



Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

**Grado en Administración y Dirección de
Empresas**

**Análisis de los determinantes del
desempleo en España**

Presentado por:

Ana Noriega de Tapia

Valladolid, 6 de Julio de 2022

RESUMEN

El mercado laboral es uno de los determinantes que comprueban el buen funcionamiento de la economía del país, siendo el desempleo un fenómeno persistente. El objetivo de este trabajo es analizar qué características personales son determinantes en el desempleo en España. La evidencia presentada en este trabajo indica que hay numerosas variables que inciden en la tasa de desempleo siendo las características personales con mayor repercusión sobre dicha tasa: la edad, sexo, nacionalidad, estado civil y nivel de estudios. Han sido analizadas a través de un análisis descriptivo basado en la Encuesta de Población Activa (INE) y un modelo Logit. Tras el estudio realizado, se ha obtenido que el perfil con mayores tasas de desempleo son mujeres, jóvenes, extranjeras y sin estudios.

ABSTRACT

The labour market is one of the determining factors that prove the proper functioning of the country's economy, being unemployment a persistent phenomenon. The aim of this work is to analyze which personal characteristics are determining factors of unemployment in Spain. The evidence presented in this work indicates that there are numerous variables that affect the unemployment rate, the personal characteristics with the greatest impact on said rate are: age, sex, nationality, marital status and educational level. They have been analyzed through a descriptive analysis based on the Active Population Survey (INE) and a Logit model. After the study carried out, it has been obtained that the profile with the highest unemployment rates are young women, foreigners and without studies.

Clasificación JEL: J64 (Desempleo: modelos, duración, incidencia y búsqueda de empleo), C35 (Modelos de regresión discreta y de elección cuantitativa).

Palabras clave: tasa de desempleo, mercado laboral, Logit, Encuesta de Población Activa.

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN	1
2. METODOLOGÍA	4
2.1 ENCUESTA DE POBLACIÓN ACTIVA (EPA)	4
2.2 MODELOS ECONOMÉTRICOS	5
2.2.1 <i>Estimación del modelo Logit</i>	9
2.2.2 <i>Interpretación del modelo Logit</i>	10
3. ANÁLISIS DESCRIPTIVO DEL DESEMPLEO	11
4. MODELO LOGIT	23
4.1 ANÁLISIS DE ESCENARIOS.	26
4.2 CURVAS DE RESPUESTA.	29
5. CONCLUSIONES	32
6. BIBLIOGRAFÍA	34

ÍNDICE DE TABLAS

Tabla 3.1: Porcentaje de personas desempleadas según sexo	12
Tabla 3.2: Porcentaje de ocupados y desempleados por grupos de edad	14
Tabla 3.3: Distribución de la población encuestada según la nacionalidad	20
Tabla 3.4: Nivel de estudios	20
Tabla 3.5: Tasa de ocupación en función del nivel de estudios.	21
Tabla 4.1: Estimación del modelo mediante el método Logit	23
Tabla 4.2: Criterio de predicciones correctas	24
Tabla 4.3: Probabilidad estimada y efecto en el desempleo de las variables dependientes	25
Tabla 4.4: Análisis de escenarios	26

ÍNDICE DE GRÁFICOS

Gráfico 1.1: Evolución de la tasa de desempleo en España	3
Gráfico 3.1: Tasa de desempleo en España del primer trimestre de 2006 al primer trimestre de 2022	11
Gráfico 3.2: Salario por hora de la mujer respecto del hombre, 2019	13
Gráfico 3.3: Tasas de paro en España por grupos de edad y sexo, 2021	16
Gráfico 3.4: Tasa de ocupados o parados en función del estado civil	17

Gráfico 3.5: Tasa de ocupados y parados en función del estado civil y sexo..	18
Gráfico 3.6: Distribución de la población encuestada según el nivel de estudios	21
Gráfico 4.1: Curvas de respuesta para la variable sexo	29
Gráfico 4.2: Curvas de respuesta para la variable estado civil	29
Gráfico 4.3: Curvas de respuesta para la variable nacionalidad.....	30
Gráfico 4.4: Curvas de respuesta para la variable nivel de estudios	31

1. INTRODUCCIÓN

A nivel conceptual, el Instituto Nacional de Estadística (INE) define la tasa de paro como: “el cociente entre el número de parados y el de activos”. Mientras que la tasa de empleo resulta del cociente entre la población que se encuentra trabajando activamente con respecto al total de la población en edad de trabajar.

Según Kirschen, E.S (1969) los objetivos coyunturales de la política económica que garantizan el equilibrio interno son: el pleno empleo y la estabilidad de precios. En este trabajo realizamos un análisis reducido de algunas características personales que afectan a la tasa de desempleo, teniendo en cuenta que no se puede extender a todas las variables que inciden en dicha tasa, como por ejemplo el salario.

El desempleo, es el desequilibrio entre la oferta y demanda de trabajo, en las condiciones y niveles de salarios existentes en un momento dado dentro de un mercado laboral. Es decir, personas en edad laboral que estando dispuestas a trabajar y en condiciones de hacerlo, no encuentran un empleo. Alcanzar una situación óptima de ocupación de los recursos humanos, no significa que la tasa de paro tenga que ser igual a cero, sino que implica la utilización óptima de los recursos productivos. De hecho, siempre existe cierto nivel de paro involuntario en todas las economías. A continuación, especificaré los cuatro tipos de desempleo:

- Paro estructural: es aquel que subsiste en el tiempo en ciertas regiones y/o ramas de actividad a consecuencia de tener una inadecuada estructura económica. Puede ser causado por la existencia de actividades productivas en regresión.
- Paro friccional: puede deberse a la rotación constante del empleo debido a los avances tecnológicos y necesidades cambiantes, y/o debido a la falta de homogeneidad en cuanto a la ubicación del empleo y la cualificación requerida.
- Paro estacional: vinculado a actividades que demandan mano de obra solo en periodos temporales, como la agricultura, turismo etc.

- Paro cíclico: tipo de desempleo propio de periodos de crisis económica, por la contracción del consumo de bienes y servicios, inversión y la caída de la producción en fases recesivas.

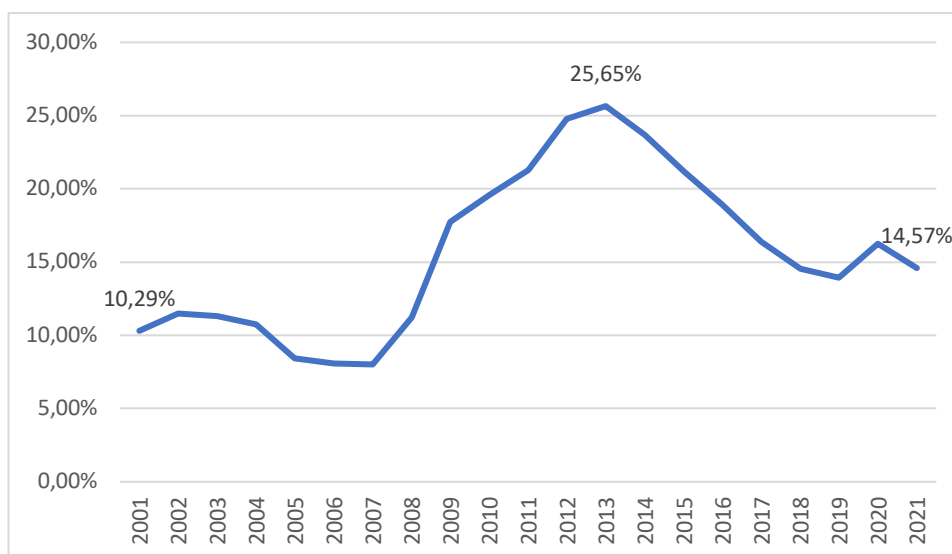
He seleccionado este tema ya que el desempleo es un problema con importantes repercusiones económicas, sociales, personales e incluso familiares. Incide con gran impacto en la salud mental de las personas, por ello, es fundamental estudiar qué factores agravan dicha tasa. Personalmente, como alumna centrada en poder iniciarse en el mercado laboral, el interés por este tema se incrementa teniendo en cuenta las elevadas tasas de desempleo juvenil. La tasa de población ocupada en el tercer trimestre del 2021 entre 16 y 24 años alcanza un 67,99%, dato muy inferior de las personas mayores de 25 con un 81,53%.

Los tres objetivos relacionados con el mercado laboral español, incluidos en la Estrategia de Europa 2020 son: aumentar las tasas de empleo entre 20 y 64 años, para crear más y mejores puestos de trabajo, coordinando a los gobiernos con iniciativas para la promoción del empleo. El segundo objetivo es reducir el abandono escolar a menos de un 10% y en último lugar elevar hasta el 40% la tasa de población de entre 30 y 34 años que finaliza los estudios superiores.

A pesar de que, en 2021, año estudiado, las tasas de desempleo han disminuido, el desempleo constituye hoy en día uno de los grandes retos que afronta la Unión Europea. Esta estrategia establecía lograr una tasa de empleo del 75% en la población activa, sin embargo, el Ministerio de Trabajo en el Observatorio de seguimiento evaluó en el último trimestre del 2020 un porcentaje de ocupados de 66,2%, lejos del objetivo.

Según el Servicio Público de Empleo Estatal (SEPE) en España durante el mes de febrero de 2021, la cifra de parados ascendía a 4.008.789 dato más alto tras la crisis del COVID 19. El comportamiento cíclico de la oferta laboral es uno de los temas más tradicionales de la economía laboral. En el *Gráfico 1.1* podemos observar la progresión histórica de la tasa de desempleo en España. Destaca como se acentúa la curva tras las crisis de la economía española acontecidas en el 2008 con el fin de la burbuja del ladrillo y en 2020 consecuencia de la pandemia mundial, viéndose afectadas las tasas de forma muy negativa.

Gráfico 1.1: Evolución de la tasa de desempleo en España



Fuente: EPA.

Este estudio tiene como objetivo analizar los determinantes del desempleo en España. En concreto, como ciertas características personales afectan a la probabilidad de estar desempleado. Las variables estudiadas son: el sexo, edad, estado civil, nacionalidad y nivel de estudios. Para ello se va a realizar un análisis descriptivo condicionado y un estudio de las características del desempleo mediante un modelo Logit.

En primer lugar, se explicará la metodología utilizada, tanto de la encuesta utilizada para el tratamiento de los datos (Encuesta de Población Activa) como los modelos econométricos planteados, eligiendo el más adecuado (Modelo Logit). A continuación, se realizará un análisis descriptivo de todas las variables, elaborando una serie de gráficos y tablas que sintetizan los resultados obtenidos en la encuesta.

