



Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Grado en Economía

Determinantes del Desempleo

Presentado por:

Sonia Martínez Cañibano

Tutelado por:

M^a Dolores de Prada Moraga

Valladolid, 30 de Junio de 2015

ÍNDICE

1. INTRODUCCIÓN	2
2. EL DESEMPLEO EN ESPAÑA	3
2.1 SITUACIÓN ACTUAL DEL DESEMPLEO	5
2.2 EVOLUCIÓN DEL DESEMPLEO EN ESPAÑA	6
2.2.1 Factores explicativos del desempleo	9
3. METODOLOGÍA.....	11
4. ESTIMACIÓN DEL MODELO.....	16
4.1 VARIABLES	16
4.2 EL MODELO	19
5. RESULTADOS.....	22
6. CONCLUSIONES.....	31
7. BIBLIOGRAFIA	34

1. INTRODUCCIÓN

La historia económica se ha centrado en estudiar la economía desde muchos puntos de vista. Cuantiosos economistas¹ han aportado sus ideas a esta ciencia, pero no hay unanimidad entre ellos acerca de lo que es la economía. A lo largo de los años el concepto de economía ha ido cambiando y cada ecónomo ha definido esta disciplina en el contexto de su época.

Todos ellos tienen en común que la economía es una ciencia que integra todas las acciones humanas que refieren a la producción, distribución o comercialización de bienes y servicios, dando lugar a una riqueza o beneficio que sostiene un estado de bienestar.

La economía no es una ciencia perfecta y en ocasiones, ya sea por fuerzas externas o por el propio funcionamiento de la misma, surgen desajustes que ocasionan situaciones inestables y crisis. Estos pueden ser corregidos mediante políticas fiscales y monetarias aplicadas por los gobiernos de cada país, pero en ocasiones esto no basta para solventarlo.

El desempleo es una perturbación muy común y su eliminación es complicada. Según la etapa de la economía en que nos encontremos este efecto es mayor o menor, pero nunca se elimina del todo, por eso expertos del tema resaltan que una tasa de desempleo inferior al 5% es lo óptimo para la economía porque siempre existirá un porcentaje de desempleo friccional y estructural.

La situación en España con unas tasas de desempleo superiores al 24%, muy por encima de la media europea, hacen que el desempleo sea el principal problema de la economía española. Sería, por tanto, muy interesante tratar de explicar qué factores son los determinantes del desempleo.

El objetivo de esta investigación es estudiar este desequilibrio desde una perspectiva microeconómica. Se pretende estimar el efecto que tienen determinadas características individuales sobre la probabilidad de estar

¹Eckelund, R.B, & Hébert, R.F. (1992). *Historia de la teoría económica y de su método*. McGraw-Hill, México.

Rodríguez Braun, C. (1997). *Grandes economistas*. Pirámide, Madrid.

desempleado. En concreto analizaremos como ciertas variables exógenas, como el sexo, la nacionalidad, la educación y la edad, influyen en la probabilidad de que un individuo esté desempleado. Asimismo, compararemos el efecto que tienen sobre la probabilidad de estar desempleado diferentes situaciones personales.

Se es consciente de las limitaciones de este estudio, ya que ciertas variables fundamentales para la explicación del desempleo, como el salario, no han podido ser analizadas por falta de datos. Aun así, se cree que este trabajo puede arrojar cierta luz en la explicación de los factores individuales que afectan al desempleo.

La organización del trabajo ha sido la siguiente:

En primer lugar se analiza la situación el desempleo en España, resaltando la situación actual y la evolución que ha sufrido este desequilibrio desde 1977 a 2014.

En segundo lugar se explica la metodología a utilizar para estimar este tipo de modelos. La característica de la variable dependiente, que solo toma los valores 0 y 1, hace que el análisis de regresión clásico no se pueda utilizar y por tanto se necesita un método específico de estimación.

En tercer lugar se lleva a cabo la estimación del modelo utilizando la metodología anterior. A continuación se analizan los resultados obtenidos en la estimación, haciendo hincapié en una serie de escenarios. Éstos tratan de determinar la probabilidad de estar desempleado para unas determinadas características establecidas de las diferentes variables explicativas.

Por último se tratará de extraer unas conclusiones de todo lo expuesto anteriormente.

2. EL DESEMPLEO EN ESPAÑA

El desempleo comienza cuando se produce un exceso de oferta, es decir, cuando la cantidad ofrecida es mayor que la cantidad demandada, esto es cuando el número de trabajadores disponibles para trabajar excede a lo que realmente necesitan las empresas.

En la actualidad podemos encontrar dos tipos de desempleo:

- Desempleo Friccional: Es un tipo de desempleo que se suele producir en varias ocasiones. Una de ellas puede ser cuando se abandona el puesto de trabajo para encontrar otro más acorde con las características de la persona; otra es la nueva incorporación al mercado de trabajo, esta es muy común hoy en día ya que existe un desfase temporal entre la incorporación al mercado laboral y la finalización de los estudios. Otro caso, es cuando las empresas desarrollan su actividad en un determinado periodo de tiempo, provocando que el resto del tiempo no necesiten mano de obra.

Estudios empíricos² demuestran que existe un cierto nivel de paro friccional, pero que en el largo plazo debería reducirse.

- Desempleo Estructural: Este tipo de desempleo ocurre cuando lo que demandan las empresas no se corresponde con las necesidades de los trabajadores, esto puede ser por la localización del trabajo o por las diferentes capacidades que puede tener el empleado.

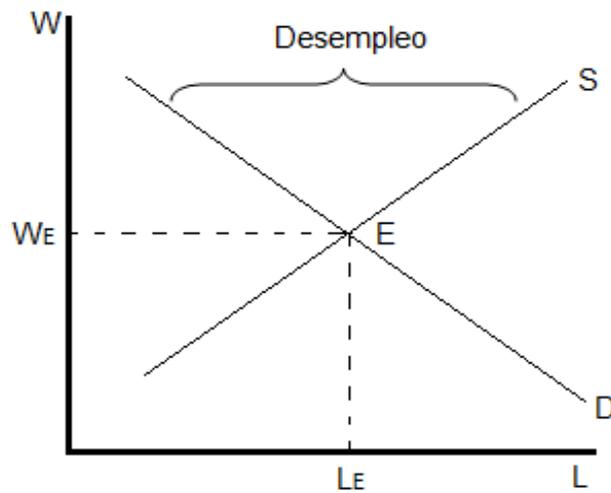
La reducción de este tipo de paro es más complicada, ya que existe un período prolongado de tiempo entre la búsqueda y la ocupación en un empleo.

La diferencia entre ambos tipos de desempleos reside en el tipo de duración, ya que el primero es un paro de corta duración y el segundo de larga duración más complicado de eliminar.

En el gráfico 2.1 recoge una situación de desempleo teóricamente.

² McConnell, C.R., Brue, S.L., & Macpherson, D.A. (2003). *Economía Laboral*. McGraw-Hill, Madrid.

Gráfico 2.1: Situación de desempleo



Fuente: Elaboración propia

2.1 SITUACIÓN ACTUAL DEL DESEMPLEO

El desempleo en la actualidad es uno de los desequilibrios que afecta a todos los países, y en mayor medida al caso español. Las devastadoras cifras macroeconómicas muestran una situación delicada, es por eso que la población española señala el desempleo como principal preocupación. Desde el inicio de la última crisis en el año 2008, las tasas de desempleo no han dejado de crecer lo que supone un grave problema para la economía española ya que, las personas ven minoradas sus rentas, lo que provoca que disminuya el consumo de bienes secundarios. Estos bienes, que no son de primera necesidad, son menos demandados ocasionando que muchas empresas tenga que cesar su actividad.

Esta situación es un gran inconveniente puesto que España es uno de los países que más desarrollado tiene su sector terciario de manera que necesita una mayor mano de obra.

La EPA³ recoge los datos de la población activa. En este último año y en comparación con los anteriores en esta etapa de crisis, la tasa de paro para el

³ EPA: Encuesta de población activa, elaborada por el INE (Instituto Nacional de Estadística) desde 1964. Refleja la situación del mercado laboral incluyendo una serie de variables que afectan a este como son el sexo, edad, educación...etc.

colectivo global de la población ha disminuido 6,3 puntos porcentuales⁴. Esta reducción es un dato esperanzador pero dista mucho de corregir el problema.

