

COMUNICACIÓN

Título: La duración y la probabilidad de salida del desempleo. Análisis de los factores determinantes mediante la aplicación de modelos de supervivencia.

Autores: Helena Corrales Herrero y Beatriz Rodríguez Prado

Dirección: Departamento de Economía Aplicada (Estadística y Econometría). Facultad de Ciencias Económicas y Empresariales (Universidad de Valladolid), Avda. Valle Esgueva, 6, 47011 Valladolid.

Correo electrónico: helena@eae.uva.es , bprado@eco.uva.es

Introducción

El desempleo existente en una economía es uno de los principales problemas a los que se enfrentan las sociedades actuales y es considerado por muchos como una infrautilización de los recursos humanos disponibles. Es por ello que numerosos esfuerzos se han orientado a entender y solucionar el problema del paro. Desde un punto de vista académico, la mayor parte de los trabajos se ha dirigido a explicar su evolución, composición y determinantes, y menor énfasis se ha puesto en el estudio de los determinantes de la duración del desempleo, a pesar de ser una cuestión clave para entender el fenómeno del desempleo.

Las razones que fundamentan esta menor atención sobre la duración del desempleo hay que buscarlas en aspectos técnicos. En primer lugar, la escasez de información estadística para estudiar la duración del desempleo, que debe incluir el historial laboral de los individuos y, en segundo lugar, la dificultad para implementar en la práctica los modelos estadísticos que teóricamente se han desarrollado para modelizar este tipo de variables (modelos de supervivencia).

Actualmente parte de estos problemas se ha solventado y ya se han publicado algunos trabajos dedicados a estudiar la duración del desempleo en España. Sin embargo en la comunidad de Castilla y León todavía es un tema inexplorado. Por ello, el objetivo de este trabajo es analizar los determinantes de la duración del desempleo en Castilla y León y de la probabilidad de salir del mismo.

La salida del desempleo, desde el punto de vista de la teoría económica (modelos de búsqueda), viene determinada por dos probabilidades: la probabilidad de recibir una oferta de empleo y la probabilidad de aceptarla. La probabilidad de recibir una oferta de empleo depende de factores relacionados con las características personales del individuo que pueden afectar a la intensidad de la búsqueda de un empleo (la edad, el sexo, el nivel de estudios, etc), de factores cíclicos y de factores institucionales (políticas de empleo dirigidas a la contratación de determinados colectivos). Por otro lado, la probabilidad de que una vez recibida una oferta se acepte, está relacionada con el salario reserva de cada individuo, que viene determinado a su vez por sus características personales y por la situación cíclica de la economía.

El objetivo de este trabajo es cuantificar la influencia de cada uno de estos factores, sobre la probabilidad de salir del desempleo y realizar un análisis pormenorizado de la duración del desempleo.

El trabajo se estructura del siguiente modo. El apartado 1 describe los modelos de supervivencia, que son la base de la modelización estadística de la duración del desempleo y del trabajo empírico realizado. Los apartados 2 y 3 se dedican a la aplicación práctica propiamente dicha. El apartado 2 describe los datos utilizados en el análisis y las principales variables incorporadas y el apartado 3 incluye los resultados. Finalmente, aparece un apartado de conclusiones.

1.- Metodología estadística: modelos de supervivencia.

El análisis de la duración del desempleo, desde un punto de vista estadístico, consiste en caracterizar su distribución a partir de la estimación de la función de supervivencia ($S(t, X)$) y de la función de riesgo ($h(t, X)$). La primera informa sobre la probabilidad de que la duración del desempleo sea superior a un determinado número de meses y la función de riesgo cuantifica la probabilidad de que un individuo deje de estar parado condicionada a que ha estado desempleado un determinado número de meses.

La estimación de ambas funciones puede llevarse a cabo siguiendo métodos no paramétricos, como el de Kaplan-Meier, que utilizan solamente la información de la duración del desempleo y/o métodos paramétricos que permiten introducir en la especificación de la función de riesgo un vector X de factores determinantes tanto de la probabilidad de recibir una oferta de trabajo como de la probabilidad de aceptarla.

Dentro de los métodos paramétricos, el modelo más utilizado es el modelo de riesgos proporcionales propuesto por Cox (1968). El modelo de Cox se caracteriza por separar, de forma clara, los efectos de la duración (t) y el efecto del resto de factores recogidos en las variables explicativas (X), al especificar la función de riesgo mediante el producto de dos funciones, cada una dependiente de uno solo de estos aspectos,

$$h(t; x) = h_0(t) \cdot e^{x\beta}$$

A la función $h_0(t)$ se la conoce con el nombre de función de riesgo base pues expresa el comportamiento de la función de riesgo para un individuo de referencia, esto

es, cuando las variables X toman el valor 0. Por su parte, el término $e^{X\beta}$ permite determinar el efecto diferencial de los distintos factores recogidos en X sobre la probabilidad de abandonar el desempleo.

Adicionalmente, un aspecto de interés de la función de riesgo es que su comportamiento con respecto a t (duración) permite determinar lo que se denomina la *dependencia de la duración*, es decir, determinar si la probabilidad de que un individuo salga del desempleo es menor, mayor o igual a medida que su duración aumenta.

La especificación original de este modelo partía del supuesto de que la duración es una variable que se puede observar de forma continua. Sin embargo, en nuestro caso, y dadas las características de la fuente estadística utilizada (EPA), la duración sólo se observa de forma discreta. Esto nos ha llevado a utilizar los llamados modelos discretos de riesgos proporcionales. Una documentación más extensa de estos modelos la podemos encontrar en Prentice y Gloecker (1978), Nickell (1979), Meyer (1990), Kiefer (1990) y Jenkins (1995).

Dado que los individuos sólo pueden ocupar un estado o dejarlo, los modelos aquí especificados están ampliamente relacionados con los modelos de elección discreta binarios. Bastaría con transformar los datos de tal forma que partiendo de una observación por individuo obtengamos tantas observaciones para cada individuo como veces éste está sujeto al riesgo de salir del desempleo, teniendo en cuenta que la unidad de tiempo elegida para la duración es el mes, y estimar un modelo de elección discreta con estos datos transformados.

2.- Datos, variables utilizadas y descripción de la muestra

El análisis de la duración del desempleo realizado en este estudio abarca todo un ciclo económico que va del segundo trimestre de 1992 al segundo trimestre del año 2000 y los datos utilizados corresponden a los ficheros anonimizados de la Encuesta de Población Activa de los segundos trimestres de cada uno de esos años.

Más concretamente, para llevar a cabo el estudio hemos utilizado las preguntas retrospectivas del cuestionario de la EPA, las cuales proporcionan información sobre la

situación de los entrevistados un año antes de la realización de la encuesta y permiten reconstruir la duración del desempleo para los individuos de la muestra considerada¹.

Así, la muestra con la que hemos trabajado corresponde a los individuos que en la encuesta del segundo trimestre de cada año responden a la pregunta de *¿En qué situación se encontraba hace exactamente un año? buscando empleo*, para el periodo 1992-1998, y *parado (disponible y buscando empleo)*, en la EPA de 1999 y 2000 y que un año después siguen desempleados o han encontrado un empleo. De este modo, todos los individuos seleccionados parten de una situación de desempleo en el año previo a la encuesta (selección *stock*). Su situación laboral en el momento de realizar la encuesta nos permite distinguir entre aquellos que están empleados (han encontrado un empleo), y los que siguen desempleados. Dado que a partir del cuestionario de la EPA es imposible saber cuando un desempleado pasa a la inactividad y que en este estudio nos centramos en la salida del desempleo hacia el empleo, hemos prescindido de los individuos que abandonan el mercado laboral.

La selección de la muestra ha condicionado el proceso de construcción de la duración del desempleo. Antes de describir como se ha llevado a cabo ese proceso, hemos de señalar que las variables de la EPA que informan sobre duración, se han homogeneizado de modo que en todas ellas la unidad de medida es el mes. Así, contamos con variables que nos miden los meses que un trabajador lleva en el empleo actual, los meses desde que dejó el último empleo un parado con experiencia laboral o los meses que un desempleado con o sin experiencia laboral lleva buscando un empleo. Teniendo en cuenta este hecho, la duración del desempleo de los individuos de la muestra utilizada se ha construido del siguiente modo.

- Para los individuos que durante el año de seguimiento encuentran un empleo, hemos calculado el tiempo que están desempleados tras la selección como la diferencia entre los doce meses que se les ha seguido y el tiempo que llevan en el empleo actual. Estas observaciones son *no censuradas*.
- Para aquellos individuos que siguen parados al final del periodo, el valor de la duración es igual a doce meses, correspondiendo estas observaciones a *duraciones censuradas*².

