

GRADO EN ECONOMÍA

Curso 2022/2023

Impacto de la COVID-19 en el Mercado Laboral

Autora: Janire González Martínez

Director: Javier Fernández-Macho

Universidad: Universidad del País Vasco/Euskal Herriko Unibertsitatea

Bilbao, a 25 de septiembre de 2023



ÍNDICE

1. INTRODUCCIÓN Y OBJETIVOS	3
1.1 Contexto de la situación	3
1.2 Objetivos	5
2. BASE DE DATOS, VARIABLES Y METODOLOGÍA	6
2.1 Base de datos y variables contempladas	6
2.1.1 Variables endógenas	7
2.1.2 Variables explicativas	7
2.3. Metodología	11
3. ESTIMACIÓN Y ANÁLISIS DE LOS MODELOS ECONÓMICOS	14
3.1 Modelo con situación de empleo como variable dependiente	14
3.1.1 Análisis de los resultados del primer modelo	17
3.2 Modelo con tipo de contrato como variable dependiente	20
3.2.1 Análisis de los resultados del segundo modelo	24
4. CONCLUSIONES	27
5. BIBLIOGRAFÍA:	29

RESUMEN:

El presente trabajo estudia el impacto de la crisis mundial provocada por la COVID-19 en el mercado laboral español. Para llevarlo a cabo, se analiza el efecto de la pandemia sobre la situación del empleo y los contratos laborales en 2020 en función de características personales de los individuos. Se establecen dos modelos econométricos con microdatos extraídos de la Encuesta de Población Activa (EPA) que indican la forma en la que diferentes variables influyen en la situación laboral de los individuos. Como resultado se determina que aunque la COVID-19 ha afectado más a individuos con determinadas características, todos los ciudadanos han resultado perjudicados en mayor o menor medida.

Palabras clave: mercado laboral, COVID-19, España, modelo econométrico.

ABSTRACT:

This paper studies the impact of the global crisis caused by COVID-19 on the Spanish labour market. To do so, the effect of the pandemic on the employment situation and employment contracts in 2020 is analysed according to individuals' personal characteristics. Two econometric models are established with microdata obtained from the Encuesta de Población Activa (EPA) that indicate how different variables influence the employment status of individuals. As a result, it is determined that although COVID-19 has affected more individuals that have certain characteristics, all citizens have been negatively affected to a greater or lesser extent.

Keywords: labour market, COVID-19, Spain, econometric model.

1. INTRODUCCIÓN Y OBJETIVOS

1.1 Contexto de la situación

La pandemia de la COVID-19 iniciada en Wuhan, China, el 12 de diciembre de 2019, no tardó en extenderse y afectar todos los Estados del mundo. La crisis derivada de esta emergencia tuvo un origen sanitario, pero las consecuencias procedentes de ésta también fueron de carácter económico y social. Como resultado, la actividad económica mundial fue reducida drásticamente en un periodo de apenas tres meses, a medida que numerosos gobiernos imponían medidas de confinamiento con el fin de frenar su expansión.

En España, dado que para el 13 de marzo de 2020 el número de casos de COVID-19 contabilizados por el Ministerio ascendió a 4.209, incluyendo a más de 100 personas fallecidas, el gobierno español decretó el estado de alarma ese mismo día y se limitó la libre circulación de los ciudadanos a varios actos esenciales. Este evento ocasionó que las empresas se vieran obligadas a reducir drásticamente o incluso suspender sus actividades y con ello, se produjeron cambios en la oferta y demanda laboral. Asimismo, el empleo en España aún no se había recuperado completamente de la crisis económica de 2008, por lo que la nueva crisis provocada por la pandemia de la COVID-19 supuso un duro golpe. Debido a esto, el mercado laboral fue uno de los grandes afectados por la pandemia y con el fin de minimizar su impacto se introdujeron varias políticas de intervención. Las medidas que se adoptaron en el ámbito laboral estaban mayormente dirigidas a favorecer la conservación del empleo y a ayudar económicamente a los trabajadores afectados por la pandemia.

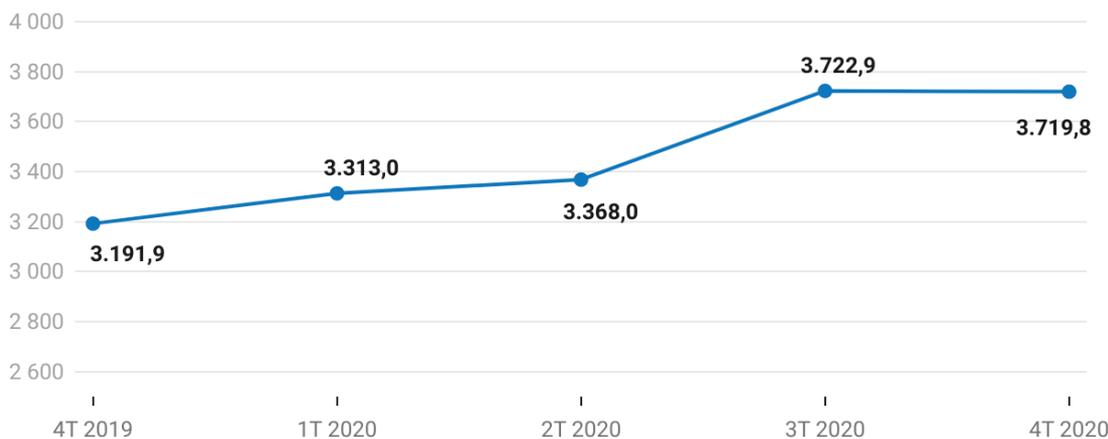
En cuanto a las medidas de conservación de empleo, las primeras que se establecieron estuvieron dirigidas a los sectores del turismo, comercio y hostelería vinculados a la actividad turística, entre las que se encontraba la aprobación de una bonificación del 50% de las cuotas empresariales a la Seguridad Social. Otras medidas consistieron en declarar el trabajo a distancia como método de trabajo preferente antes de recurrir a la detención de las actividades de las empresas; la imposibilidad de finalizar un contrato por motivos relacionados a la COVID-19; o la entrega de permisos especiales a aquellas personas que trabajan por cuenta ajena (y que cumplan ciertos requisitos), para que se les permitiera mantener el derecho a obtener la retribución que les correspondía en caso de haber trabajado con normalidad.

Las medidas se fueron introduciendo paulatinamente debido al desconocimiento que existía acerca del comportamiento del virus y la rapidez con la que éste actuaba, lo cual ocasionó que las políticas de intervención siempre fueran establecidas a una menor velocidad a la que la COVID-19 se expandía. Por tanto, es posible que no se tuvieran en cuenta todos los aspectos necesarios a imponer en las medidas o que fueran considerados demasiado tarde. Si atendemos a los datos publicados por el Instituto Nacional de Estadística (INE), se conoce que a finales de 2019 la tasa de paro se encontraba en un 13,78% y el número de desempleados se situaba en 3.191.900. Sin

embargo, a causa de la pandemia, 2020 finalizó con una tasa de paro que subió a un 16,13% y el número de desempleados creció en 527.900 personas.

Número de parados T4 2019-T4 2020

Datos en miles de personas.



En el primer trimestre de 2020 el tamaño muestral de la EPA en las semanas 11 a 13 fue inferior al de otros trimestres.

Gráfico: Elaboración propia • Fuente: INE • Creado con Datawrapper

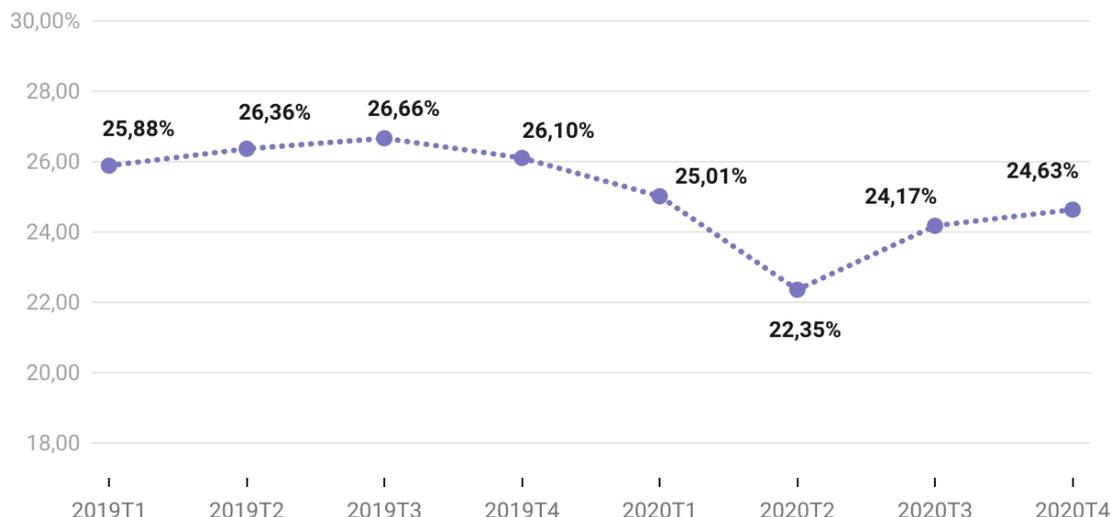
Sin las políticas de intervención es posible que estas cifras fueran mayores, mas es innegable que a pesar de los esfuerzos realizados, la COVID-19 ha perjudicado la vida laboral de miles de personas.

Dentro del mercado de trabajo, uno de los ámbitos más afectados fue el de los contratos temporales. En base a las cifras registradas por Eurostat, en el tercer y cuarto trimestre de 2020, España era el segundo país con mayor temporalidad de Europa, con una tasa que alcanzaba un 24,7% al terminar el año. Los contratos temporales llevan tiempo estando presentes en España, habiéndose creado más de 20,3 millones de estos contratos en 2019, en base a las estadísticas publicadas por el Servicio Público de Empleo Estatal (SEPE), de entre los cuales varios millones tienen una duración de un solo día. En 2020, esta cifra se redujo a 14,3 millones como consecuencia de la pandemia.