Los datos de desempleo son elaborados por dos fuentes distintas: la Encuesta de Población Activa (EPA), publicada por el INE trimestralmente, y los datos de registros administrativos que publica el Servicio Público de Empleo Estatal (SEPE). La diferencia entre ambos es que el INE ofrece datos del paro estimado y el SEPE ofrece los datos de paro registrado. Ésto hace que el INE sobreestime las cifras de desempleo, porque contabiliza como parados todos los trabajadores de la economía sumergida, y el SEPE subestime los datos de paro, porque no todos los parados se registran como tales en las oficinas de empleo. A la hora de comparar los datos con otros países o analizar los del propio país se utilizan los datos obtenidos en la EPA ya que los criterios de obtención de los datos son similares a los de otros países de la UE.

2.2 EVOLUCIÓN DEL DESEMPLEO EN ESPAÑA

El desempleo español ha sufrido cambios durante el periodo de 1977 a 2014, por eso es necesario analizar los cambios que provocan esta evolución.

Los datos utilizados para el estudio son los elaborados por la EPA, que como se explica anteriormente son los más idóneos⁵. Para ver la evolución que desarrolla la tasa de paro, hay que dividir el periodo en épocas de recesión y

⁴ Tasa de Variación: $X_t - X_{t-1} / X_{t-1}$. Comparando el año 2014 con el 2013, la tasa de paro ha disminuido 6,3%.

⁵ Los datos elaborados por la EPA para el periodo de 1964 a 2014 han sufrido una serie de cambios de metodología. La primera publicación de esta encuesta se realizó en 1964, hasta 1968 la encuesta fue trimestral, pero en 1969 hasta 1974 pasó a ser semestral y en 1975 volvió a cambiar a trimestral. En 1987, la encuesta cambió para adaptarse a las recomendaciones internacionales, recalculando los datos de los anteriores años.

En el año 1999, se convierte en una encuesta continua realizándose a lo largo de las 13 semanas de cada trimestre y no de las 12 como se venía haciendo hasta este año. En 2002, se realiza una nueva definición del paro creando una separación entre las series de parados y activos.

La última modificación se realizó en 2005, introduciendo un nuevo cuestionario realizado por encuesta telefónica. Las variables anuales fueron recogidas en una submuestra representativa que se realizaba mediante el promedio del año, en vez del trimestre como se hacía anteriormente. El último cambio de este consistió en recalcular los datos del periodo 1996-2004 con la nueva base de población instaurada, el fin es intentar mantener la homogeneidad.

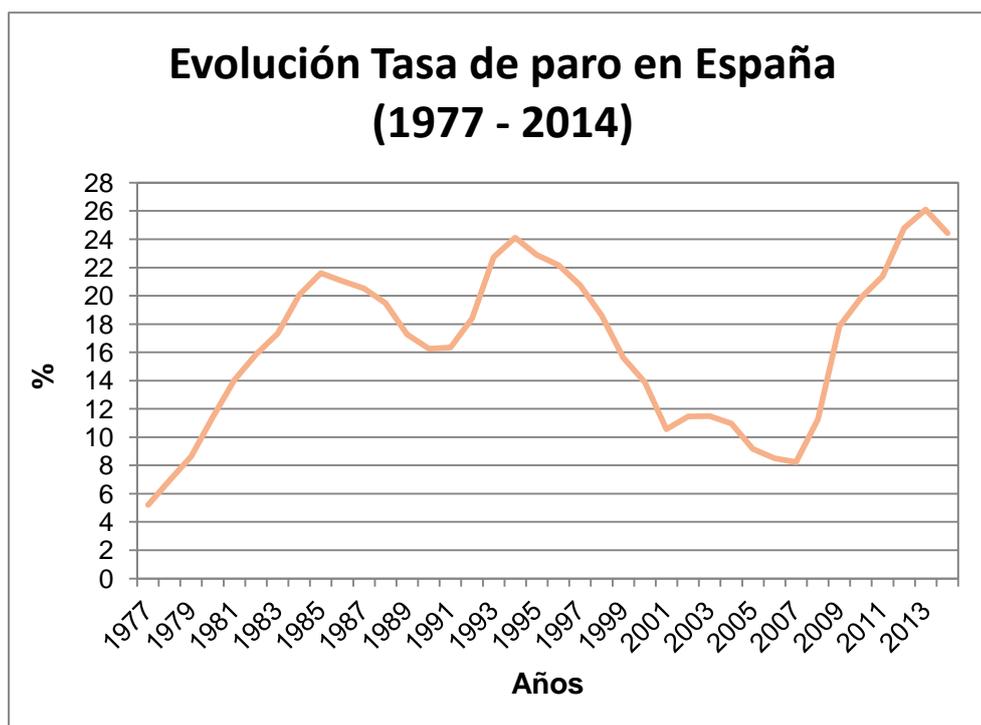
de auge económico. Cuando se cambia de fase o ciclo es porque, al menos, tiene que haber un crecimiento del 2%.

Las etapas que se dan en este periodo son:

- Desde 1975 hasta 1984: fase recesiva
- Desde 1985 hasta 1991: fase expansiva
- Desde 1992 hasta 1993: fase recesiva
- Desde 1994 hasta 2007: fase expansiva
- Desde 2008 hasta 2014: fase recesiva

En el gráfico 2.2 se muestra la evolución de la tasa de paro en España desde el año 1977 hasta el 2014.

Gráfico 2.2: Evolución Tasa de Paro en España para el periodo 1977-2014



Fuente: Elaboración propia a partir de datos del INE

La evolución que sufre la tasa de desempleo para años anteriores a 1975 es una situación de equilibrio porque había tasas por debajo del 5%. En estos años no aumentó el número de activos ya que muchas personas emigraban al extranjero en busca de mejores condiciones. Lo que explica estas

tasas tan bajas, es el aumento de la productividad que mantuvo la línea de población ocupada constante.

A partir del año 1975 el equilibrio en el mercado español se rompe hasta 1984, esto puede ser debido a la caída intensa de la tasa de ocupación, tras la crisis que afectó al país durante 10 años, provocó que las tasas de actividad disminuyeran por las malas expectativas generadas causando un efecto desanimo a la hora de buscar empleo. En esta época aumentó el número de personas en edad de trabajar a causa de un aumento de las tasas de natalidad y del retorno de los inmigrantes que no encontraron trabajo. En el exterior, se producía un aumento de las materias primas y el petróleo, ocasionando dos crisis que hicieron que muchas empresas acabaran en quiebra.

A la vez que había una fuerte crisis, los pactos de la Moncloa que pretendían actuar como una fuerza de expansión, intentado reducir la fuerte inflación y estabilizando el proceso de transición democrático.

La fase expansiva de la economía española duró de 1985 a 1991, lo más destacable fue el incremento de la tasas de ocupación por la creación del contrato temporal y el contrato en prácticas que ayudaron a que se creara empleo. En esta fase la oferta de trabajo aumenta debido a que se incorporan al mercado de trabajo los nacidos por el "baby boom" de los años 60 además de la incorporación de la mujer.

La tasa de actividad es pro-cíclica, es decir, va acorde al ciclo y como es una fase expansiva las ganas de trabajar aumentan no solo en las mujeres, sino en el resto de colectivos.

De 1992 a 1993, la actividad económica se desacelera y con ella la demanda de trabajo que hace que se destruya empleo.

La fase expansiva del ciclo económico completo se produjo de 1994 a 2007 donde hay cambio de tendencia en la demanda con una fuerte creación de empleo. Este cambio se produce por los pactos de Toledo⁶ y por la creciente llegada de inmigrantes.

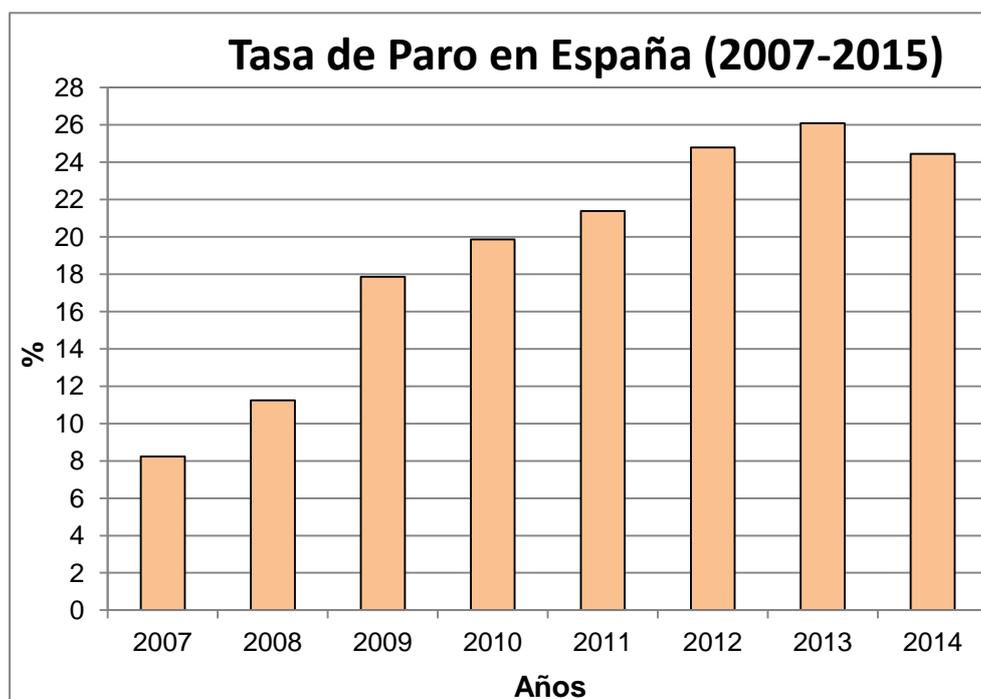
⁶ El pacto de Toledo surgió en 1995, está integrada por distintos parlamentarios de los partidos políticos. Su propósito es mantener la sostenibilidad del sistema de pensiones.