¹ La EPA de flujos que recoge la parte común de la muestra rotatoria que configura la EPA no incluye la Comunidad Autónoma, por la que no es posible utilizarla para este análisis.

El objetivo del trabajo es encontrar los factores determinantes de la duración del desempleo y, por tanto, de la probabilidad de encontrar un empleo, partiendo de que la duración finalmente observada toma valores comprendidos entre 1 y 12 meses. Así, los individuos que encuentran empleo tienen duraciones de 1 a 12 meses, mientras que los que siguen desempleados su duración es mayor a 12. Dado que se desconoce cuál es la duración para estos últimos individuos, constituyen datos censurados en 12 meses.

Por último, antes de pasar a aplicar las técnicas propias de los modelos de supervivencia vamos a describir las principales características de los datos con los que vamos a trabajar.

En primer lugar, la muestra final está constituido por 92832 individuos en el caso de España, de los cuales 7840 corresponden a Castilla y León. Todos los individuos se caracterizan por partir de una situación de desempleo y se distinguen por que algunos tienen empleo un año después mientras que otros siguen en la misma situación. La Tabla 1 recoge el número de individuos que, tanto en España como en Castilla y León, partiendo de una situación de desempleo, han encontrado empleo (datos no censurados) o siguen desempleados (datos censurados).

	Castilla y León						España					
	Sexo				Total		Sexo				Total	
	Mujer		Varón		n	% col.	Mujer		Varón		n	% col.
	n	% col.	n	% col.			n	% col.	n	% col.		
Censura	3052	71.5%	1890	52.9%	4942	63.0%	31388	68.1%	25350	54.2%	56738	61.1%
No censura	1218	28.5%	1680	47.1%	2898	37.0%	14688	31.9%	21406	45.8%	36094	38.9%

Tabla 1: Distribución de los individuos según la censura

Como se desprende de la tabla, el porcentaje de los que siguen desempleados es casi el doble de los que encuentran empleo, siendo esa diferencia ligeramente más pronunciada en Castilla y León. Estos datos globales encubren las diferencias que

² Una cuestión a tener en cuenta es que algunos de los individuos que en la encuesta contestan que hace un año estaban desempleados y que actualmente siguen en esa situación, han podido tener un periodo de empleo intermedio durante el año, de modo que presentarían dos periodos de desempleo, cuya duración es imposible de determinar a partir del cuestionario de la EPA. Estos individuos se han identificado a partir de los meses que dicen estar buscando empleo, que en su caso será menor a 12 y se han eliminado del análisis.

existen por género. Mientras el porcentaje de hombres que encuentra empleo y el porcentaje de los que siguen desempleados tiende a aproximarse, la divergencia de esas cifras para las mujeres se acentúa. Así, el porcentaje de mujeres de Castilla y León que se colocan es de sólo el 28.5% frente al 47% que se registra para los hombres.

A pesar de estas discrepancias entre hombres y mujeres hemos de señalar que, tal y como aparece en la Tabla 2, apenas se observan diferencias en la duración media del periodo de desempleo de aquellos individuos que encuentran empleo. Tanto para Castilla y León como para el total nacional y, tanto para hombres como para mujeres, la duración media se cifra entorno a los 7 meses y medio, siendo el grado de dispersión en todos los ámbitos similar.

	Castilla y León			España		
	Total	Varón	Mujer	Total	Varón	Mujer
Toda la muestra	10,3709 (2,8214)	9,9224 (3,0548)	10,7459 (2,5507)	10,3479 (2,7686)	10,0713 (2,9038)	10,6285 (2,5944)
Muestra no censurada	7,5928 (3,0842)	7,5851 (3,0841)	7,6034 (2,9992)	7,7508 (2,9460)	7,7873 (2,9658)	7,6975 (2,9161)

(*) Las desviaciones estándar entre paréntesis

Tabla 2: Duración media del desempleo

La elección de los posibles factores determinantes de la duración del desempleo y, en consecuencia, de la probabilidad de encontrar un empleo se ha visto condicionada por la disponibilidad de datos, siendo los factores finalmente considerados los siguientes.

- Sexo del individuo. En la estimación se recoge mediante una variable que toma el valor 1 si es hombre y 0 si es mujer.
- Edad del individuo. Esta variable continua se ha agrupado en cuatro intervalos que comprenden las siguientes edades: de 16 a 29 años, de 30 a 44 años, de 45 a 54 años y 55 años o más.
- Nivel de estudios del individuo (estudios2). La información que proporciona la EPA se ha agrupado en cinco grupos: analfabetos y sin estudios, estudios primarios y FPI, bachillerato y FPII, estudios universitarios medios y FPIII y estudios universitarios superiores. Con esta clasificación se trata de recoger el número de años que cada individuos ha dedicado a su formación académica y, por lo tanto, su grado de cualificación.

- Estado civil del individuo. A este respecto se ha distinguido a las personas casadas de las que se encuentran en otra situación.
- Cabeza de familia. Se distingue si el individuo es cabeza de familia frente a cualquier otra situación.
- Cargas familiares. Se diferencia entre los individuos que tienen hijos de aquellos que no los tienen.
- Numero de hijos que tiene un individuo.
- Tamaño de la familia. Recoge el número de miembros que conviven en el hogar familiar.
- Porcentaje de parados en el hogar familiar.
- Porcentaje de ocupados en el hogar familiar.
- Situación cíclica de la economía. A partir de la evolución de la tasa de paro mensual a nivel nacional y de la tasa de variación del PIB mensual a nivel nacional se han definido dos variables ficticias³. Coyun1 que toma el valor 1 en los años con menor crecimiento económico y mayor tasa de paro (años 1992-1995) y Coyun2 que toma el valor 1 en aquellos que presentan el inicio de la fase ascendente del ciclo (años 1996-1998).

Las variables sexo, edad, y nivel de estudios tratan de recoger características personales que pueden influir tanto en la intensidad de la búsqueda como en el salario reserva de los individuos. Las variables familiares, como el estado civil, el ser cabeza de familia, el número de hijos, el tamaño de la familia y el porcentaje de parados y ocupados recogen condicionamientos externos que pueden generar una mayor o menor presión, según sea de la variable de la que se hable, sobre el individuo a la hora de buscar empleo o de fijar su salario de reserva. Finalmente, dos variables ficticias que recogen la situación cíclica de la economía y el grado de competencia por los puestos vacantes existentes en el mercado de trabajo.

La caracterización de los individuos que constituyen la muestra según las variables personales y familiares consideradas aparece en la Tabla 3, tanto para Castilla y León como España.

³ Agradecemos a José Luis Rojo García la mensualización de las series trimestrales de activos, parados y del PIB nacional.

		Castilla y León		España	
				Recuento	% col.
		Recuento	% col.		
Sexo	Mujer	4270	54.5%	46076	49.6%
	Varón	3570	45.5%	46756	50.4%
	Tota	7840	100.0%	92832	100.0%
Edad	16 a 29 años	4023	51.3%	47648	51.3%
	30 a 44 años	2725	34.8%	30874	33.3%
	45 a 54 años	766	9.8%	9919	10.7%
	55 años y más	326	4.2%	4391	4.7%
	Tota	7840	100.0%	92832	100.0%
Nivel de estudios	Analfabetos y sin estudios	207	2.6%	6540	7.0%
	Estudios primarios y FPI	4970	63.4%	61396	66.1%
	Bachillerato y FPII	1459	18.6%	14719	15.9%
	Estudios universitarios medios y FPIII	606	7.7%	5104	5.5%
	Estudios universitarios superiores	598	7.6%	5073	5.5%
	Tota	7840	100.0%	92832	100.0%
Estado civil	Otra situación	4776	60.9%	54135	58.3%
	Casados	3064	39.1%	38697	41.7%
	Tota	7840	100.0%	92832	100.0%
Cabeza de familia	Otra situación	6463	82.4%	71846	77.4%
	Cabeza de familia	1377	17.6%	20986	22.6%
	Tota	7840	100.0%	92832	100.0%
Cargas familiares	Sin hijos	5213	66.5%	59744	64.4%
	Con hijos	2627	33.5%	33088	35.6%
	Tota	7840	100.0%	92832	100.0%

Tabla 3: Descripción de las variables explicativas según ámbito territorial

La descripción de los datos muestra que hay una distribución más o menos proporcional entre hombres y mujeres, si bien ese reparto es menos equitativo en el caso de Castilla y León. Por otra parte, más del 50% de los individuos son jóvenes entre 16 y 29 años de edad, lo que determina a su vez que el porcentaje de personas casadas sea sólo del 41.7% en España y del 39.1% en Castilla y León. Ligeramente por debajo del porcentaje de personas casadas se encuentra el porcentaje de individuos con cargas familiares (con hijos), que se cifra en un 33% para Castilla y León, y un 35% en el total nacional. Asimismo, el porcentaje de individuos que son cabeza de familia es tan sólo de un 17.6% en Castilla y León, cinco puntos por debajo de la media nacional.

