

# PRINCIPALES RASGOS DEL DESEMPLEO MASCULINO EN ESPAÑA

A la hora de diseñar políticas de empleo, se ha de tener en cuenta a quiénes han de ir dirigidas para que sean eficientes. En este trabajo de **Jaime García, Clemente Polo y José Luis Raymond**, se utilizan métodos econométricos para modelos de elección discreta a fin de conocer cuál es la probabilidad de que un individuo con determinadas características personales esté en paro. Este enfoque elimina los posibles sesgos que se derivan de utilizar los porcentajes de paro publicados en la Encuesta de Población Activa como medida de dicha probabilidad y, a la vez, permite determinar cuáles son los grupos mayormente afectados por el paro.

## 1. INTRODUCCION

**S**EGUN la Encuesta de Población Activa (EPA), en el tercer trimestre de 1984 unos 2,7 millones de españoles estaban en paro, lo que representa una tasa de paro ligeramente superior al 20 por 100. Bajo un planteamiento muy simplista, esto significa que si el desempleo estuviese distribuido en forma equitativa, si la duración fuese de un año, y suponiendo que en el año considerado el flujo de nuevos desempleados fuese igual al de personas que dejan de estar en paro (1), cada persona perteneciente a la población activa estaría desempleada cada cinco años (2). Pero es bien sabido que éste no es el caso, ya que una buena parte de la población activa no sufre problemas de desempleo a lo largo de su vida laboral. El problema del desempleo está concentrado sobre unos grupos específicos, por lo que sería de interés analizar si el que estos grupos sean los más afectados se debe a que la probabilidad de entrar en el paro es mayor que la media, o

si, una vez en el paro, la duración del período de desempleo es mayor que la media. Los resultados de este tipo de análisis podrían ser utilizados a la hora de diseñar políticas que pretendan hacer más equitativa la situación de los individuos de una comunidad. En base a esto, el propósito del presente trabajo es tratar de determinar cuáles son esos grupos en el caso de la economía española, con las limitaciones derivadas del tipo de información disponible. Estas limitaciones serán objeto de discusión a lo largo de esta sección.

Este tipo de análisis, ya efectuado en otros países, por ejemplo Gran Bretaña (Nickell [1980]), requiere disponer de datos individuales sobre características personales y situación laboral. A la hora de efectuar el presente trabajo, sólo se disponía de una base de datos proveniente de la Encuesta de Presupuestos Familiares (EPF) correspondiente al período 1980-81. Dicha encuesta es esencialmente una encuesta de consumo, por lo que los datos acerca de la situación laboral de los individuos se redu-

cen a cuál es la situación en que se encuentran, sin referencia alguna al tiempo que llevan en paro. Por ello, no es posible con esta información abordar el estudio antes mencionado en su totalidad, debiéndonos centrar en el análisis de cuáles son los grupos con mayor probabilidad de estar desempleados. La disposición en el futuro de los datos de la EPA permitiría poder abordar el estudio en toda su extensión (3). Por otra parte, la carencia de datos en la EPF sobre el número de horas trabajadas y en la EPA sobre ingresos, no permite efectuar análisis de oferta de trabajo a nivel de todo el país, siendo sólo posible efectuar estudios de participación. El uso de encuestas de carácter local sí que ha permitido análisis más estructurales de las decisiones de oferta de trabajo. A este respecto, el trabajo de Hernández-Iglesias y Riboud (1985) es un buen ejemplo.

Aunque en la EPF existe información sobre características personales para los distintos miembros de cada hogar, esta información no es homogénea, en particular en lo que hace referencia al nivel de instrucción de los integrantes del hogar. Mientras que del sustentador principal se conoce el nivel máximo de estudios terminado, del resto de miembros sólo se conoce el nivel de estudios si son estudiantes o si cursan algún tipo de estudio de forma regular y organizada en el momento de la entrevista. Dada la importancia de la variable que refleja el nivel de instrucción, se ha preferido centrar el análisis en los sustentadores principales (cabezas de familia) únicamente.

Debe destacarse que la tasa de paro entre los sustentadores principales difiere notoriamente

CUADRO N.º 1

## TASAS DE PARO EN LA EPA

	<i>Personas principales (*)</i>	<i>General</i>
1980-III .....	4,1	11,6
1984-III .....	7,1	20,5

(\*) Se ha obtenido del siguiente modo: (Personas en paro/Total de viviendas familiares) × 100.  
Por lo tanto, esta tasa no es propiamente una tasa de paro, ya que la persona principal en alguna de las viviendas puede no pertenecer a la población activa.