La crisis que estamos sufriendo ahora comienza en 2008, es una crisis mundial que se inicia en EE.UU con la quiebra de Lehman Brothers⁷ en 2007.

En esta recesión económica, la oferta de trabajadores se mantiene hasta 2012 pero la demanda de trabajo cae en picado. Desde este año, la oferta disminuye porque hay una salida de población tanto nacional como extranjera y por lo tanto los flujos migratorios vuelven a ser negativos. En 2014 la tasa de paro disminuye debido en parte a la caída de la población activa a causa de la emigración, esto es consecuencia del efecto desanimo que afecta sobre todo a la población joven y al envejecimiento de la población.

Gráficamente se recoge la situación de crisis en el siguiente gráfico.

Gráfico 2.3: Evolución de la tasa de paro para el periodo de crisis 2007-2015



Fuente: Elaboración propia a partir de datos del INE

2.2.1 Factores explicativos del desempleo

Las tasas de desempleo se han visto afectadas por una serie de factores que explican las fluctuaciones que ha sufrido a lo largo de los años.

⁷ Lehman Brothers: Banca de inversión estadounidense que gestiona todo tipo de activos especulativos.

Estudios empíricos⁸ muestran que la tasa de paro Española es mayor que la media Europea a lo largo de todo el periodo expuesto anteriormente. Para explicar los cambios en el desempleo, lo más común, es hacerlo desde el lado de la oferta y la demanda, y por otro lado de las instituciones del mercado. En los años comprendidos entre 1975 y 1984, como se ha comentado antes, se caracteriza por un aumento del desempleo, que sobre todo viene afectado por el lado de la oferta. Algunos shocks que afectan a la oferta son:

- Subida de los precios del petróleo, a partir de 1975 ya no se puede subvencionar más los precios del petróleo.
- Aumenta la eficiencia del trabajador, por lo tanto el empleo que hay es muy productivo.
- Aumento de los tipos de interés, con los pactos de la Moncloa se aplicó una política monetaria restrictiva para frenar la inflación provocando un aumento de los tipos de interés.

En 1985 a 1993, en esta etapa hay una fase expansiva que comprende de 1985 a 1991 y una fase recesiva en 1992 hasta 1993. La fase expansiva puede ser explicada por:

- Reducción de los precios del petróleo.
- Introducción de los contratos temporales, que ayudan a crear empleo.
- Incorporación de España a la UE.

La fase recesiva puede determinarse por:

- Aumento de la oferta de trabajo por la incorporación de la mujer y el "baby boom" de los años 60.
- Carente sensibilidad del salario real al desempleo.

En la etapa más expansiva de la economía española, en los años 1994 - 2007, es donde tasa de desempleo ha alcanzado el valor más bajo del todo el periodo estudiado, ya que se situaba en el 8%. Las perturbaciones que intentan explicar esta situación son:

- Caída de los tipos de interés, que provocó el boom inmobiliario y tuvo efectos positivos en la inversión y por tanto en el empleo.

⁸ García Delgado, J, & Myro, R. (2013). *Lecciones de Economía Española*. Aranzadi, Navarra.

- Liberalización de muchas actividades, se produce una mejora de la eficiencia y la innovación.
- Cuantiosos flujos de inmigración, el saldo migratorio es positivo, ya que hay una gran entrada de inmigrantes en busca de trabajo.

Todo este escenario virtuoso, causó desequilibrios en el mercado produciendo una crisis que ha elevado la tasa de paro hasta el 26% en el año 2013. Esta etapa se cree que se puede explicar desde dos puntos de vista:

- La oferta continua porque siguen llegando inmigrantes, pero en 2010 se estanca la llegada.
- Una fuerte destrucción de empleo por la caída de la demanda, ya que aumenta los tipos de interés y los precios del petróleo, créditos de mala calidad con tipos de interés elevados que hacen que no se puedan devolver, elevada temporalidad.

3. METODOLOGÍA

El objetivo de este estudio es observar que grado de influencia tienen ciertas variables independientes sobre la probabilidad de estar desempleado.

Para realizar el estudio se ha utilizado la encuesta de población activa del 2º trimestre de 2014, donde vienen recogidos una serie de características de cada persona entrevistada⁹.

El modelo trata de analizar los factores que influyen en que una persona esté desempleada. Vamos a estimar el efecto de esas variables en la probabilidad de que un individuo pueda estar desempleado. La variable dependiente es el desempleo, que tomará dos valores: 1 cuando este desempleado y 0 cuando no lo esté. Este tipo de regresión, con una variable de estas características, se conoce como modelos de variable binaria o dicotómica, o modelos de respuesta cualitativa, porque el factor a estudiar es una variable cualitativa.

⁹ La encuesta se realiza a unas 65.000 familias que aproximadamente son 180.000 personas.

Los métodos más utilizados para estudiar una regresión dicotómica son el modelo lineal de probabilidad y los modelos Logit y Probit, que describiremos brevemente a continuación.

Modelo lineal de probabilidad

Suponemos que nuestro modelo tiene la siguiente expresión:

$$Y_i = \beta' X_i + u_i$$

Con $E(u_i) = 0$

Donde $Y_i = 1$ si la persona está desempleada, y 0 si tiene empleo.

X_i = Conjunto de variables que afectan al desempleo como puede ser la edad, sexo...etc.

Entonces la probabilidad de que el suceso ocurra dado un X_i viene dado por la esperanza condicionada.

$$E(Y_i|X_i) = \beta' X_i$$

Lo que quiere decir que, dadas esas variables explicativas, la probabilidad estimada de que una persona esté desempleada viene dada por.

$$\hat{Y}_i = \hat{\beta}' X_i$$

Este método de estimación presenta varios problemas:

- Heterocedasticidad: Como Y_i solo toma dos valores, $u_i = Y_i - \beta' X_i$ tomará también dos valores $1 - \beta' X_i$ y $-\beta' X_i$. Por lo que para que la $E(u_i)=0$ la probabilidad que de adopte dichos valores será:

$$Y_i = 1 \rightarrow P(u_i = 1 - \beta' X_i) = 1 - \beta' X_i$$

$$Y_i = 0 \rightarrow P(u_i = -\beta' X_i) = \beta' X_i$$

Por consiguiente, después de realizar varias operaciones la varianza de la perturbación es:

$$Va(u_i) = E(u_i^2) = (1 - \beta' X_i)\beta' X_i = [1 - E(Y_i)]E(Y_i) = \hat{Y}_i (1 - \hat{Y}_i)$$

La varianza depende de los valores de la estimación de la variable dependiente, es decir es diferente para cada individuo por tanto presenta heterocedasticidad.

- El problema más importante que puede presentar este modelo es que la estimación $(\hat{\beta}' X_i)$ puede no estar dentro del intervalo $[0,1]$ y por tanto la interpretación del modelo como la probabilidad de que el suceso ocurra dado un X_i no es válida.

Estos problemas que plantea el modelo lineal en probabilidad, no se presentan en los modelos descritos a continuación.

Suponemos $Y_i^* = \beta' X_i + u_i$ donde Y_i^* es una variable no observable que mide la utilidad de tomar una decisión. Lo que se observa es una variable binaria asociada a ese modelo con el siguiente esquema:

$$Y_i = 1 \quad \text{si } Y_i^* > 0$$

$$Y_i = 0 \quad \text{si } Y_i^* \leq 0$$

Por consiguiente, la probabilidad en cada caso será:

$$P(Y_i = 1) = P(u_i > -\beta' X_i) = 1 - F(-\beta' X_i)$$

$$P(Y_i = 0) = P(u_i \leq -\beta' X_i) = F(-\beta' X_i)$$

Los valores observados son las realizaciones de una variable binomial con esas probabilidades, que dependerán de X_i .