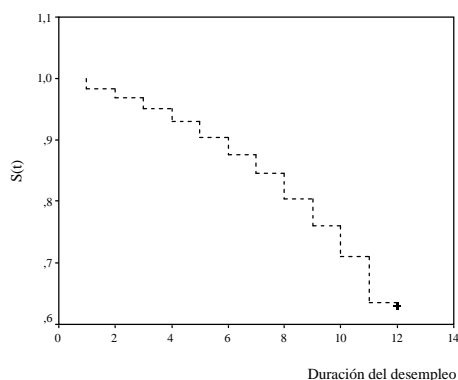
Por niveles de estudios hay un claro de predominio de individuos con estudios primarios y/o FPI, representando un 63.4% en Castilla y León y un 66.1% a nivel nacional. Asimismo, hay que destacar que en Castilla y León, el porcentaje de desempleados con bachillerato y estudios universitarios ya sean medios o superiores está dos puntos por encima de la media nacional. Por término medio las familias están formadas por 4 miembros y el número medio de hijos es 1.6. El porcentaje medio de ocupados en el hogar es del 35% y el de parados el 25%.

3.- Resultados

3.1.- Análisis univariante de la duración del desempleo: estimador de Kaplan-Meier

Una primera aproximación para estudiar la duración del desempleo consiste en estimar la función de supervivencia mediante el estimador no paramétrico de Kaplan-Meier, también llamado del producto límite (Kaplan y Meier, 1958). Los resultados, calculados a partir de las tablas de vida utilizando el programa SPSS se han representado en el siguiente gráfico:

España



Castilla y León

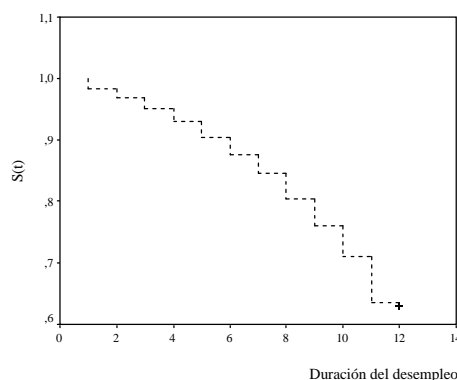


Gráfico 1: Función de supervivencia Kaplan-Meier

Como ya hemos señalado, la función de supervivencia recoge la probabilidad de que la duración del desempleo sea mayor a un número determinado de meses. Los resultados a este respecto ponen de manifiesto que tanto para España como para Castilla y León, la función de supervivencia desciende lentamente al principio, y a partir de aproximadamente el sexto mes, la caída es algo más drástica. Esto indica que los individuos que entran al desempleo, muestran mayores dificultades para salir de él en los primeros meses del mismo. Pasados estos meses, tal vez por un mayor esfuerzo en la búsqueda o por una reducción de su salario de reserva, las dificultades se van

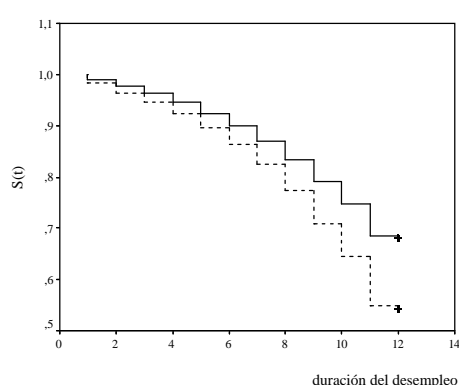
reduciendo. Por otra parte, la función de supervivencia de España está por encima de la función de supervivencia de Castilla y León, aunque esta diferencia no es significativa.

Si bien el estimador de Kaplan-Meier no permite incorporar factores como el sexo, los estudios o la edad en su estimación debido a su carácter no paramétrico, una práctica habitual en las aplicaciones, como paso preliminar en la especificación de un modelo más complejo, es la de obtener este estimador para diferentes subgrupos, definidos a partir de dichos factores. Esto nos va a permitir detectar determinadas asociaciones entre el tiempo en el desempleo (*tiempo de supervivencia*) y los factores considerados.

En el trabajo hemos realizado estas estimaciones para las distintas modalidades de los factores explicativos señalados anteriormente. Los resultados muestran que para todas las variables existen diferencias significativas entre las funciones de supervivencia, las cuales se han confirmado mediante el contraste del logaritmo del rango (Mantel y Haenszel, 1959). Asimismo, se observa que las diferencias, aún siendo significativas en ambos ámbitos geográficos, son más importantes en España que en Castilla y León, sobre todo para determinadas variables como los estudios o la variable que recoge la relación con el cabeza de familia.

Aquí, solo vamos a presentar los resultados para el factor sexo. El gráfico 2 muestra las funciones de supervivencia correspondientes a hombres y mujeres.

España



Castilla y León

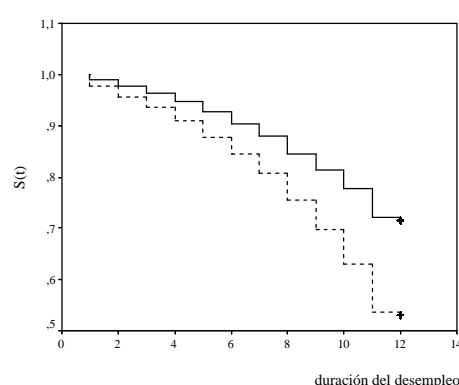


Gráfico 2: Funciones de supervivencia según sexo (---- hombres — mujeres)

Como acabamos de señalar la diferencia entre hombres y mujeres en cuanto a su función de supervivencia es más evidente en Castilla y León que a nivel nacional. Así, por ejemplo, la probabilidad de abandonar el paro en el primer año de desempleo es 14

puntos porcentuales más alta para hombres que para mujeres, 46% vs. 32%, en el caso de España, mientras que para Castilla y León, la diferencia es de 19 puntos (47% vs. 29%). En ambos casos, la función de supervivencia de las mujeres se encuentra siempre por encima de la de los hombres, no llegándose a cruzar en ningún momento, sino más bien alejándose aún más, lo que parece constatar que las mujeres tienen mayores problemas a la hora de abandonar el desempleo.

Junto con la función de supervivencia, el otro pilar del análisis de supervivencia es la función de riesgo. El estimador no paramétrico de esta función aparece en la Tabla 4 y se representa a su vez en el Gráfico 3.

	1	2	3	4	5	6	7	8	9	10	11	12
España	,0142 (,0004)	,0156 (,0004)	,0176 (,0004)	,0212 (,0005)	,0270 (,0006)	,0315 (,0006)	,0395 (,0007)	,0536 (,0008)	,0685 (,0010)	,0741 (,0010)	,1225 (,0014)	,0077
C y L	,0165 (,0015)	,0158 (,0014)	,0181 (,0016)	,0221 (,0017)	,0272 (,0019)	,0319 (,0021)	,0360 (,0023)	,0506 (,0028)	,0561 (,0030)	,0671 (,0034)	,1112 (,0046)	,0087

Tabla 4: Función de riesgo Kaplan-Meier

La característica más sobresaliente de esta función es su crecimiento lento pero continuado tanto para España como para Castilla y León. A la vista de estos resultados, se puede decir que la probabilidad instantánea de salir del desempleo en un periodo t crece con el tiempo, esto es, cuanto más tiempo se está desempleado más posibilidades hay de salir de esa situación (dependencia positiva). Asimismo, se observa un fuerte decrecimiento de la función de riesgo en el último valor de t observado ($t=12$), que se debe a que el número de individuos que encuentran empleo a los 12 meses de estar desempleados es bastante menor que en cualquier otra duración. En nuestra opinión este resultado se debe al cuestionario de la EPA y no ha un hecho real, por lo que los resultados para $t=12$ deben valorarse con mucha cautela. Téngase en cuenta que si realmente se diera ese resultado la entrada en la situación de parado de larga duración (más de un año) pudiera ser ciertamente perjudicial para la consecución de un empleo. Por otra parte, como consecuencia de la censura no podemos predecir que pasará transcurrido un año, pero en nuestra opinión la probabilidad de salir del desempleo seguirá una trayectoria creciente hasta llegar a un máximo, posiblemente hacia los 24

meses (parado de muy larga duración), a partir del cual empezará a caer progresivamente.

