto al mejor modelo: 25.44
El modelo 5 tiene un BIC de 1839.532275, diferencia de BIC respecto al mejor modelo: 34.00
El modelo 6 tiene un BIC de 1900.831094, diferencia de BIC respecto al mejor modelo: 95.30
El modelo 7 es el mejor modelo con un BIC mínimo de 1805.53
```


Por tanto, el modelo propuesto más adecuado es el número 7, que corresponde a aquel formado por “Country” y “Year” como efectos aleatorios. El segundo mejor modelo es el número 4, que incluye como efecto fijo “Country” y “Year”, y el tercer mejor modelo es el número 2, que incluye “Country” como efecto fijo.

En caso de que se hiciese la selección del mejor modelo usando el criterio de AIC, se obtendrían los mismos tres mejores modelos. La diferencia está en que el mejor modelo sería el 4, es decir, aquel formado por “Country” y “Year” como efectos fijos.

```
Resultados de comparación de modelos utilizando el método del AIC:
El modelo 1 tiene un AIC de 1913.26, diferencia de AIC respecto al mejor modelo: 226.52
El modelo 2 tiene un AIC de 1732.44, diferencia de AIC respecto al mejor modelo: 45.71
El modelo 3 tiene un AIC de 1840.37, diferencia de AIC respecto al mejor modelo: 153.64
El modelo 4 es el mejor modelo con un AIC mínimo de 1686.73
El modelo 5 tiene un AIC de 1803.47, diferencia de AIC respecto al mejor modelo: 116.74
El modelo 6 tiene un AIC de 1864.77, diferencia de AIC respecto al mejor modelo: 178.04
El modelo 7 tiene un AIC de 1765.86, diferencia de AIC respecto al mejor modelo: 79.13
```

A continuación, se identifican las tendencias comunes encontradas entre los 3 mejores modelos.

3) Tendencias comunes entre los mejores modelos

En la Tabla 5 se puede observar las similitudes entre los tres mejores modelos, en función de la significancia de las variables con un intervalo de confianza del 95%. Se sigue el mismo criterio utilizado en la previa Tabla 4.

	GDP	CPI	Ret. Age	SecEd	TerEd	GPempl	GPed
Modelo 2	-	-		+	+		+
Modelo 4	-	-		+			+
Modelo 7	-	-					

Tabla 5: Similitudes entre significancia de las variables para los tres mejores modelos

Como se puede observar, la variable Retirement Age no es representativa para ninguno de los tres mejores modelos. Como bien se ha mencionado antes, esto indica que no existe evidencia suficiente para afirmar que la edad de jubilación presenta una relación significativa con el desempleo juvenil, que es lo que se pretendía investigar en el presente Trabajo Fin de Grado.

Las variables GDP y CPI son significativas para los 3 mejores modelos, además de que el signo del coeficiente es negativo para todos. Una vez más, esto confirma lo

esperado por el estudio literario. En cuanto al valor de su coeficiente, los coeficientes de las 2 variables en los 3 modelos poseen un valor muy bajo (menor que la unidad). Esto indica una relación moderada casi leve, ya que, cuanto más cerca esté el valor del coeficiente de 0, menor será la fuerza entre variables. Aunque se observa una relación débil en los 3 modelos debido a los bajos valores de los coeficientes, se explica a continuación el posible motivo de la relación inversa entre variable dependiente e independiente.

Una relación inversa entre el GDP y el paro juvenil puede explicarse a que un incremento en el GDP refleja un crecimiento económico y un aumento en la actividad económica. Esto suele generar un mayor número de oportunidades de empleo, lo que provocaría en una disminución de la tasa de desempleo juvenil.

Una relación inversa entre la inflación y el desempleo juvenil puede explicarse debido a que, en situaciones de elevada tasa de desempleo y una gran disponibilidad de trabajadores, se genera una mayor competencia laboral, lo que reduce el poder de negociación de los trabajadores. Como resultado, las empresas pueden ofrecer salarios más bajos debido a la disponibilidad de mano de obra, lo que reduce los costes laborales. Esto, a su vez, permite a las empresas disminuir el precio de los bienes y servicios que ofrecen (menor inflación), y se consigue fomentar la demanda que ha sido afectada negativamente por el aumento del desempleo.

En cuanto al resto de las variables. Las variables SecEd y GPed son representativas únicamente para los mejores modelos con efectos fijos (2 y 4).