La función de verosimilitud asociada será:

$$L = \prod_{Y_i=0} F(-\beta' X_i) * \prod_{Y_i=1} [1 - F(-\beta' X_i)] = \prod_{i=1}^n [F(-\beta' X_i)]^{1-Y_i} * [1 - F(-\beta' X_i)]^{Y_i}$$

La forma de la función F depende de las hipótesis que se hagan sobre u. Si la función de distribución de u es logística estamos ante un modelo Logit y si la función distribución es la de una normal será un modelo Probit.

En el caso del modelo Logit, la función de distribución de u es la logística, es decir

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

Si $z = -\beta' X_i$

$$F(-\beta' X_i) = \frac{e^{-\beta' X_i}}{1 + e^{-\beta' X_i}}$$

Si multiplicamos el numerador y denominador por $e^{\beta' X_i}$ obtenemos:

$$F(-\beta' X_i) = \frac{1}{e^{\beta' X_i} + 1}$$

$$1 - F(-\beta' X_i) = \frac{e^{\beta' X_i}}{1 + e^{\beta' X_i}}$$

La función de verosimilitud será:

$$L = \prod_{i=1}^n \left(\frac{1}{1 + e^{\beta' X_i}} \right)^{1-Y_i} * \left(\frac{e^{\beta' X_i}}{1 + e^{\beta' X_i}} \right)^{Y_i} = \frac{e^{\beta' \sum X_i Y_i}}{\prod_{i=1}^n (1 + e^{\beta' X_i})}$$

El logaritmo de la función de verosimilitud es:

$$\ln L = \beta' \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n \log(1 + e^{\beta' X_i})$$

El siguiente paso para calcular el estimador máximo verosímil (MV), es maximizar el logaritmo de la función de verosimilitud, para ello derivamos dicha función con respecto del parámetro.

$$\frac{d \ln L}{d \beta} = \sum X_i Y_i - \sum \frac{e^{\beta' X_i}}{1 + e^{\beta' X_i}} X_i = 0$$

La ecuación que obtenemos es una función no lineal en los parámetros, por tanto hace falta un método iterativo para hallar la solución.

Una vez que se obtienen los $\hat{\beta}$ se puede calcular la probabilidad estimada de que la i-ésima observación valga uno como:

$$\hat{P}_i = \frac{e^{\hat{\beta}' X_i}}{1 + e^{\hat{\beta}' X_i}}$$

Para un individuo con las características X_i , se puede calcular la probabilidad estimada de que esté desempleado. Al ser una función de distribución dicha probabilidad estará siempre comprendida entre 0 y 1. Por tanto, este tipo de modelos soluciona el principal problema del modelo lineal de probabilidad que no garantiza que la probabilidad estimada esté entre 0 y 1.

En el caso de tener un modelo Probit, lo que cambiaría con respecto al modelo anterior es que la función de distribución de u se supone normal. El método de estimación es igual al del modelo Logit, con los cambios oportunos en la función de distribución

El modelo Logit tiene una función de distribución logística, mientras que el modelo Probit tiene una función de distribución normal. Ambas distribuciones son próximas salvo en las colas, los resultados obtenidos no son muy diferentes de un modelo a otro excepto que la muestra sea muy grande que pueden aparecer desigualdades.

Los estimadores de ambos modelos no son directamente compensables porque para el Probit suponemos que $\sigma = 1$ y para el logit la varianza es $\frac{n^2}{3}$ con lo que para comparar los resultados del logit hay que multiplicar por $\frac{\sqrt{3}}{n}$ el modelo probit. Amemiya¹⁰ propone que se multiplique por 0,625, de esta manera se aproxima de la mejor forma las distribuciones.

Una característica fundamental de estos modelos es que los β_j ya no se interpretan como en el modelo de regresión clásico, ya no representan el cambio que se produce en la variable dependiente al incrementar una unidad la variable independiente. En este caso, el efecto que un incremento en la X_i produce sobre la probabilidad de estar desempleado es:

$$\frac{\partial E(Y_i|X_i, \beta)}{\partial X_{ij}} = f(-X_i'\beta)\beta_j$$

Que depende de las características del individuo, es por tanto diferente para cada individuo. Sin embargo, la dirección de dicho efecto la señala únicamente el signo del β correspondiente.

Dependiendo del modelo que tengamos nos quedará una expresión u otra de dicho efecto, en el caso del modelo Logit será:

¹⁰ Amemiya, T. (1981). Qualitative Response Models: A. Survey. *Journal of economic literature*, 19, 4, pp. 481-536.

$$\frac{\partial E(Y_i|X_i, \beta)}{\partial X_{ij}} = \frac{e^{X_i' \beta}}{(1 + e^{X_i' \beta})} \beta_j$$

Para calcular el efecto medio que tiene una variable sobre la probabilidad de estar desempleado se pueden seguir dos caminos, o bien se calcula ese valor para los valores medios de las variables independientes, o bien se calcula la media de los efectos para cada individuo.

4. ESTIMACIÓN DEL MODELO

4.1 VARIABLES

Las variables que afectan al desempleo son muchas, pero para realizar un estudio más preciso se seleccionan aquellas que pueden aportar información más relevante a este desequilibrio.

La variable dependiente es la población desempleada, que tal como se define en la Encuesta de Población Activa son todas las personas de 16 o más años que cumplan unas condiciones:

- Sin trabajo, es decir, que no hayan tenido un empleo ni por cuenta propia ni ajena en la semana de referencia¹¹.
- En busca de trabajo, que hayan tomado iniciativa de buscar empleo por cuenta propia o ajena.
- Disponibles para trabajar, en condiciones de comenzar a hacerlo en un plazo de dos semanas a partir de la semana de referencia.

Además, también se consideran paradas las personas que se han quedado sin empleo, están disponibles para trabajar y que no lo buscan por que ya han encontrado uno y se incorporarán en los tres meses posteriores a la semana de referencia.

En estas condiciones, cuando una persona reúna las tres características, a la vez, ó la condición anterior, estará desempleada.

¹¹ La semana de referencia es la semana anterior a la entrevista según este fijado en el calendario.

La variable desempleo debe recoger las condiciones anteriores, es por eso que hay que identificarlas en la encuesta de población activa, ya que como se comentó antes, ésta recoge mucha información. En este caso, el INE tiene construida una variable derivada que recoge la actividad económica de los encuestados cuando la persona tiene 16 años o más, es decir, la situación en que se encuentra cada persona en cuanto al empleo. Esta variable recoge tanto la población económicamente activa como la inactiva.

La población activa son las personas de 16 o más años que tienen un empleo, que están disponibles para ello y los que lo buscan activamente.

La población inactiva son las personas de 16 o más años que no desarrollan una actividad económica, no están disponibles para trabajar y no quieren buscar empleo. Este campo es prescindible para el estudio ya que como no desarrollan una actividad, no nos sirve para construir la variable desempleados.

La etiqueta AOI en la encuesta de población activa, recoge la situación anterior descrita mediante una serie de códigos.

Tabla 4.1: Clasificación de la actividad económica

Códigos	AOI
03	Ocupados subempleados por insuficiencia de horas
04	Resto de ocupados
05	Parados que buscan el primer empleo
06	Parados que han tenido un empleo antes
07	Inactivos pertenecientes al grupo de los desanimados
08	Inactivos desanimados y activamente potenciales
09	Resto de Inactivos

Fuente: Elaboración propia a partir de datos de la EPA

En este estudio lo que nos interesa es lo relacionado con el desempleo, es por eso, que se necesitan los datos referentes a los códigos 5 y 6, que

agrupan la opción de estar parado, y los datos con códigos 3 y 4 que reflejan la condición de tener un empleo. Para realizar el estudio hay que excluir los datos relativos a los individuos inactivos.

La variable desempleo será pues una variable ficticia que toma el valor 1 si el individuo está desempleado y el valor 0 si no está desempleado. En la muestra hay un promedio de 24,48% de desempleados.