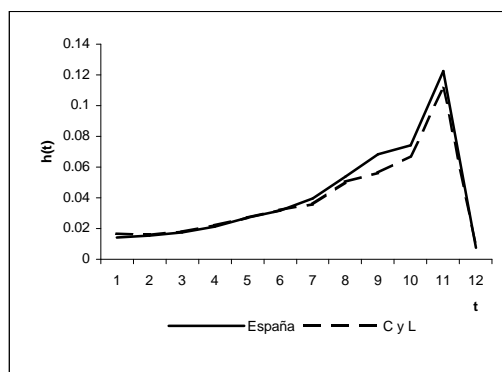


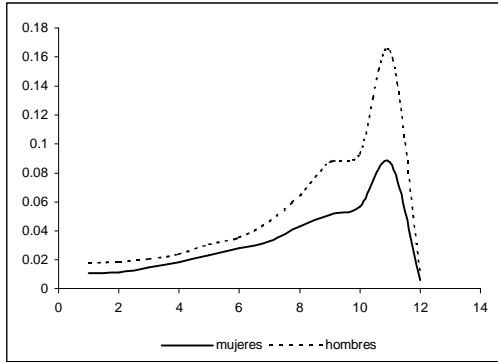
Gráfico 3: Función de riesgo Kaplan-Meier

Otra conclusión extraída de los datos es que la función de riesgo es ligeramente superior en el primer tramo para la población castellano-leonesa frente a la población española, pero a continuación se invierten los términos. Esto indicaría que mantendríamos una pequeña ventaja en cuanto a las posibilidades de salida del desempleo en los primeros meses de la duración del mismo pero esta situación ventajosa rápidamente se pierde e incluso las diferencias se amplían.

En cualquier caso podemos observar que la probabilidad de salir del desempleo condicionada a la duración que se lleva en el mismo no es muy elevada, alcanzando el valor máximo a los 11 meses en Castilla y León y España y cifrándose en tan solo un 11% y 12% respectivamente. De nuevo, señalamos que en nuestra opinión esa probabilidad seguirá una trayectoria creciente alcanzando su verdadero máximo una vez transcurrido los doce meses. Por otra parte, también hay que considerar que el modo de selección de la muestra (selección *stock*: individuos que en un momento dado están desempleados) tiende a aumentar la probabilidad de seleccionar a individuos con duraciones de desempleo más largas y que tienen, por tanto, menores posibilidades de salir del desempleo⁴.

⁴ El modelo estimado (aplicable a muestras *stock*) permite prescindir de la información sobre cuánto tiempo lleva desempleado un individuo en el momento en el que es seleccionado, por lo que muchos de los individuos pueden tener duraciones de desempleo previas a la selección muy diversas.

España



Castilla y León

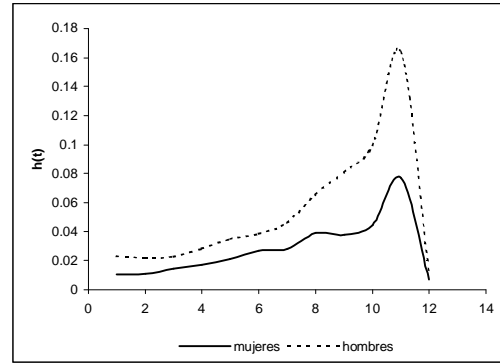


Gráfico 4: Funciones de riesgo Kaplan-Meier

En cuanto a las diferencias según sexo, observamos que el comportamiento de la función de riesgo para los hombres en España y Castilla y León mantiene la misma forma, esto es, ligeramente plana durante los primeros meses, con un crecimiento fuerte a partir del sexto mes hasta aproximadamente el noveno mes, donde de nuevo da un fuerte tirón para alcanzar su valor máximo. En el caso de las mujeres castellano-leonesas, el crecimiento a partir del sexto mes es más suave, y su posición final está por debajo claramente de la situación a nivel nacional.

3.2.- Análisis multivariante de la duración: modelo discreto de riesgos proporcionales

El método de Kaplan-Meier proporciona estimaciones útiles, tanto de la función de supervivencia como de la función de riesgo, que nos pueden ayudar a identificar los factores que afectan a la duración del desempleo, pero no proporciona estimaciones exactas debido a la falta de control sobre otras variables correlacionadas. Es por ello que es necesario recurrir a métodos paramétricos.

La estimación de la función de riesgo se ha llevado a cabo partiendo de la siguiente parametrización de la función de riesgo

$$h(t; x) = 1 - e^{-e^{x\beta + \theta(t)}} \Leftrightarrow \ln(-\ln(1 - h(t; x))) = x\beta + \theta(t)$$

que, siguiendo el planteamiento de Jenkins (1995) puede estimarse como un modelo de elección discreta, siempre que se preparen los datos del modo adecuado. Esta preparación de los datos implica desdoblarse las observaciones, de modo que el número total de datos pasa a ser de 960612 en el caso de España y 81308 para Castilla y León.

La elección de la función de riesgo base, $\theta(t)$, (la probabilidad de salir del desempleo en cada mes con independencia de las variables explicativas) puede realizarse utilizando distintas funciones. La opción que hemos seguido en este trabajo ha sido no imponer ningún tipo de estructura y dejar que sean los datos los que determinen esas probabilidades. Esto se consigue incorporando 12 variables ficticias, una por cada mes en el que el individuo puede salir del desempleo. Estas variables, denominadas mes1, mes2, ..., mes12 toman el valor 1 en aquellos individuos para los que su duración del desempleo ha sido igual al mes correspondiente y 0 para el resto.

El análisis de los resultados obtenidos lo hemos dividido en dos partes. En primer lugar, vamos a señalar cuales son las variables explicativas utilizadas indicando cuál es el efecto esperado sobre la probabilidad de encontrar un empleo. En segundo lugar, presentaremos los modelos estimados para Castilla y León y España.

3.2.1- Variables explicativas

A continuación se explican los principales efectos de las variables incluidas en el modelo. Hemos de señalar que la elección de las variables explicativas se ha visto condicionada por la disponibilidad de datos y que hay algunas factores con los que no se ha podido contar como la prestación por desempleo, es decir, si los individuos mientras están en la situación de desempleo perciben o no prestación. Las variables han sido agrupadas en características personales, características socio-demográficas y características del mercado.

Características individuales

Dentro de las variables que describen características personales y tras barajar diferentes combinaciones, el modelo final contiene: la edad del individuo, el nivel de estudios y el estado civil.

Edad: La edad puede tener efectos ambiguos sobre la probabilidad de salir del desempleo. Por un lado, en general, la edad considerada como una proxy del grado de experiencia, tendría un efecto positivo de tal forma que a medida que la edad aumenta, las posibilidades de abandonar el desempleo son mayores debido a una mayor experiencia. Sin embargo, y por otra parte, las personas de mayor edad se caracterizan por tener una menor movilidad geográfica y profesional, y por una menor adaptación a los cambios.

Nivel de estudios: Respecto al nivel de estudios, se espera que a medida que se completan niveles de estudios, las posibilidades de salir del desempleo sean mayores dado que las personas que reciben un mayor nivel de educación acumulan mayor capital humano. Por otra parte, de nuevo podemos encontrar un resultado contrario ya que los individuos, al incorporarse más tarde al mercado de trabajo, quieren ver recompensado este esfuerzo adicional, con un salario mayor. Eso significa que su salario de reserva, aquél a partir del cual una oferta recibida es aceptada, será más elevado y, por lo tanto, las posibilidades de abandonar el desempleo menores.

Estado civil: El hecho de estar casado puede tener un efecto diferenciado sobre la probabilidad de salir del desempleo según el sexo del individuo. Así, mientras que en los hombres se espera que el efecto sea positivo, en las mujeres el efecto puede ser negativo. Esto es debido a que mientras los empresarios pueden preferir hombres casados por ser más estables, pueden también ejercer una discriminación negativa sobre las mujeres casadas por el hecho de poder tener cargas familiares. Por otra parte, de cara a la intensidad de la búsqueda, las mujeres casadas pueden verse menos presionadas por considerarse segundas perceptoras de renta, no siendo ese el comportamiento cuando hablamos de hombres.

Características socio-demográficas

La situación familiar, en concreto su estructura y composición, pueden afectar de forma importante la búsqueda de un empleo. De ahí que las variables que a continuación comentamos recojan aspectos que están relacionados con las características de la familia a la que pertenece el individuo. De nuevo y tras una selección minuciosa, la lista de variables incluidas en el modelo son: el tamaño de la familia, el porcentaje de parados y de ocupados dentro de la familia, la existencia de cargas familiares y, en el caso de que existan cargas, el número de hijos.