Al estudiar la evolución de la tasa de temporalidad mediante los datos del INE, se determina que incluso antes del comienzo de la pandemia, España disponía de porcentajes elevados que fueron reducidos durante 2020 a causa de la crisis sanitaria.

Evolución tasa de temporalidad en España

Porcentajes elaborados a partir de las cifras de asalariados publicadas por el INE.



En el primer trimestre de 2020 el tamaño muestral de la EPA en las semanas 11 a 13 fue inferior al de otros trimestres.

Gráfico: Elaboración propia • Fuente: INE • Creado con Datawrapper

A pesar de las medidas tomadas en 2020 que no permitían la terminación de los contratos por razones relacionadas a la COVID-19, esta norma no se aplicaba a los contratos ya establecidos antes de la pandemia y que tenían su terminación fijada en algún momento de aquel año. Esto provocó que los individuos a los que les finalizó el contrato temporal por causas ajenas al virus se encontraran con mayores dificultades de encontrar un empleo de nuevo, debido a la reducción de ofertas de trabajo causada por la pandemia. Esta situación puede verse reflejada en el gráfico anterior, que muestra la reducción de la tasa de temporalidad al comienzo del confinamiento. No obstante, en el tercer trimestre, junto a la llegada de la época turística, este porcentaje volvió a incrementarse debido a la necesidad de nuevo personal en varios sectores en los que los contratos temporales son abundantes, como el turismo y la hostelería.

1.2 Objetivos

En base a los datos expuestos, mediante este trabajo se quiere determinar de manera específica el efecto provocado por la COVID-19 en la situación de empleo y los contratos en España en función de varias características de la población. Por medio de un conjunto de variables que describa a los individuos, se tratará de analizar cómo condicionan estas cualidades al estado en el que se pueden encontrar los individuos en el mercado laboral. Para llevarlo a cabo, primero se describirán las variables contempladas utilizando datos de la Encuesta de Población Activa (EPA) y seguido se estimarán dos modelos econométricos que ayuden a analizar qué atributos eran relevantes a la hora de encontrar un trabajo u obtener una mayor estabilidad laboral antes de la pandemia y observar el efecto que tuvo la COVID-19 en éstos.

2. BASE DE DATOS, VARIABLES Y METODOLOGÍA

2.1 Base de datos y variables contempladas

Las variables a utilizar en los modelos econométricos provienen de los microdatos correspondientes a la Encuesta de Población Activa (EPA). Esta encuesta es una investigación realizada por el Instituto Nacional de Estadística (INE) de manera trimestral desde 1991, con el fin de obtener información acerca de la situación del mercado laboral. La EPA nos proporciona datos relacionados con la condición laboral del individuo encuestado durante la semana de referencia, además de características que describen al individuo, tales como el género, la edad, las horas trabajadas semanalmente, etc.

Debido a que el objetivo de este estudio es analizar los efectos a corto plazo que la crisis de la COVID-19 provocó en el mercado laboral español, se han utilizado los microdatos correspondientes a los trimestres 3 y 4 de 2019 y 2020. Se han tomado los datos del año 2020 en lugar de los del 2021 para comprobar las consecuencias que tuvo el confinamiento en el mercado laboral antes de finalizar el año y determinar la efectividad de las medidas impuestas por los gobiernos para combatir los efectos de la pandemia. Las muestras iniciales de 2019 tienen 163.365 (trimestre 3) y 163.152 observaciones (trimestre 4), mientras que las de 2020 tienen 146.986 (trimestre 3) y 147.296 (trimestre 4).

Con el fin de estudiar el impacto de la crisis sanitaria a corto plazo, se ha decidido establecer dos modelos econométricos. El primero utiliza los datos del cuarto trimestre de 2019 y 2020, y a través de este modelo se quiere analizar la relevancia que tuvo la COVID-19 en la situación de empleo de los individuos el año que ocurrió el confinamiento. Por otro lado, con la ayuda de un segundo modelo se quiere determinar el estado de los contratos indefinidos y temporales a partir de verano de 2020 y con ello comprobar si los contratos temporales fueron los que sufrieron en mayor medida las consecuencias del confinamiento. Por lo tanto, se utilizarán los datos del tercer trimestre de 2019 y 2020, que corresponden a los meses de verano y son los más cercanos al fin del confinamiento en España, que concluyó el 21 de junio de 2020.

Para ajustar las muestras de datos al objetivo de este trabajo, en primer lugar, se han filtrado los datos por edad del individuo y se han eliminado las observaciones que pertenecen a individuos menores de 16 años y mayores de 65 años, y así tener en cuenta únicamente a las personas en edad de trabajar. Asimismo, a fin de que no surgiera ningún problema cuando el programa a utilizar tuviera que leer los datos disponibles, se han realizado varias modificaciones leves en los datos de algunas variables, como rellenar con ceros algunas preguntas que los individuos no hubieran respondido.

Por último, se han juntado las observaciones de 2019 y 2020 (las de los terceros trimestres para el primer modelo y las de los cuartos trimestres para el segundo) y se

ha creado una nueva variable llamada *postCovid*, que indica si los datos pertenecen a 2019 o 2020. Cuando las observaciones correspondan al año 2019, la variable *postCovid* tomará el valor 0 y cuando correspondan al año 2020, el valor 1.

Como resultado, en el primer modelo ha quedado una muestra con 178.670 observaciones y en el segundo, una muestra con 86.282.

2.1.1 Variables endógenas

Primer modelo

La variable endógena TRABAJO indica si los individuos estuvieron trabajando durante la semana de referencia de la encuesta o si por el contrario no lo hicieron, lo cual engloba a los individuos que se encontraban buscando trabajo, estudiando o que decidieron no participar en el mercado laboral por razones personales.

Se ha escogido esta variable como endógena para el primer modelo debido a que es adecuada para mostrar cómo han afectado las diferentes variables a la situación de empleo de los ciudadanos después del confinamiento y una vez establecida la “normalidad”.

Los datos correspondientes a la variable TRABAJO provienen de la variable TRAREM de la EPA, que toma el valor 1 si los individuos realizaron algún trabajo remunerado durante la semana anterior a la que se realizó la encuesta y el valor 0 en caso contrario.

Segundo modelo

La variable endógena CONTRATO señala el tipo de contrato laboral del que disponía el individuo la semana anterior a la que se realizó la encuesta. Los datos han sido extraídos de la variable DUCON1 de la EPA, que toma el valor 1 cuando el individuo tiene un contrato indefinido y 0 cuando tiene un contrato temporal.

Mediante esta variable se quiere analizar la manera en la que fueron afectados los contratos por la COVID-19. Se comprobará si la crisis sanitaria influyó de manera negativa a la estabilidad laboral de los individuos.

2.1.2 Variables explicativas

Se han elegido 6 variables explicativas para la estimación del primer modelo y 7 para la estimación del segundo. De las 6 variables utilizadas en el primer modelo, 5 también han sido aplicadas en el segundo, por lo que se comenzará describiendo todas las variables que ambos modelos tienen en común y después se describirán las variables que utilizarán individualmente.

Las variables explicativas usadas en los dos modelos planteados son las siguientes:

- La variable GÉNERO es una variable cualitativa que indica el género de los individuos. Toma el valor 1 si la persona es hombre y 0 si es mujer. Se considera una variable relevante dado que se podrá comprobar si ser hombre o mujer afecta a las probabilidades de obtener un trabajo o de disponer de una mayor estabilidad laboral.

- La variable EDAD ha sido conseguida a partir de la eliminación de las observaciones que contienen a los individuos menores de 16 años y mayores de 65 que provienen de la variable DEDAD de la EPA. Es una variable cualitativa que agrupa a los individuos que tengan entre 16-64 años de edad en tramos de 4 años, por tanto, se han creado diez agrupaciones diferentes. Las variables ficticias creadas a partir de esta variable indican lo siguiente:

EDAD_16	Toma el valor 1 cuando el individuo tiene 16-19 años
EDAD_20	Toma el valor 1 cuando el individuo tiene 20-24 años
EDAD_25	Toma el valor 1 cuando el individuo tiene 25-29 años
EDAD_30	Toma el valor 1 cuando el individuo tiene 30-34 años
EDAD_35	Toma el valor 1 cuando el individuo tiene 35-39 años
EDAD_40	Toma el valor 1 cuando el individuo tiene 40-44 años
EDAD_45	Toma el valor 1 cuando el individuo tiene 45-49 años
EDAD_50	Toma el valor 1 cuando el individuo tiene 50-54 años
EDAD_55	Toma el valor 1 cuando el individuo tiene 55-59 años
EDAD_60	Toma el valor 1 cuando el individuo tiene 60-64 años

Donde $EDAD_{16i} + EDAD_{20i} + EDAD_{25i} + EDAD_{30i} + EDAD_{35i} + EDAD_{40i} + EDAD_{45i} + EDAD_{50i} + EDAD_{55i} + EDAD_{60i} = 1 \quad i=1,2,\dots,N \quad \forall i$. Con el fin de evitar problemas de multicolinealidad, al estimar los modelos se dejará fuera la variable ficticia EDAD_40.

Se quiere contrastar que la edad de los individuos es un factor importante que, hasta cierta edad, afecta de manera positiva a la probabilidad de que un individuo esté trabajando, pero que, a partir de dicha edad, tiene un efecto negativo y las posibilidades de conseguir un trabajo disminuyen.