Por último, se realizará una estimación del modelo Logit, calculando el efecto marginal de las variables, analizando una serie de escenarios con características concretas y se efectuarán unas curvas de respuesta para todas las variables en función de la edad. Finalmente se expondrán las conclusiones halladas con el estudio realizado.

2. METODOLOGÍA

2.1 Encuesta de Población Activa (EPA)

El objetivo de este trabajo es realizar un estudio sobre qué características personales son determinantes en el desempleo entre la población española. Para ello, he utilizado como fuente la Encuesta de Población Activa (EPA) realizada por el Instituto Nacional de Estadística (INE). La EPA comenzó a realizarse en 1964, desde entonces, forma parte de una investigación continua y trimestral que va dirigida a una muestra de 65.000 familias, es decir, unas 200.000 personas.

Los resultados se obtienen mes y medio después de finalizar el trabajo de campo. En este estudio he analizado el tercer trimestre del 2021. Para poder comprender mejor los resultados, cabe recordar definiciones clave sobre el mercado de trabajo:

- **Activos:** personas de 16 o más años que durante la semana anterior en que se realiza la encuesta suministran mano de obra para la producción de bienes y servicios o están disponibles y en condiciones de incorporarse a dicha producción. Dentro de esta definición encontramos a:
 - **Ocupados:** siendo de 16 o más años que durante la semana de referencia han estado trabajando al menos una hora a cambio de una retribución o quien teniendo trabajo han estado ausentes temporalmente por vacaciones, enfermedad... Dentro de este grupo quedan incluidos trabajadores por cuenta propia y asalariados.
 - **Parados:** personas de 16 o más años que durante la semana de referencia han estado sin trabajo, disponibles para trabajar y buscando activamente empleo.
- **Inactivos:** conjunto de ciudadanos de 16 años o más, que cumplen las condiciones para trabajar y que no se incorporan al mercado laboral. Forman este colectivo, estudiantes, jubilados, personas que se dedican a las labores del hogar, incapacidad permanente etc.

Para abordar este estudio he analizado únicamente a las personas activas (ocupados y parados) estudiando la influencia que ejercen variables independientes sobre el estado laboral de la población. Las variables independientes elegidas son: edad, sexo, estado civil, nacionalidad y nivel de estudios.

AOI es la variable que clasifica a todas las personas de 16 y más años por relación con la actividad económica según criterios OIT. Con ella analizaremos a la población activa, es decir la ocupada y desempleada con los parámetros que indica la *Tabla 2.1*. Por lo tanto, simplificaremos los valores en:

$$Y_i (\text{DESEMPLEADO}) = \begin{cases} 1 & \text{si es DESEMPLEADO} \\ 0 & \text{si es OCUPADO} \end{cases}$$

Tabla 2.1: Explicación de valores de la variable AOI

ACTIVOS	OCUPADOS (0)	3- Ocupados subempleados por insuficiencia de horas
		4- Resto de ocupados
	DESEMPLEADOS (1)	5- Parados que buscan primer empleo
		6- Parados que han trabajado antes
INACTIVOS	7- Inactivos 1: desanimados	
	8- Inactivos 2: junto con los desanimados forman los activos potenciales	
	9- Inactivos 3: resto de inactivos	

Fuente: elaboración propia.

2.2 Modelos econométricos

Para analizar la probabilidad de estar desempleado y qué variables independientes afectan en mayor proporción a esta situación laboral, voy a utilizar modelos econométricos de elección discreta o binaria que tratan de explicar una variable cualitativa, en este caso el desempleo, a través de otras variables cualitativas y cuantitativas.

Existen varios modelos econométricos a aplicar cuando tenemos variables dicotómicas entre ellos pueden realizar este análisis: el modelo lineal de probabilidad, modelo Logit y Probit.

Por lo tanto, analizaré las características de los modelos para determinar el más adecuado para este estudio:

○ **MODELO LINEAL DE PROBABILIDAD:**

El modelo lineal de probabilidad es un modelo de regresión lineal donde la forma escalar del modelo es $Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_i$ para $i = 1 \dots N$

Las variables explicativas del modelo o regresores son X_i : edad, sexo, estado civil, nacionalidad y nivel de estudios. Y_i es una variable dependiente dicotómica en concreto variable de Bernoulli definida como:

$$Y_i = \begin{cases} 1 & \text{si es DESEMPLEADO} & p_i = p(Y_i = 1) \\ 0 & \text{si es OCUPADO} & 1 - p_i = p(Y_i = 0) \end{cases}$$

$$E(Y_i | X) = p_i (Y = 1 | X) = p(x_i)$$

$$Var(Y_i | X) = p_i(x) [1 - p_i(x)]$$

A partir de la especificación lineal del modelo y asumiendo que $E(\varepsilon_i) = 0$

Por tanto, $E(Y | X) = p(x) = (\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)$

Sin embargo, la estimación por MCO del modelo lineal de probabilidad presenta una serie de problemas que imposibilita su uso para realizar este análisis:

- La distribución de las perturbaciones no es una normal. Solo puede tomar dos valores.
- Presenta heteroscedasticidad: $Var(\varepsilon_i) = p_i(1 - p_i)$
- El principal problema es que la variable $\hat{Y}_i = \hat{p}_i$ puede tomar valores tanto en el intervalo $[0, 1]$ como fuera de él.

○ **MODELO LOGIT Y PROBIT**

La estimación e interpretación de los modelos lineales de probabilidad plantean una serie de problemas que han llevado a la búsqueda de otros modelos que permitan estimaciones más fiables de las variables dicotómicas, además la formulación de los modelos Logit y Probit garantiza que la probabilidad estimada se encuentre entre el intervalo (0,1). Para ello se considera que la variable depende de una función de distribución acumulada que, por lo tanto, está comprendida entre 0 y 1.

Partimos de una variable dicotómica, este tipo de variable se caracteriza por que solo puede tomar dos valores, que habitualmente son cero como ausencia de una característica (ocupado) y uno como presencia (desempleado). En este caso, es una variable Bernoulli que sigue una distribución de probabilidad discreta que toma el valor 1 para la probabilidad de éxito p_i y 0 para la probabilidad de fracaso $1 - p_i$.

$$Y_i = \begin{cases} 1 & \text{si es DESEMPLEADO} \\ 0 & \text{si es OCUPADO} \end{cases} \quad \begin{matrix} p_i = p(Y_i = 1) \\ 1 - p_i = p(Y_i = 0) \end{matrix}$$

Estos modelos, relacionan la variable endógena Y_i con las variables explicativas X_{ki} a través de una función de distribución.

$$Y_i = F(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}) + u_i = F(X'_i \beta) + u_i$$

$i = 1 \dots N$ donde $0 \leq F(X'_i \beta) \leq 1$

Siendo:

$$X'_i = (1, X_{1i}, \dots, X_{ki}) \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

Por lo tanto:

$$E(Y_i | X) = p(Y_i = 1 | X) = p_i = F(X'_i \beta)$$

$$p(Y_i = 0 | X) = (1 - p_i) = 1 - F(X'_i \beta)$$

El modelo Probit asume que F es una función de distribución normal N (0,1). Sin

embargo, para este trabajo hemos optado por el modelo Logit que asume F como una función de distribución logística de parámetros $\alpha = 0$, $\beta = 1$, siendo la esperanza cero y la desviación típica $\frac{\pi}{\sqrt{3}} = 1,814$.

El modelo Logit parte de la siguiente función de distribución acumulada logística estándar:

$$F(t) = \frac{e^t}{1+e^t} = \frac{1}{1+e^{-t}}$$

Aplicándolo a nuestro modelo, sería:

$$Y_i = \frac{e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}}} + u_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki})}} + u_i \quad \text{siendo } i = 1 \dots n$$

El valor medio de la función de Y_i que coincide con la probabilidad de que la función tome el valor 1 es:

$$E(Y_i) = p(Y_i = 1) = p_i = \frac{e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}}}$$

Si toma el valor cero:

$$p(Y_i = 0) = (1 - p_i) = \frac{1}{1 + e^{(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki})}}$$

El *odd* de un suceso, se define como el cociente de probabilidad de ocurrencia entre sus posibilidades de no ocurrencia. Indica la proporción que hay entre la probabilidad de estar desempleado (p_i) y no estarlo ($1 - p_i$). Siendo:

$$odd = \frac{p_i}{1 - p_i} = e^{\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}}$$

Por tanto, el modelo Logit se puede considerar un modelo lineal donde la variable dependiente es el logaritmo neperiano de una oportunidad (*odd*), para que los valores estimados caigan dentro del rango (0,1), se obtiene la expresión:

$$\ln(odd) = \ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}$$

Debido a las dificultades para interpretar el efecto marginal y que muchas de las variables son cualitativas, el efecto marginal en la práctica se calcula como la razón entre las probabilidades con el cociente denominado Odd ratio.

El interés de esta medida cobra importancia a la hora de comparar las ventajas de diferentes valores de la variable explicativa, calculándose el cociente entre odds. De esta forma se compara la situación de la observación “i” con la de la observación “j” (que suele ser la de referencia), el cociente entre odds mide cuanto se incrementa el odd, es decir la oportunidad entre la probabilidad de que pase un suceso o no pase.

Si calculamos el cociente entre el odd de un individuo cuando incrementa en una unidad la variable X_j manteniendo el resto constante, obtenemos dicha expresión. Para calcularle consideramos el cociente entre dos odds:

$$odds\ ratio = \frac{\frac{p_i}{(1-p_i)}}{\frac{p_j}{(1-p_j)}} = \frac{O(\beta_0 + \beta_1 X_{1i} + \dots + \beta_j (X_{ji} + 1) + \dots + \beta_k X_{ki})}{O(\beta_0 + \beta_1 X_{1i} + \dots + \beta_j X_{ji} + \dots + \beta_k X_{kj})} = e^{\beta_j}$$

Por lo tanto, el Odds ratio es: e^{β_j} . Su interpretación para las variables cuantitativas es: si aumenta X_j en una unidad, permaneciendo el resto de los regresores constantes el ratio del cociente de probabilidades varía en e^{β_j} . Es decir, representa el cambio producido en el odds ante el cambio en una unidad de una variable explicativa.