El signo positivo del coeficiente de SecEd establece que, a mayor porcentaje de personas que han completado la educación secundaria (primaria, ESO, bachillerato y/o formación profesional de 1 año), mayor tasa de desempleo juvenil. Esta relación es coherente con la realidad actual ya que la mayoría de los trabajos demandan profesionales cada vez más cualificados y con mayor nivel educativo. La falta de una formación adicional más allá de la ESO puede limitar las oportunidades de empleo para los jóvenes, ya que se encontrarían en desventaja frente a aquellos que han obtenido niveles educativos superiores.

Por otro lado, el signo positivo del coeficiente de GPed establece que, a mayor gasto público destinado a educación (sin especificar a qué sector de educación se

refiere), mayor tasa de desempleo juvenil. Esta relación no resulta obvia, ya que contradice las expectativas convencionales. La razón de esta asociación positiva podría deberse a que un mayor gasto público en educación podría implicar una mayor financiación de programas educativos, becas, investigación y desarrollo.... Esto, a su vez, podría generar una mayor oferta de jóvenes cualificados, lo que supondría mayor competencia en el mercado laboral y, por tanto, mayor tasa de desempleo juvenil. Otro factor que podría influir en esta asociación positiva es un desajuste entre la oferta y la demanda en el mercado laboral. Esto se debe a que, si el gasto público en educación no se alinea con las necesidades del mercado laboral, pueden surgir desajustes entre las habilidades adquiridas por los jóvenes y las demandas del empleo, lo que resultaría en una mayor tasa de desempleo juvenil.

La variable TerEd únicamente es representativa con coeficiente positivo para el modelo 2 (“Country” como efecto fijo). Esto también contradice las expectativas convencionales, ya que normalmente se espera que a medida que aumenta el porcentaje de la población con niveles educativos superiores, las probabilidades de estar desempleado disminuyan. Esta asociación positiva podría deberse al aumento de la competencia en el mercado laboral a medida que cada vez más personas adquieren niveles educativos superiores. Por lo que, si la oferta de empleo no crece en la misma proporción que el número de personas cualificadas, y existe una demanda limitada de puestos de trabajo, los jóvenes se ven obligados a competir entre sí. Por tanto, aumenta la tasa de desempleo juvenil.

Es importante mencionar que, para entender adecuadamente el significado del signo de estos coeficientes, sería necesario un estudio más profundo y comprobar si dicho signo se debe a limitaciones del modelo, al filtrado de datos de la variable, o a razones similares a las que se han comentado anteriormente.

9. CONSTRUCCIÓN DEL MODELO INCREMENTAL

Como se ha visto en el apartado anterior, la variable Retirement Age no es representativa para ninguno de los tres mejores modelos utilizados. Por lo tanto, no se dispone de suficiente evidencia para afirmar que la edad de jubilación tiene un impacto significativo en el desempleo de los jóvenes.

Por tanto, considerando que la variable Retirement Age no proporciona una explicación significativa para el desempleo juvenil, el siguiente paso en este Trabajo de Fin de Grado consiste en identificar las variables que influyen de manera más relevante en el desempleo de los jóvenes. Para ello, se selecciona el modelo más adecuado a partir de los resultados obtenidos previamente (modelo 7 según el criterio BIC), y se lleva a cabo un enfoque incremental. De esta manera, al ir añadiendo variables de forma incremental, se puede identificar de manera sistemática y gradual qué variables contribuyen significativamente al desempleo juvenil, y cuáles no.

1) Modelo incremental mediante BIC

El procedimiento para implementar el modelo incremental es el siguiente. Se construye un modelo por variable independiente, que contenga únicamente la variable dependiente, la independiente de interés, y los términos aleatorios.

$$YouthUnemployment \sim 1 + vars\{i\} + (1|Country) + (1|Year)$$

Esta aproximación nos permite identificar qué variable es más significativa, y cuál tiene un mayor poder explicativo sobre la variable dependiente. A la hora de comparar modelos, se seguirá utilizando el BIC debido a la teoría previamente expuesta.

A continuación, se presentan los valores de BIC para cada modelo formado por el desempleo juvenil y cada variable, ordenados de menor a mayor.