- La variable CCAA indica la comunidad autónoma en la que reside el individuo y sus datos han sido extraídos de la variable de la EPA con el mismo nombre. Se han creado las siguientes variables ficticias, correspondientes a cada Comunidad Autónoma:

CCAA_ANDALUCIA	Toma el valor 1 si el individuo reside en Andalucía
CCAA_ARAGON	Toma el valor 1 si el individuo reside en Aragón
CCAA_ASTURIAS	Toma el valor 1 si el individuo reside en Asturias
CCAA_BALEARES	Toma el valor 1 si el individuo reside en Islas Baleares
CCAA_CANARIAS	Toma el valor 1 si el individuo reside en Canarias
CCAA_CANTABRIA	Toma el valor 1 si el individuo reside en Cantabria
CCAA_CASTLEON	Toma el valor 1 si el individuo reside en Castilla y León
CCAA_CASTMAN	Toma el valor 1 si el individuo reside en Castilla La Mancha
CCAA_CATALU	Toma el valor 1 si el individuo reside en Cataluña
CCAA_VALENCIA	Toma el valor 1 si el individuo reside en Comunidad Valenciana
CCAA_EXTR	Toma el valor 1 si el individuo reside en Extremadura
CCAA_GALICIA	Toma el valor 1 si el individuo reside en Galicia
CCAA_MADRID	Toma el valor 1 si el individuo reside en Madrid
CCAA_MURCIA	Toma el valor 1 si el individuo reside en Murcia
CCAA_NAVARRA	Toma el valor 1 si el individuo reside en Navarra
CCAA_PV	Toma el valor 1 si el individuo reside en País Vasco
CCAA_RIOJA	Toma el valor 1 si el individuo reside en La Rioja
CCAA_CEUTA	Toma el valor 1 si el individuo reside en Ceuta
CCAA_MELILLA	Toma el valor 1 si el individuo reside en Melilla

Donde $CCAA_ANDALUCIA_i + CCAA_ARAGON_i + CCAA_ASTURIAS_i + CCAA_BALEARES_i + CCAA_CANARIAS_i + CCAA_CANTABRIA_i + CCAA_CASTLEON_i + CCAA_CASTMAN_i + CCAA_CATALU_i + CCAA_VALENCIA_i + CCAA_EXTR_i + CCAA_GALICIA_i + CCAA_MADRID_i + CCAA_MURCIA_i + CCAA_NAVARRA_i + CCAA_PV_i + CCAA_RIOJA_i + CCAA_CEUTA_i + CCAA_MELILLA_i = 1 \quad i=1,2,\dots,N \quad \forall i$.
Con el fin de evitar problemas de multicolinealidad, al estimar los modelos se dejará fuera la variable ficticia CCAA_ANDALUCIA.

Se ha considerado adecuado estudiar la relevancia que tiene el lugar de residencia de los individuos con la probabilidad de estar trabajando, dado que, si bien todas las comunidades forman parte del mismo país, es posible que existan numerosas diferencias entre ellas. Varias comunidades tienen culturas o costumbres que difieren entre sí y que podrían afectar al comportamiento de los individuos a la hora de entrar al mercado laboral. Asimismo, algunas comunidades autónomas tienen una mayor dependencia en algunos sectores económicos que otras, y no todas presentan la misma cantidad de ofertas de trabajo. Por ello, se quiere contrastar que el lugar de residencia es influyente tanto en las posibilidades de obtener un empleo como en el tipo de contrato que pueden llegar a tener los individuos.

- La variable EST_CIVIL informa acerca del estado civil del individuo. Sus datos provienen de la variable ECIV1 de la EPA y están divididos en las siguientes categorías:

EST_SOLTERO	Variable ficticia que toma el valor 1 si el individuo está soltero
EST_CASADO	Variable ficticia que toma el valor 1 si el individuo está casado
EST_VIUDO	Variable ficticia que toma el valor 1 si el individuo está viudo
EST_DIVORC	Variable ficticia que toma el valor 1 si el individuo está divorciado

Donde $EST_SOLTERO_i + EST_CASADO_i + EST_VIUDO_i + EST_DIVORC_i = 1$
 $i=1,2,\dots,N \forall i$. Con el fin de evitar problemas de multicolinealidad, al estimar los modelos se dejará fuera la variable ficticia EST_SOLTERO.

Quiere observarse el efecto que tiene el estado civil de los individuos con la situación de empleo de éstos. En concreto, se quiere analizar si estar casado o soltero es influyente a la hora de encontrar un empleo.

- La variable N_FORMACION ha sido formateada a partir de la variable NFORMA de la EPA, que indica el nivel de estudios del individuo. Los datos de la variable NFORMA están divididos en 7 categorías, pero en los ajustes realizados para este trabajo, 4 de las categorías han sido agrupadas únicamente en dos. Se han unido los niveles “Analfabetos” y “Educación primaria incompleta” en una sola categoría, al igual que “Primera etapa de educación secundaria” y “Segunda etapa de educación secundaria. Orientación general”. Como resultado, se ha dividido la variable N_FORMACION en 5 categorías:

N_SINESTUDIOS	Toma el valor 1 cuando el individuo no ha finalizado educación primaria
N_EDPRIMARIA	Toma el valor 1 cuando el individuo ha cursado hasta educación primaria
N_EDSECUN	Toma el valor 1 cuando el individuo ha cursado hasta educación secundaria general
N_FP	Toma el valor 1 cuando el individuo ha cursado ed secundaria con formación profesional
N_EDSUP	Toma el valor 1 cuando el individuo ha cursado hasta educación superior

Donde $N_SINESTUDIOS_i + N_EDPRIMARIA_i + N_EDSECUN_i + N_FP_i + N_EDSUP_i = 1$
 $i=1,2,\dots,N \forall i$. Con el fin de evitar problemas de multicolinealidad, al estimar los modelos se dejará fuera la variable ficticia N_EDPRIMARIA.

A través de esta variable se quiere verificar que un mayor nivel de estudios aumenta las posibilidades de que un individuo esté trabajando a medida que dicho nivel aumenta. Además, se quiere revelar si la formación está relacionada con la estabilidad laboral.

Siguiendo con la variable que se utilizará exclusivamente en el primer modelo, se encuentra:

- La variable PREVIO, cuyos datos provienen de la variable EMPANT de la EPA, que señala si los individuos que no trabajaron la semana de referencia han realizado algún trabajo con anterioridad. La variable toma el valor 1 si se ha trabajado en una ocasión previa y 0 si no. Para que el programa Gretl pudiera interpretar los datos de manera adecuada, se han completado las celdas vacías correspondientes a las personas que sí

tienen un empleo con el valor 1. Con esta variable se quiere determinar si no haber trabajado previamente influye de manera negativa a las probabilidades de encontrar un trabajo.

Las variables que únicamente se han añadido al segundo modelo están relacionadas con el puesto de trabajo que tienen los individuos, por lo que las preguntas de la encuesta correspondientes a las siguientes variables solo fueron respondidas por los individuos que tenían un empleo.

- La variable SECTOR representa la actividad principal que realiza el individuo, la cual ha sido formateada a partir de la variable ACT1 de la EPA después de agrupar las actividades en relación a los sectores de la actividad económica. A partir de ello se han conformado unas variables ficticias agrupadas en tres categorías:

SECTOR_PRIM	Variable ficticia que toma valor 1 si el individuo trabaja en el sector primario
SECTOR_SEC	Variable ficticia que toma valor 1 si el individuo trabaja en el sector secundario
SECTOR_TERC	Variable ficticia que toma valor 1 si el individuo trabaja en el sector terciario

Donde $SECTOR_PRIM_i + SECTOR_SEC_i + SECTOR_TERC_i = 1 \quad i=1,2,\dots,N \quad \forall i$. Con el fin de evitar problemas de multicolinealidad, al estimar el modelo se dejará fuera la variable ficticia SECTOR_PRIM.

Se quiere contrastar que la probabilidad de tener un contrato indefinido es mayor en unos sectores que en otros, además de comprobar qué sectores son los que han resultado más afectados por la pandemia.

- La variable JORNADA es otra variable cualitativa, que señala si el individuo trabaja a jornada completa o parcial. Toma el valor 1 cuando el individuo tiene una jornada completa y 0 cuando tiene una jornada parcial. La finalidad de añadir esta variable es analizar si la crisis sanitaria ha influido en el tipo de jornada laboral que pueden llegar a tener los individuos.

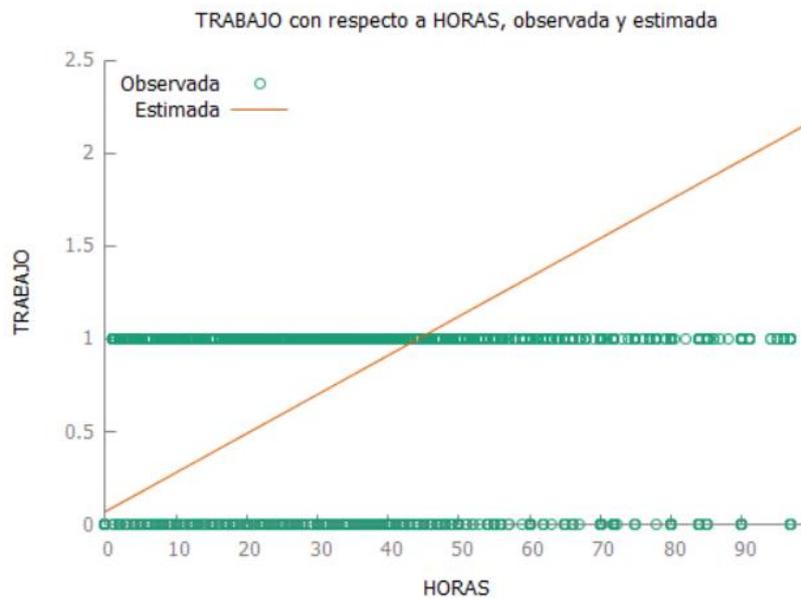
2.3. Metodología

Dado que se desea conocer la relación entre una variable dependiente dicotómica y varias variables explicativas, no es adecuado utilizar un modelo de regresión lineal para estimar los modelos pensados.