Cargas familiares: La variable que al final se incluyó en el modelo queda definida como una variable dicotómica donde 0 recoge la no existencia de cargas familiares (familias que no tienen hijos) y 1, familias con hijos. Cabe suponer que los individuos que tienen hijos, se encuentran más presionados en la búsqueda de un empleo y esta búsqueda, tal vez más activa, les conduce a que sus posibilidades de encontrarlo sean mayores, de ahí el signo positivo. En el caso de las mujeres el signo no es tan claro, pues también actúa un efecto contrapuesto, debido a que el tiempo que se

dedica al cuidado de los hijos las puede hacer mostrarse menos activas en la búsqueda de empleo. También hemos considerado otras variables que tenían en cuenta la edad de los hijos. En concreto, hemos considerado los hijos menores de 3 años e hijos menores de 6 años. En ambos casos, los resultados no mejoran y en algunos casos empeoran por lo que finalmente hemos considerado sólo si los individuos tienen o no tienen hijos conviviendo en el hogar familiar.

Número de hijos: Consideramos que no sólo es importante a la hora de explicar la tasa de salida del desempleo, la tenencia o no de hijos (cargas familiares), sino también el número de éstos. En principio, el efecto de esta variable tendría la misma dirección que la ya comentada cargas familiares. Sin embargo, también hay que tener en cuenta que las características socioeconómicas y culturales de las familias con muchos y pocos hijos son muy diferentes, por lo que el efecto de esta variable sobre la probabilidad de encontrar un empleo es ambiguo. Posiblemente, esta variable no sólo recoja el efecto del número de hijos, sino las diferentes tipologías de familias. La pregunta que queda sin respuesta es qué tipo de familias suelen ser las que tienen mayor número de hijos.

Porcentaje de ocupados en la familia y porcentaje de parados en la familia: Con estas dos variables se recogen dos efectos contrapuestos, por un lado, el tener más miembros de la familia que trabajan, supone ampliar el número de contactos a la hora de buscar trabajo, pero también puede incidir negativamente ya que se está sometido a una menor presión ante la no urgencia de incrementar los recursos económicos de la familia. La interpretación de la otra variable es justamente al contrario.

Tamaño: El signo esperado es ambiguo. Por una parte, cuanto mayor es el número de miembros que conviven en el hogar mayor puede ser la presión para encontrar un empleo, pero a su vez mayores son las posibles fuentes de ingresos, lo que puede reducir la intensidad de la búsqueda y aumentar el salario reserva. Asimismo, creemos que tras esta variable se recogen otros factores que nuestros datos no permiten tener en cuenta como la renta familiar o factores culturales.

Cabeza de familia: Con esta variable se pretende recoger la mayor presión en cuanto a la búsqueda de un empleo que conlleva ser el responsable de mantener a la familia. En un modelo de oferta de trabajo familiar en el que las decisiones se toman conjuntamente, algunos miembros de la familia tienen menor probabilidad de trabajar si

el cabeza de familia está empleado. Por otro lado, el que el cabeza de familia trabaje supone la existencia de una red de contactos que facilita la posibilidad de encontrar trabajo a otros miembros de la familia. La variable no resultó significativa inicialmente por lo que se excluyó del modelo.

Situación económica

Dado que nuestro estudio abarca un periodo amplio (1992-2000), consideramos que las posibilidades de salir del desempleo se ven también afectadas por las condiciones del mercado y la situación cíclica de la economía. Para recoger estos dos aspectos se han considerado las dos variables ficticias definidas anteriormente, Coyun1 y Coyun2. El signo esperado de Coyun1 es negativo por la relación existente entre esta variable y las variables a partir de las cuales se ha construido: tasa de paro y tasa de variación del PIB.

Tasa de paro mensual: Esta variable trataría de recoger la competencia existente por los puestos vacantes en el mercado de trabajo, por lo que cuanto mayor sea esta tasa, menor es la probabilidad de salir del desempleo. El signo esperado, por tanto, es negativo.

Tasa de variación del PIB mensual: Trata de recoger la situación cíclica de la economía y, por tanto, se espera un signo positivo. En los periodos de menor crecimiento económico hay menos empleo y menores posibilidades de salir del desempleo.

3.2.2.- Modelos estimados

El modelo finalmente estimado, con el programa Eviews 3.0, se recoge en la Tabla 5 e incluye los resultados para España y Castilla y León.

Dado que una hipótesis adicional del trabajo es que los determinantes de la probabilidad de salir del desempleo son distintos para hombres y mujeres, o más precisamente que los determinantes, aún pudiendo ser los mismos, tienen un grado de influencia diferente, hemos reestimado las ecuaciones anteriores para la muestra de hombres y mujeres en los dos ámbitos geográficos analizados. La Tabla 6 recoge los resultados para la muestra de hombres y mujeres de España y la Tabla 7 los que corresponde a Castilla y León.

Finalmente, para facilitar la interpretación de los coeficientes del modelo, se presentan los efectos proporcionales de las variables explicativas sobre la tasa de riesgo, calculados como $e^{\beta} - 1$. Estos valores para todos los modelos estimados aparecen en la Tabla 8. En esa tabla sólo se recogen los datos de las variables cuyo coeficiente es significativamente distinto de cero en cada uno de los modelos.

La interpretación de los efectos proporcionales es la siguiente. Si la variable considerada es una variable dicotómica como por ejemplo el sexo (1=hombre, 0=mujer) el valor obtenido de 0.14 para España indica que la probabilidad de salir del desempleo de los hombres es un 14% superior a la de las mujeres. Si la variable cualitativa tiene varias modalidades y, por tanto, se recoge con varias variables ficticias, el efecto indica la porcentaje de variación de la probabilidad de salir del desempleo de los individuos con modalidad igual a 1 frente a los individuos cuya variable ficticia no está en el modelo. Así, por ejemplo, la edad está agrupada en 4 intervalos pero en los modelos sólo se incluyen tres variables dicotómicas, correspondientes a los 3 últimos grupos, por lo que el grupo de referencia es el de los individuos mas jóvenes (Edad1). De este modo, los coeficientes y los efectos proporcionales recogen el efecto diferencial frente a ese grupo de referencia. Concretamente el efecto proporcional de -0.029 para la variable Edad2 en España, indica que la probabilidad de salir del desempleo de un individuo con edad entre 30 y 44 años es un 2.8% menor a la que tienen los individuos con edad entre 16 y 29 años. Finalmente si la variable es métrica como por ejemplo el tamaño de la familia o el porcentaje de ocupado, el efecto proporcional recoge en cuánto varía la probabilidad de salir del desempleo por cada unidad adicional de la variable. Así, por ejemplo, el efecto proporcional de 0.012 asociado al tamaño familiar en España, indica que por cada miembro adicional en el hogar familiar, la probabilidad de salir del desempleo aumenta en un 1.2%.

Dado que la Tabla 8 de efectos proporcionales resume el signo de las variables en cada uno de los modelos, la significación de cada variable, es decir, si su coeficiente es distinto de cero, y el efecto sobre la tasa de riesgo, es decir, sobre la probabilidad de salir del desempleo, vamos a centrarnos en ella a la hora de comentar los resultados obtenidos.

El primer resultado que se obtiene es que la probabilidad de salir del desempleo es mayor para los hombres que las mujeres, tanto a nivel nacional como en Castilla y

León. En el caso de España la probabilidad de salir del desempleo de los hombres es un 14% mayor, mientras que en Castilla y León alcanza el 17.1%.

La probabilidad de salir del desempleo se reduce con la edad, lo que indica que domina el efecto de la menor adaptación a los cambios y deterioro del capital humano sobre la mayor experiencia. Aunque este es el resultado que se registra a nivel nacional, para Castilla y León y cuando se distingue entre hombres y mujeres, los resultados cambian. Así, en el caso de Castilla y León, los individuos con edad entre 45 y 55 años no presentan diferencias con los más jóvenes (la variable Edad3 no es significativa). Más evidente es el resultado que se obtiene para las mujeres, tanto en Castilla y León como en España, para las cuáles la edad no influye en la probabilidad de salir del desempleo.

La probabilidad de salir del desempleo aumenta con el nivel de estudios, es decir, con el grado de cualificación de los individuos, a nivel nacional, aunque no se aprecian diferencias significativas en la probabilidad de salir de desempleo de los individuos con estudios universitarios medios y superiores. Los resultados para Castilla y León, son un tanto diferentes, pues lo que se observa es que no hay variaciones significativas en la probabilidad de salir del desempleo de los individuos con distintos niveles de estudios, salvo los que tienen estudios primarios o FPI cuya probabilidad de salir del desempleo es 4.8% más alta que para el resto de los individuos. Sorprende el resultado que se obtiene para las mujeres de Castilla y León para las cuales la probabilidad de salir del desempleo no depende del nivel de estudios que se haya alcanzado. A nivel nacional, las mujeres con estudios primarios o FPI tienen una probabilidad de salir del desempleo un 3.34% más baja que la que tienen las que cuentan con estudios universitarios y las mujeres con bachillerato o FPPII también tienen una probabilidad de salir del desempleo menor, cifrándose la diferencia en un 3.7%.