Mientras que, en el caso de las variables cualitativas carece de sentido interpretarlo con el incremento de una unidad de la variable, por lo que la ratio se interpreta como: el odd cuando vale 0 y cuando vale 1, permaneciendo el resto constante. Por ejemplo: mide como cambia la probabilidad de estar desempleado si pasas: de soltero a casado o de hombre a mujer.

2.2.1 Estimación del modelo Logit

La estimación que se realiza para estos modelos, no lineales en los parámetros, puede ser por mínimos cuadrados no lineales o máxima verosimilitud. En muestras grandes estos estimadores son consistentes, asintóticamente normales y eficientes. No se estiman por mínimo cuadrados no lineales ya que

no garantizan estas propiedades.

Por lo tanto, la estimación se realiza a través del modelo de máxima verosimilitud. Selecciona los valores que maximizan el logaritmo de la verosimilitud. La función de densidad de Y_i dado X'_i se describe como:

$$f(Y_i; \beta) = [F(X'_i\beta)]^{Y_i}[1 - F(X'_i\beta)]^{1-Y_i}$$

La función de verosimilitud es:

$$L(\beta) = \prod_{i=1}^n f(Y_i; \beta) = \prod_{i=1}^n [F(X'_i\beta)]^{Y_i} [1 - F(X'_i\beta)]^{1-Y_i}$$

El logaritmo neperiano de la función de verosimilitud:

$$\ln(L(\beta)) = \sum (Y_i \ln(F(X'_i\beta))) + (1 - Y_i) \ln [1 - F(X'_i\beta)]$$

Para concluir, el estimador máximo verosímil de β , $\hat{\beta}_{MV}$ es una función no lineal en β y hay que emplear algoritmos de optimización no lineal, por lo que no hay una expresión analítica.

$$\max_{\beta} \ln(L(\beta))$$

2.2.2 Interpretación del modelo Logit

La interpretación de los parámetros en un modelo Logit y Probit no es como la de los parámetros en el modelo lineal de probabilidad.

- El modelo lineal de probabilidad es una función de esperanza condicional lineal, $E(Y_i|X) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}$, el parámetro de la pendiente β_j es la derivada de $E(Y_i|X_j)$ con respecto a X_j , es decir β_j . Por lo tanto β_j es el efecto sobre el valor esperado de Y_i , de la variación en una unidad de X_j , manteniendo el resto constante.
- En el modelo Logit, la derivada no es constante: varía con X . Ya que los parámetros no miden:
 - En cuanto varía la probabilidad, si se incrementa en una unidad una variable explicativa, el resto ceteris paribus.
 - El efecto marginal de una variable explicativa.

Por lo tanto, en este modelo, el cambio en la probabilidad de que $Y_i = 1$ o valor esperado de $(Y | X) : E(Y | X) = p(x) = F(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki})$ si se incrementa en una unidad $X_{ji} : E(Y_i)^{+1} = p_i^{+1} = F(\beta_0 + \beta_1 X_{1i} + \dots + \beta_j (X_{ji} + 1) + \dots + \beta_k X_{ki})$ no es constante. La función de distribución no es constante porque depende de cada valor de X_j .

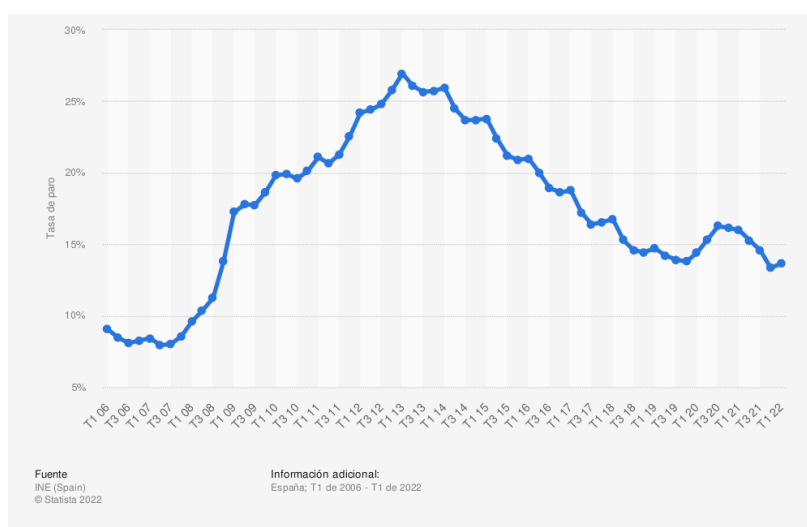
$$\frac{dp_i}{dX_{ji}} = f(X_i' \beta) \beta_j$$

En cuanto al efecto marginal, en la expresión anterior $f(X_i' \beta)$ siempre es positivo, por lo que el signo de β_j nos mide la dirección del efecto, si es negativo al aumentar X_j la probabilidad disminuye, mientras que si es positivo aumenta.

3. ANÁLISIS DESCRIPTIVO DEL DESEMPLEO

En este estudio se han analizado 60.843 observaciones, de las cuales 52.593 están ocupadas, es decir un 86,44%. Mientras que el 13,56% de la muestra estudiada están desempleadas. Como observamos en el *Gráfico 3.1* tras el repunte de la tasa de desempleo en el primer trimestre del 2020, en los últimos trimestres está siguiendo una tendencia decreciente estabilizándose en torno a un 13%.

Gráfico 3.1 Tasa de desempleo en España del primer trimestre de 2006 al primer trimestre de 2022



Fuente: Statista, INE 2022.

Para analizar el desempleo en profundidad he seleccionado una serie de variables de diversa naturaleza y con gran impacto en el estudio del paro. A continuación, pasaré a describirlas:

- **Sexo**

En esta primera variable, examinaré si hay diferencias entre la población ocupada y desempleada según su género. Para ello, he realizado un análisis descriptivo condicionado calculando el porcentaje de paro entre mujeres y hombres.

La EPA le asigna el valor 1 para los varones y 6 para mujeres. La respuesta a esta encuesta ha sido de un 51,88% de hombres y un 48,13% de mujeres. Sin embargo, para este estudio he convertido la variable en:

$$X_i (HOMBRE) = \begin{cases} 1 & \text{si es HOMBRE} \\ 0 & \text{si es MUJER} \end{cases}$$

Tabla 3.1: Porcentaje de personas desempleadas según sexo

	MUJER	HOMBRE
OCUPADA	84,49%	88,25%
DESEMPLEADA	15,51%	11,75%

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

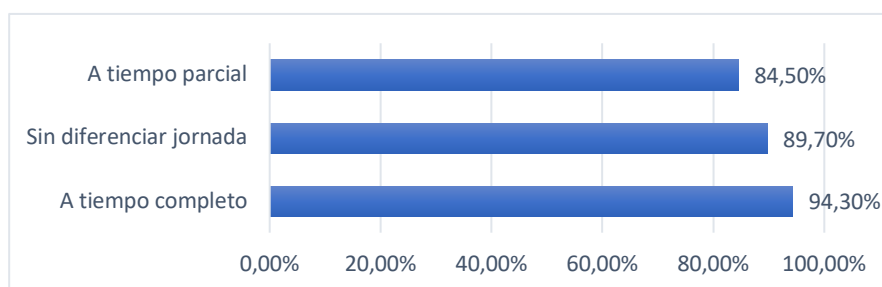
En la *Tabla 3.1* se observa una clara diferencia de casi cuatro puntos porcentuales de la situación laboral en función del género, mientras que los hombres poseen una tasa de desempleo del 11,75%, las mujeres alcanzan el 15,51%.

Si bien, las mujeres han luchado durante varias décadas por participar en el mercado de trabajo, a partir de los 80 y con la consolidación de la democracia, se fueron integrando más al mercado laboral. Tras este acontecimiento histórico, se inició un crecimiento sostenido de la tasa de actividad femenina, especialmente con el incremento de la participación laboral de las mujeres casadas y madres (*Treviño et al, 2007*). Esto ha permitido equiparar las tasas de participación, aunque no en términos de calidad del empleo.

Tradicionalmente, las mujeres poseían un rol en el hogar en el que: realizar las tareas del domicilio, limpiar, cocinar, cuidar de los niños y atender a su familia eran sus actividades principales. A pesar de que esta concepción ha ido cambiando, en la actualidad sigue siendo una de las causas que justifican que haya una menor tasa de mujeres ocupadas en el mercado laboral. Por lo tanto, las mujeres y el desempleo tienen una relación creciente.

En ocasiones, el aumento de la tasa de actividad femenina en los últimos años se ha atribuido en parte a un incremento del paro masculino. Es decir, mujeres cuyos hogares son mantenidos por el hombre deciden incorporarse al mercado laboral para mitigar la caída de ingresos. De hecho, Addabbo et al. (2013) estiman que, en nuestro país, cuando una pareja pierde el trabajo, la tasa de actividad de las mujeres aumenta un 21%, mientras que desciende un 0,7% para los hombres.

Gráfico 3.2 Salario por hora de la mujer respecto del hombre, 2019



Fuente: Encuesta Anual de Estructura Salarial, ¹ INE.

El *Gráfico 3.2.* refleja otro factor de la desigualdad laboral: las diferencias salariales por razón de género. Para analizar el concepto de ‘brecha salarial de género’², es necesario considerar los salarios por hora y distinguir entre los tipos de jornada. En trabajos a tiempo completo, los salarios por hora de las mujeres en 2019 fueron el 94,3 % respecto de los hombres, es decir, cobraron un 5,7% menos. Al considerar los trabajos a tiempo parcial, el porcentaje disminuye al 84,5%.

¹ Publicación ‘Mujeres y hombres en España: salarios, ingresos, cohesión social (2019)

² Diferencia entre el salario medio percibido por hombres y mujeres al desempeñar su actividad laboral.

Una de las principales razones de esta diferencia salarial es el tipo de jornada: mientras que para los hombres la jornada a tiempo parcial es algo coyuntural, normalmente durante los primeros años laborales. Para las mujeres se convierte en algo estructural, llegando a cobrar un 15,5% menos que los hombres en 2019.

Además, los empleos con jornada parcial tienden a estar peor remunerados por hora. Siendo un 25% de mujeres asalariadas con jornadas a tiempo parcial, mientras que, este tipo de contratos le afecta únicamente a un 7% de hombres.

- **Edad**

La edad es una variable cuantitativa continua, para trabajar con ella he seleccionado a los encuestados mayores de 16 años, según el Artículo 6 del Estatuto de los Trabajadores, supone la edad mínima de trabajo en España.