```
1: BIC para Y y la variable GDP: 1791.4743
2: BIC para Y y la variable CPI: 1792.5752
3: BIC para Y y la variable GPed: 1806.0767
4: BIC para Y y la variable GPempl: 1806.7695
5: BIC para Y y la variable RetirementAge: 1807.0699
6: BIC para Y y la variable SecEd: 1807.9595
7: BIC para Y y la variable TerEd: 1808.0617
```

Por lo que, el modelo incremental empieza a construirse a partir de la variable de GDP. Si se hubiese usado AIC en vez de BIC, se hubiese llegado a la misma conclusión.

```
1: AIC para Y y la variable GDP: 1773.4453
2: AIC para Y y la variable CPI: 1774.5462
3: AIC para Y y la variable GPed: 1788.0477
4: AIC para Y y la variable GPempl: 1788.7405
5: AIC para Y y la variable RetirementAge: 1789.0409
6: AIC para Y y la variable SecEd: 1789.9305
7: AIC para Y y la variable TerEd: 1790.0327
```

Por tanto, se sigue el siguiente planteamiento para construir el modelo incremental. En primer lugar, se analiza el modelo inicial y se obtiene el BIC para dicho modelo base. Posteriormente, se prueba la adición de una variable adicional y se comprueba cuánto mejora el coeficiente BIC. Esto se repite con todas las variables disponibles, una a una. Se seleccionará el modelo que muestre una mejora en el BIC en comparación con el modelo anterior. Este proceso se repite hasta que no se observe una mejora significativa en el modelo, o cuando se introduzcan variables con p-valores no significativos. De esta manera se identifican las variables que mejor explican el desempleo juvenil desde un enfoque incremental.

Dicho planteamiento se resume en la siguiente Tabla 6. En las celdas se muestra el valor correspondiente a la diferencia entre el coeficiente BIC del modelo al que se le ha añadido la variable *i* correspondiente y el coeficiente BIC del mejor modelo. Para la iteración 1, el BIC del mejor modelo será el BIC del modelo base.

Por tanto, dado que se busca reducir el BIC, la variable seleccionada para entrar en el modelo incremental será aquella cuya diferencia sea mayor y negativa. Aparece coloreada de color amarillo. Una vez seleccionada para entrar en la siguiente iteración, se borra de las variables disponibles, y dicha casilla aparece para las siguientes iteraciones en color gris.

Por lo que, partiendo del modelo base ($Y_{YouthUnemployment} \sim 1 + GDP + (1|Country) + (1|Year)$) cuyo BIC es 1791,474, se obtiene:

BIC	CPI	Ret. Age	SecEd	TerEd	GPempl	GPed
Iteración 1	-10,45	5,10	5,04	2,83	2,96	5,54
BIC actualizado del mejor modelo: 1791,474 – 10,45 = 1781,029						
Iteración 2		5,53	5,13	4,18	4,02	4,93

Tabla 6: Variaciones del coeficiente BIC en el modelo incremental

Como se puede observar, ninguna variable reduce el BIC del mejor modelo (1781,029). Si examinamos los p-valores de las variables restantes al incluirlas en el modelo incremental, observamos que ninguna de ellas es significativa con un nivel de confianza del 95%. Esto se muestra en la siguiente Tabla 7:

p-valor	CPI	Ret. Age	SecEd	TerEd	GPempl	GPed
Iteración 1	0,000043	0,476	0,371	0,0758	0,104	0,788
Iteración 2		0,781	0,422	0,216	0,208	0,39

Tabla 7: P-valor para el modelo incremental

Por lo tanto, al no reducir el BIC y no ser variables significativas, no se produce ninguna mejora en el modelo durante la segunda iteración, y el proceso se detiene.

Por ello, el modelo final que mejor explica la variable desempleo juvenil es el siguiente:

$$YouthUnemployment \sim 1 + GDP + CPI (1|Country) + (1|Year)$$

2) Modelo incremental mediante AIC

Por otro lado, resulta de interés analizar los resultados obtenidos si el criterio de elección del mejor modelo hubiese sido usando el AIC. El objetivo de este enfoque es determinar si se obtendrían resultados similares utilizando AIC, que, como se ha visto antes, por lo general es más adecuado para la predicción, en comparación con BIC, que se normalmente se utiliza para seleccionar el modelo más verdadero en función de los datos disponibles.

Por tanto, usando el mismo código que para la parte anterior se obtienen los siguientes resultados que aparecen resumidos en la siguiente Tabla 8.

Por lo que, partiendo del modelo base ($YouthUnemployment \sim 1 + GDP + (1|Country) + (1|Year)$) cuyo AIC es 1773,445, se obtiene:

AIC	CPI	Ret. Age	SecEd	TerEd	GPempl	GPed
Iteración 1	-14,05	1,50	1,43	-0,77	-0,64	1,93
AIC actualizado del mejor modelo: $1773,445 - 14,05 = 1759,394$						
Iteración 2		1,92	1,53	0,57	0,42	1,33

Tabla 8: Variaciones del coeficiente AIC en el modelo incremental

Como se puede observar, ninguna variable reduce el AIC del mejor modelo (1759,394). Como en el caso anterior, si examinamos los p-valores de las variables

restantes al incluirlas en el modelo incremental, observamos que ninguna de ellas es significativa con un nivel de confianza del 95%. Esto se muestra en la siguiente Tabla 9:

p-valor	CPI	Ret. Age	SecEd	TerEd	GPempl	GPed
Iteración 1	0,000043	0,476	0,371	0,0758	0,104	0,788
Iteración 2		0,781	0,422	0,216	0,208	0,39

Tabla 9: P-valor para el modelo incremental

Por lo tanto, no se produce ninguna mejora en el modelo durante la segunda iteración, y el proceso se detiene. Destaca que los resultados obtenidos son los mismos que los obtenidos para el apartado anterior en el que se siguió el criterio BIC. Por lo que, al igual que en caso anterior, el modelo final que mejor explica la variable desempleo juvenil es el siguiente:

$$YouthUnemployment \sim 1 + GDP + CPI (1|Country) + (1|Year)$$

10. ANÁLISIS DE LOS RESULTADOS DEL MODELO INCREMENTAL

Como se ha podido comprobar en el apartado anterior, los modelos finales obtenidos mediante los criterios BIC y AIC coinciden. Para ambos criterios, el mejor modelo es $YouthUnemployment \sim 1 + GDP + CPI + (1|Country) + (1|Year)$.

Al no incluir la variable Retirement Age en el modelo, podemos concluir que no hay evidencia de que esta variable explique el paro juvenil. Por tanto, se puede afirmar que un aumento en la edad de jubilación no necesariamente conlleva un aumento en el desempleo juvenil.

Los resultados del modelo incremental se muestran a continuación:

```

Model fit statistics:
  AIC      BIC      LogLikelihood      Deviance
  1759.4   1781      -873.7              1747.4

Fixed effects coefficients (95% CIs):
  Name      Estimate      SE      tStat      DF      pValue      Lower      Upper
  {'(Intercept)'}      31.836      3.0819      10.33      269      2.7586e-21      25.768      37.903
  {'GDP' }      -0.00023059      5.4758e-05      -4.211      269      3.4708e-05      -0.00033839      -0.00012278
  {'CPI' }      -0.98079      0.23584      -4.1588      269      4.3071e-05      -1.4451      -0.51647

Random effects covariance parameters (95% CIs):
Group: Country (17 Levels)
  Name1      Name2      Type      Estimate      Lower      Upper
  {'(Intercept)'}      {'(Intercept)'}      {'std'}      7.7641      5.4847      10.991