Dependent Variable: SALIDA					
Method: ML – Binary Extreme Value					
Sample: 1 960612					
Included observations: 960612					
Convergence achieved after 5 iterations					
Covariance matrix computed using second derivatives					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
SEXO	0.131174	0.003940	33.28949	0.0000	
EDAD2	-0.029239	0.005333	-5.482206	0.0000	
EDAD3	-0.032493	0.008216	-3.954918	0.0001	
EDAD4	-0.111202	0.011880	-9.360655	0.0000	
ESTUD11	-0.059634	0.011641	-5.122731	0.0000	
ESTUD12	-0.035267	0.008318	-4.239717	0.0000	
ESTUD13	-0.030166	0.009154	-3.295350	0.0010	
ESTUD14	-0.011455	0.011000	-1.041392	0.2977	
ECIVC	-0.023785	0.006484	-3.668392	0.0002	
CARGAS	0.054764	0.009867	5.550024	0.0000	
NUMHIJOS	-0.007664	0.003665	-2.090932	0.0365	
OCUPADOP	0.507334	0.010050	50.48095	0.0000	
PARADOP	-1.423323	0.013587	-104.7594	0.0000	
TAMANO	0.012021	0.001418	8.476326	0.0000	
COYUN1	-0.175606	0.005481	-32.03715	0.0000	
COYUN2	-0.178317	0.005542	-32.17537	0.0000	
MES1	-1.416819	0.015254	-92.88104	0.0000	
MES2	-1.398613	0.015245	-91.74467	0.0000	
MES3	-1.350659	0.015097	-89.46660	0.0000	
MES4	-1.288748	0.014986	-85.99832	0.0000	
MES5	-1.206087	0.014826	-81.34832	0.0000	
MES6	-1.140003	0.014794	-77.05999	0.0000	
MES7	-1.062390	0.014726	-72.14524	0.0000	
MES8	-0.935880	0.014609	-64.05970	0.0000	
MES9	-0.824929	0.014594	-56.52686	0.0000	
MES10	-0.764920	0.014636	-52.26440	0.0000	
MES11	-0.519961	0.014575	-35.67491	0.0000	
MES12	-1.193528	0.016625	-71.79066	0.0000	
Mean dependent var					
S.D. dependent var					
S.E. of regression					
Sum squared resid					
Log likelihood					
Avg. log likelihood					
Obs with Dep=0					
Obs with Dep=1					
Total obs					
960612					

Dependent Variable: SALIDA					
Method: ML - Binary Extreme Value					
Sample(adjusted): 246 960574 IF CASTILLA Y LEON					
Included observations: 81308 after adjusting endpoints					
Convergence achieved after 6 iterations					
Covariance matrix computed using second derivatives					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
SEXO	0.158009	0.014099	11.20751	0.0000	
EDAD2	-0.039071	0.018629	-2.097340	0.0360	
EDAD3	-0.027692	0.029801	-0.929219	0.3528	
EDAD4	-0.159932	0.046897	-3.410297	0.0006	
ESTUD11	-0.009178	0.051748	-0.177361	0.8592	
ESTUD12	0.047752	0.026323	1.814052	0.0697	
ESTUD13	0.002365	0.029177	0.081044	0.9354	
ESTUD14	0.036071	0.034268	1.052614	0.2925	
ECIVC	0.002421	0.025362	0.095451	0.9240	
CARGAS	0.058285	0.037832	1.540651	0.1234	
NUMHIJOS	-0.027169	0.014401	-1.886631	0.0592	
OCUPADOP	0.517675	0.035687	14.50599	0.0000	
PARADOP	-1.683981	0.050574	-33.29729	0.0000	
TAMANO	0.011240	0.005065	2.219187	0.0265	
COYUN1	-0.110715	0.020535	-5.391611	0.0000	
COYUN2	-0.122903	0.020953	-5.865547	0.0000	
MES1	-1.466215	0.052219	-28.07814	0.0000	
MES2	-1.482056	0.052562	-28.19640	0.0000	
MES3	-1.418298	0.052121	-27.21174	0.0000	
MES4	-1.368392	0.051816	-26.40864	0.0000	
MES5	-1.286981	0.051143	-25.16417	0.0000	
MES6	-1.213062	0.051114	-23.73269	0.0000	
MES7	-1.167029	0.051133	-22.82330	0.0000	
MES8	-1.019564	0.050399	-20.22992	0.0000	
MES9	-0.973293	0.050406	-19.30899	0.0000	
MES10	-0.867338	0.050502	-17.17422	0.0000	
MES11	-0.636925	0.049900	-12.76396	0.0000	
MES12	-1.247035	0.057146	-21.82190	0.0000	
Mean dependent var	0.035642	S.D. dependent var		0.185398	
S.E. of regression	0.175441	Akaike info criterion		0.244090	
Sum squared resid	2501.756	Schwarz criterion		0.247294	
Log likelihood	-9895.217	Hannan-Quinn criter.		0.245071	
Avg. log likelihood	-0.121700				
Obs with Dep=0	78410	Total obs		81308	
Obs with Dep=1	2898				

Tabla 5: Modelo de riesgos proporcionales discreto para España y Castilla y León

Dependent Variable: SALIDA Method: ML - Binary Extreme Value Sample(adjusted): 1 960233 IF SEXO= HOMBRE Included observations: 470895 after adjusting endpoints Convergence achieved after 5 iterations Covariance matrix computed using second derivatives					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
EDAD2	-0.049879	0.007249	-6.880611	0.0000	
EDAD3	-0.090241	0.010678	-8.451374	0.0000	
EDAD4	-0.215506	0.014077	-15.30959	0.0000	
ESTUD11	-0.063801	0.016243	-3.927866	0.0001	
ESTUD12	-0.013565	0.013106	-1.035055	0.3006	
ESTUD13	-0.007152	0.014346	-0.498509	0.6181	
ESTUD14	-0.015863	0.018083	-0.877240	0.3804	
ECIVC	0.081065	0.009589	8.454189	0.0000	
CARGAS	0.080644	0.013645	5.909972	0.0000	
NUMHIJOS	-0.016164	0.004744	-3.407425	0.0007	
OCUPADOP	0.490553	0.013127	37.37039	0.0000	
PARADOP	-1.277770	0.017484	-73.08245	0.0000	
TAMANO	0.008681	0.001848	4.698363	0.0000	
COYUN1	-0.158214	0.007228	-21.88866	0.0000	
COYUN2	-0.156203	0.007351	-21.24844	0.0000	
MES1	-1.341261	0.020801	-64.48106	0.0000	
MES2	-1.326148	0.020805	-63.74269	0.0000	
MES3	-1.295451	0.020723	-62.51189	0.0000	
MES4	-1.246235	0.020620	-60.43915	0.0000	
MES5	-1.160053	0.020416	-56.82168	0.0000	
MES6	-1.102512	0.020446	-53.92243	0.0000	
MES7	-1.005110	0.020281	-49.55941	0.0000	
MES8	-0.870784	0.020171	-43.17076	0.0000	
MES9	-0.730459	0.020170	-36.21593	0.0000	
MES10	-0.674911	0.020284	-33.27244	0.0000	
MES11	-0.381701	0.020264	-18.83631	0.0000	
MES12	-1.136478	0.022955	-49.50900	0.0000	
Mean dependent var	0.045458	S.D. dependent var		0.208307	
S.E. of regression	0.197243	Akaike info criterion		0.302467	
Sum squared resid	18318.99	Schwarz criterion		0.303101	
Log likelihood	-71188.09	Hannan-Quinn criter.		0.302647	
Avg. log likelihood	-0.151176				
Obs with Dep=0	449489	Total obs		470895	
Obs with Dep=1	21406				