La EPA recoge a los individuos en tramos de 5 años, sin embargo, yo he creado intervalos de 8 años que comprenden personas desde los 16 hasta los 64 años, para poder analizar el impacto que genera la edad en el mercado laboral y demostrar si existen diferencias entre población ocupada por rangos de edad.

Tabla 3.2: Porcentaje de ocupados y desempleados por grupos de edad

EDAD	16-24	25-34	35-44	45-54	55-64
Nº encuestados	4136	8654	15233	18642	13314
% sobre encuestados	6,9%	14,43%	25,4%	31,08%	22,2%
OCUPADOS	67,99%	81,53%	89,13%	89,52%	87,43%
DESEMPLEADOS	32,01%	18,47%	10,87%	10,48%	12,47%

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

En España, en torno a un 15% de la población es joven, entre los 16 y 29 años. Sin embargo, la cifra de personas mayores de 65 años es de un 19,95%³ por lo que se trata, de una población envejecida. Esto se debe principalmente a que España presenta una tasa de natalidad muy baja, de tan solo el 7,19 por mil

³ Dato obtenido de Datosmacro (2020)

habitantes y la esperanza de vida de sus habitantes es de 82,3 años, por encima de la media en comparación con el resto del mundo.

Una vez determinada la población joven como porcentaje de la población total, es necesario analizar las principales características de los jóvenes en términos de participación en el mercado laboral.

Como observamos en la *Tabla 3.2* el porcentaje de población desempleada se dispara en el rango de 16 a 24 años con un 32% mientras que a los 25 años decrece hasta un 18%, tendencia que sigue hasta los 45-54 años que se estabiliza en un 10%. Mientras que a partir de los 55 años crece dos puntos porcentuales.

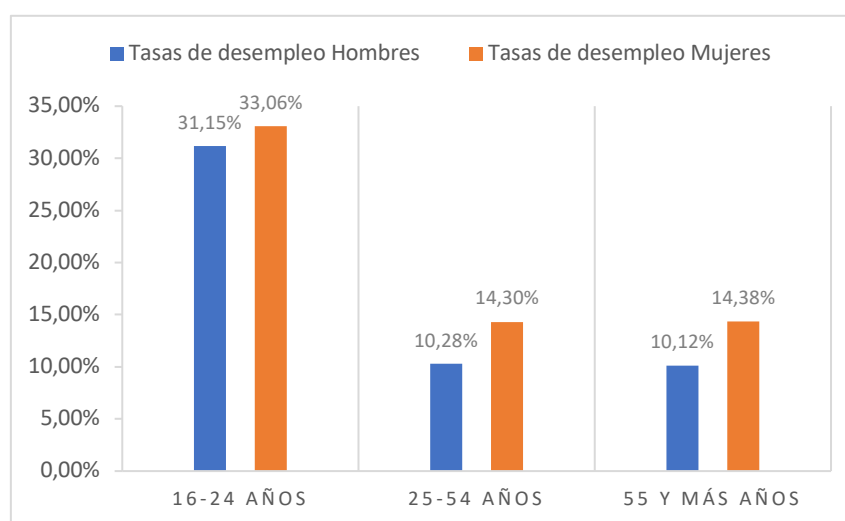
De hecho, García Viña, Jordi (2020) vincula la tasa de desempleo juvenil con la informalidad laboral, es decir, cuando quieren acceder al primer puesto de trabajo se suelen dar limitaciones en su empleabilidad en cuanto a falta de educación o bajos niveles de riqueza. Los salarios en personas jóvenes suelen ser inferiores al salario de personas adultas y los contratos suelen ser en mayor medida temporales.

En este sentido se manifiesta García, J. Ramón (2011) explicando que las razones del paro juvenil en España hay que buscarlas en las insuficiencias del sistema educativo y mercado laboral. Entre ellos, el abandono escolar temprano y polarización en la educación. Además, cabe destacar la fragmentación del mercado laboral e ineficacia de las políticas activas de empleo.

Según el Informe del mercado de Trabajo de los Jóvenes, SEPE (2021), las cinco ramas de actividad con mayor volumen de ocupados en el rango de 16 a 29 son: comercio al por mayor y al por menor, reparación de vehículos de motor, industria manufacturera, hostelería, actividades sanitarias y de servicios sociales y educación. En estos empleos, se acentúa la escasa calidad del empleo, predicando con mayor porcentaje de contratos temporales, jornadas parciales, imposibilidad de acceder a formación de la empresa o de carrera profesional, así como de bajos salarios o en algunos casos de actividades de economía sumergida.

Por último, a partir de los 55 años, se incrementa ligeramente la tasa de desempleo. Un desempleado de 23 años es diferente a uno con 50 o más, ya que las oportunidades, circunstancias, necesidades, son diferentes. Pero en muchas ocasiones esta diferencia es creada por un hándicap derivado de la edad, en vez de tener en cuenta el valor añadido que da la experiencia al empleo. Según Rubio Arribas, Fco. Javier (2013) Los principales problemas a los que se enfrentan las personas mayores de 50 años con el empleo son: los largos periodos de desempleo, la dificultad de insertabilidad y falta de motivación.

Gráfico 3.3 Tasas de paro en España por grupos de edad y sexo, 2021.



Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

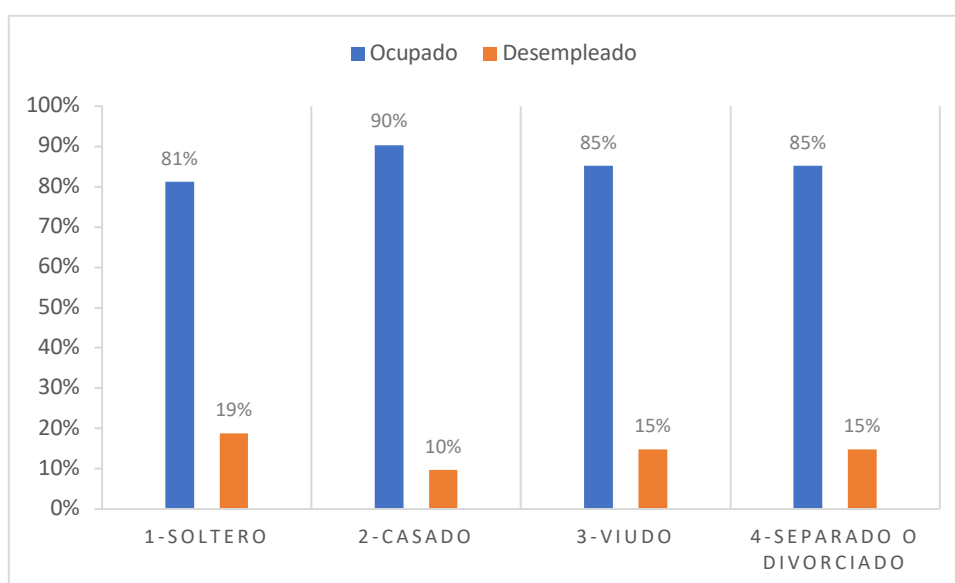
En el *Gráfico 3.3* observamos como los hombres poseen tasas de desempleo inferiores a las de las mujeres en todos los rangos de edad. La tasa de parados se dispara como hemos comentado anteriormente en el rango de población más joven, llegando a un 33,06% en mujeres entre 16 y 24 años. Mientras que, a partir de los 25 años decrece hasta un 10,28% en hombres y un 14,3% en mujeres. En este tramo y personas mayores de 55 años observamos tasas de desempleo muy similares y que poseen la diferencia más acentuada entre sexos, de cuatro puntos porcentuales, a pesar de estar en el mismo intervalo de edad. Por último, las tasas de desempleo más bajas se dan en los varones mayores de 55 años con un 10,12%.

- **Estado Civil**

Otra variable dependiente seleccionada que influye en la tasa de desempleo es el estado civil, los parámetros estudiados son:

1. Soltero
2. Casado
3. Viudo
4. Separado o divorciado

Gráfico 3.4: Tasa de ocupados o parados en función del estado civil.

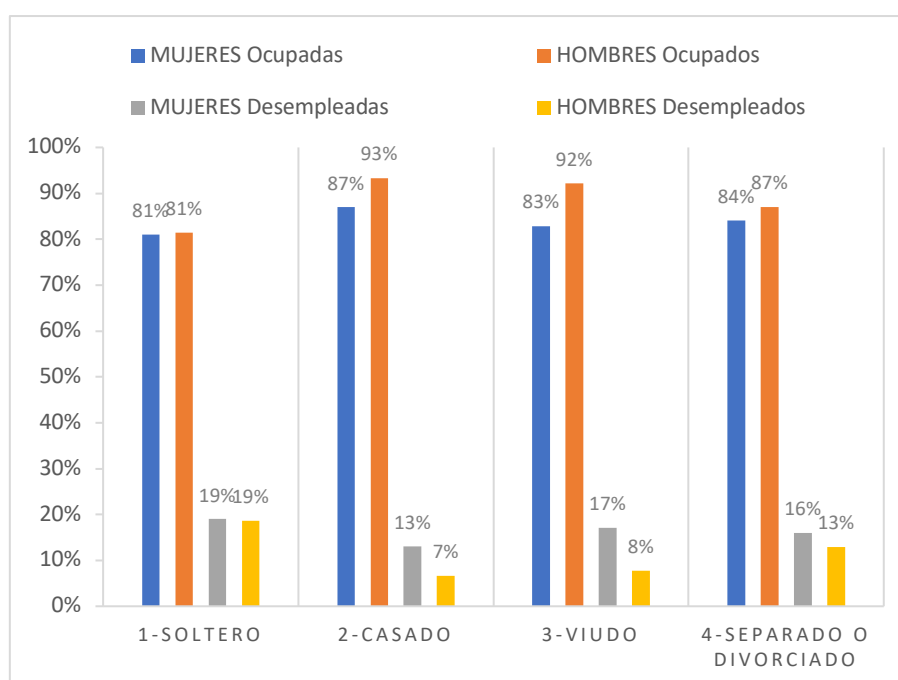


Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

En el *Gráfico 3.4* he analizado la tasa de ocupados y parados según el estado civil. Destaca que las personas solteras presentan el índice más alto de parados con un 19%, mientras que el porcentaje más bajo de desempleo se da en el colectivo de casados con un 10%. Los separados, divorciados y viudos se aproximan más a la tasa media, con un 15%.

El grupo de jóvenes con edad comprendida entre los 16 y 24 años como he analizado anteriormente es el más castigado por la falta de empleo y teniendo en cuenta que la edad media del primer matrimonio está en torno a los 30 años, se crea una relación directamente proporcional entre la alta tasa de desempleo de los solteros y la población joven.