Group: Year (16 Levels)
  Name1      Name2      Type      Estimate      Lower      Upper
  {'(Intercept)'}      {'(Intercept)'}      {'std'}      3.4731      2.3133      5.2143

```

Group: Error			
Name	Estimate	Lower	Upper
{'Res Std'}	5.0409	4.6083	5.5142

Se puede ver que los p-valores de las 2 variables que definen el desempleo juvenil (GDP y CPI) son aproximadamente nulos. Por tanto, estas 2 variables son significativas, y, por lo tanto, existe una relación estadísticamente significativa entre el GDP y la inflación con el desempleo juvenil. En cuanto al signo de su coeficiente, es negativo. Esto concuerda con lo visto en el estudio previo de la literatura, ya que es lo que establece la ley de Okun (producción) y la Ley de Phillips (inflación-precios).

La varianza estimada asociada al efecto aleatorio “Country” es de valor 7,7641, y su intervalo de confianza del 95% es [5,4847, 10,991]. Dado que este intervalo no incluye el valor 0, el término de efectos aleatorios “Country” es significativo. Por otro lado, la varianza asociada al efecto aleatorio “Year” es de valor 3,4731 con un intervalo de confianza del 95% de [2,3133, 5,2143]. Dado que este intervalo no incluye al 0, el término de efectos aleatorios “Year” es significativo. Esto quiere decir que los efectos aleatorios de "Country" y "Year" capturan las diferencias para cada país y cada año que no se explican por las variables independientes incluidas en el modelo.

11. CONCLUSIONES

En este presente Trabajo de Fin de Grado se ha estudiado el envejecimiento de la población y su impacto en el mercado laboral. El objetivo principal era estudiar si retrasar la edad de jubilación podría ser una solución viable para afrontar los desafíos demográficos y económicos derivados del envejecimiento de la población. Esta cuestión genera preocupación en cuanto a las oportunidades laborales de los jóvenes, ya que se cree que, si los trabajadores seniors no se jubilan, no dejan espacio para que entren en el mercado laboral los más jóvenes. Por tanto, se ha estudiado cuantitativamente si existe verdaderamente una relación entre la edad de jubilación y la tasa de desempleo juvenil.

En primer lugar, los resultados indican que la variable de Retirement Age no es representativa en la mayoría de los modelos analizados, lo que indica que no existe evidencia suficiente para afirmar una relación entre la edad de jubilación y el desempleo juvenil. En los modelos donde la variable de edad de jubilación sí es representativa, su coeficiente presenta un signo negativo. Esto indica que, en caso de existir dicha

relación, esta sería inversa. Es decir, a mayor edad de jubilación, menor tasa de desempleo juvenil.

Estos resultados sugieren que prolongar la vida laboral de los seniors no perjudica necesariamente las oportunidades de empleo para los jóvenes. Por tanto, al retrasar la edad de jubilación, nos encontraríamos frente a una posible solución para el problema de las pensiones. Adicionalmente, al sacar provecho de todas las ventajas que una sociedad longeva ofrece, se lograría impulsar la economía del país.

Sin embargo, es importante tener en cuenta que estos resultados se basan en el contexto específico estudiado, y pueden estar influenciados por diversos factores. Por ejemplo, es posible que los resultados también se vean influenciados por algunas limitaciones de los modelos empleados, así como por otras variables incluidas en el modelo, la forma en que se han filtrado y la posible omisión de variables relevantes.

Por otro lado, al observar que la edad de jubilación no explica el desempleo juvenil, se procedió a analizar qué variables podrían estar relacionadas. Para ello se llevó a cabo un modelo incremental. Los resultados mostraron que el PIB y la inflación desempeñan un papel significativo en la explicación del paro juvenil, que es lo que explica la Ley de Okun y la Curva de Phillips.

Como conclusión final, estos resultados abren nuevas perspectivas para futuras políticas que busquen promover un equilibrio entre el mercado laboral de trabajadores seniors y jóvenes, dejando a un lado los prejuicios existentes, y sacando provecho de las ventajas que este cambio demográfico supone.

12. BIBLIOGRAFÍA

- Alisa, M. (2015). Relationship between inflation and unemployment: a theoretical discussion about the Philips curve. *Journal of International Business and Economics*, 89-97.
- Bates, D. (2005). Fitting linear mixed models in R. *R News*.
- Bertoni, M., & Brunello, G. (2020). Does a Higher Retirement Age Reduce Youth Employment? *Economic Policy*, 325-372.
- Böheim, R., & Nice, T. (2019). The effect of early retirement schemes on youth employment. *IZA World of Labor*.
- Bollen, K. A., & Brand, J. E. (2010). A General Panel Model with Random and Fixed Effects: A Structural Equations Approach. *Social Forces*, 1-34.
- Chakrabarti, A., & Ghosh, J. K. (2011). AIC, BIC and Recent Advances in Model Selection. In *Philosophy of Statistics* (pp. 583-605).
- España, B. A. (2022, octubre 26). Cómo impacta la inflación en el empleo y en los salarios. *BBVA Asset Management*.
- Expansión. (2021). *La natalidad se reduce en España en 2021*. Retrieved from Datosmacro: <https://datosmacro.expansion.com/demografia/natalidad/espana>
- INE. (2022). *Proyecciones de Población 2022-2072*.
- Martín Román, Á. (2002). Ley de Okun, paro registrado y paro EPA. *Boletín económico del ICE, Información Comercial Española*, 11-16.
- Mestres Domènech, J. (2019). El envejecimiento de la población española y su impacto macroeconómico. *CaixaBank Research*(161), 100-111.
- Mitrakos, T., Tsakloglou, P., & Cholezas, I. (2010). Determining factors of youth unemployment in Greece with emphasis on tertiary education graduates. *Economic Bulletin*, pp. 21-62.
- Mohammed, E. A., Naugler, C., & Far, B. H. (2015). Chapter 32 - Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics.

- In Q. Nam Tran, & H. Arabnia, *Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology* (pp. 577-602).
- Mukaka, M. (2012). Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research . *Malawi Medical Journal*, 69-71.
- Portet, S. (2020). A primer on model selection using the Akaike Information Criterion. *Infectious Disease Modelling*, 111-128.
- Radosavljevic, D., & Pitulic, S. (2023). Application of nonlinear regression in recognizing distribution of signals in wireless channels. *Proceedings of the Estonian Academy of Sciences*, 72-105.
- Sánchez-Cabezudo Bayón, E., & del Olmo García, F. (2021). *Trabajadores senior y políticas socialmente responsables de Age Management*. Instituto Universitario de Análisis Económico y Social.
- Schmidheiny, K. (2016). Panel Data: Fixed and Random Effects. *Short Guides to Microeconomics*.
- Scott, A. J. (2021a). The longevity economy. *The Lancet Healthy Logevity*, e828-e835.
- Scott, A. J. (2021b). The longevity society. *The Lancet Healthy Longevity*, e820-e827.
- Torres-Reyna, Ó. (2007). Panel Data Analysis Fixed and Random Effects using Stata. *Data & Statistical Services, Princeton University*.
- Velázquez Orihuela, D. (2015). El efecto del gasto público en el ciclo económico: una visión alternativa. *Estudios Económicos (México, D. F.)*, 30(1), pp. 99-140.
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion. *Psychological Methods*, 228-243.

13. ANEXO