Fuente: Encuesta de Población Activa.

de la tasa de paro para la totalidad de la población. Ello es debido a la correlación (positiva) existente entre ser sustentador principal y estar empleado, en virtud de cómo se define en la EPF al sustentador principal (4). Ello queda patente en el cuadro número 1, en el que se presentan las tasas de paro totales y para el subgrupo de personas principales de las viviendas para los terceros trimestres de los años 1980 y 1984. Deben destacarse las ligeras diferencias entre el concepto de sustentador principal en la EPF y el de persona principal en la EPA (5).

El Instituto Nacional de Estadística (INE) ha contrastado la fiabilidad de las estimaciones del paro a través de la EPF comparándolas con las obtenidas de la EPA (véase INE [1985]), siendo bastante similares las cifras que se deducen a partir de ambas encuestas. Por otra parte, de los 23.972 hogares que componen la muestra de la EPF, en 1.058 de ellos el sustentador principal estaba en paro, lo que representa un 4,4 por 100, cifra similar a la correspondiente a las personas principales en paro según la EPA. Por ello, el uso de la EPF para analizar los rasgos característicos del paro en Es-

paña parece adecuado. Dado que en el análisis efectuado se han excluido aquellos sustentadores principales que forman parte de la población inactiva (rentistas, jubilados, estudiantes, amas de casa), la tasa de paro en las distintas muestras utilizadas es ligeramente superior al 6 por 100.

En la monografía del INE sobre el paro y sus repercusiones sobre el nivel de vida (INE [1985]) antes mencionada, se presenta un análisis descriptivo de la tipología de los hogares con sustentador principal en paro. En el presente estudio, el interés se extiende a las características propias del sustentador principal (edad, nivel de instrucción, condición de emigrante, entre otras). Adicionalmente, se hace uso de modelos de elección discreta para determinar los rasgos característicos del paro entre los sustentadores principales. En esta clase de modelos, la variable endógena es de tipo discreto (o cualitativa); en nuestro caso, estar o no estar en paro. Con respecto al simple cálculo del porcentaje de individuos con una determinada característica que están en paro, este tipo de análisis ofrece la ventaja, al igual que los modelos de regresión, de

que se trata, en cierto modo, de un análisis de correlaciones parciales, es decir, que el efecto de cada característica sobre la variable endógena se obtiene manteniendo constantes todas las demás. Esto no ocurre cuando simplemente se calculan porcentajes. El único aspecto no contrastable con la información disponible es si los resultados obtenidos con la EPF 1980-81 sirven para explicar los rasgos característicos del paro en la actualidad. La información de la EPA, y el que sea solamente algo más de cuatro años la diferencia temporal entre el momento en que se efectuó la encuesta y la actualidad, hace pensar que las conclusiones que se deriven de este análisis siguen siendo válidas en el momento actual (6).

Por lo tanto, el propósito del presente trabajo se centra en determinar los rasgos característicos del paro entre los sustentadores principales de los hogares españoles. En la sección 2 se ilustra, para un sustentador tipo, cuál es el efecto sobre la probabilidad de estar empleado (o desempleado) de distintos rasgos característicos como la edad, el nivel de instrucción, la condición de reciente emigrante, la comunidad autónoma de residencia y el número de habitantes del municipio en que se reside. Estos resultados corresponden a la estimación de distintos modelos de elección discreta que se detallan en el apéndice 2 de este trabajo. Las variables utilizadas, así como los criterios para definir las muestras empleadas, están descritos en el apéndice 1. En la sección 3 se discute la viabilidad y efectos de distintas políticas tendentes a hacer más equitativa la situación de los parados en España. En la última sección se exponen las

conclusiones y algunas líneas de investigación de interés en este campo, caso de disponer de la información adecuada.