Gráfico 3.5: Tasa de ocupados y parados en función del estado civil y sexo.



Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

A continuación, en el *Gráfico 3.5* analizaré esta variable en relación con el género. Observamos como no existe diferencia en la tasa de desempleo para ambos sexos en la soltería, etapa comúnmente asociada a la juventud, sin embargo, esa diferencia entre sexos sí que se nota cuando pasas a cualquier otro estado civil.

La población soltera alcanza un 19% de desempleados para ambos géneros, como he mencionado anteriormente la edad media del primer matrimonio en España está en torno a los 30 años, eso y con una tasa de desempleo del 32,01% de la población entre 16 y 24 años (dato que disminuye a 18,47% entre 25 y 34 años) se puede determinar que hay un vínculo directo entre la población soltera, joven y la tasa de desempleo, tanto en hombres como mujeres la soltería eleva la precariedad laboral.

Desde hace tiempo, el mercado laboral se caracteriza por una profunda segmentación entre los mayores de 30 años, con contratos fijos y una notable protección frente al despido, y los jóvenes con altas tasas de empleo temporal y grandes facilidades para ser despedidos, lo que genera una fuerte rotación en puestos de trabajo (García-Pérez y Muñoz-Bullón, 2011).

Mientras, los casados con un 13% entre mujeres y 7% varones poseen las menores tasas de desempleo. El segundo lugar es para los separados o divorciados, con un 16% para mujeres y 13% para hombres. Por último, las mujeres viudas presentan un 17% mientras que en los hombres es de un 8% la tasa de desempleo.

Según González-Val, Rafeel y Marcén, Miriam (2020) los resultados de su estudio revelan que la situación laboral de la mujer juega un papel en la decisión de ruptura en España. Ya que la probabilidad de separación o divorcio disminuyen cuando las mujeres no trabajan.

Por último, he creado una variable ficticia denominada estado civil separando a los solteros con el valor 0 y concentrando los valores 2,3 y 4 en casados. Cuyo resultado es:

$$X_I(CASADO) = \begin{cases} 1 & \text{si es CASADO} \\ 0 & \text{si SOLTERO} \end{cases}$$

- **Nacionalidad**

Con respecto a esta variable, la encuesta de población activa la recoge con tres valores:

1. Española
2. Española y doble nacionalidad
3. Extranjera

Para dicho estudio he traspasado los valores, agrupando el 1 y 2 en un solo valor, de la siguiente forma:

$$X_I(ESPAÑOL) = \begin{cases} 1 & \text{si es ESPAÑOL} \\ 0 & \text{si EXTRANJERO} \end{cases}$$

Tabla 3.3 Distribución de la población encuestada según la nacionalidad

	ESPAÑÓLES	EXTRANJEROS
Población encuestada	87%	13%
Ocupados	87,07%	77,55%
Desempleados	13,93%	22,45%

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

En esta encuesta la población española presenta un 87% del total y la población extranjera un 13%. De la cual, un 77,55% de los extranjeros tienen un empleo mientras que el 22,45% de la población son desempleados. Se observa una clara diferencia con respecto a la población española *Tabla 3.3* con una tasa de ocupados del 87% y una tasa de desempleados del 13%.

- **Nivel de estudios**

El nivel de estudios es una variable cualitativa, se subdivide en más de una variable explicativa. En la EPA se recoge como viene explicado en la *Tabla 3.4*.

Tabla 3.4 Nivel de estudios

Código	Descripción
AN	Analfabetos.
P1	Educación primaria incompleta.
P2	Educación primaria.
S1	Primera etapa de educación secundaria.
SG	Segunda etapa de educación secundaria.
SP	Segunda etapa de educación secundaria.
SU	Educación superior.

Fuente: EPA, INE.

Sin embargo, para la realización de este trabajo he definido variables ficticias para cada nivel de estudios. Dividiéndolas en cuatro: hasta estudios primarios (incluye: AN, P1, P2), primera etapa de educación secundaria, es decir, estudios

secundarios (hasta 4º ESO, S1), estudios superiores (bachillerato y grado medio, SG y SP) y estudios universitarios (universidad, SU). Para realizar de forma correcta el modelo, elimino Estudios primarios, por lo que será la variable de referencia.

$$X_i (\text{ESTUDIOS SECUNDARIOS}) = \begin{cases} 1 & \text{hasta 4º ESO} \\ 0 & \text{Otros} \end{cases}$$

$$X_i (\text{ESTUDIOS SUPERIORES}) = \begin{cases} 1 & \text{hasta Bachillerato y Grado Medio} \\ 0 & \text{Otros} \end{cases}$$

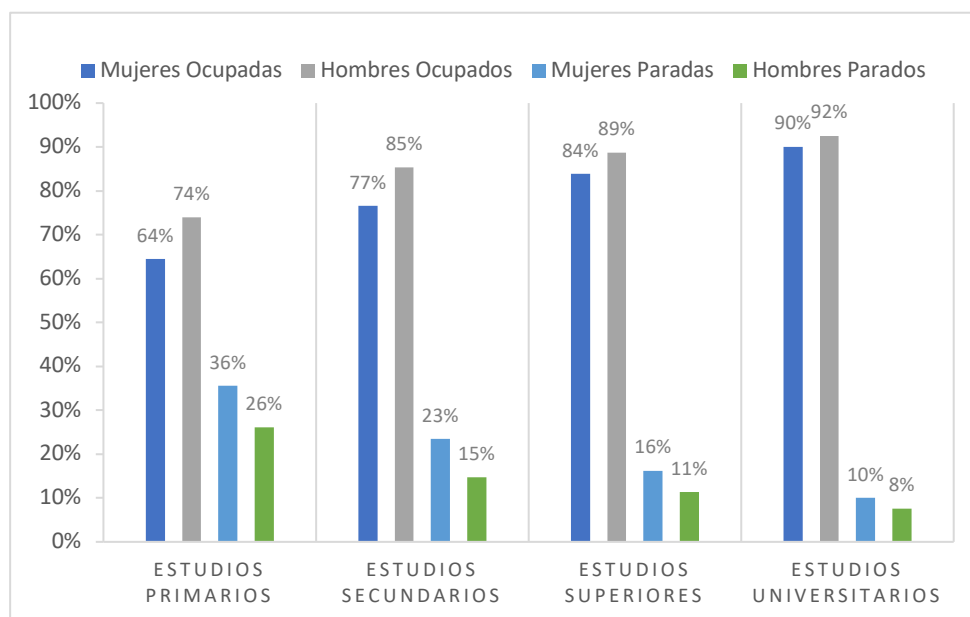
$$X_i (\text{ESTUDIOS UNIVERSITARIOS}) = \begin{cases} 1 & \text{Formación Universitaria} \\ 0 & \text{Otros} \end{cases}$$

Tabla 3.5 Tasa de ocupación en función del nivel de estudios.

	Estudios Primarios	Estudios Secundarios	Estudios Superiores	Estudios Universitarios
Ocupados	70,23%	81,82%	86,38%	91,14%
Parados	29,77%	18,18%	13,62%	8,86%

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

Gráfico 3.6 Distribución de la población encuestada según el nivel de estudios.



Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

Tanto la *Tabla 3.5* como el *Gráfico 3.6* muestran que cuanto mayor es el nivel educativo, mayor es la tasa de empleo y menor la tasa de desempleo. Además,

cuanta mayor sea la formación, mejor será el nivel salarial. Los datos difundidos por la EPA muestran que el grupo con mayor tasa de paro son los que tienen formación básica (29,77%), mientras que el grupo con mayor tasa de actividad es el que tiene estudios universitarios (91,14%). Además, según el estudio de la OCDE Panorama de la Educación 2019, los trabajadores con estudios universitarios tardan menos en encontrar trabajo y ganan de promedio un 57% más de los que solo han cursado la enseñanza secundaria.

Se observa una clara diferencia en todos los niveles educativos de la tasa de desempleo entre mujeres y hombres. La diferencia más notoria se da entre la población con estudios primarios, las mujeres poseen una tasa de desempleo en diez puntos porcentuales mayor que la de los hombres. A medida que se alcanza un mayor nivel de estudios, esta diferencia decrece. Si poseen estudios secundarios la probabilidad de estar desempleado en las mujeres es de un 23% mientras que los hombres de un 15%. Hasta llegar a la menor diferencia de tan solo dos puntos porcentuales entre mujeres y hombres con estudios universitarios (10% y 8% respectivamente)

En España, según Alonso-Villar y Del Río (2007) son los varones jóvenes los que abandonan relativamente pronto sus estudios para iniciarse prematuramente en el mercado laboral enfocándose hacia trabajos físicos que no precisen de una elevada cualificación.

Sin embargo, con los estudios universitarios se observa un comportamiento análogo al resto de Europa, es mayor el número de graduadas mujeres que hombres. Teniendo una sobrerrepresentación las mujeres en carreras de Ciencias de la Salud, Educación, Jurídico-Social, Humanidades y Artes, mientras que hay una mínima representación en estudios relacionados con Ingenierías.

Por lo tanto, teniendo en cuenta estas diferencias en cuanto al nivel de estudios alcanzado entre ambos sexos. Se puede justificar el importante diferencial en las tasas de desempleo favoreciendo a los hombres.

Por último, las condiciones de trabajo de la población varían en función de la eficacia del sistema educativo. En concreto, la posibilidad de que los jóvenes

consigan empleos estables y bien remunerados depende de aspectos como la cantidad y la calidad de la educación recibida.

4. MODELO LOGIT

El modelo a estimar es el siguiente:

$$Y_i = F(\beta_0 + \beta_1 \text{EDAD}_i + \beta_2 \text{HOMBRE}_i + \beta_3 \text{ESPAÑOL}_i + \beta_4 \text{CASADO}_i + \beta_5 \text{ESTUDIOS SECUNDARIOS}_i + \beta_6 \text{ESTUDIOS SUPERIORES}_i + \beta_7 \text{ESTUDIOS UNIVERSITARIOS}_i) + u_i$$

Para poder trabajar con los datos, he utilizado el programa EViews 12 seleccionado el método 'Binary Logit' para emplear el modelo adecuado.