En cuanto a las variables exógenas, hay muchas variables que influyen de manera directa en la situación de estar parado. Para estudiar el efecto que tienen se recogen solo las más significativas, estas son:

- Edad: es una variable continua. En la encuesta viene recogido con el mismo nombre. El promedio de la muestra es de 40 años, siendo la edad mínima de 16 años y la máxima de 65 años.
- Nacionalidad: Establece la nacionalidad de la persona encuestada. Al recodificar las variables, se modifica la serie y obtenemos que 1 es ser español y 0 extranjero. En la encuesta viene reflejado con la etiqueta NAC. Un 93,35% de las personas de la encuesta tienen nacionalidad española.
- Nivel de Estudios: Refleja el nivel de educación máximo adquirido por la persona. En este caso el nivel viene recogido por:
 - AN: Analfabeto
 - P1: Educación primaria incompleta
 - P2: Educación primaria
 - S1: 1ª Etapa de educación secundaria
 - S2: 2ª Etapa de educación secundaria
 - SP: 2ª Etapa de educación secundaria profesional
 - SU: Educación superior

Cuando se recodifica la variable, el valor 1 refleja poseer ese nivel y 0 no poseerlo. Las variables están agrupadas respecto a la muestra, ahora

poseer la característica de sin estudios será pertenecer al grupo de los analfabetos y al de educación primaria incompleta. También la educación secundaria comprende la 2º etapa de la educación secundaria y la 2º etapa de educación profesional.

El 1,75% de la muestra no tienen estudios, el 7,57% tienen educación primaria, el 31,84% 1º etapa de la educación secundaria, el 9,91% su máximo nivel de estudios es la educación secundaria y el 36,07% tienen educación superior.

En la encuesta la variable viene recogida con el rótulo NFORM.

- Sexo: Viene recogida en la encuesta con ese nombre. Al generar la variable dummy tenemos que 1 corresponde a la situación de ser hombre y 0 a la situación de ser mujer. El 52,8% de la muestra son hombre lo que quiere decir que hay un 47,2% de mujeres.

En la elección de las variables hay que ser consciente de que hay otras variables importantes como el salario que afecta a la probabilidad de estar desempleado pero de las que por falta de datos no han sido introducidas.

4.2 EL MODELO

El punto de partida será un modelo simple que conste de cinco variables exógenas, la edad, el sexo, la nacionalidad, la educación secundaria y la educación superior, para tratar de estimar el efecto que tienen sobre el desempleo.

Este modelo se estima mediante el método Logit, ya que como se comentó antes, este tipo de modelo nos permite obtener el efecto que tienen las variables sobre la probabilidad de estar desempleado. Para ello utilizaremos el programa Eviews que nos permite estimar el modelo, ver la bondad de ajuste de la estimación y el efecto que tiene cada variable explicativa sobre el desempleo.

Los pasos para estimar el modelo anterior descrito son:

1. El primer paso es definir el modelo y estimarlo por el método Logit.
Cuando tenemos definido el modelo con su variable dependiente y las

variables independientes, seleccionamos el método por el cual lo queremos estimar, en este caso como hemos dicho tenemos que escoger el sistema binario y dentro de este elegimos el procedimiento Logit.

2. El segundo paso es medir la bondad de ajuste. En este tipo de modelos de respuesta binaria, podemos utilizar el criterio de predicciones correctas que radica en observar el porcentaje de veces que el valor de la variable dependiente, en este caso el desempleo, coincide con su predicción. También se puede ver la diferencia entre el modelo estimado y el modelo cuando solo exista el término constante. Este efecto refleja lo que ganas o pierdes estimando el modelo. En todo caso la probabilidad de éxito será cuando sea mayor que 0,5.
3. La tercera fase consiste en ver el efecto de las variables explicativas sobre el desempleo. Esto se ve mediante el efecto parcial de una variable sobre la probabilidad de respuesta.

$$\frac{\partial E(Y_i|X_i, \beta)}{\partial X_{ij}} = f(-X_i'\beta)\beta_j$$

El coeficiente de cada variable viene dado cuando estimamos la ecuación en el primer paso. Como ya se ha comentado el efecto varía con el individuo para hallar el efecto lo primero que hay que hacer es calcular $(-X_i'\beta)$, una vez que se tienen se calcula la función de densidad de una distribución logística evaluada en esos valores y por último se multiplica por los β_j pertinentes de cada variable. Una vez que tenemos esa serie de números se hace es la media para ver el efecto medio de esa variable sobre el desempleo.

Este resultado es interesante cuando tenemos una variable continua. Ante variables discretas, la derivada no tiene mucho sentido. Para medir el efecto que poseer una determinada característica tiene sobre la

probabilidad de ocurrencia del suceso, se calcula la probabilidad de estar desempleado cuando la variable toma el valor 1 y el caso contrario cuando toma el valor 0, fijando el valor del resto de las variables en la media. La diferencia entre ambas probabilidades nos mide el efecto de poseer esa determinada característica. Algunos autores indican¹², que el resultado obtenido por este procedimiento y calculando la derivada, explicada anteriormente proporciona resultados muy similares, aunque señalan que no sería del todo correcto analizarlo mediante la derivada. En este caso vamos a ver las dos opciones para comprobar que existe muy poca diferencia.

4. La última fase es ver la curva de respuesta, que consiste en calcular la probabilidad estimada de estar desempleado en función de las variaciones en una variable continua. En nuestro caso, podremos analizar el efecto de la edad, que es la única variable continua, sobre la probabilidad de estar desempleado manteniendo las otras variables ficticias constantes en media.

Se han realizado unas curvas de respuesta para dos situaciones diferentes de las variables ficticias, como por ejemplo como varía la edad en la probabilidad de desempleo para hombre y mujer, viendo gráficamente la diferencia entre ambos. De la misma manera se hace para la nacionalidad y la educación.

También es posible plantear una serie de escenarios fijando la edad y creando una serie de circunstancias para las demás variables. Ahora no obtendremos una curva de respuesta sino que nos dará una probabilidad de ocurrencia. Esto es interesante porque permite comparar diferentes casos posibles reales observando cómo cambia la probabilidad al variar una característica. Al igual que antes tenemos que ir creando escenarios pero en este caso la diferencia es que las variables dummy tomarán valor en cada escenario que puede ser, o bien 1 o bien 0, representando una situación real en lugar de tomar el valor medio que no tiene interpretación.

¹² Greene, W.H. (1999). *Análisis Económico*. Prentice-Hall, Madrid.

5. RESULTADOS

Una vez explicado la estimación del modelo, vamos a ver los resultados que se obtienen.

La estimación del modelo viene recogida en la siguiente tabla.

Tabla 5.1: Estimación del modelo mediante el método Logit

Dependent Variable: DESEMPLEO				
Method: ML - Binary Logit (Quadratic hill climbing)				
Date: 06/27/15 Time: 00:22				
Sample: 1 78683				
Included observations: 78683				
Convergence achieved after 4 iterations				
Covariance matrix computed using second derivatives				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	1.094033	0.041906	26.10714	0.0000
EDAD	-0.038275	0.000766	-49.93728	0.0000
SEXO	-0.198600	0.017311	-11.47226	0.0000
NACIONALIDAD	-0.287557	0.031484	-9.133450	0.0000
EDUSECUNDARIA	-0.382803	0.028717	-13.33017	0.0000
EDUSUPERIOR	-1.024692	0.020348	-50.35794	0.0000
McFadden R-squared	0.060908	Mean dependent var		0.244792
S.D. dependent var	0.429967	S.E. of regression		0.415342
Akaike info criterion	1.045440	Sum squared resid		13572.50
Schwarz criterion	1.046147	Log likelihood		-41123.16
Hannan-Quinn criter.	1.045656	Restr. log likelihood		-43790.34
LR statistic	5334.365	Avg. log likelihood		-0.522644
Prob(LR statistic)	0.000000			
Obs with Dep=0	59422	Total obs		78683
Obs with Dep=1	19261			

Fuente: Elaboración propia a partir de datos de la EPA

En la tabla se puede ver cómo el $\hat{\beta}_j$ edad, sexo, nacionalidad, educación secundaria y educación superior tiene un signo negativo, lo cual quiere decir que ante un aumento de cualquiera de las variables tiene un efecto contrario en el desempleo, es decir, por ejemplo al aumentar la edad hace que sea menos probable que exista desempleo. Como ya hemos comentado la magnitud de los $\hat{\beta}_j$ no mide el efecto en la probabilidad de estar desempleado.

Todas las variables son significativas tanto individualmente como conjuntamente. El R^2 McFadden es muy bajo, esto puede ser debido a que el número de observaciones es muy elevado.

Para medir la bondad de ajuste acudimos al criterio de predicciones correctas, lo cual vemos en la siguiente tabla.

Tabla 5.2: Criterio de predicciones correctas

Expectation-Prediction Evaluation for Binary Specification						
Equation: EQ01						
Date: 06/22/15 Time: 18:10						
Success cutoff: C = 0.5						
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	58592	18128	76720	59422	19261	78683
P(Dep=1)>C	830	1133	1963	0	0	0
Total	59422	19261	78683	59422	19261	78683
Correct	58592	1133	59725	59422	0	59422
% Correct	98.60	5.88	75.91	100.00	0.00	75.52
% Incorrect	1.40	94.12	24.09	0.00	100.00	24.48
Total Gain*	-1.40	5.88	0.39			
Percent Gain**	NA	5.88	1.57			

Fuente: Elaboración propia a partir de datos de la EPA

En la tabla se puede ver como de las 78.683 observaciones, 59.725 están bien explicadas en el modelo, lo cual significa que el modelo predice adecuadamente un 75,91% de las observaciones alcanzando para el caso de los individuos empleados el 98,60%.