El objetivo de la regresión lineal es intentar predecir el valor de una variable dependiente continua (Y) a través de un modelo lineal que describa la relación entre la variable

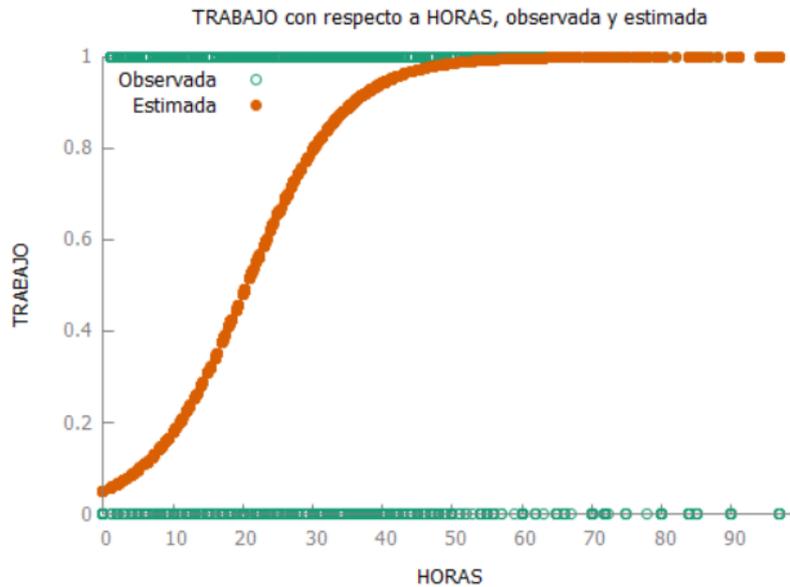
dependiente y las variables explicativas. Por ello, la regresión lineal se ajusta a una línea recta que minimiza la distancia entre el valor predicho y el valor observado.

Si utilizamos este método de estimación con una variable dependiente que tenga dos únicos valores (0 y 1), veremos que, mientras estimar el modelo sí es posible, los valores predichos de Y pueden llegar a ser mayores que 1 y menores que 0. Utilizando como ejemplo el objeto de estudio, si realizamos una estimación mediante una regresión lineal que tenga como variable dependiente la variable TRABAJO y como variable independiente la variable HORAS, se observa que más de la mitad de los valores de Y estimados son mayores que 1.



Esto se debe a que, mediante la regresión lineal, se está prediciendo un valor absoluto, que puede salirse del rango entre 0 y 1. Sin embargo, este problema se soluciona a través de la regresión logística, ya que no se intenta predecir el valor de la variable dependiente (Y) a través de las variables explicativas (X_i), sino que se trata de predecir la *probabilidad* de que ocurra Y a partir de los valores de las variables X_i . Las probabilidades siempre oscilan entre 0 y 1, donde la probabilidad de que un evento ocurra con certeza es 1 y que dicho evento sea poco probable que ocurra, 0.

Si realizamos una regresión logística utilizando los datos anteriores, observamos que esta vez los valores predichos sí fluctúan entre 0 y 1.



La regresión lineal es adecuada cuando la variable dependiente es continua, como predecir el salario de un empleado, donde el valor predicho puede ser un número real. En cambio, cuando la variable dependiente es discreta, es más preciso emplear la regresión logística.

Después de demostrar que el mejor método de estimación para los dos modelos planteados es la regresión logística, se va a explicar de qué trata dicho método.

La regresión logística es un modo de regresión que permite estimar la probabilidad de una variable discreta en función de varias variables explicativas. Por ello, el método consiste en establecer una función logística de las variables explicativas que permita organizar a los individuos en una de las categorías establecidas por los valores de la variable dependiente. Es posible escribir el modelo de regresión logística como:

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

No obstante, debido a que se pretende expresar la probabilidad de que cierto evento ocurra en función de determinadas variables, después de realizar varias operaciones matemáticas, se llega a la siguiente expresión:

$$P(Y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}$$

Donde la variable dependiente se denomina por Y y las k variables explicativas se denotan mediante X_1, X_2, \dots, X_k

Dentro de los métodos Logit, existe la regresión logística binaria (la variable dependiente tiene dos categorías) y la multinomial (la variable dependiente tiene dos categorías o

más). Entonces, dado que las variables dependientes escogidas para ambos modelos son variables cualitativas dicotómicas, se utilizará la regresión logística binaria para la estimación de ambos modelos.

3. ESTIMACIÓN Y ANÁLISIS DE LOS MODELOS ECONOMETRÍCOS

3.1 Modelo con situación de empleo como variable dependiente

El primer modelo plantea descubrir cuáles son las probabilidades de que un individuo esté trabajando y el efecto que ha tenido la COVID-19 sobre las variables. Para mostrar dicho efecto se han incluido variables de interacción en el modelo siguiendo la siguiente fórmula: *variable de interacción*=*postCovid***variable explicativa*, siendo *postCovid* la variable que indica a qué año pertenecen los datos.

La categoría base establecida en el modelo es la de un individuo de género masculino (GENERO), de 40 a 45 años (EDAD_40), soltero (EST_SOLTERO), que reside en Andalucía (CCAA_ANDALUCIA) y que dispone de estudios de educación primaria (N_EDPRIM). En un principio, se ha especificado el modelo 1 con todas las variables mencionadas en el apartado 2.1.2 junto con las variables de interacción correspondientes. No obstante, al estimar el modelo, el programa Gretl ha avisado de que se ha eliminado la variable PREVIO, debido a que $\text{Prob}(\text{TRABAJO} = 0 \mid \text{PREVIO} = 0) = 1$. Esto significa que la probabilidad de que los individuos que no han trabajado anteriormente no tengan un puesto de trabajo es igual a uno y, por tanto, no es correcto insertar esta variable. Por ello, se ha estimado el modelo de nuevo a través de una regresión logística binaria, sin añadir la variable PREVIO ni la variable de interacción Int_PREVIO, que mostraba el efecto de la COVID-19 en la variable PREVIO.

Como resultado se ha conseguido la siguiente estimación:

Modelo 1

Probabilidad de estar trabajando explicada por las variables contempladas.

Categoría base: Individuo de género masculino, de 40 a 45 años, soltero, que reside en Andalucía y dispone de estudios de educación primaria.

Modelo 1: Logit, usando las observaciones 1-178669

Variable dependiente: TRABAJO

Desviaciones típicas basadas en el Hessiano

Impacto de la COVID-19 en el Mercado Laboral

	<i>Coefficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Pendiente*</i>	<i>valor p</i>	
const	-2.70039	0.0815051	-33.13		<0.0001	***
GENERO	-0.376493	0.0283665	-13.27	-0.0929311	<0.0001	***
EDAD_16	-0.976272	0.0879411	-11.10	-0.236740	<0.0001	***
EDAD_20	0.168000	0.0640617	2.622	0.0412036	0.0087	***
EDAD_25	0.347241	0.0658932	5.270	0.0839627	<0.0001	***
EDAD_30	-0.0624285	0.0620852	-1.006	-0.0154865	0.3146	
EDAD_35	-0.109670	0.0558519	-1.964	-0.0272412	0.0496	**
EDAD_45	0.0404525	0.0528679	0.7652	0.00999422	0.4442	
EDAD_50	-0.0972484	0.0525957	-1.849	-0.0241400	0.0645	*
EDAD_55	-0.382476	0.0531720	-7.193	-0.0952453	<0.0001	***
EDAD_60	-0.791025	0.0566816	-13.96	-0.194749	<0.0001	***
EST_CASADO	-0.0357763	0.0367372	-0.9738	-0.00885280	0.3301	
EST_VIUDO	0.0220282	0.110089	0.2001	0.00544504	0.8414	
EST_DIVORC	-0.161257	0.0589820	-2.734	-0.0401148	0.0063	***
N_SINESTUDIOS	-0.449395	0.127118	-3.535	-0.111873	0.0004	***
N_EDSECUN	0.128512	0.0627551	2.048	0.0317802	0.0406	**
N_FP	0.0888452	0.0730320	1.217	0.0218961	0.2238	
N_EDSUP	0.202725	0.0642968	3.153	0.0499659	0.0016	***
CCAA_ARAGON	0.0459984	0.0721607	0.6374	0.0113569	0.5238	
CCAA_ASTURIAS	-0.235618	0.0858496	-2.745	-0.0587147	0.0061	***
CCAA_BALEARES	-0.279915	0.0885589	-3.161	-0.0697812	0.0016	***
CCAA_CANARIAS	-0.105972	0.0723036	-1.466	-0.0263308	0.1427	
CCAA_CANTABRIA	-0.355450	0.0923363	-3.850	-0.0886063	0.0001	***
CCAA_CASTLEON	0.106558	0.0545018	1.955	0.0262356	0.0506	*
CCAA_CASTMAN	0.0840771	0.0598852	1.404	0.0207204	0.1603	
CCAA_CATALU	0.0573575	0.0533132	1.076	0.0141578	0.2820	
CCAA_VALENCIA	0.0787400	0.0576837	1.365	0.0194122	0.1722	
CCAA_EXTR	0.0321639	0.0783376	0.4106	0.00794663	0.6814	
CCAA_GALICIA	-0.246233	0.0516235	-4.770	-0.0613033	<0.0001	***
CCAA_MADRID	0.0583363	0.0649438	0.8983	0.0143951	0.3690	
CCAA_MURCIA	0.0442064	0.0798288	0.5538	0.0109150	0.5797	
CCAA_NAVARRA	-0.270168	0.0905725	-2.983	-0.0673497	0.0029	***
CCAA_PV	-0.0881615	0.0682834	-1.291	-0.0218926	0.1967	
CCAA_RIOJA	0.0120909	0.105743	0.1143	0.00299021	0.9090	
CCAA_CEUTA	-0.0936224	0.194695	-0.4809	-0.0232597	0.6306	