Tabla 4.1 Estimación del modelo mediante el método Logit

Dependent Variable: DESEMPLEADOS
Method: ML - Binary Logit (Newton-Raphson / Marquardt steps)
Date: 06/17/22 Time: 20:42
Sample: 1 60843
Included observations: 59979
Convergence achieved after 7 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
EDAD	-0.019208	0.001205	-15.94387	0.0000
HOMBRE	-0.503081	0.025011	-20.11455	0.0000
ESPAÑOL	-0.356645	0.042760	-8.340584	0.0000
CASADO	-0.552994	0.029597	-18.68434	0.0000
E_SECUNDARIOS	-0.698091	0.048633	-14.35436	0.0000
E_SUPERIORES	-1.213873	0.050840	-23.87642	0.0000
E_UNIVERSITARIOS	-1.704403	0.050225	-33.93503	0.0000
C	1.021695	0.071224	14.34489	0.0000
McFadden R-squared	0.064802	Mean dependent var		0.136798
S.D. dependent var	0.343637	S.E. of regression		0.334173
Akaike info criterion	0.746756	Sum squared resid		6697.070
Schwarz criterion	0.747957	Log likelihood		-22386.85
Hannan-Quinn criter.	0.747129	Deviance		44773.70
Restr. deviance	47876.18	Restr. log likelihood		-23938.09
LR statistic	3102.483	Avg. log likelihood		-0.373245
Prob(LR statistic)	0.000000			
Obs with Dep=0	51774	Total obs		59979
Obs with Dep=1	8205			

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

Como observamos en la *Tabla 4.1*, todos los coeficientes $\hat{\beta}_j$ de las variables dependientes estudiadas, presentan el signo esperado: negativo. Es decir, cuando las variables cambian su valor de 0 a 1, se reduce la posibilidad de estar desempleado. Por otro lado, todas las variables independientes son significativas al tener un p-valor igual a 0.

En el modelo Logit, no existe un valor que nos indique la ‘variabilidad explicada’ como sí sucede en los modelos lineales. Sin embargo, el valor McFadden R-squared, refleja la bondad de ajuste del modelo a la variable que se quiere explicar, en este caso el desempleo. Cuanto más cercano sea a 0, peor es el ajuste y peor su capacidad predictiva, como en este caso 0,065. Cifra que puede ser consecuencia del elevado número de observaciones: 59979, inferior a la que hemos analizado, ya que solo tiene consideración las que han respondido a todas las consideraciones de la encuesta.

Tabla 4.2 Criterio de predicciones correctas

Expectation-Prediction Evaluation for Binary Specification
Equation: UNTITLED
Date: 06/17/22 Time: 21:05
Success cutoff: C = 0.135594892

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	32301	2961	35262	0	0	0
P(Dep=1)>C	19473	5244	24717	51774	8205	59979
Total	51774	8205	59979	51774	8205	59979
Correct	32301	5244	37545	0	8205	8205
% Correct	62.39	63.91	62.60	0.00	100.00	13.68
% Incorrect	37.61	36.09	37.40	100.00	0.00	86.32
Total Gain*	62.39	-36.09	48.92			
Percent Gain**	62.39	NA	56.67			

	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	45086.36	6687.64	51774.00	44691.43	7082.57	51774.00
E(# of Dep=1)	6687.64	1517.36	8205.00	7082.57	1122.43	8205.00
Total	51774.00	8205.00	59979.00	51774.00	8205.00	59979.00
Correct	45086.36	1517.36	46603.72	44691.43	1122.43	45813.85
% Correct	87.08	18.49	77.70	86.32	13.68	76.38
% Incorrect	12.92	81.51	22.30	13.68	86.32	23.62
Total Gain*	0.76	4.81	1.32			
Percent Gain**	5.58	5.58	5.58			

*Change in "% Correct" from default (constant probability) specification
**Percent of incorrect (default) prediction corrected by equation

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

La bondad de ajuste también la podemos comprobar en la *Tabla 4.2*, a través del criterio de predicciones correctas. Con ello, se pretende determinar las ocasiones en que el valor de la variable dependiente (desempleo) concuerda con la predicción. Para poder calcularla se determina un umbral de probabilidad estimada, de forma predeterminada es de 0,5 sin embargo, en este caso le he cambiado a un 13,55% debido al gran desequilibrio de la muestra entre 0 y 1.

Observamos como el número total de encuestados ha sido de 59979. Existe una clasificación correcta del 62,39% para los empleados (32301), mientras que ha pronosticado erróneamente la situación de desempleo en 2961 observaciones, acertando 5244 observaciones (63,91%).

En la *Tabla 4.3* se recoge la probabilidad estimada y el efecto en el desempleo de las variables dependientes. Para calcular esta probabilidad, se han sustituido los valores en el modelo estimado por 0 y 1 en cada caso, manteniendo el resto de las variables en la media. El efecto en el desempleo se ha calculado como la diferencia entre la probabilidad estimada cuando es 1 y cuando es 0, midiendo la influencia en la probabilidad de estar desempleado al pasar de cero a uno.

Tabla 4.3 Probabilidad estimada y efecto en el desempleo de las variables dependientes

Variable	Valor	Probabilidad estimada	Efecto en el desempleo
Sexo	Hombre=1	0,096185	-0,053475
	Mujer =0	0,14966	
Nacionalidad	Español y doble nacionalidad=1	0,116935	-0,042139
	Extranjero= 0	0,159074	
Estado civil	Casado=1	0,099433	-0,061602
	Soltero=0	0,161035	
Estudios Secundarios	Hasta 4ºEso=1	0,187167	-0,129219
	E. Primarios=0	0, 316386	
Estudios superiores	Hasta Bachillerato y Grado Medio=1	0,120861	-0,195525
	E. Primarios=0	0, 316386	
Estudios Universitarios	Formación Universitaria=1	0,077641	-0,238745
	E. Primarios=0	0, 316386	

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

Se observa que la variable SEXO, cuando toma el valor 0=Mujer, la probabilidad estimada de estar desempleada es de 0,14966. Mientras que cuando es 1=Hombre la probabilidad estimada disminuye a 0,096185. Es decir, la diferencia entre ambos valores es de -5,3475 puntos porcentuales lo que indica que: los hombres tienen un 5,3475 puntos porcentuales menos de probabilidad de estar desempleados que las mujeres.

En el caso de NACIONALIDAD las personas extranjeras tienen mayor probabilidad de estar desempleadas, en concreto 4,2139 puntos porcentuales

más que la población española y de doble nacionalidad. En la variable ESTADO CIVIL, la probabilidad de que una persona soltera este desempleada es 6,1602 puntos porcentuales mayor que las casadas.

En la variable ESTUDIOS, observamos una relación directamente proporcional entre el nivel de estudios y estar ocupado. Es decir, cuanto mayor sea el nivel de estudios, menor será la probabilidad de estar desempleado. La diferencia de la probabilidad de estar desempleado si tiene estudios secundarios respecto de hasta tener estudios primarios es de -0,129219, si tiene estudios superiores respecto de hasta tener estudios primarios es de -0,195525 y si tiene estudios universitarios respecto de hasta poseer estudios primarios la probabilidad de desempleo se reduce en 0,238745.

4.1 Análisis de escenarios.

En la siguiente tabla (*Tabla 4.4*), analizaré escenarios aleatorios que determinan características de varias personas, para analizar qué individuos con características concretas son los que tienen la mayor y menor probabilidad de estar desempleados.

Tabla 4.4 Análisis de escenarios

EDAD	SEXO	E. CIVIL	NACIONALIDAD	ESTUDIOS	PROB
20	MUJER	SOLTERA	ESPAÑOLA	Primarios	0,569764
20	MUJER	SOLTERA	ESPAÑOLA	Secundario	0,397186
20	MUJER	SOLTERA	ESPAÑOLA	Superior	0,282319
20	MUJER	SOLTERA	ESPAÑOLA	Universitarios	0,194111
20	HOMBRE	SOLTERO	ESPAÑOL	Primarios	0,444679
20	HOMBRE	SOLTERO	ESPAÑOL	Secundario	0,284900
20	HOMBRE	SOLTERO	ESPAÑOL	Superior	0,192155
20	MUJER	SOLTERA	EXTRANJERA	Primarios	0,654195
20	MUJER	SOLTERA	EXTRANJERA	Secundario	0,484865
20	MUJER	SOLTERA	EXTRANJERA	Superior	0,359775
20	HOMBRE	SOLTERO	EXTRANJERO	Primarios	0,533562
20	HOMBRE	SOLTERO	EXTRANJERA	Superior	0,253615
30	MUJER	SOLTERA	ESPAÑOLA	Primarios	0,522187
30	MUJER	SOLTERA	ESPAÑOLA	Superior	0,245072

30	HOMBRE	SOLTERO	ESPAÑOL	Superior	0,164084
30	MUJER	SOLTERA	EXTRANJERA	Primarios	0,609557
30	MUJER	SOLTERA	EXTRANJERA	Superior	0,316820
30	HOMBRE	SOLTERO	EXTRANJERA	Primarios	0,485596
30	MUJER	SOLTERA	ESPAÑOLA	Universitarios	0,165813
30	HOMBRE	SOLTERO	ESPAÑOL	Universitarios	0,107294
30	MUJER	SOLTERA	EXTRANJERA	Universitarios	0,221154
30	HOMBRE	SOLTERO	EXTRANJERO	Universitarios	0,146536
40	MUJER	CASADA	ESPAÑOLA	Universitarios	0,086221
40	HOMBRE	CASADO	ESPAÑOL	Universitarios	0,053975
40	MUJER	CASADA	EXTRANJERA	Primario	0,425649
40	HOMBRE	CASADO	EXTRANJERA	Primario	0,309447
40	MUJER	CASADA	ESPAÑOLA	Primarios	0,341578
40	HOMBRE	CASADO	ESPAÑOL	Primarios	0,238785
40	MUJER	CASADA	EXTRANJERA	Secundarios	0,269391
40	HOMBRE	CASADO	EXTRANJERO	Secundarios	0,182307
50	MUJER	CASADA	ESPAÑOLA	Universitarios	0,072242
50	MUJER	SOLTERA	ESPAÑOLA	Superiores	0,181053
50	MUJER	SOLTERA	ESPAÑOLA	Universitarios	0,119228
50	HOMBRE	SOLTERO	ESPAÑOL	Universitarios	0,075659
50	MUJER	SOLTERA	EXTRANJERA	Universitarios	0,162041
50	HOMBRE	SOLTERO	EXTRANJERA	Universitarios	0,104687
50	MUJER	CASADA	ESPAÑOLA	Primarios	0,299779
50	HOMBRE	SOLTERO	ESPAÑOL	Primarios	0,310359
50	MUJER	SOLTERA	EXTRANJERA	Secundarios	0,345969
50	HOMBRE	SOLTERO	EXTRANJERO	Secundarios	0,242341
60	MUJER	CASADA	ESPAÑOLA	Primarios	0,261066
60	MUJER	CASADA	ESPAÑOLA	Secundarios	0,149500
60	MUJER	CASADA	ESPAÑOLA	Superiores	0,094979
60	MUJER	CASADA	ESPAÑOLA	Universitarios	0,060379
60	HOMBRE	CASADO	ESPAÑOL	Primarios	0,176025
60	HOMBRE	CASADO	ESPAÑOL	Secundarios	0,096076
60	HOMBRE	CASADO	ESPAÑOL	Superiores	0,059671
60	HOMBRE	CASADO	ESPAÑOL	Universitarios	0,037402
60	MUJER	CASADA	EXTRANJERA	Universitarios	0,084078

60	HOMBRE	CASADO	EXTRANJERA	Universitarios	0,052587
60	MUJER	CASADA	EXTRANJERA	Secundarios	0,200707
60	HOMBRE	CASADA	EXTRANJERO	Secundarios	0,131820

Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021.