La ganancia de estimar el modelo explicado por las variables edad, sexo, nacionalidad y educación con respecto a un modelo donde solo hay término constante sería de un 0.39%. Se ha comentado que hay ciertas variables importantes en la explicación de los determinantes del desempleo que no han sido incluidas en el modelo, lo que podría explicar que la ganancia no sea muy grande.

En el efecto parcial explicado anteriormente, podemos ver el efecto que provoca en el desempleo un cambio en una unidad de la variable explicativa. En este tipo de modelo solo tiene sentido explicar mediante la derivada cuando la variable es continua como se ha comentado antes, en este caso solo es la edad y esta al incrementarse una unidad hace que la probabilidad de desempleo disminuya un 0,00643. El efecto de las variables dicotómicas no

tiene mucho sentido porque al tomar solo valores 0 y 1 no refleja en realidad el efecto que se produce, ya que la media de estos valores no es muy significativa pero aun así es una buena aproximación, como se ha comentado anteriormente. Así mismo al pasar de hombre a mujer el desempleo disminuye un 0.034218, también si se cambia de español a extranjero el desempleo disminuye un 0,049544. Al igual ocurre con la educación, ya que al incrementarse esta el desempleo disminuye un 0,065955 en el caso de la educación secundaria y un 0,176548 en el caso de la educación superior.

Otra forma de ver el efecto de la variable dicotómica es calcular la probabilidad estimada cuando esa variable toma el valor 1 y compararlo con la probabilidad cuando toma el valor 0. El resto de variables se fija es su media como se comentó antes. La tabla siguiente recoge dichos efectos

Tabla 5.3: Probabilidad estimada de una variable dicotómica

Variables ficticias	Probabilidad	Efecto en el desempleo
Sexo	Sexo=0 → p= 0,246050	-0,034964
	Sexo=1 → p= 0,211086	
Nacionalidad	Nac=0 → p= 0,277641	-0,053857
	Nac=1 → p= 0,223784	
Educación Secundaria	E.Sec=0 → p= 0,233854	-0,061563
	E.Sec=1 → p= 0,172291	
Educación Superior	E.Sup=0 → p= 0,298391	-0,165964
	E.Sup=1 → p= 0,132427	

Fuente: Elaboración propia a partir de datos de la EPA

La diferencia en las probabilidades estimadas cuando una variable, por ejemplo el sexo toma el valor 1, que corresponde con ser hombre, y cuando toma el valor 0 que pertenece a ser mujer, nos mide el efecto diferencial que provoca en la probabilidad de estar desempleado el ser hombre. En este caso dicho efecto es de una disminución del 3,4964%, esto quiere decir que la

probabilidad de estar desempleado en caso de una mujer es mayor que si fuera un hombre, y cuantificamos dicho efecto en un 3,4964%. Si ahora analizamos el efecto que tiene la nacionalidad, se puede ver que la diferencia entre ser un individuo español y otro extranjero, hace que la probabilidad de estar desempleado disminuya un 5,3857%. Al igual que en el caso anterior la probabilidad de estar desempleado de una persona española es menor que la de una persona extranjera.

De igual modo ocurre con la educación. En cuanto a la educación secundaria los valores 0 y 1 reflejan no poseer ese nivel de educación y si poseerlo, respectivamente. El cambio de un individuo que tiene estudios secundarios a una persona que no los tenga hace que la probabilidad de desempleo disminuya un 6,1563%. Si ahora nos fijamos en la educación superior, se puede ver como el efecto es mayor, es decir, el paso de una persona con estudios a una persona sin estudios superiores hace que la probabilidad de desempleo disminuya un 16,5964%.

Por tanto, vemos que los efectos de las variables ficticias sobre la probabilidad de estar desempleado son negativos, esto quiere decir que cuando se posee esa característica la probabilidad de que una persona esté desempleada disminuye. El efecto más significativo en la probabilidad de estar desempleado se puede ver en la educación y más concretamente en la educación superior, ya que la probabilidad de una persona con estudios superiores reduce el desempleo en un 16,5964%.

Al comparar los resultados anteriores con los calculados al hacer la derivada vemos como son muy parecidos por lo tanto si hubiera realizado solo con la derivada el ajuste sería bueno.

Para el cálculo de las probabilidades de la tabla anterior, se fijan los valores del resto de las variables en su media, lo que a veces hace que sea difícil su interpretación. Por este motivo se ha planteado calcular la probabilidad de estar desempleado en una serie de escenarios. Lo que se pretende es estimar dicha probabilidad para determinadas características del individuo. Los escenarios propuestos a continuación muestran para una determinada edad, en diferentes situaciones en cuando a sexo, nacionalidad y educación, la

probabilidad de que un individuo tipo con esas características esté desempleado.

Tabla 5.4: Escenarios

Edad	Sexo	Nacionalidad	E. Superior	\widehat{P}_i
23	H	Español	Si	0,157102
23	H	Extranjero	Si	0,199026
23	H	Español	No	0,341806
23	M	Español	No	0,387779
35	H	Español	Si	0,105340
35	H	Extranjero	Si	0,135674
35	H	Español	No	0,247022
35	M	Español	No	0,285782
60	H	Español	Si	0,043267
60	H	Extranjero	Si	0,056863
60	H	Español	No	0,111904
60	M	Español	No	0.133213

Fuente: Elaboración propia

Como se puede apreciar en la tabla, los cambios que se producen en la probabilidad de estar desempleado cuando modificamos solo una característica en el individuo tipo, miden el efecto, entre poseer esa característica y no poseerla en el desempleo, para un individuo con el resto de las características individuales fijas.

Se puede ver como hay escenarios muy significativos que reflejan porcentajes muy altos en el desempleo. El porcentaje más alto de desempleo sería para el escenario en que el individuo tenga 23 años, sea una mujer española que posea estudios secundarios en este caso la probabilidad de estar desempleado sería de un 38,7779%. En el mismo caso anterior pero siendo

hombre la probabilidad de estar desempleado disminuye, con respecto a ser mujer, pero aun así la probabilidad es de 34,1806%. Entonces entre los jóvenes españoles cuyos estudios máximos sea la educación secundaria, las mujeres tienen un 4,5973% más de estar desempleadas.

Si se analiza este caso para otro grupo de edad como puede ser el de 35 años, vemos que la probabilidad disminuye el 10% aproximadamente en ambos casos, ya que cuando el individuo es mujer la probabilidad sería de un 28,5782% y de un 24,7022% cuando es hombre. Sin embargo, para el grupo de edad de 60 años, las probabilidades de este tipo de escenario reducen considerablemente ya que representan el 13,3213% cuando la persona es mujer y el 11,1904% cuando la persona es hombre.

En contra podemos ver que el escenario donde el individuo tiene 23 años es hombre español y tiene como nivel máximo estudios superiores, la probabilidad de estar desempleado es de un 15,7102% y cuando la persona es extranjero con las mismas características que el individuo español la probabilidad de estar desempleado es de un 19,9026%. En este tipo de escenarios el cambio de la nacionalidad reflejaría una variación entorno al 4% de la probabilidad de desempleo. Si comparamos este caso con la edad de 35 y 60 años, vemos como la probabilidad de desempleo disminuye notablemente, cuando el individuo tiene 35 años la probabilidad de estar desempleado es de 10,5340% cuando la persona es española y un 13,5674% es una persona extranjera. El individuo tiene 60 años la probabilidad de desempleo con las mismas características que antes la probabilidad de estar desempleado sería de un 4,3267% y un 5,6863% para un hombre español y otro extranjero, respectivamente. En esta evolución vemos como el cambio de nacionalidad refleja cada vez una menor variación según se va incrementado la edad.

Por otro lado, se puede comparar la diferencia entre un hombre de 23 años español con estudios superiores, y otro con estudios secundarios, la diferencia entre ambos en la probabilidad de estar desempleado sería de un 18,4704%. Si se compara este mismo escenario para un individuo de 35 años la probabilidad de estar desempleado es de 14,1682%, ahora si el individuo tiene 60 años la diferencia sería de 6,8637%. Luego se puede ver que al ir

aumentado la edad la probabilidad de estos individuos de estar desempleados disminuye.