Dependent Variable: SALIDA Method: ML - Binary Extreme Value Sample(adjusted): 4147 960612 IF SEXO= MUJER Included observations: 489717 after adjusting endpoints Convergence achieved after 6 iterations Covariance matrix computed using second derivatives					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
EDAD2	-0.009077	0.008101	-1.120503	0.2625	
EDAD3	-0.003325	0.013335	-0.249307	0.8031	
EDAD4	0.012018	0.023993	0.500894	0.6164	
ESTUD11	-0.019473	0.018598	-1.047064	0.2951	
ESTUD12	-0.038521	0.010942	-3.520363	0.0004	
ESTUD13	-0.038735	0.012058	-3.212467	0.0013	
ESTUD14	-0.010235	0.013993	-0.731483	0.4645	
ECIVC	-0.128407	0.009124	-14.07396	0.0000	
CARGAS	0.031020	0.014809	2.094691	0.0362	
NUMHIJOS	0.008003	0.005883	0.136513	0.8914	
OCUPADOP	0.582532	0.016020	36.36323	0.0000	
PARADOP	-1.564523	0.021871	-71.53484	0.0000	
TAMANO	0.016352	0.002234	7.318302	0.0000	
COYUN1	-0.187799	0.008472	-22.16703	0.0000	
COYUN2	-0.197772	0.008497	-23.27685	0.0000	
MES1	-1.413280	0.022924	-61.65104	0.0000	
MES2	-1.386599	0.022865	-60.64219	0.0000	
MES3	-1.314340	0.022457	-58.52659	0.0000	
MES4	-1.236190	0.022218	-55.63823	0.0000	
MES5	-1.157061	0.021964	-52.68039	0.0000	
MES6	-1.081410	0.021844	-49.50674	0.0000	
MES7	-1.026049	0.021846	-46.96836	0.0000	
MES8	-0.907798	0.021631	-41.96780	0.0000	
MES9	-0.830955	0.021607	-38.45815	0.0000	
MES10	-0.761476	0.021598	-35.25641	0.0000	
MES11	-0.565219	0.021490	-26.30089	0.0000	
MES12	-1.152498	0.024572	-46.90207	0.0000	
Mean dependent var	0.029993	S.D. dependent var		0.170568	
S.E. of regression	0.161535	Akaike info criterion		0.210035	
Sum squared resid	12777.72	Schwarz criterion		0.210647	
Log likelihood	-51401.87	Hannan-Quinn criter.		0.210208	
Avg. log likelihood	-0.104962				
Obs with Dep=0	475029	Total obs		489717	
Obs with Dep=1	14688				

Tabla 6: Modelo de riesgos proporcionales discreto para España. Hombres y Mujeres

Dependent Variable: SALIDA Method: ML - Binary Extreme Value Sample(adjusted): 246 960148 IF CASTILLA Y LEON, HOMBRE Included observations: 35423 after adjusting endpoints Convergence achieved after 5 iterations Covariance matrix computed using second derivatives					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
EDAD2	-0.079496	0.026445	-3.006113	0.0026	
EDAD3	-0.110552	0.039441	-2.802968	0.0051	
EDAD4	-0.320895	0.054227	-5.917635	0.0000	
ESTUD11	0.085135	0.064279	1.324464	0.1853	
ESTUD12	0.108851	0.041064	2.650785	0.0080	
ESTUD13	0.063649	0.046149	1.379203	0.1678	
ESTUD14	0.082014	0.058235	1.408326	0.1590	
ECIVC	0.118487	0.038240	3.098534	0.0019	
CARGAS	0.149084	0.055754	2.673966	0.0075	
NUMHIJOS	-0.047161	0.019885	-2.371721	0.0177	
OCUPADOP	0.480666	0.047231	10.17692	0.0000	
PARADOP	-1.481135	0.066665	-22.21770	0.0000	
TAMANO	0.004617	0.006790	0.679990	0.4965	
COYUN1	-0.090495	0.027254	-3.320387	0.0009	
COYUN2	-0.118955	0.028181	-4.221047	0.0000	
MES1	-1.372509	0.071477	-19.20197	0.0000	
MES2	-1.402118	0.071513	-19.60646	0.0000	
MES3	-1.378125	0.071942	-19.15608	0.0000	
MES4	-1.306429	0.071291	-18.32543	0.0000	
MES5	-1.212458	0.070498	-17.19849	0.0000	
MES6	-1.187377	0.070870	-16.75424	0.0000	
MES7	-1.115670	0.070734	-15.77280	0.0000	
MES8	-0.957537	0.069925	-13.69380	0.0000	
MES9	-0.853354	0.069634	-12.25490	0.0000	
MES10	-0.734151	0.070254	-10.44989	0.0000	
MES11	-0.469219	0.069921	-6.710677	0.0000	
MES12	-1.235610	0.082200	-15.03184	0.0000	
Mean dependent var					
S.E. of regression					
Sum squared resid					
Log likelihood					
Avg. log likelihood					
Obs with Dep=0					
Obs with Dep=1					
Total obs					
35423					

Dependent Variable: SALIDA Method: ML - Binary Extreme Value Sample(adjusted): 4437 960574 IF CASTILLA Y LEON, MUJER Included observations: 45885 after adjusting endpoints Convergence achieved after 6 iterations Covariance matrix computed using second derivatives					
Variable	Coefficient	Std. Error	z-Statistic	Prob.	
EDAD2	-0.003886	0.027221	-0.142751	0.8865	
EDAD3	0.002577	0.048484	0.053152	0.9576	
EDAD4	0.013698	0.104951	0.130514	0.8962	
ESTUD11	-0.172145	0.111715	-1.540932	0.1233	
ESTUD12	0.008242	0.034832	0.236634	0.8129	
ESTUD13	-0.038343	0.038058	-1.007502	0.3137	
ESTUD14	-0.003959	0.043023	-0.092021	0.9267	
ECIVC	-0.104116	0.035350	-2.945269	0.0032	
CARGAS	0.019367	0.054001	0.358640	0.7199	
NUMHIJOS	-0.015661	0.021545	-0.726892	0.4673	
OCUPADOP	0.615742	0.056013	10.99278	0.0000	
PARADOP	-1.883625	0.079208	-23.78075	0.0000	
TAMANO	0.016212	0.007752	2.091212	0.0365	
COYUN1	-0.113143	0.031523	-3.589256	0.0003	
COYUN2	-0.118869	0.031708	-3.748870	0.0002	
MES1	-1.463278	0.076994	-19.00504	0.0000	
MES2	-1.446557	0.077727	-18.61076	0.0000	
MES3	-1.340787	0.075953	-17.65291	0.0000	
MES4	-1.317500	0.075656	-17.41440	0.0000	
MES5	-1.250020	0.074623	-16.75103	0.0000	
MES6	-1.120039	0.073951	-15.14572	0.0000	
MES7	-1.094079	0.074399	-14.70556	0.0000	
MES8	-0.966269	0.073206	-13.19926	0.0000	
MES9	-0.975716	0.073542	-13.26741	0.0000	
MES10	-0.876212	0.073340	-11.94719	0.0000	
MES11	-0.664059	0.072205	-9.196885	0.0000	
MES12	-1.142169	0.080779	-14.13946	0.0000	
Mean dependent var					
S.E. of regression					
Sum squared resid					
Log likelihood					
Avg. log likelihood					
Obs with Dep=0					
Obs with Dep=1					
Total obs					
45885					

Tabla 7: Modelo de riesgos proporcionales discreto para Castilla y León. Hombres y Mujeres

El estado civil del individuo, concretado en el hecho de estar o no casado, presenta un efecto diferente sobre la probabilidad de salir del desempleo según sea el sexo del individuo. Así, mientras que para los hombres de Castilla y León y España el efecto es positivo para las mujeres es negativo. Más concretamente los hombres casados tienen una probabilidad de salir del desempleo frente a los que no lo están un 8.4% más alta en el caso de España y un 12.5% en el caso de Castilla y León. Por su parte, las mujeres casadas tienen una probabilidad de salir del desempleo frente a las que no lo están un 12.1% más baja en el caso de España y un 9.9% en Castilla y León.

	España (total)	España (hombres)	España (mujeres)	Castilla y León (total)	Castilla y León (hombres)	Castilla y León (mujeres)
SEXO	0.140			0.171		
EDAD2	-0.029	-0.049		-0.038	-0.076	
EDAD3	-0.032	-0.086			-0.105	
EDAD4	-0.105	-0.194		-0.148	-0.275	
ESTUD11	-0.058	-0.062				
ESTUD12	-0.035		-0.038	0.049	0.115	
ESTUD13	-0.030		-0.038			
ESTUD14						
ECIVC	-0.024	0.084	-0.121		0.126	-0.099
CARGAS	0.056	0.084	0.032		0.161	
NUMHIJOS	-0.008	-0.016		-0.027	-0.046	
OCUPADOP	0.661	0.633	0.791	0.678	0.617	0.851
PARADOP	-0.759	-0.721	-0.791	-0.814	-0.773	-0.848
TAMANO	0.012	0.009	0.016	0.011		0.016
COYUN1	-0.161	-0.146	-0.171	-0.105	-0.087	-0.107
COYUN2	-0.163	-0.145	-0.179	-0.116	-0.112	-0.112

Tabla 8: Valores de $\beta - 1$ en los seis modelos estimados

Las cargas familiares tienen un efecto positivo sobre la probabilidad de salir del desempleo, excepto para las mujeres de Castilla y León, lo que indica que los hijos ejercen una mayor presión en la intensidad de búsqueda de un empleo y en la fijación del salario reserva. En España los individuos con hijos tiene un 5.6% más de posibilidades de encontrar un empleo que los que no los tienen. Esa diferencia en el caso de Castilla y León se cifra en un 6%. Por sexo, los hombres con hijos ya sean de España o de Castilla y León presentan una ventaja significativamente más elevada que

las mujeres, que como ya hemos señalado en el caso de Castilla y León no presentan ninguna frente a las mujeres sin hijos.