```
% Primera parte: Análisis de correlación entre variables
filename = 'DatosTFGfinal.xlsx';
países = {'Austria', 'Belgium', 'Estonia', 'Finland', 'France',
'Germany', 'Greece', 'Ireland', 'Italy', 'Latvia', 'Lithuania',
'Luxembourg', 'Netherlands', 'Portugal', 'Slovakia', 'Slovenia',
'Spain'} ;
T={'B2:B17", "C2:C17", "D2:D17", "E2:E17", "F2:F17", "G2:G17", "H2:H17",
", "I2:I17", "J2:J17", "K2:K17", "L2:L17", "M2:M17", "N2:N17", "O2:O17",
", "P2:P17", "Q2:Q17", "R2:R17"};

Rs1=cell(1,17);
matriz_países=cell(1,17);
matriz_vars_concatenado=[];

for i = 1:length(países)
    % Leer los datos en el rango especificado

Y=readcell("DatosTFGfinal.xlsx", "Sheet", "Parojuvenil", "Range", sp
rintf("%s", T{i}));
    GDP=
readcell("DatosTFGfinal.xlsx", "Sheet", "GDP", "Range", sprintf("%s"
, T{i}));
    CPI=
readcell("DatosTFGfinal.xlsx", "Sheet", "CPI", "Range", sprintf("%s"
, T{i}));
    RetirAge=
readcell("DatosTFGfinal.xlsx", "Sheet", "Edadjubilacion", "Range", s
printf("%s", T{i}));
    ObEd=
readcell("DatosTFGfinal.xlsx", "Sheet", "Educacionbasica", "Range",
sprintf("%s", T{i}));
```

```

    SecEd=
    readcell("DatosTFGfinal.xlsx","Sheet","Educacionsecundaria","Range",sprintf("%s", T{i}));
    TerEd=
    readcell("DatosTFGfinal.xlsx","Sheet","Educacionsuperior","Range",sprintf("%s", T{i}));
    GPempl=
    readcell("DatosTFGfinal.xlsx","Sheet","GPempleoMLMP","Range",sprintf("%s", T{i}));
    GPed=
    readcell("DatosTFGfinal.xlsx","Sheet","GPeducacion","Range",sprintf("%s", T{i}));
  
```

%Concatenar las variables para formar 1 matriz por país

```
vars=horzcat(Y,GDP,CPI,RetirAge,ObEd,SecEd,TerEd,GPempl,GPed);
```

%Pasar los valores de celdas a matrices

```
matriz_vars = cell2mat(vars);
```

%Comprobación que los datos se han importado bien desde

Excel

```
matriz_paises{i}=matriz_vars;
```

%Correlación de la matriz por país

```
Rs1{i}=corr(matriz_vars);
```

%Representacion grafica de la matriz de correlacion de cada país

%Nombre de los ejes

```
XDisplayLabels =
```

```
{'Y','GDP','CPI','Ret.Age','ObEd','SecEd','TerEd','GPempl','GPed'};
```

```

YDisplayLabels =
{'Y', 'GDP', 'CPI', 'Ret.Age', 'ObEd', 'SecEd', 'TerEd', 'GPempl', 'GPed
'};

%Representación gráfica
figure
heatmap(Rs1{i}, 'Colormap', redbluecmap, 'ColorLimits', [-1
1], 'XDisplayLabels', XDisplayLabels, 'YDisplayLabels',
YDisplayLabels)
title(sprintf('Correlation Matrix: %s', paises{i}));
colorbar

%Matriz global concatenando todos los países, tal que se
mantengan las
%columnas (indicadores) pero se aumente la franja temporal
de años

matriz_vars_concatenado=[matriz_vars_concatenado;matriz_vars];
end

%Matriz de correlacion global
Rs_global=corr(matriz_vars_concatenado)
%Nombre de los ejes
XDisplayLabels =
{'Y', 'GDP', 'CPI', 'Ret.Age', 'ObEd', 'SecEd', 'TerEd', 'GPempl', 'GPed
'};
YDisplayLabels =
{'Y', 'GDP', 'CPI', 'Ret.Age', 'ObEd', 'SecEd', 'TerEd', 'GPempl', 'GPed
'};
%Representación gráfica
figure

```

```
heatmap(Rs_global, 'Colormap', redbluecmap, 'ColorLimits', [-1  
1], 'XDisplayLabels', XDisplayLabels, 'YDisplayLabels',  
YDisplayLabels)  
title(sprintf('Correlation Global Matrix'));  
colorbar  
  
%Para mostrar las matrices de correlacion en pantalla  
for i = 1:numel(Rs1)  
    % Accede a la matriz de correlación i-ésima de la celda Rs  
    matriz_correlacion = Rs1{i};  
    % Muestra la matriz en pantalla  
    fprintf('Correlation matrix: %s\n', paises{i});  
    disp(matriz_correlacion);  
end  
  
% Segunda parte: Panel Data  
n_paises=17;  
n_years=16;  
  
%Crear vector de identificación de países  
id_pais_num=repelem(1:n_paises, n_years)';  
%Crear vector de nombres de países en orden correspondiente  
paises={'Austria', 'Belgium', 'Estonia', 'Finland', 'France',  
'Germany', 'Greece', 'Ireland', 'Italy', 'Latvia', 'Lithuania',  
'Luxembourg', 'Netherlands', 'Portugal', 'Slovakia', 'Slovenia',  
'Spain'};  
paises_rep = repelem(paises, n_years)';  
%Crear vector de identificación de países con nombres en lugar  
de números  
id_pais=cellstr(num2str(id_pais_num));  
id_pais(ismember(id_pais_num, 1:n_paises))=paises_rep;
```