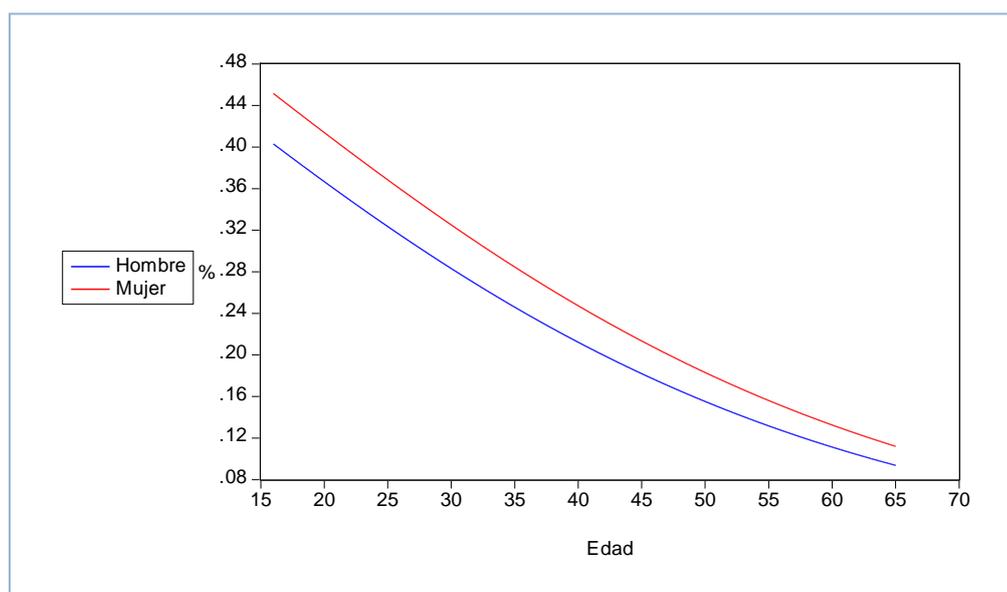
Como conclusión podemos ver que al ir aumentándose la edad la probabilidad de estar desempleado para una gran variedad de escenarios disminuye.

Otra forma de ver cómo afectan las variables a la probabilidad de desempleo es mediante la curva de respuesta que mide el efecto sobre la probabilidad de estar desempleado cuando hay una variación en la variable continua. Como solo hay una variable continua todas las curvas calculadas se representan frente a la edad. Lo que se ha hecho ha sido calcular dichas curvas para diferentes situaciones de las variables cualitativas, para así poder medir si la respuesta de la edad es la misma, o por el contrario cambia, según los valores que toma la variable ficticia. Las diferencias entre ambas curvas nos medirán el efecto de poseer una determinada cualidad sobre la probabilidad del desempleo para los diferentes valores de la edad.

La primera curva de respuesta recoge la probabilidad de estar desempleado según la edad para las dos situaciones hombre y mujer.

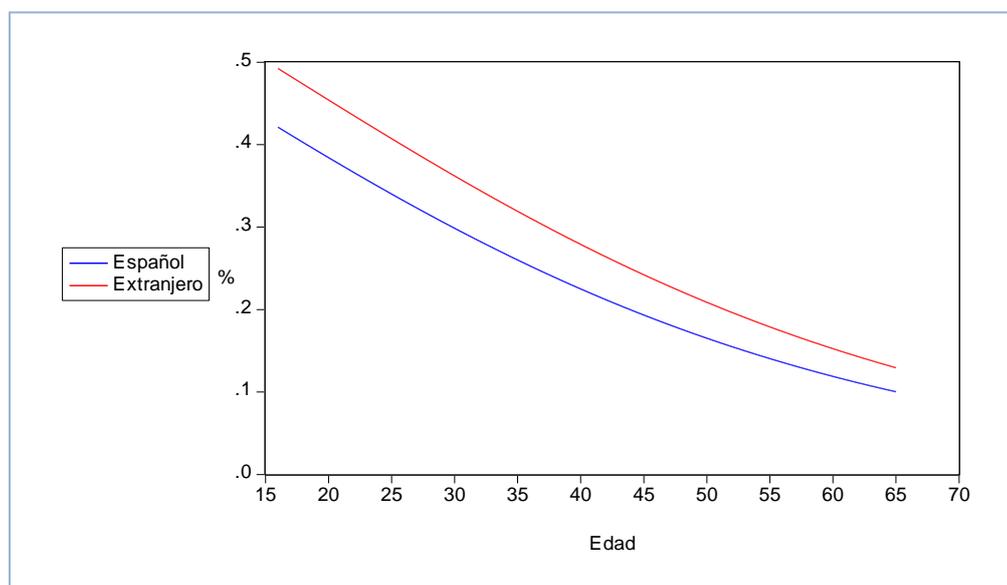
En la siguiente curva de respuesta podemos ver que las diferencias entre el sexo son pequeñas, para los primeros niveles de edad la diferencia es un poco más acentuada pero según pasan los años estas se van reduciendo. La probabilidad de desempleo de una mujer siempre será mayor que la del hombre.

Gráfico 5.1: Curva de respuesta de la edad y el sexo



Fuente: Elaboración propia a partir de datos de la EPA

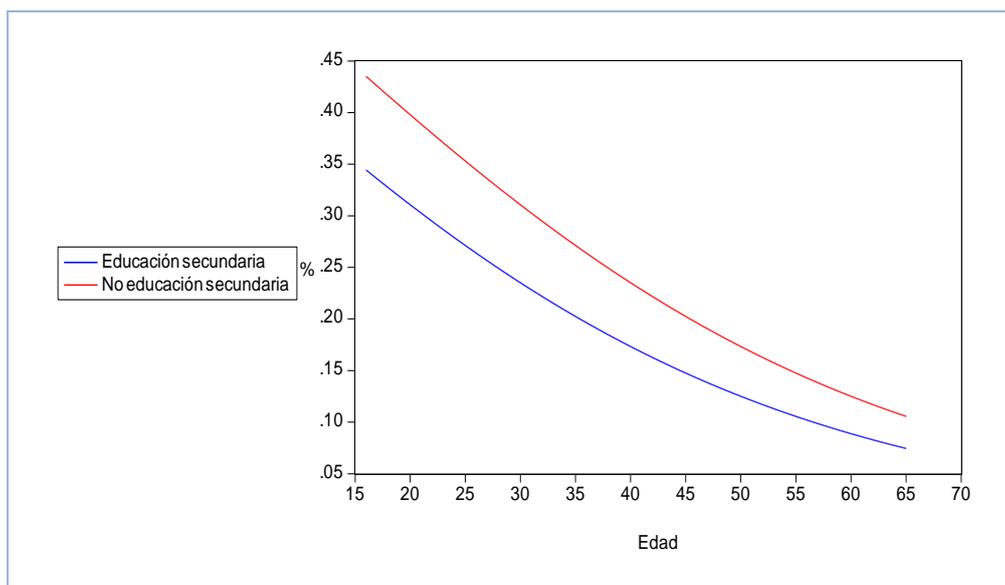
Gráfico 5.2: Curva de respuesta de la edad y la nacionalidad



Fuente: Elaboración propia a partir de datos de la EPA

En el gráfico 5.2 se observa que las diferencias en cuanto a la nacionalidad son más acentuadas en las edades tempranas y medias, reduciéndose en el último tramo de edad. En este caso la probabilidad de estar desempleado es mayor para una persona extranjera que para una persona nacional.

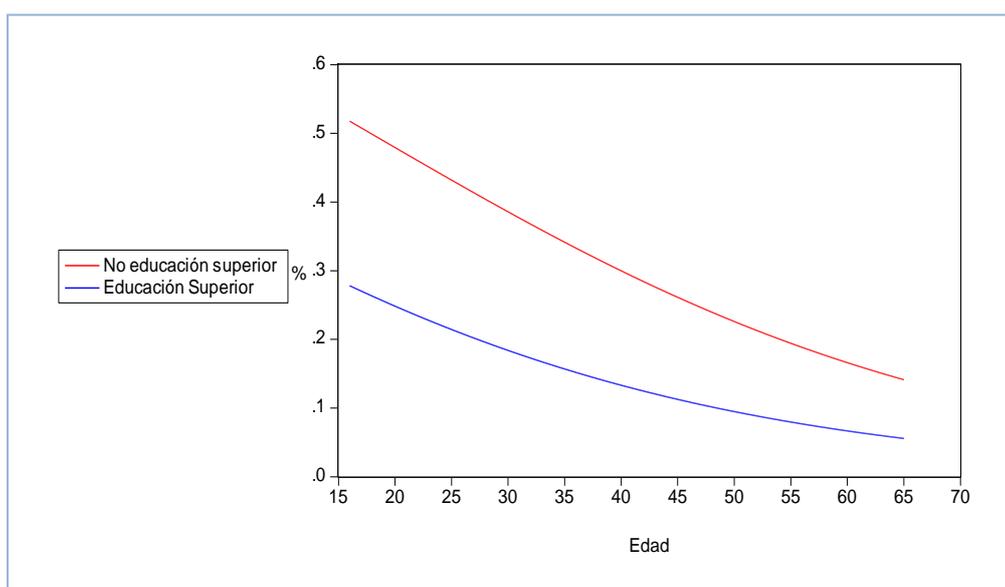
Gráfico 5.3: Curva de respuesta de la edad y la educación secundaria



Fuente: Elaboración propia a partir de datos de la EPA

En esta curva de respuesta se puede observar como las diferencias aumentan con respecto al nivel de educación. La diferencia es bastante elevada en los primeros años haciéndose más pequeña cuando se va incrementando la edad. La probabilidad de estar desempleado sin tener estudios secundarios es bastante más alta, que teniendo estudios. Estas diferencias se notan más en los años de incorporación al mercado de trabajo.