Impacto de la COVID-19 en el Mercado Laboral

CCAA_MELILLA	0.0272608	0.196021	0.1391	0.00673636	0.8894	
HORAS	0.147478	0.000915084	161.2	0.0364947	<0.0001	***
postCovid	-0.111066	0.118479	-0.9374	-0.0274843	0.3485	
Int_GENERO	0.0790772	0.0398724	1.983	0.0195226	0.0473	**
Int_EDAD_16	-0.409937	0.138738	-2.955	-0.102124	0.0031	***
Int_EDAD_20	-0.148969	0.0911435	-1.634	-0.0370600	0.1022	
Int_EDAD_25	-0.122737	0.0933147	-1.315	-0.0305150	0.1884	
Int_EDAD_30	-0.0449855	0.0879710	-0.5114	-0.0111541	0.6091	
Int_EDAD_35	-0.0609230	0.0783627	-0.7774	-0.0151142	0.4369	
Int_EDAD_45	-0.0464599	0.0740301	-0.6276	-0.0115189	0.5303	
Int_EDAD_50	-0.0151366	0.0735722	-0.2057	-0.00374814	0.8370	
Int_EDAD_55	-0.0544908	0.0738694	-0.7377	-0.0135141	0.4607	
Int_EDAD_60	0.0267577	0.0789888	0.3388	0.00661322	0.7348	
Int_EST_CASADO	0.102712	0.0510461	2.012	0.0253389	0.0442	**
Int_EST_VIUDO	0.00301908	0.159303	0.01895	0.000746986	0.9849	
Int_EST_DIVORC	0.0894161	0.0820349	1.090	0.0220211	0.2757	
Int_N_SINESTUDIOS	0.173813	0.188238	0.9234	0.0425405	0.3558	
Int_N_EDSECUN	0.0223371	0.0927938	0.2407	0.00552400	0.8098	
Int_N_FP	0.122114	0.106005	1.152	0.0300190	0.2493	
Int_N_EDSUP	0.311421	0.0946342	3.291	0.0759386	0.0010	***
Int_CCAA_ARAGON	-0.225500	0.101415	-2.224	-0.0561908	0.0262	**
Int_CCAA_ASTURIAS	-0.0935154	0.119979	-0.7794	-0.0232315	0.4357	
Int_CCAA_BALEARES	-0.303239	0.122613	-2.473	-0.0756094	0.0134	**
Int_CCAA_CANARIAS	-0.343306	0.101581	-3.380	-0.0855861	0.0007	***
Int_CCAA_CANTABRIA	0.307312	0.132176	2.325	0.0743549	0.0201	**
Int_CCAA_CASTLEON	-0.0860288	0.0767971	-1.120	-0.0213615	0.2626	
Int_CCAA_CASTMAN	-0.0262331	0.0846539	-0.3099	-0.00649930	0.7566	
Int_CCAA_CATALU	-0.119983	0.0762770	-1.573	-0.0298239	0.1157	
Int_CCAA_VALENCIA	-0.0249851	0.0822481	-0.3038	-0.00618973	0.7613	
Int_CCAA_EXTR	0.144440	0.111224	1.299	0.0354348	0.1941	
Int_CCAA_GALICIA	-0.0861425	0.0725599	-1.187	-0.0213880	0.2352	
Int_CCAA_MADRID	-0.0490448	0.0915025	-0.5360	-0.0121627	0.5920	
Int_CCAA_MURCIA	-0.104213	0.113323	-0.9196	-0.0258979	0.3578	
Int_CCAA_NAVARRA	-0.0418971	0.129324	-0.3240	-0.0103878	0.7460	
Int_CCAA_PV	-0.154715	0.0973103	-1.590	-0.0385012	0.1119	
Int_CCAA_RIOJA	0.0257730	0.145339	0.1773	0.00636931	0.8592	

Int_CCAA_CEUTA	0.00956517	0.281545	0.03397	0.00236584	0.9729
Int_CCAA_MELILLA	-0.0860379	0.280941	-0.3062	-0.0213701	0.7594
Int_HORAS	-0.00897116	0.00128141	-7.001	-0.00222000	<0.0001 ***

Media de la vble. dep.	0.541269	D.T. de la vble. dep.	0.498295
R-cuadrado de McFadden	0.654645	R-cuadrado corregido	0.654045
Log-verosimilitud	-42559.63	Criterio de Akaike	85267.27
Criterio de Schwarz	86014.17	Crit. de Hannan-Quinn	85488.18

*Evaluado en la media

Número de casos 'correctamente predichos' = 164910 (92.3%)

f(beta|x) en la media de las variables independientes = 0.498

Contraste de razón de verosimilitudes: Chi-cuadrado(73) = 161350 [0.0000]

		Predicho	
		0	1
Observado	0	73117	8844
	1	4915	91793

3.1.1 Análisis de los resultados del primer modelo

Comenzando por la bondad de ajuste del modelo, nos centraremos en el número de casos correctamente predichos. Cuanto mayor sea ese número de casos, más explicativo es el modelo y, por tanto, se puede considerar que las variables explicativas predicen de manera correcta el modelo. En este caso, la estimación cuenta con un 92,3% de casos correctamente predichos, por ende, puede considerarse que el modelo cuenta con un buen ajuste.

En relación al **género** de los individuos, los resultados señalan que se trata de una variable significativa, debido a que su p-valor se encuentra muy por debajo de 0,05. Su coeficiente nos indica la diferencia en las probabilidades de estar trabajando de un hombre respecto de una mujer manteniendo los demás regresores constantes. Por tanto, al ser negativo, revela que las mujeres tenían más probabilidades de estar trabajando que los hombres.

La **edad** se muestra como una característica que muestra una gran relevancia al tratar con las probabilidades de tener un empleo. Se puede observar a través de los resultados obtenidos que, los individuos de entre 16 y 20 años de edad son los que menos probabilidades tienen de estar trabajando con respecto a la franja de edad tomada como categoría base (40-45 años). Sin embargo, se trata de un resultado que se encuentra dentro de las expectativas, dado que una gran parte de estos individuos deciden continuar con su formación y no muestran deseos de entrar en el mercado laboral hasta finalizar sus estudios. Seguido se encuentran las variables ficticias EDAD_20 y EDAD_25, que muestran que para los individuos de entre 20 y 29 años de edad, las probabilidades de estar trabajando en relación con la franja de edad base aumentan. Las dos siguientes franjas de edades, de 30 a 39 años, en cambio, muestran una disminución de las probabilidades en comparación con los individuos de entre 40 y 44 años, y que como se puede apreciar, vuelven a aumentar a partir de los 40 años. Por último, cabe destacar que a partir de la franja de edad 45-50 años, a medida que la edad de los individuos aumenta y la edad de jubilación se acerca, las probabilidades de estar trabajando disminuyen de manera significativa.

Una de las características que no ha demostrado demasiada significación es el **estado civil** de los individuos, a excepción de la variable ficticia EST_DIVORC. No obstante, las demás variables ficticias que no se han mostrado relevantes son clases que pertenecen a la variable cualitativa EST_CIVIL, por lo que no se deberían eliminar de la estimación del modelo. En relación a los individuos divorciados, los resultados señalan que frente a los individuos que están solteros, el logaritmo de las probabilidades de estar trabajando cambia a $-0,161257$ y, por tanto, sus probabilidades de estar trabajando disminuyen. Al observar la pendiente, que nos muestra el efecto marginal de la variable, es posible confirmar que dichas probabilidades decrecen concretamente en un 4%.

Los resultados verifican que una variable influyente en las probabilidades es el **nivel de formación académica**. Dado que como referencia se ha seleccionado a un individuo que disponga de estudios primarios, se puede observar que poseer niveles de estudios superiores a éstos aumentan las probabilidades de estar trabajando, siendo la educación superior la que en mayor medida incrementa estas probabilidades (específicamente, en un 4,99%). Por otro lado, no tener estudios resulta en que disminuyan alrededor de un 11% en relación a tener estudios primarios. Dentro de las categorías de la variable cualitativa N_FORMACION, N_FP es la única variable ficticia que no se muestra relevante. No obstante, tiene establecido un coeficiente con signo positivo. Entonces, a pesar de ello, es posible concluir que si el nivel de formación aumenta, las probabilidades de estar trabajando también lo hacen.

En cuanto a la **localización geográfica**, los resultados obtenidos sugieren que el lugar de residencia puede mostrar cierta relevancia. Tomando como base un individuo que reside en Andalucía, podemos advertir que las comunidades autónomas que aumentan las posibilidades de estar trabajando son Aragón, Castilla y León, Castilla La-Mancha, Cataluña, Valencia, Extremadura, Madrid, Murcia, La Rioja y Melilla. Sin embargo, sus

correspondientes variables ficticias no se muestran relevantes en el modelo, por tener p-valores mayores a 0,05. Por ello, centrándonos únicamente en las variables ficticias que son significativas, se puede concluir que, en relación a vivir en Andalucía, tener el lugar de residencia en Asturias, Islas Baleares, Cantabria, Galicia y Navarra influye de manera negativa a la probabilidad de estar trabajando. Dicha probabilidad puede llegar a disminuir desde un 5,8% hasta un 8,8%, en caso de que el individuo resida en Asturias o en Cantabria, respectivamente.

La variable HORAS, que indica el **número de horas semanales dedicadas a trabajar**, ha resultado mostrar una gran relevancia en el modelo. A través de la estimación se le ha asociado un coeficiente positivo, que indica que para cada cambio en una unidad en HORAS, la probabilidad de estar trabajando (en comparación con no estar trabajando) aumenta en un 0,147478.

Una de las variables clave de este objeto de estudio, la **variable postCovid**, tiene asociado un efecto marginal negativo de -0.0274843, por tanto, en términos generales, la probabilidad de que el individuo de referencia esté trabajando en el cuarto trimestre de 2020, después del comienzo de la pandemia, disminuye alrededor de un 2,74% con respecto al cuarto trimestre de 2019. Por ello, si bien se puede deducir que las medidas adoptadas para intentar frenar los efectos negativos de las restricciones surtieron efecto y lograron que muchas personas que se encontraban en una situación de precariedad mantuvieran sus puestos de trabajo, es necesario destacar que muchas personas sí se vieron afectadas.