```
%Crear un vector de identificadores de año
id_year=repmat(2005:2020, 1, n_paises)';

%Importo las observaciones para cada año-pais e indicador
load('matrizdatosfinal.mat', 'matriz_vars_concatenado');
%Quito la columna 5 correspondiente a educación básica
matriz_vars_concatenado(:,5)=[];

%PanelData
%Crear matriz de datos en formato panel
matriz_panel = num2cell(reshape(matriz_vars_concatenado,
n_years*n_paises, []));
%Agregar titulo a la matriz
titulos = {'Country', 'Year', 'YouthUnemployment', 'GDP', 'CPI',
'RetirementAge','SecEd', 'TerEd', 'GPempl', 'GPed'};
%Agregar columnas de identificación de país y año
matriz_panel_celda=
[titulos;id_pais,num2cell(id_year),matriz_panel];

%Paso la celda PanelData a tabla
tabla = cell2table(matriz_panel_celda(2:end,:), 'VariableNames',
titulos);

%Exportar tabla a Excel
filename = 'matrizdatosPanel2.xlsx';
writetable(tabla, filename)

%Abrir el archivo en Excel
winopen(filename);

%% Tercera parte: Construcción del Modelo
%Crear vector con los nombres de las variables independientes
```



```
vars = {'GDP', 'CPI', 'RetirementAge', 'SecEd', 'TerEd',  
'GPempl', 'GPed'};  
  
%Convertir la variable "Year" en categórica  
tabla.Year = categorical(tabla.Year);  
  
fprintf('Datos del modelo de regresión lineal completo. Modelo  
básico:\n');  
%CERO  
modelo_mixto0 = fitlme(tabla, 'YouthUnemployment ~ 1 + GDP + CPI  
+ RetirementAge + SecEd + TerEd + GPempl + GPed')  
fprintf('Datos del modelo de regresión lineal completo. Modelo  
fixed effects:\n');  
%UNO  
fprintf('Opción de modelo 1: siendo fixed effect "Country".  
\n');  
modelo_mixto1 = fitlme(tabla, 'YouthUnemployment ~ 1 + GDP + CPI  
+ RetirementAge + SecEd + TerEd + GPempl + GPed + Country')  
%DOS  
fprintf('Opción de modelo 2: siendo fixed effect "Year". \n');  
modelo_mixto2 = fitlme(tabla, 'YouthUnemployment ~ 1 + GDP + CPI  
+ RetirementAge + SecEd + TerEd + GPempl + GPed + Year')  
%TRES  
fprintf('Opción de modelo 3: siendo fixed effect "Country y  
"Year". \n');  
modelo_mixto3 = fitlme(tabla, 'YouthUnemployment ~ 1 + GDP + CPI  
+ RetirementAge + SecEd + TerEd + GPempl + GPed + Country +  
Year')  
  
fprintf('Datos del modelo de regresión lineal completo. Modelo  
random effects:\n');  
%CUATRO
```

```
fprintf('Opción de modelo 1:\n');
modelo_mixto4 = fitlme(tabla, 'YouthUnemployment ~ 1 + GDP + CPI
+ RetirementAge + SecEd + TerEd + GPempl + GPed + (1|Country)')
%CINCO
fprintf('Opción de modelo 2:\n');
modelo_mixto5 = fitlme(tabla, 'YouthUnemployment ~ 1 + GDP + CPI
+ RetirementAge + SecEd + TerEd + GPempl + GPed + (1|Year)')
%SEIS
fprintf('Opción de modelo 3:\n');
modelo_mixto6 = fitlme(tabla, 'YouthUnemployment ~ 1 + GDP + CPI
+ RetirementAge + SecEd + TerEd + GPempl + GPed + (1|Year) +
(1|Country)')

% Cuarta parte (1): Elección del modelo mediante BIC
%Creación de una celda que contenga los modelos
modelos={modelo_mixto0, modelo_mixto1, modelo_mixto2,
modelo_mixto3, modelo_mixto4, modelo_mixto5, modelo_mixto6};
num_modelos=7;

%Cálculo del BIC para cada modelo
BIC_vals = zeros(1, num_modelos);
for i = 1:num_modelos
    BIC_vals(i) = modelos{i}.ModelCriterion.BIC;
end

%Elección del mejor modelo, aquel con menor BIC
[min_BIC, mejor_modelo_idx] = min(BIC_vals);

%Resultados
fprintf('Resultados de comparación de modelos utilizando el
método del BIC:\n');
for i = 1:num_modelos
```

```
    if i == mejor_modelo_idx
        fprintf('El modelo %d es el mejor modelo con un BIC
mínimo de %.2f\n', i, BIC_vals(i));
    else
        fprintf('El modelo %d tiene un BIC de %f, diferencia de
BIC respecto al mejor modelo: %.2f\n', i, BIC_vals(i),
BIC_vals(i) - min_BIC);
    end
end

% Cuarta parte (2): elección del modelo mediante AIC
% Cálculo del AIC (Akaike Information Criterion) para cada modelo
AIC_vals2 = zeros(1, num_modelos);
for i = 1:num_modelos
    AIC_vals2(i) = modelos{i}.ModelCriterion.AIC;
end

% Elección del mejor modelo, aquel con menor AIC
[min_AIC, mejor_modelo_idx] = min(AIC_vals2);