Gráfico 5.4: Curva de respuesta de la edad y la educación superior



Fuente: Elaboración propia a partir de datos de la EPA

En este gráfico se muestran las diferencias en cuanto a la educación superior, se puede ver que la diferencia es muy desmesurada en todo el periodo pero haciéndose más pronunciada en los primeros años. La probabilidad de estar desempleado sin educación superior es mucho más elevada que teniendo estudios superiores. En este caso las curvas no son paralelas ya que a más estudios la curva tiene un alto nivel de desempleo en los primeros años que luego se va reduciendo paulatinamente. Esta interpretación tiene sentido a partir de la finalización de estudios superiores, es decir, a partir de 22 ó 23 años que es cuando comienzan a incorporarse al mercado laboral.

Como conclusión podemos decir que las diferencias en la probabilidad de desempleo en las variables sexo y nacionalidad son bajas, es decir la diferencia entre ser hombre y mujer, y ser español o extranjero son insignificantes. Cuando observamos la educación las diferencias en la probabilidad de desempleo entre poseer ese nivel de educación y no poseerlo son suficientemente altas como para que suponga un fuerte incremento en la probabilidad de estar desempleado dejar de formarse. En el caso de la educación superior, la probabilidad de desempleo de un individuo con ese nivel de formación es considerablemente más baja que un individuo que no posea ese nivel y dicho efecto es más acentuado entre los jóvenes.

6. CONCLUSIONES

Después de ver las características del desempleo y su evolución, se ha intentado determinar que características individuales afectan al desempleo. Para ello se ha estimado un modelo Logit que mide la probabilidad del desempleo en función de unas variables como son la edad, el sexo, la nacionalidad y la educación. La bondad de ajuste de este tipo de modelos se mide mediante el criterio de predicciones correctas, en este caso el modelo predice correctamente el 75,91% de las observaciones.

En el modelo estimado se observa que todas las variables explicativas afectan negativamente a la probabilidad de estar desempleado, esto quiere

decir que ante cambios en estas variables la probabilidad de estar desempleado disminuye. El efecto de las variables dicotómicas se ve mediante la probabilidad del suceso, asignando una probabilidad al valor 0 y otra probabilidad al valor 1, la diferencia entre ambos es el efecto que tendrá sobre la probabilidad del desempleo. Se puede extraer que hay diferencias entre ambos valores para las variables estudiadas, ya que en el sexo las mujeres tienen más probabilidad de desempleo que los hombres. En la nacionalidad, una persona extranjera tiene más probabilidad de estar desempleada que otra que sea española. En cuanto a la educación, tanto secundaria como superior, la condición de no poseer ese nivel de educación hace que aumente la probabilidad de desempleo. El mayor grado de influencia que se obtiene sobre el desempleo lo tiene la educación y más concretamente la educación superior, reflejando una reducción de la probabilidad de desempleo un 16,5964% en el caso de poseer este nivel de educación.

Los escenarios planteados surgen por las limitaciones que presenta la probabilidad del suceso ya que las demás variables quedan fijas en su media, de aquí se obtiene que ante cambios en la edad la probabilidad de estar desempleado disminuye, por tanto para unas determinadas características al incrementarse la edad hacen que la probabilidad de estar desempleado disminuya. Además del resultado anterior, la probabilidad más reducida de estar desempleado está en el caso de que una persona sea hombre español y posea estudios secundarios y superiores. Este resultado también se corresponde cuando el individuo es una mujer, pero la diferencia en el sexo hace que normalmente la condición de ser mujer tenga una mayor probabilidad de estar desempleada. La conclusión más llamativa que se puede extraer es la diferencia entre un hombre de 23 años español con estudios secundarios y superiores, y otro con estudios solo secundarios, es del 18,4704%. Esto quiere decir que cuando menos nivel de estudios más aumenta la probabilidad de estar desempleado. Al aumentar la edad la probabilidad entre los individuos anteriores disminuyen siendo la probabilidad de desempleo de un 14,1682% a los 35 años y de un 6,8637% a los 60 años. Estos resultados reflejan que la brecha en cuanto a la educación es decreciente pero en edades más avanzadas es la que mayor efecto tiene sobre el desempleo.

Gráficamente podemos ver la probabilidad de estar desempleado para una variación de la variable continua, esto se conoce como curva de respuesta. Las curvas representan los dos valores que puede tomar la variable dicotómica, siendo la diferencia entre ambas lo que aumenta la probabilidad de desempleo de una característica a otra. Al analizar el sexo se ve como las desigualdades en la probabilidad de estar desempleado entre hombre y mujer son muy poco representativas ya que las curvas están próximas y la diferencia entre ambas es pequeña. Cuando se incrementa la edad esta distancia se va haciendo más corta, aunque la probabilidad de desempleo de la mujer es mayor que la del hombre. En cuanto a la nacionalidad la diferencia entre ser español o extranjero es más alta en los primeros años de la muestra disminuyendo, al igual que en el sexo, en los últimos tramos de la edad.

Cuando observamos las diferencias en educación, tanto secundaria como superior, las diferencias son más acentuadas haciendo que la probabilidad de estar desempleado sin tener estudios sea considerablemente más elevada que teniendo un nivel de educación. Al igual que en las otras dos variables dummy, al aumentar la edad estas diferencias se reducen, pero en el caso de la educación superior aún son bastante elevadas.

La conclusión general que se puede extraer de esta investigación es que al incrementarse la edad hace que la probabilidad de estar desempleado disminuya. Cuando se observan las variables ficticias todas ellas coinciden que poseer esa característica hace que disminuya la probabilidad de desempleo. Estas diferencias en las variables son más marcadas en edades jóvenes, lo que quiere decir que la probabilidad de desempleo en este intervalo de edad es mayor que en cualquier otro. Así mismo, las diferencias en educación son notables en el desempleo haciendo que esta diferencia sea notable también en edades avanzadas.

7. BIBLIOGRAFIA

Amemiya, T. (1981). Qualitative Response Models: A. Survey. *Journal of economic literature*, 19, 4, pp. 481-536.

Campbell R, M. (2003). *Economía Laboral* . McGraw-Hill, Madrid.

De la Fuente, A. (2012). Series enlazadas de los principales agregados nacionales de la EPA 1964-2009. *BBVA* , pp. 1-78.

Eckelund, R.B, & Hébert, R.F. (1992). *Historia de la teoría económica y de su método*. McGraw-Hill, México.

García Delgado, J., & Myro, R. (2013). *Lecciones de Economía Española*. Aranzadi, Navarra.

García, J., Fernández Villaverde, J., & Cabrales, A. (2009). *La crisis de la Economía Española*, FEDEA , Madrid.

Greene, W.H. (1999). *Análisis Económico*. Prentice-Hall, Madrid.

Gujarati, D. N. (2010). *Econometría* . McGraw.Hill, Bogotá.

Gujarati, D. N. (2006). *Principios de Econometría*. McGraw-Hill, Madrid.

Instituto Nacional de Estadística. Disponible en <http://www.ine.es/> [Último acceso: 03/04/2015].

Novales, A., & Sebastian, C. (1990). *El Paro en España: características, causas y medidas*, FEDEA, Madrid.

Pérez López, C. (2012). *Econometría Básica: Aplicaciones con EViews, STATA, SAS y SPSS*. Garceta, Madrid.

Rodríguez Braun, C. (1997). *Grandes economistas*. Pirámide, Madrid.

Sánchez Robles, B. (2006). La Economía: Concepto y Método. *Universidad de Cantabria* , pp. 1-64.