Por último, se analizarán las **variables de interacción**, que fueron introducidas con el fin de estudiar el efecto de la COVID-19 en las variables. Mediante la estimación del modelo, se ha observado que solamente algunas de estas variables son significativas. No obstante, las que no disponen de una gran relevancia pertenecen a variables cualitativas, por tanto, no sería adecuado eliminarlas. Cabe destacar que el análisis de las variables de interacción difiere ligeramente de la interpretación de las variables anteriores, dado que sus resultados miden el efecto que ha tenido la crisis sanitaria entre 2019 y 2020. Por ello, se espera que éstos sean negativos y así verificar que la COVID-19 ha sido perjudicial en la mayoría de los aspectos.

En primer lugar, es posible comprobar que la variable de interacción que mide el efecto de la crisis sanitaria en el género de los individuos es bastante significativa. Debido a que la categoría base de este modelo es un hombre, se puede apreciar que en el último trimestre de 2020, los hombres tenían más probabilidades de estar trabajando que a finales de 2019 en comparación con las mujeres. En concreto, un 1,95% más.

Respecto a las variables de interacción relacionadas con la edad de los individuos, se puede observar que, si bien la mayoría de dichas variables tienen asociado un efecto marginal negativo, solo se muestra relevante Int_EDAD_16, que representa el rango de edad de 16 a 19 años. Al analizar la pendiente se verifica que en relación a los individuos de entre 40-45 años de edad, el efecto de la COVID-19 ha provocado que la probabilidad

de estar trabajando disminuya en un 10% para las personas que tienen entre 16 y 19 años.

Siguiendo con las variables relacionadas con el estado civil de los individuos, se percibe que a causa de la pandemia, la probabilidad de estar trabajando de los individuos casados ha aumentado en un 2,53% en relación a los individuos solteros. Por otro lado, es importante mencionar que los individuos estén viudos o divorciados no ha mostrado relevancia.

Al observar el efecto de la pandemia en el nivel de formación de los individuos, únicamente se muestra significativa la obtención de estudios superiores. Respecto a las personas que tienen estudios primarios, la crisis sanitaria ha provocado que las probabilidades de trabajar aumenten en un 7,59% si se obtiene una educación superior.

Dependiendo del lugar de residencia, las probabilidades han sido reducidas en la mayoría de las comunidades autónomas en relación al grupo de referencia, Andalucía. El efecto de la COVID-19 solamente aparece relevante en cuatro comunidades, de las cuales tres han obtenido un efecto negativo y una un efecto positivo. Si los individuos viven en Aragón, en las Islas Baleares o en las Islas Canarias, sus probabilidades de trabajar respecto a Andalucía disminuyen en un 5,61%, un 7,56% y un 8,55% respectivamente. En cambio, si residen en Cantabria, dichas probabilidades aumentan en un 7,43%.

Para finalizar el análisis del primer modelo, falta estudiar el efecto marginal de la variable de interacción de HORAS, la única variable cuantitativa. Sus resultados nos indican que por cada unidad de hora trabajada, la probabilidad de estar trabajando disminuye en un 0,02%.

3.2 Modelo con tipo de contrato como variable dependiente

Después de haber estudiado las probabilidades de que un individuo esté trabajando, se quiere analizar el efecto de la COVID-19 en la estabilidad laboral de aquellos individuos que han encontrado trabajo (que dispongan de un contrato indefinido o temporal) y cómo influyen las características utilizadas en el anterior modelo. Asimismo, se quiere estudiar si el tipo de jornada laboral y el sector económico en el que los individuos están trabajando son relevantes en su estabilidad laboral, por lo que se añadirán en el modelo dos variables adicionales que representen dichas características.

El segundo modelo plantea descubrir cuáles son las probabilidades de que un individuo tenga un contrato indefinido o temporal. La categoría base establecida es la de un individuo de género masculino (GENERO), de 40 a 45 años (EDAD_40), soltero (EST_SOLTERO), que reside en Andalucía (CCAA_ANDALUCIA), que ha cursado

educación primaria (N_EDPRIMARIA), que trabaja en el sector primario (SECT_PRIM) y que trabaja a jornada completa (JORNADA).

Al igual que en el primer modelo estimado, se han introducido variables de interacción que muestran el efecto de la pandemia en las variables. Como resultado, se ha obtenido la siguiente estimación:

Modelo 2

Probabilidad de tener un contrato indefinido explicada por las variables contempladas.

Categoría base: Individuo de género masculino, de 40 a 45 años, soltero, que reside en Andalucía, ha cursado educación primaria, trabaja en el sector primario y trabaja a jornada completa.

Modelo 2: Logit, usando las observaciones 1-86282

Variable dependiente: CONTRATO

Desviaciones típicas basadas en el Hessiano

	<i>Coficiente</i>	<i>Desv. Típica</i>	<i>z</i>	<i>Pendiente*</i>	<i>valor p</i>	
const	-1.09439	0.0922458	-11.86		<0.0001	***
GENERO	0.138528	0.0231971	5.972	0.0311551	<0.0001	***
EDAD_16	-2.72006	0.160595	-16.94	-0.549454	<0.0001	***
EDAD_20	-1.63894	0.0574472	-28.53	-0.387734	<0.0001	***
EDAD_25	-0.840480	0.0465416	-18.06	-0.202581	<0.0001	***
EDAD_30	-0.313670	0.0433320	-7.239	-0.0729755	<0.0001	***
EDAD_35	-0.0435653	0.0397341	-1.096	-0.00982701	0.2729	
EDAD_45	0.0928784	0.0377200	2.462	0.0206247	0.0138	**
EDAD_50	0.176097	0.0388234	4.536	0.0387011	<0.0001	***
EDAD_55	0.288858	0.0418807	6.897	0.0623858	<0.0001	***
EDAD_60	0.279434	0.0509541	5.484	0.0600964	<0.0001	***
EST_CASADO	0.290022	0.0271622	10.68	0.0654020	<0.0001	***
EST_VIUDO	0.403269	0.108925	3.702	0.0840853	0.0002	***
EST_DIVORC	0.164640	0.0457681	3.597	0.0360755	0.0003	***
N_SINESTUDIOS	-0.292260	0.120738	-2.421	-0.0682533	0.0155	**
N_EDSECUN	0.400084	0.0546073	7.327	0.0885082	<0.0001	***
N_FP	0.472318	0.0614870	7.682	0.0989154	<0.0001	***
N_EDSUP	0.688684	0.0556683	12.37	0.151477	<0.0001	***

Impacto de la COVID-19 en el Mercado Laboral

CCAA_ARAGON	0.485438	0.0568620	8.537	0.100221	<0.0001	***
CCAA_ASTURIAS	0.180614	0.0718746	2.513	0.0393497	0.0120	**
CCAA_BALEARES	0.352374	0.0643779	5.474	0.0744398	<0.0001	***
CCAA_CANARIAS	0.272122	0.0589652	4.615	0.0583904	<0.0001	***
CCAA_CANTABRIA	0.263415	0.0746179	3.530	0.0565363	0.0004	***
CCAA_CASTLEON	0.232027	0.0430788	5.386	0.0504058	<0.0001	***
CCAA_CASTMAN	0.255798	0.0475752	5.377	0.0552168	<0.0001	***
CCAA_CATALU	0.524037	0.0419920	12.48	0.108804	<0.0001	***
CCAA_VALENCIA	0.291421	0.0459440	6.343	0.0626091	<0.0001	***
CCAA_EXTR	-0.0195726	0.0626403	-0.3125	-0.00440481	0.7547	
CCAA_GALICIA	0.178178	0.0406980	4.378	0.0390617	<0.0001	***
CCAA_MADRID	0.732151	0.0538186	13.60	0.144390	<0.0001	***
CCAA_MURCIA	0.178201	0.0624111	2.855	0.0388631	0.0043	***
CCAA_NAVARRA	0.479730	0.0771267	6.220	0.0986865	<0.0001	***
CCAA_PV	0.461836	0.0581804	7.938	0.0957392	<0.0001	***
CCAA_RIOJA	0.379010	0.0842636	4.498	0.0795241	<0.0001	***
CCAA_CEUTA	0.818285	0.189564	4.317	0.155029	<0.0001	***
CCAA_MELILLA	0.0238483	0.172488	0.1383	0.00533103	0.8900	
HORAS	-0.0275403	0.00155961	-17.66	-0.00617994	<0.0001	***
JORNADA	1.05090	0.0440313	23.87	0.252561	<0.0001	***
SECTOR_SEC	1.10151	0.0539578	20.41	0.219124	<0.0001	***
SECTOR_TERC	1.06820	0.0517527	20.64	0.249885	<0.0001	***
postCovid	1.35736	0.138536	9.798	0.292239	<0.0001	***
Int_GENERO	-0.0300864	0.0336253	-0.8948	-0.00676674	0.3709	
Int_EDAD_16	-0.0366917	0.252444	-0.1453	-0.00828087	0.8844	
Int_EDAD_20	-0.0569510	0.0865092	-0.6583	-0.0128890	0.5103	
Int_EDAD_25	-0.0121646	0.0695762	-0.1748	-0.00273464	0.8612	
Int_EDAD_30	-0.00167501	0.0644915	-0.02597	-0.000375958	0.9793	
Int_EDAD_35	-0.0114356	0.0583338	-0.1960	-0.00257027	0.8446	
Int_EDAD_45	-0.115856	0.0544744	-2.127	-0.0263904	0.0334	**
Int_EDAD_50	-0.139202	0.0558452	-2.493	-0.0318019	0.0127	**
Int_EDAD_55	-0.249318	0.0597858	-4.170	-0.0577298	<0.0001	***
Int_EDAD_60	-0.365460	0.0713351	-5.123	-0.0858255	<0.0001	***
Int_EST_CASADO	-0.150762	0.0395610	-3.811	-0.0341951	0.0001	***
Int_EST_VIUDO	0.113095	0.165596	0.6830	0.0249054	0.4946	
Int_EST_DIVORC	-0.0627625	0.0664445	-0.9446	-0.0142123	0.3449	