% Resultados
fprintf('Resultados de comparación de modelos utilizando el
método del AIC:\n');
for i = 1:num_modelos
    if i == mejor_modelo_idx
        fprintf('El modelo %d es el mejor modelo con un AIC
mínimo de %.2f\n', i, AIC_vals2(i));
    else
        fprintf('El modelo %d tiene un AIC de %.2f, diferencia
de AIC respecto al mejor modelo: %.2f\n', i, AIC_vals2(i),
AIC_vals2(i) - min_AIC);
    end
end
```

end

```
% Quinta parte (1): modelo incremental parte 1. BIC
%Nombre variable objetivo
variableObjetivo = 'YouthUnemployment';

%Vector para almacenar los resultados
resultados=cell(1, length(vars));

%Vector para almacenar los valores de AIC
valoresBIC=zeros(1, length(vars));

%Ajuste para cada par de variables
for i = 1:length(vars)
    % Construir el término del modelo
    terminoModelo = [variableObjetivo ' ~ 1 + ' vars{i} ' +
(1|Country) + (1|Year)'];

    %Ajuste del modelo mixto
    fprintf('Modelo formado por la variable Y y la variable
%s:\n', vars{i});
    modelo_mixto = fitlme(tabla, terminoModelo)

    %Obtención de los valores de BIC
    valoresBIC(i) = modelo_mixto.ModelCriterion.BIC;
    fprintf('BIC para el modelo formado por la variable Y y la
variable %s es: %f \n', vars{i}, valoresBIC(i))
end

%Para buscar qué modelo es el que tiene menor BIC:
% Buscar el valor mínimo de BIC
[valorMinimo2, indiceMinimo2] = min(valoresBIC);
```

```
% Mostrar el valor mínimo de BIC y el modelo correspondiente
fprintf('El valor mínimo de BIC es: %.4f\n', valorMinimo2);
fprintf('El modelo correspondiente es el formado por la variable
Y y la variable %s\n', vars{indiceMinimo2});
```

```
%Clasificación de los valores de BIC de menor a mayor
[valoresBICOrdenados, indicesOrdenados] = sort(valoresBIC);
```

```
%Mostrar los pvalores ordenados y los modelos correspondientes
fprintf('Valores de BIC ordenados de menor a mayor:\n');
for i = 1:length(vars)
    indiceVariable = indicesOrdenados(i);
    fprintf('%d: BIC para Y y la variable %s: %.4f\n', i,
vars{indiceVariable}, valoresBICOrdenados(i));
end
```

```
% Quinta parte (2): modelo incremental parte 1. AIC
%Nombre variable objetivo
variableObjetivo = 'YouthUnemployment';
```

```
%Vector para almacenar los resultados
resultados=cell(1, length(vars));
```

```
%Vector para almacenar los valores de AIC
valoresAIC=zeros(1, length(vars));
```

```
%Ajuste para cada par de variables
for i = 1:length(vars)
    % Construir el término del modelo
    terminoModelo = [variableObjetivo ' ~ 1 + ' vars{i} ' +
(1|Country) + (1|Year)'];
```

```
%Ajuste del modelo mixto
fprintf('Modelo formado por la variable Y y la variable
%s:\n', vars{i});
modelo_mixto = fitlme(tabla, terminoModelo)

%Obtención de los valores de AIC
valoresAIC(i) = modelo_mixto.ModelCriterion.AIC;
fprintf('AIC para el modelo formado por la variable Y y la
variable %s es: %.4f \n', vars{i}, valoresAIC(i))
end

%Para buscar qué modelo es el que tiene menor AIC:
% Buscar el valor mínimo de AIC
[valorMinimo2, indiceMinimo3] = min(valoresAIC);

% Mostrar el valor mínimo de AIC y el modelo correspondiente
fprintf('El valor mínimo de AIC es: %.4f\n', valorMinimo2);
fprintf('El modelo correspondiente es el formado por la variable
Y y la variable %s\n', vars{indiceMinimo3});

%Clasificación de los valores de AIC de menor a mayor
[valoresAICOrdenados, indicesOrdenados2] = sort(valoresAIC);

%Mostrar los valores de AIC ordenados
fprintf('Valores de AIC ordenados de menor a mayor:\n');
for i = 1:length(vars)
    indiceVariable = indicesOrdenados2(i);
    fprintf('%d: AIC para Y y la variable %s: %.4f\n', i,
vars{indiceVariable}, valoresAICOrdenados(i));
end
```

```
% Sexta parte (1): modelo incremental parte 2 (en función de lo  
obtenido en la parte 1 y seleccionado mediante BIC)
```

```
baseModelFormula = 'YouthUnemployment ~ 1 + GDP';  
baseModelFormulaWithRandomEffects = [baseModelFormula ' +  
(1|Country) + (1|Year)'];  
baseModel = fitlme(tabla, baseModelFormulaWithRandomEffects);  
bestModel = baseModel;  
bestModel2 = baseModelFormula;  
  
bestBIC = baseModel.ModelCriterion.BIC;  
fprintf('El BIC para el modelo base [YouthUnemployment ~ 1 + GDP  
+ (1|Country) + (1|Year)] es de %f\n',  
baseModel.ModelCriterion.BIC);
```

```
%Variables del modelo, que son las mismas que el vector vars  
pero quitando GPD que aparece en el modelo base  
variables = {'CPI', 'RetirementAge', 'SecEd', 'TerEd', 'GPempl',  
'GPed'};
```

```
%Variables que aún no se han utilizado en el modelo  
variablesLeft = variables;
```

```
%PrevAIC inicial para poder inicializar el bucle  
prevBIC = Inf;
```

```
while ~isempty(variablesLeft) && bestBIC < prevBIC  
    %Vector para almacenar los valores de BIC  
    aicValues = zeros(length(variablesLeft), 1);  
  