Comentaré las probabilidades resaltadas, en concreto las rojas pertenecen a valores máximos de probabilidad de desempleo. Mientras que las verdes representan las probabilidades mínimas.

Las tasas máximas de desempleo con un 0,654195 y 0,609557 respectivamente lo poseen mujeres de 20 y 30 años, solteras, extranjeras y con estudios primarios. Les siguen con un 0,569764 y 0,522187, mujeres de 20 y 30 años solteras, españolas y con estudios primarios.

Los hombres de forma generalizada poseen menor tasa de desempleo que las mujeres, sin embargo, observamos como comparten las mismas características con las mujeres que presentan las mayores tasas de desempleo. Es decir, las probabilidades mayores (0,533562 y 0,485596 respectivamente) son hombres de 20 y 30 años, solteros, extranjeros y con estudios primarios.

Por otro lado, el escenario con la menor probabilidad de desempleo con un 0,037402 es de hombres con 60 años, casados, españoles y con estudios universitarios. Le siguen con un 0,052587: hombres de 60 años casados, extranjeros y con estudios universitarios. Muy cercano a este valor, con un 0,053975: hombres de 40 años, casados, españoles y con estudios universitarios.

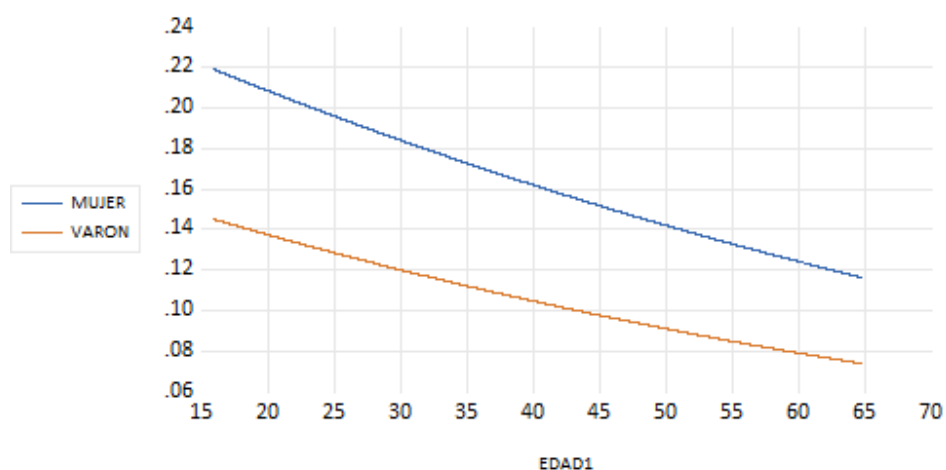
La menor tasa de desempleo de las mujeres es de un 0,060379 (0,02 puntos porcentuales mayor que los hombres) compartiendo las mismas características: 60 años, española, casada y con estudios superiores y con un 0,072242 el mismo escenario, pero con mujeres de 50 años.

Para finalizar y tras el análisis de los escenarios planteados, podemos concluir que el desempleo está muy vinculado a las variables estudiadas. De forma global, los más afectados por el desempleo son personas extranjeras, jóvenes y sin estudios, en concreto, se agrava siendo mujer y soltera.

4.2 Curvas de respuesta.

Por último, realizaré las curvas de respuesta para ver cómo afectan dichas variables a la probabilidad de desempleo de forma gráfica. Cada variable se representa en un mismo gráfico que mide el efecto sobre la probabilidad de estar desempleado cuando varía una variable continua, en este caso la edad.

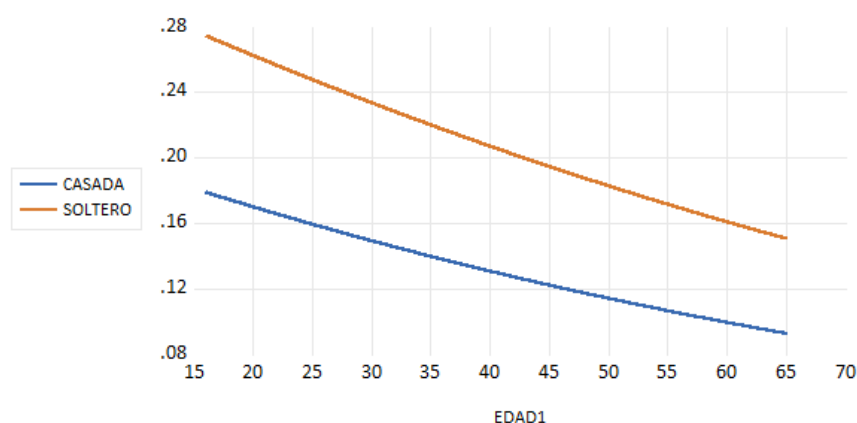
Gráfico 4.1 Curvas de respuesta para la variable sexo



Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021

El *Gráfico 4.1* representa el efecto esperado en la variable sexo, las mujeres presentan una probabilidad de estar desempleado mayor que los hombres, para cualquier rango de edad. En la población joven la diferencia es más significativa, mientras que a partir de los 40 años decrece progresivamente, encontrándose siempre con una tasa de desempleo mayor en las mujeres.

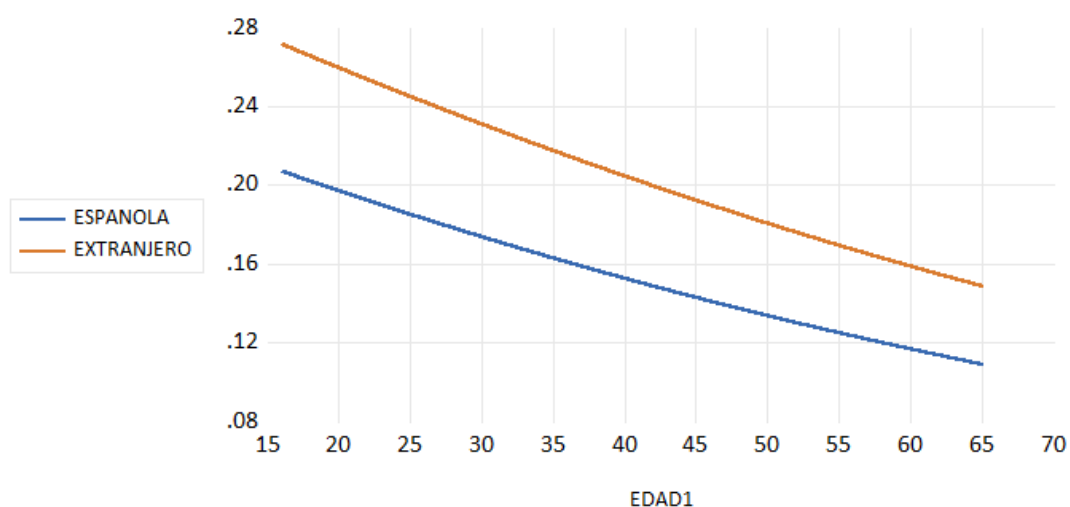
Gráfico 4.2 Curvas de respuesta para la variable estado civil



Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021

Por otro lado, el *Gráfico 4.2* representa la variable estado civil con respecto a la edad. Como se observa, la población soltera presenta una mayor probabilidad de desempleo que la casada. La probabilidad de estar desempleado disminuye a lo largo de los años, siguiendo la misma tendencia que el gráfico comentado anteriormente respecto a la variable sexo, en los primeros años la diferencia es más significativa y a medida que aumentan los años decrece progresivamente.

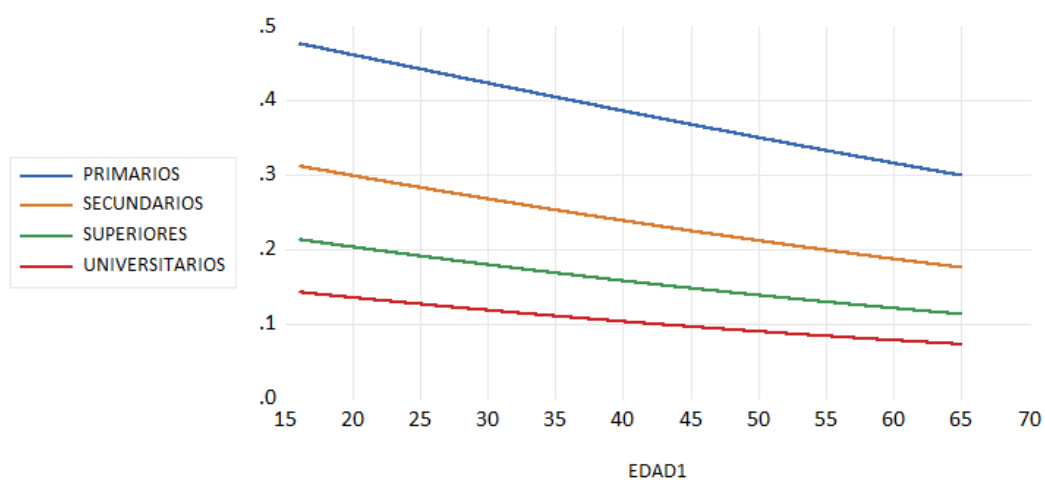
Gráfico 4.3 Curvas de respuesta para la variable nacionalidad



Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021

El *Gráfico 4.3* representa las curvas de respuesta de la nacionalidad respecto de la edad, en este caso la población extranjera tiene una probabilidad mayor de estar desempleada de la española. La probabilidad de estar desempleado disminuye cuanto mayor sea la persona encuestada.