La influencia del número de hijos sobre la probabilidad de encontrar un empleo es negativa, en todos los ámbitos, si bien en el caso de las mujeres, ya sean de España o de Castilla y León, el efecto es nulo. Cuantitativamente el efecto es bastante más elevado en Castilla y León que en España. Mientras en España por cada miembro adicional que tiene una familia la probabilidad de encontrar un empleo se reduce en un 0.7%, en Castilla y León lo hace en un 2.6%. Como ya señalamos previamente el efecto de esta variable tendría que ser la misma que la de cargas familiares. Sin embargo, posiblemente, esta variable no sólo recoja el efecto del número de hijos, sino las diferentes tipologías socioeconómicas de las familias con pocos y muchos hijos.

La influencia tanto del porcentaje de ocupados y del porcentaje de parados en cada hogar es significativa en los dos ámbitos espaciales y, tanto para hombres como para mujeres. A medida que aumenta el número de ocupados la probabilidad de encontrar trabajo crecen, mientras que con el porcentaje de parados ocurre lo contrario. Como ya indicamos previamente parece que la situación laboral familiar tiende a afianzarse y así el pertenecer a una familia con mayor porcentaje de ocupados eleva la probabilidad de encontrar trabajo para su miembros. Lo contrario ocurre en las familias con un porcentaje elevado de parados. Como ya hemos señalado esto puede estar motivado por el hecho de que la principal forma de búsqueda de empleo son los contactos familiares y personales.

El número de miembros que conviven en el hogar familiar tiene una influencia positiva en la probabilidad de encontrar un empleo. Solamente aparece como una variable no significativa en el caso de los hombres de Castilla y León. Parece que predomina el efecto de que cuanto mayor es el número de miembros que conviven en el hogar mayor es la presión para encontrar un empleo, que la que implica que cuantas más fuentes de renta existan en la familia menor puede ser la intensidad de la búsqueda y mayor el salario reserva. En cualquier caso la influencia en términos cuantitativos no es muy elevada. Así, en Castilla y León por cada miembro adicional en el hogar la probabilidad de encontrar un empleo aumenta en un 1.1%.

Las dos variables que recogen la influencia del ciclo muestran el efecto esperado y así, cuanto menos expansiva es la situación económica, la probabilidad de encontrar

un empleo es menor. Más concretamente, en Castilla y León, la probabilidad de salir del desempleo en los años de crisis (1992-1995) es un 10.4% menor que en los años de mayor expansión (1999-2000). Un efecto similar se obtiene para los años intermedios. Contrariamente al resto de los factores, en España el efecto es más pronunciado registrándose en los años de crisis económica una probabilidad de salir de desempleo un 19.1% menor. Asimismo, se observa que el efecto del ciclo es mayor en mujeres que en hombres, tanto en España como en Castilla y León.

4.- Conclusiones

En este trabajo hemos pretendido analizar la duración del desempleo y la probabilidad de salir del mismo. Para ello hemos utilizado como fuente estadística los ficheros anonimizados de los segundos trimestres de la Encuesta de Población Activa de los años comprendidos entre 1992 y 2000. Estos datos permiten saber cuál es la situación laboral del individuo en el momento de la encuesta y un año antes de que ésta se realice, lo que posibilita calcular la duración del desempleo de cada individuo a lo largo de ese año. La unidad básica de observación ha sido el mes y los resultados permiten determinar cuál es la probabilidad de salir del desempleo cuando se lleva 1 mes, 2 meses,,12 meses desempleado. Esto constituye la función de riesgo.

La estimación de la función de riesgo se ha llevado a cabo mediante dos procedimientos, uno no paramétrico (Kaplan-Meier) y otro paramétrico (modelos de riesgos proporcionales discretos). El primero da una idea preliminar de la forma de la función, necesaria para la especificación del segundo tipo de modelos.

Los resultados de los modelos estimados permiten determinar cómo variables personales (sexo, edad, nivel de estudios y estado civil), familiares (cargas familiares, número de hijos, número de miembros que conviven en el hogar, porcentaje de ocupados y de parados en el hogar) y socioeconómicas (tasa de paro y tasa de variación del PIB mensual), influyen en la probabilidad de salir del desempleo.

El análisis se ha realizado para Castilla y León y España, en un intento de vislumbrar si existen diferencias significativas entre lo que ocurre en la comunidad y a nivel nacional. Asimismo, dado que una hipótesis adicional del trabajo es que los determinantes de la probabilidad de salir del desempleo son distintos para hombres y mujeres, o más precisamente que los determinantes, aun pudiendo ser los mismos,

tienen un grado de influencia diferente, hemos reestimado los modelos para la muestra de hombres y mujeres en los dos ámbitos geográficos analizados.

Los principales resultados obtenidos a nivel general muestran que la probabilidad de salir del desempleo es mayor en hombres que en mujeres, disminuye con la edad y crece con el nivel de estudios. El efecto de estar casado es diferente en hombres y mujeres, para los hombres el efecto es positivo mientras que para las mujeres es negativo.

La estructura familiar tiene también influencia. Por una parte, el hecho de tener hijos actúa positivamente, pero a medida que crece el número de hijos el efecto se invierte. El porcentaje de ocupados y parados en el hogar tiene un efecto contrapuesto. Mientras que la probabilidad de salir del desempleo crece con el porcentaje de ocupados en la familia, se reduce con el porcentaje de parados. Por su parte, el número de miembros que conviven en el hogar familiar influye positivamente.

El pulso de la economía influye en la probabilidad de salir del desempleo, ya que variables relacionadas con la tasa de paro mensual y la tasa de variación del PIB mensual tienen un efecto significativo.

Estos resultados, con pequeñas diferencias, se obtienen tanto a nivel nacional como en Castilla y León, si bien las diferencias más destacables se observan entre hombres y mujeres. Los hechos más resaltables pueden resumirse en que la edad y el nivel de estudios tienen mucha menor influencia en las mujeres, de Castilla y León y que en ellas tampoco influyen las cargas familiares.

Podemos concluir diciendo que este es un primer estudio en el que se trata de abordar la problemática del desempleo en Castilla y León desde el punto de vista de la duración del mismo, lo que esperamos que sea de utilidad a la hora de instrumentar medidas encaminadas a aumentar las posibilidades que tienen los individuos para encontrar un empleo o a atender a colectivos en los que esas posibilidades son de entrada menores.

Referencias Bibliográficas

- (1) AHN, N., DE LA RICA, S. Y UGIDOS, A. (1999): Willingness to move for work and unemployment duration in Spain, *Economica*, 66, pág. 335-357.
- (2) AHN, N. Y GARCÍA PÉREZ, I. (1999): *Unemployment duration and workers' wage aspiration*, documento de trabajo de FEDEA, 99-20.

- (3) ALBA RAMÍREZ, A. (1999): Explaining the transitions out of unemployment in Spain: the effect of unemployment insurance, *Applied Economics*, 31.
- (4) BOVER, O. Y GÓMEZ, R. (1999): *Nuevos resultados sobre la duración del desempleo: el paro de larga duración y la salida a un empleo fijo*, documento de trabajo del Servicio de Estudios del Banco de España, nº 9903.
- (5) COX, D. R. (1972): Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society*, serie B, 34, pág. 187-220.
- (6) DEVINE, T. J. Y KIEFER, N. M. (1991): *Empirical Labor Economics*, Oxford University Press.
- (7) JENKINS, S. P. (1995): Easy estimation methods for discrete-time duration models, *Oxford Bulletin of Economics and Statistics*, 57 (1), pág. 129-138.
- (8) KAPLAN, E. L. Y MEIER, P. (1958): Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53, pág. 457-481.
- (9) KIEFER, N. M. (1990): Econometric methods for grouped duration data, en *Panel Data and Labor Market Studies*, Elsevier Science Publishers.
- (10) LANCASTER, T. (1979): Econometric methods for the duration of unemployment, *Econometrica*, 47 (4), pág. 939-956.
- (11) MEYER, B. D. (1990): Unemployment insurance and unemployment spells, *Econometrica*, 48, pág. 757-782.
- (12) NICKELL, S. (1979): Estimating the probability of leaving unemployment, *Econometrica*, 47 (5), pág. 1249-1266.
- (13) PRENTICE, R. L. Y GLOECKER, L. A. (1978): Regression analysis of grouped survival data with applications to breast cancer data, *Biometrics*, 34, pág. 57-67.