    %El BIC del modelo base  
    prevBIC=bestBIC;
```

```
fprintf('----- \n');

for i = 1:length(variablesLeft)
    %Construir el modelo con la variable adicional
    fprintf('\n');
    currentFormula = [char(bestModel2) ' + '
char(variablesLeft{i}) ' + (1|Country) + (1|Year)'];
    fprintf('Modelo %s: \n', currentFormula);
    currentModel = fitlme(tabla, currentFormula);
    currentBIC = currentModel.ModelCriterion.BIC;
    bicValues(i) = currentBIC;
    fprintf('El BIC para el modelo que incluye %s es de
%f\n', variablesLeft{i}, currentBIC);
    % Calcular el incremento de BIC con respecto al mejor
modelo
    bicIncrement = currentBIC - prevBIC;
    fprintf('El incremento de BIC con respecto al mejor
modelo anterior es de %.2f\n', bicIncrement);

    fprintf('----- \n');
end

% Obtener el índice del modelo con el menor BIC
[~, bestIndex] = min(bicValues);

% Verificar si se mejora el BIC
if bicValues(bestIndex) < bestBIC %En caso de que mejore el
BIC
    % La formula del modelo se actualiza añadiendo la nueva
% variable (la de menor BIC)
    bestVariable = variablesLeft{bestIndex};
```



```
bestModel2 = [char(bestModel2) ' + '  
char(bestVariable)];  
    %Añado los términos aleatorios a la fórmula  
bestModelFormula = [bestModel2 ' + (1|Country) +  
(1|Year)'];  
    %Calculo el modelo  
bestModel = fitlme(tabla, bestModelFormula);  
bestBIC = bestModel.ModelCriterion.BIC;  
fprintf('El modelo se ha actualizado con la variable %s  
(BIC = %f)\n', bestVariable, bestBIC);  
    % Eliminar la variable utilizada del vector de variables  
restantes  
variablesLeft(bestIndex) = [];  
else  
    fprintf('No se mejora el BIC con las variables  
restantes. El proceso se ha detenido.\n');  
    break;  
end  
end  
  
fprintf('El mejor modelo encontrado es el que incluye las  
variables: %s\n', bestModel.Formula);  
fprintf('con un BIC de %f\n', bestBIC);  
  
% Sexta parte (2): modelo incremental parte 2 (en función de lo  
obtenido en la parte 1 y seleccionado mediante AIC)  
baseModelFormula_2 = 'YouthUnemployment ~ 1 + GDP';  
baseModelFormulaWithRandomEffects_2 = [baseModelFormula_2 ' +  
(1|Country) + (1|Year)'];  
baseModel_2 = fitlme(tabla,  
baseModelFormulaWithRandomEffects_2);  
bestModel_2 = baseModel_2;
```

```
bestModel2_2 = baseModelFormula_2;
bestAIC = baseModel_2.ModelCriterion.AIC;
fprintf('El AIC para el modelo base [YouthUnemployment ~ 1 + GDP
+ (1|Country) + (1|Year)] es de %f\n',
baseModel_2.ModelCriterion.AIC);

%Variables del modelo, que son las mismas que el vector vars
pero quitando la variable GPempl que aparece en el modelo base
variables_2 = {'CPI', 'RetirementAge', 'SecEd', 'TerEd',
'GPempl', 'GPed'};

%Variables que aún no se han utilizado en el modelo
variablesLeft_2 = variables_2;

%PrevAIC inicial para poder inicializar el bucle
prevAIC = Inf;

while ~isempty(variablesLeft_2) && bestAIC < prevAIC
    %Vector para almacenar los valores de AIC
    aicValues = zeros(length(variablesLeft_2), 1);

    %El AIC del modelo base
    prevAIC=bestAIC;

    fprintf('----- \n');

    for i = 1:length(variablesLeft_2)
        %Construir el modelo con la variable adicional
        fprintf('\n');
        currentFormula_2 = [char(bestModel2_2) ' + '
char(variablesLeft_2{i}) ' + (1|Country) + (1|Year)'];
        fprintf('Modelo %s: \n', currentFormula_2);
```

```
currentModel_2 = fitlme(tabla, currentFormula_2)
currentAIC = currentModel_2.ModelCriterion.AIC;
aicValues(i) = currentAIC;
fprintf('El AIC para el modelo que incluye %s es de
%f\n', variablesLeft_2{i}, currentAIC);
% Calcular el incremento de AIC con respecto al mejor
modelo
aicIncrement = currentAIC - prevAIC;
fprintf('El incremento de AIC con respecto al mejor
modelo anterior es de %.2f\n', aicIncrement);

fprintf('----- \n');
end

% Obtener el índice del modelo con el menor AIC
[~, bestIndex_2] = min(aicValues);

% Verificar si se mejora el BIC
if aicValues(bestIndex_2) < bestAIC %En caso de que mejore
el AIC
% La formula del modelo se actualiza añadiendo la nueva
% variable (la de menor AIC)
bestVariable_2 = variablesLeft_2{bestIndex_2};
bestModel2_2 = [char(bestModel2_2) ' + '
char(bestVariable_2)];
%Añado los términos aleatorios a la fórmula
bestModelFormula_2 = [bestModel2_2 ' + (1|Country) +
(1|Year)'];
%Calculo el modelo
bestModel_2 = fitlme(tabla, bestModelFormula_2);
bestAIC = bestModel_2.ModelCriterion.AIC;
```

```
fprintf('El modelo se ha actualizado con la variable %s  
(AIC = %f)\n', bestVariable_2, bestAIC);  
    % Eliminar la variable utilizada del vector de variables  
restantes  
    variablesLeft_2(bestIndex_2) = [];  
else  
    fprintf('No se mejora el AIC con las variables  
restantes. El proceso se ha detenido.\n');  
    break;  
end  
end  
  
fprintf('El mejor modelo encontrado es el que incluye las  
variables: %s\n', bestModel_2.Formula);  
fprintf('con un AIC de %f\n', bestAIC);
```