Gráfico 4.4 Curvas de respuesta para la variable nivel de estudios



Fuente: elaboración propia a partir de los datos de la EPA, INE 3º trimestre 2021

Finalmente, el análisis de las curvas de respuesta para la variable nivel de estudios, respecto de la edad (*Gráfico 4.4*) nos muestra como las personas con estudios primarios poseen la mayor probabilidad de desempleo, esta diferencia en la probabilidad se reduce considerablemente en el caso de los estudios secundarios y superiores. Mientras que las personas que tienen estudios universitarios consiguen la probabilidad menor de estar desempleado manteniéndose casi constante a lo largo de los años.

Una característica común que poseen todas las variables estudiadas en las curvas de respuesta es que, durante los primeros años, las curvas conservan las probabilidades más elevadas de desempleo, además que presentan las mayores diferencias entre las curvas, sin embargo, todas siguen una tendencia decreciente cuanto mayor sea la edad.

5. CONCLUSIONES

El objetivo general del estudio buscaba conocer los determinantes del desempleo en España y podemos concluir, que las variables analizadas: sexo, estado civil, nacionalidad, edad y nivel de estudios, provocan un gran impacto en la probabilidad de estar desempleado. Además, el signo de todos los coeficientes de las variables analizadas es el esperado.

En primer lugar, he hecho un análisis descriptivo del desempleo realizando una investigación a través de la bibliografía consultada, para posteriormente formalizar un análisis econométrico a través del modelo Logit, que ha reafirmado las conclusiones obtenidas en el apartado anterior y ha podido determinar que todas las variables influyen en la probabilidad de estar desempleado.

La tasa de desempleo calculada a través de la Encuesta de Población Activa del Instituto Nacional de Estadística para el tercer trimestre de 2021 es de un 13,56%. De forma global he observado que hasta el 2020 seguía una tendencia decreciente, que se vio truncada en ese mismo año como consecuencia del COVID-19. Sin embargo, en el 2021 esta tasa ha comenzado a decrecer y estabilizarse por debajo de la media 15,9%⁴.

Respecto a la variable sexo he comprobado como las mujeres presentan una tasa de desempleo mayor que los hombres (15,51% y 11,75% respectivamente) causa de ello son razones culturales, el rol femenino en la vida social y económica, peores condiciones laborales, tipo de jornada y diferencia salarial entre otras.

Además, los jóvenes poseen las tasas más elevadas de paro en torno a un 32% entre 16 y 24 años. Mientras que a partir de esta edad sigue una tendencia decreciente llegando a un 12% en personas mayores de 55 años. Vinculando las tasas de desempleo juvenil con la informalidad laboral, limitaciones para acceder a un primer puesto de trabajo, contratos temporales y escasa calidad del empleo

⁴ Media calculada desde el primer trimestre de 2001 hasta el primer trimestre de 2022. Datos obtenidos de la sección de prensa de la EPA, INE.

Con el estudio del estado civil he concluido que las personas casadas tienen una menor tasa de desempleo. Por otro lado, la población extranjera posee mayores problemas para la integración en el mercado laboral, presentando una tasa de desempleo del 22,45% con casi diez puntos porcentuales de diferencia con la población española (13%).

Adicionalmente, se ha podido concluir que el nivel de estudios tiene una gran incidencia en la tasa de desempleo. Las personas con estudios primarios poseen tasas de desempleo muy elevadas llegando a un 29,77%, mientras que poseer estudios secundarios la disminuye a 18,18%, superiores a un 13,62% y universitarios 8,86%. La formación, no solo afecta a la tasa de desempleo, sino también al nivel salarial, condiciones de trabajo y tiempo de búsqueda de empleo entre otras.

El análisis descriptivo se respalda con el modelo Logit obteniendo las mismas conclusiones. Los coeficientes de todas las variables dependientes poseen signo negativo, es decir, cuando pasan de cero a uno, la probabilidad de estar desempleado disminuye. Se han calculado los efectos marginales sobre la probabilidad de desempleo y se concluye que tener algún tipo de educación contra no tener es la variable que más reduce el desempleo, la mayor diferencia se obtiene en el caso de: poseer estudios universitarios respecto de hasta poseer estudios primarios reduciendo la probabilidad de estar desempleado en 23,8745 puntos porcentuales.

En cuanto al sexo, los hombres tienen menos probabilidad de estar desempleados que las mujeres, en concreto 5,3475 puntos porcentuales. Las personas extranjeras tienen mayor probabilidad de estar desempleada (4,2139 puntos porcentuales más). Mientras que una persona casada disminuye la probabilidad de estar desempleada en 6,1602 puntos porcentuales respecto de estar soltera.

Por tanto, el desempleo es un fenómeno económico muy sensible al ciclo y coyuntura económica, que no solo tiene repercusiones en el nivel de riqueza de las personas, sino también consecuencias económicas, sociales e incluso familiares y personales. Por último, extrapolar los escenarios analizados, las

personas más afectadas por el desempleo son: personas extranjeras, jóvenes y sin estudios, concretamente, se incrementa cuando son mujeres y solteras.

6. BIBLIOGRAFÍA

Addabbo, T. Rodríguez-Modroño, P. y Gálvez-Muñoz, L. (2013), Gender and the Great Recession: Changes in labour supply in Spain, Department of Economics (DEMB), University of Modena and Reggio Emilia, Department of Economics "Marco Biagi".

Datosmacro (2020): "Pirámide de población". Disponible en <https://datosmacro.expansion.com/demografia/estructura-poblacion/espana> [consulta 3/4/2022]

García Viña, J. (2020). La situación del empleo de los jóvenes y una propuesta de políticas públicas para afrontar el desempleo juvenil. Revista latinoamericana de derecho social, 65-94. <https://doi.org/10.22201/ij.24487899e.2020.30.14072>

García, J. Ramón (2011). "Desempleo juvenil en España. Causas y soluciones", Documentos de Trabajo 1130, Banco BBVA, Departamento de Estudios Económicos.

Goldberger, A.S. (2001): Introducción a la Econometría, Ariel, Barcelona, cap. 17.

González-Val, R. y Marcén, M. (2020). Desempleo, ruptura de las parejas y género en España: Unemployment, Marital Breakdown and Gender in Spain. Reis: Revista Española de Investigaciones Sociológicas, 171, pp. 145–158. <https://www.jstor.org/stable/26976882>

INE (2019): "Mujeres y hombres en España", Instituto Nacional de Estadística, Madrid.

Instituto Nacional de Estadística. Disponible en <http://www.ine.es/> [consulta 15/2/2022]

Jaén, M. E. F. (2019). FEIR 45: Regresión logística. FEIR 45: Regresión logística. <https://gauss.inf.um.es/feir/45/#:%7E:text=La%20interpretaci%C3%B3n%20del%20ods%2Dratio,medida%20del%20tama%C3%B1o%20del%20efecto.> [consultado 01/06/2022]

Observatorio igualdad y empleo (2022): "Herramientas propuestas para acabar con la brecha salarial". Disponible en

<https://www.observatorioigualdadyempleo.es/herramientas-propuestas-para-acabar-con-la-brecha-salarial/> [consulta 2/4/2022]

Quesada, A. (2010): “Variables ficticias”. Disponible en https://www.um.es/econometria/adeydcho/transpar/07-ficticias_Alfonso.pdf [consultado 10/4/2022]

Rubio Arribas, FJ (2013): “Tengo 50 años ¿y qué? Dificultades y estrategias en el retorno al mercado de trabajo”. *Nómadas*, vol. 40, núm. 4. Disponible en: <https://www.redalyc.org/articulo.oa?id=18153270006> [consulta 3/4/2022]

Ruiz, R. (2019). Los trabajadores con estudios universitarios tardan menos en encontrar trabajo. *La Razón*. <https://www.larazon.es/sociedad/los-trabajadores-con-estudios-universitarios-tardan-menos-en-encontrar-trabajo-NG24880793/> [consultado 03/06/2022]

Salvador Padilla (2020): “Modelos Logit y Probit”. Disponible en <https://economipedia.com/definiciones/modelos-logit-y-probit.html> [consulta: 26/3/2022]

Secretaría Confederal de Mujeres e Igualdad de CCOO (2021): “Cuentas claras para acabar con la brecha salarial”. Disponible en <https://www.ccoo.es/d6358833d528a8efd7ef9ec0c12e8fc3000001.pdf> [consulta 1/4/2022]

SEPE (2020): “Informe del Mercado de Trabajo de los Mayores de 45 años”. Disponible en <https://sepe.es/HomeSepe/que-es-el-sepe/comunicacion-institucional/publicaciones/publicaciones-oficiales/listado-pub-mercado-trabajo/informe-mercadotrabajo-estatal-mayores45.html> [consulta 4/4/2022]

SEPE (2021): “Informe del Mercado de Trabajo de los Jóvenes”. Disponible en <https://sepe.es/HomeSepe/que-es-el-sepe/comunicacion-institucional/publicaciones/publicaciones-oficiales/listado-pub-mercado-trabajo/informe-mercadotrabajo-estatal-jovenes.html> [consulta 4/4/2022]

Subdirección general de estadística y análisis sociolaboral (2021): “La situación de las mujeres en el mercado de trabajo 2021”. Disponible en https://www.mites.gob.es/es/sec_trabajo/analisis-mercado-trabajo/situacion-mujeres/index.htm [consulta 2/4/2022]

Todo FP.(2021): "¿Cómo influye el nivel de estudios en la tasa de desempleo en España?. Disponible en <https://www.todofp.es/orientacion-profesional/busca-empleo-entrenate/mercado-laboral/la-fp-y-el-mercado-laboral/nivel-estudios.html> [consultado 11/4/2022]

Villar, A. (2015). No es país para jóvenes. Universidad Pablo de Olavide e Ivie. <https://www.upo.es/diario/wp-content/uploads/2015/01/no-es-pais-para-jovenes.pdf>

Vista de Análisis del rendimiento académico mediante un modelo logit. (2010). Vista de Análisis del rendimiento académico mediante un modelo logit. <http://revistas.ubiobio.cl/index.php/RI/article/view/56/3297> [consultado 02/06/2022]<http://revistas.ubiobio.cl/index.php/RI/article/view/56/3297>

Yáñez-Contreras, M. y Cano-Hernández, K. del C. (2011). Determinantes del Desempleo: Una revisión de la literatura. Panorama Económico, 19(19), 135–148. <https://doi.org/10.32997/2463-0470-vol.19-num.19-2011-352>