Impacto de la COVID-19 en el Mercado Laboral

Int_N_SINESTUDIOS	0.278567	0.188341	1.479	0.0594932	0.1391	
Int_N_EDSECUN	0.0329176	0.0830429	0.3964	0.00736174	0.6918	
Int_N_FP	0.116625	0.0925028	1.261	0.0257149	0.2074	
Int_N_EDSUP	-0.00395908	0.0842881	-0.04697	-0.000888729	0.9625	
Int_CCAA_ARAGON	-0.113393	0.0818105	-1.386	-0.0258678	0.1657	
Int_CCAA_ASTURIAS	-0.0375843	0.104232	-0.3606	-0.00848265	0.7184	
Int_CCAA_BALEARES	-0.120251	0.0960305	-1.252	-0.0274669	0.2105	
Int_CCAA_CANARIAS	0.0771503	0.0892804	0.8641	0.0170998	0.3875	
Int_CCAA_CANTABRIA	-0.232259	0.106416	-2.183	-0.0538416	0.0291	**
Int_CCAA_CASTLEON	0.0133116	0.0624093	0.2133	0.00298128	0.8311	
Int_CCAA_CASTMAN	-0.103539	0.0695170	-1.489	-0.0235807	0.1364	
Int_CCAA_CATALU	-0.0499684	0.0620802	-0.8049	-0.0112924	0.4209	
Int_CCAA_VALENCIA	-0.104248	0.0664835	-1.568	-0.0237400	0.1169	
Int_CCAA_EXTR	-0.187717	0.0913638	-2.055	-0.0432635	0.0399	**
Int_CCAA_GALICIA	-0.00430447	0.0591115	-0.07282	-0.000966497	0.9419	
Int_CCAA_MADRID	-0.0531499	0.0774740	-0.6860	-0.0120204	0.4927	
Int_CCAA_MURCIA	0.0528055	0.0904502	0.5838	0.0117507	0.5594	
Int_CCAA_NAVARRA	-0.290485	0.109839	-2.645	-0.0678076	0.0082	***
Int_CCAA_PV	-0.131048	0.0850776	-1.540	-0.0299700	0.1235	
Int_CCAA_RIOJA	-0.0328810	0.117882	-0.2789	-0.00741597	0.7803	
Int_CCAA_CEUTA	-0.387457	0.268418	-1.443	-0.0914602	0.1489	
Int_CCAA_MELILLA	-0.364070	0.249715	-1.458	-0.0857345	0.1449	
Int_HORAS	-0.0479476	0.00253699	-18.90	-0.0107593	<0.0001	***
Int_JORNADA	0.671599	0.0673578	9.971	0.146714	<0.0001	***
Int_SECTOR_SEC	-0.0316337	0.0801719	-0.3946	-0.00712612	0.6932	
Int_SECTOR_TERC	-0.125087	0.0771396	-1.622	-0.0282502	0.1049	

Media de la vble. dep.	0.650773	D.T. de la vble. dep.	0.476729
R-cuadrado de McFadden	0.094220	R-cuadrado corregido	0.092787
Log-verosimilitud	-50562.07	Criterio de Akaike	101284.1
Criterio de Schwarz	102033.4	Crit. de Hannan-Quinn	101513.0

*Evaluado en la media

Número de casos 'correctamente predichos' = 60674 (70.3%)

f(beta'x) en la media de las variables independientes = 0.477

Contraste de razón de verosimilitudes: Chi-cuadrado(79) = 10519 [0.0000]

		Predicho	
		0	1
Observado	0	9296	20836
	1	4772	51378

3.2.1 Análisis de los resultados del segundo modelo

Cabe recordar que a través de esta estimación se quiere observar el efecto a corto plazo de la pandemia en el tipo de contrato de los individuos una vez se comenzaron a levantar las restricciones (verano de 2020). Por ello, se han utilizado los datos del tercer trimestre de 2020 junto con los del tercer trimestre de 2019, y así poder comparar los resultados con un periodo de tiempo en el que no había empezado la crisis sanitaria.

La anterior estimación tiene un 70,3% de casos correctamente predichos, que puede considerarse una capacidad de predicción alta para poder tener en cuenta los resultados.

En primer lugar, a través de las variables no interactivas, se estudiará el efecto de las variables en las probabilidades de tener un contrato indefinido antes del comienzo de la pandemia, y después se finalizará analizando el efecto que tuvo la pandemia en dichas variables.

Comenzando a analizar las variables no interactivas, se puede observar que el **género** de los individuos tiene una gran relevancia en las probabilidades de tener una mayor estabilidad laboral. Debido a que esta variable tiene asociado un coeficiente con signo positivo, los hombres tienen más probabilidades que las mujeres de tener un contrato indefinido. Al estudiar su pendiente, comprobamos que un 3,11% más.

Respecto a la **edad** de los individuos, se aprecia que tener menos de 35 años ejerce un efecto negativo en la probabilidad de tener un contrato indefinido, en relación a los individuos que tienen entre 40-45 años. No obstante, a partir de dicha edad, la probabilidad incrementa a medida que aumentan los años que tenga el individuo.

Otra característica que afecta a las probabilidades de manera significativa es el **estado civil** de los individuos. EST_CASADO, EST_VIUDO y EST_DIVORC se presentan como significativas y todas tienen asociado un efecto positivo. Esto significa que, en relación con los individuos solteros, aquellos que estén casados, viudos o divorciados tienen mayores probabilidades de tener contratos indefinidos. Específicamente, las probabilidades se incrementan en un 6,54%, un 8,40% y un 3,60%.

Entre las variables que se anticipaba que tuvieran una gran relevancia está el **nivel de formación** y, como era de esperar, la estimación ha demostrado que en relación a tener estudios primarios, cuando el nivel de formación es mayor, las probabilidades de obtener un contrato indefinido son incrementadas. Concretamente, si los individuos disponen de educación secundaria, educación profesional o educación superior, dichas probabilidades aumentan un 8,85%, un 9,89% y un 15,14%, respectivamente. En cambio, si los individuos no disponen de estudios primarios, estas probabilidades disminuyen un 6,82%.

A excepción de Extremadura y Melilla, todas las **comunidades autónomas** han mostrado relevancia a la hora de disponer de una mayor estabilidad laboral. Con respecto a Andalucía, los lugares de residencia que han manifestado probabilidades más altas han sido Aragón, Cataluña, Madrid, Navarra y el País Vasco, siendo Madrid y Cataluña las que incrementan dichas posibilidades en mayor medida, en un 14,43% y un 10,88%.

En cuanto a la única variable cuantitativa que hay en el modelo, la variable **HORAS**, su interpretación difiere ligeramente de la de las demás variables, que son categóricas. Al fijarnos en el coeficiente, se observa que para un aumento en una unidad de dicha variable, el logaritmo de las probabilidades de tener un contrato indefinido disminuye en un -0.0275403 .

Las variables relacionadas con las características del empleo que tienen los individuos y que no se añadieron en el primer modelo, JORNADA y SECTOR, también son significativas. Por un lado, se aprecia que la variable dicotómica **JORNADA** tiene vinculado un coeficiente positivo en la probabilidad de disponer de un contrato indefinido. Dado que la categoría base es trabajar a jornada completa, este dato nos indica que si los individuos tienen una jornada laboral completa, tienen un 25,25% más de probabilidad de tener un contrato indefinido en comparación con los individuos que trabajan a jornada parcial. Por otro lado, en relación al **sector económico** en el que trabajan las personas, se aprecia que las variables ficticias SECTOR_SEC y SECTOR_TERC también tienen un efecto positivo en el objeto de estudio, al igual que la variable JORNADA. En este caso, siendo el grupo de referencia los individuos que disponen de un empleo en el sector primario, los resultados verifican que respecto a dicho grupo, si los individuos trabajan en el sector secundario o terciario, sus probabilidades de adquirir una mayor estabilidad laboral aumentan en un 21,91% y un 24,98%, respectivamente.

Para determinar el efecto general que ha tenido la COVID-19 en los tipos de contratos, nos centraremos en la variable creada al comienzo de este trabajo, **postCovid**, que indica si las observaciones pertenecen a la época pre-COVID o post-COVID. Tiene asociado un coeficiente con signo positivo, lo que señala que si los individuos estaban trabajando a finales de 2020, tenían más probabilidades de tener un contrato indefinido que uno parcial en relación a si trabajaban en 2019. Sin embargo, dicho en otras

palabras, los contratos parciales se han visto más afectados que los indefinidos. Entonces, teniendo en cuenta que España se encuentra entre los países europeos con mayor temporalidad, este dato indica que los puestos de trabajo de una gran cantidad de personas se vieron afectados por la COVID-19.

Por último, para especificar el efecto de la pandemia en aspectos concretos, se analizarán los resultados de las **variables de interacción**. Al igual que en el modelo anterior, éstas describen el efecto que ha tenido la pandemia sobre las variables entre el cuarto trimestre de 2019 y el cuarto trimestre de 2020. A primera vista, se puede apreciar que la mayoría de las variables interactivas carecen de significación, pero al ser parte de variables cualitativas, no se han eliminado de la estimación. Cabe destacar que ninguna de las variables vinculadas con la formación y el sector económico en el que trabajan los individuos se han mostrado relevantes, por lo tanto, la COVID-19 no ha afectado en gran manera las probabilidades de tener un contrato indefinido en función de dichas variables. Por otro lado, las variables que sí se han mostrado relevantes y que se analizarán a continuación, están relacionadas con la edad de los individuos, el estado civil, el lugar de residencia, las horas trabajadas y el tipo de jornada. Como era de esperar, todas estas variables tienen asociado un efecto negativo, a excepción de una, que tiene un coeficiente con signo positivo.

Comenzando por las que tienen un efecto negativo, en primer lugar, encontramos cuatro variables de interacción relacionadas con la edad de los individuos: Int_EDAD_45, Int_EDAD_50, Int_EDAD_55 e Int_EDAD_60. Examinando los resultados, se comprueba que respecto al grupo de referencia, las probabilidades de tener un contrato indefinido se reducen a partir de los 45 años. En concreto, se observa que si los individuos tienen entre 45 y 49 años, sus probabilidades disminuyen en un 2,6%; si tienen entre 50 y 54, disminuyen un 3,1%; si tienen entre 55 y 59, disminuyen un 5,7%; y si tienen entre 60 y 64, disminuyen un 8,5%.

Posteriormente se encuentran las variables de interacción correspondientes al estado civil, de entre las cuales solo muestra relevancia Int_EST_CASADO. Nos indica que, a causa de la pandemia, si los individuos están casados tienen un 3,41% menos de probabilidades de tener un contrato indefinido en comparación con los individuos solteros.

Siguiendo con las variables interactivas vinculadas con el lugar de residencia, la estimación muestra que el efecto de la COVID-19 en los contratos solo se muestra relevante en tres comunidades autónomas: Cantabria, Extremadura y Navarra. Al analizar los resultados respecto a Andalucía, la categoría base, se comprueba que la probabilidad de tener un contrato indefinido en dichas comunidades se reduce un 5,38%, un 4,32% y un 6,78%, respectivamente.

La última variable de interacción significativa a la que le ha sido asociado un coeficiente negativo es la variable Int_HORAS, que señala que por cada unidad de hora trabajada, la probabilidad de tener un contrato indefinido se reduce en un 1%.

Por otra parte se halla la única variable interactiva relevante con un coeficiente positivo, la variable Int_JORNADA. Teniendo como grupo de referencia que el individuo trabaje a jornada completa, se revela que como efecto a la COVID-19, la probabilidad de tener un contrato indefinido aumenta un 14,6% si los individuos trabajan a jornada completa en relación a si trabajan a jornada parcial.

4. CONCLUSIONES

En este trabajo se ha planteado analizar el efecto provocado por la COVID-19 en el mercado laboral español, además de extraer conclusiones acerca de cómo algunas características ejercen influencia en la situación laboral de los individuos y sus posibilidades de acceder al mercado laboral.

Utilizando como base de datos las Encuestas de Población Activa (EPA) correspondientes a 2019 y 2020 publicadas por el Instituto Nacional de Estadística (INE), se han determinado las características más adecuadas para estudiar la situación de empleo y la estabilidad laboral de los individuos. Posteriormente, mediante el conjunto de variables elegido, se han establecido dos modelos econométricos que posibilitaran la ejecución del análisis.

Las estimaciones de los modelos han demostrado que, a pesar de las medidas adoptadas en el ámbito laboral, los efectos de la COVID-19 han tenido consecuencias significativas en el mercado de trabajo. Si bien es cierto que fueron efectivas y lograron reducir los efectos negativos y que numerosos individuos que se encontraban en una situación de precariedad mantuvieran sus puestos de trabajo, no resultaron ser suficientes. Pese a los múltiples esfuerzos realizados, multitud de individuos y negocios en España resultaron económicamente afectados y tres años después, aún no les ha sido posible recuperarse.

Mediante los modelos econométricos se ha cuantificado la influencia de las características, en primer lugar, sobre la probabilidad de los individuos de acceder al mercado laboral y, en segundo lugar, sobre la probabilidad de tener un contrato indefinido. Asimismo, a través de variables de interacción, se ha analizado el efecto de la crisis sanitaria sobre dichas características.

Los resultados obtenidos en las modelizaciones concluyen que el género es una variable que sí tiene relevancia en el ámbito laboral. La brecha de participación entre hombres y mujeres se ha ido reduciendo durante los últimos años, hasta que antes de la pandemia las mujeres tenían mayores probabilidades de estar trabajando en relación a los hombres. No obstante, de los datos obtenidos se desprende que la COVID-19 ha perjudicado en mayor medida a las mujeres y ha provocado que dicha brecha vuelva a crecer. Esta situación se debe principalmente a los efectos negativos que han sufrido los contratos temporales como consecuencia de la crisis sanitaria. Las mujeres tienen

más probabilidades en comparación con los hombres de tener un contrato temporal, por lo que la pérdida de estos contratos y, por tanto, la pérdida de trabajo, ha empeorado la situación laboral de las mujeres.

Otra situación a considerar es que la crisis ha provocado que los jóvenes sean uno de los grupos de edad más perjudicados por la COVID-19. Además, el nivel de estudios requerido para tener una mejor situación laboral ha aumentado después de la pandemia, por lo que es posible que cada vez un mayor número de jóvenes prefiera esperar a terminar su formación antes de incorporarse en el mercado laboral. Por otro lado, los datos reflejan que los solteros en 2020 podían encontrarse con mayores dificultades de estar trabajando o tener contratos indefinidos que antes de que comenzara la pandemia. No obstante, cabe destacar que en España, la edad media de los individuos que contraen matrimonio por primera vez ha ido en aumento durante los últimos años, por lo que al analizar estos resultados, es preciso señalar que el grupo de solteros está principalmente formado por individuos jóvenes. Entonces, si bien los jóvenes y solteros han resultado ser de los grupos de personas más afectados por la pandemia, esto puede deberse a una tardía incorporación al mercado laboral y a la reducción de ofertas de trabajo por parte de las empresas.

En cuanto a las personas mayores de 40 años, a partir de los datos se concluye que los individuos con más probabilidades de encontrarse en una buena situación laboral son aquellos que tienen entre 40 y 50 años y que se han casado al menos una vez. Por otra parte, antes de la pandemia las personas mayores de 50 años se beneficiaban de mayores probabilidades de adquirir un contrato indefinido a medida que aumentaba la edad del individuo, pero, como se ha determinado mediante las variables de interacción, estas probabilidades han disminuido de manera significativa en 2020, lo cual es una consecuencia de la pérdida de empleo causada por la COVID-19.

En relación con la pérdida de empleo, entre las comunidades autónomas más desfavorecidas por la pandemia respecto a Andalucía se encuentran las Islas Baleares y las Islas Canarias; zonas que cuentan con mayor peso del turismo e incidencia en la movilidad. Esto ha podido ocurrir en vista de que la hostelería y el turismo fueron los sectores más afectados en 2020 y que incluso a día de hoy están en proceso de recuperación.

Como conclusión final, se quiere señalar que aunque la COVID-19 haya ocasionado peores consecuencias en unos aspectos más que en otros, es innegable que todos los ciudadanos se han visto perjudicados en mayor o menor medida, sin importar sus características personales. Los cambios necesarios para retomar el estado en el que se encontraba el mercado laboral no se producirán de manera inmediata, pero por medio de ayudas y nuevas medidas, a largo plazo el empleo y las empresas lograrán reponerse de esta crisis.

5. BIBLIOGRAFÍA:

Covid: todas las fechas destacadas del año de la pandemia en España. (2021, March

13). Redacción Médica. <https://www.redaccionmedica.com/secciones/sanidad-hoy/covid-espana-todas-fechas-destacadas-ano-pandemia-4910>

INE - Instituto Nacional de Estadística. (n.d.). *Sección prensa / Encuesta de Población Activa (EPA).* INE - Instituto Nacional De Estadística.

https://www.ine.es/prensa/epa_tabla.htm

Statistics | Eurostat. (n.d.). *Temporary employees as a percentage of the total number of employees, by sex and age (%)*. Eurostat.

https://ec.europa.eu/eurostat/databrowser/view/lfsq_etpga/default/table?lang=en

De Empleo Estatal, S. P. (n.d.). *Resumen de datos estadísticos.* Servicio Público De Empleo Estatal. <https://www.sepe.es/HomeSepe/que-es-el-sepe/estadisticas/datos-avance/datos.html>

INE - Instituto Nacional de Estadística. (n.d.). *Ocupados por tipo de contrato o relación laboral de los asalariados, sexo y tipo de jornada(3961).* INE.

<https://www.ine.es/jaxiT3/Tabla.htm?t=3961&L=0>

Medidas económicas y sociales en la crisis del COVID-19. (n.d.).

<https://www.lamoncloa.gob.es/consejodeministros/Paginas/enlaces/210420-enlace-medidas.aspx>

DatosRTVE, D. F. /. (2021, December 23). *Radiografía de la temporalidad en España.*

RTVE.es. <https://www.rtve.es/noticias/20211223/temporalidad-empleo-espana/2103964.shtml>

Fita Ortega, F. (n.d.). *La pandemia COVID-19, mercado de trabajo y medidas laborales en España.* https://www.cielolaboral.com/wp-content/uploads/2021/04/fita_noticias_cielo_n3_2021.pdf

Jaén, M. E. F. (n.d.). *FEIR 45: Regresión logística*.

https://gauss.inf.um.es/feir/45/#1_introducci%C3%B3n

Jingles. (2021, December 9). *Why Linear Regression is not suitable for Classification*.

Medium. <https://towardsdatascience.com/why-linear-regression-is-not-suitable-for-binary-classification-c64457be8e28>

Aguayo Canela, M. (2007). *Cómo hacer una Regresión Logística con SPSS “paso a*

paso”. (I). [http://metodos-avanzados.sociales.uba.ar/wp-](http://metodos-avanzados.sociales.uba.ar/wp-content/uploads/sites/216/2014/03/Regres_log_AGUAYO-otros.pdf)

[content/uploads/sites/216/2014/03/Regres_log_AGUAYO-otros.pdf](http://metodos-avanzados.sociales.uba.ar/wp-content/uploads/sites/216/2014/03/Regres_log_AGUAYO-otros.pdf)

Moral Peláez, I. (n.d.). *Modelos de regresión: lineal simple y regresión logística*.

<https://revistaseden.org/files/14-CAP%2014.pdf>

LaValley, M. P. (2008, May 6). *Logistic Regression*. *Circulation*.

<https://www.ahajournals.org/doi/full/10.1161/CIRCULATIONAHA.106.682658>

Dolores, F. P. M. (2000, December 1). *La regresión logística: una herramienta versátil*.

Nefrología. <https://www.revistanefrologia.com/es-la-regresion-logistica-una-herramienta-versatil-articulo-X0211699500035664>