## 2. GRUPOS MAYORMENTE AFECTADOS POR EL PARO

Tal y como se indicó en la introducción, el estudio separado de los dos componentes que caracterizan el desempleo, es decir, la probabilidad de entrar en el paro y la duración de los períodos de desempleo, no puede ser llevado a cabo, dada la carencia de información acerca de la duración de los períodos de desempleo en la EPF. Lo que sí puede analizarse es la probabilidad de estar desempleado, que no es más que el efecto conjunto de los dos componentes antes mencionados. En este sentido, vamos a considerar qué características afectan la probabilidad de estar en paro, teniendo en cuenta que dicha situación puede deberse a una decisión voluntaria del individuo (en este caso, el sustentador principal) o a razones ajenas al mismo (despido, reestructuración de plantilla o razones de salud) (7).

Los que son despedidos o se ven afectados por la reestructuración de las plantillas son aquellos cuya productividad es baja comparada con el salario que reciben o esperan recibir. En otras palabras, se tratará de individuos con escaso capital humano. Así pues, sería de esperar que los mayormente afectados fueran el personal poco cualificado y los más jóvenes. De ahí el interés en tener entre los factores explicativos la edad y el nivel de instrucción, aunque este último no

recoja con exactitud el grado de cualificación de los individuos. En este sentido, hubiese sido útil disponer de variables que nos indicaran el tiempo que el individuo lleva en la empresa, si ha estado en paro anteriormente y por cuanto tiempo, o el tipo de trabajo que desarrolla actualmente o desarrollaba antes de estar en paro (8). Aun cuando puedan poseer capital humano, las personas que están hacia el final de su vida laboral también se verán afectados por el desempleo involuntario, ya que su productividad esperada es decreciente y suelen recibir salarios más elevados.

Los que pasan a formar parte del grupo de parados en forma voluntaria, mayormente, son aquellos que no tienen importantes obligaciones de tipo económico o que esperan en futuros empleos obtener ingresos mayores, o bien que los ingresos que perciben cuando están en paro les compensan suficientemente. De nuevo los jóvenes parecen formar parte importante de este grupo, así como aquellos que poseen escaso capital humano y reciben una baja remuneración, ya que el nivel de ingresos a recibir en sus nuevos empleos es probable que exceda el que actualmente tienen. De hecho, la evidencia que se posee de otros países parece ser acorde a lo anterior. Por ejemplo, García y Stern (1985) obtienen, con datos británicos, que son aquellos con escaso capital humano, y los jóvenes, los que tienen incrementos porcentuales mayores en sus ingresos tras un período de desempleo.

La composición familiar recoge, en cierto modo, en qué medida las obligaciones económicas del individuo son importantes, aunque quizá esta especificación

se hubiese podido ver completada con una variable que indicase el estado civil del individuo y otra que indicase el número (o la existencia) de perceptores de ingresos en el hogar. Finalmente, las variables que hacen referencia a la comunidad autónoma y al tamaño del municipio en que el sustentador principal reside, hacen de variables *proxy* de la situación de la demanda de trabajo en la zona. En este sentido, hubiese sido de interés disponer de información sobre el total de puestos vacantes existentes en el área de residencia. Probablemente, sea conveniente introducir las tasas de paro por comunidades, nivel de cualificación y/o distribución por edades para captar estos efectos de la demanda. Conocer el sector económico (o la tasa de paro del mismo) en el que el individuo trabaja (o trabajaba) también puede ser de utilidad a la hora de determinar los rasgos característicos del paro.

En lo referente a la duración de los períodos de desempleo, ésta depende de la probabilidad de recibir ofertas y del salario de reserva, que es aquel mediante el cual se determina si la oferta es aceptada o no. Entre los determinantes de la duración esperada, bajo un planteamiento basado en la teoría de la búsqueda (9), cabe destacar todo aquello que afecta a los costes de estar desempleado, el horizonte temporal de las decisiones del individuo, el nivel de cualificación de éste, la tasa de descuento que actualiza los ingresos futuros y posibles determinantes de las ofertas que recibe el individuo.

Finalmente, la condición de emigrante se ha incluido para tratar de determinar si existe un efecto discriminatorio con res-

pecto a los recién llegados (en los últimos años) al municipio de residencia. Hubiese sido de interés cualificar esa emigración según el lugar de procedencia, pero dicha información, aunque disponible en la EPF, no estaba contenida en la base de datos con que se realizó este trabajo.

El tipo de modelo econométrico utilizado en este trabajo está descrito en el apéndice 2. En el cuadro A.2 de dicho apéndice se presentan las estimaciones del modelo de participación para tres submuestras. En la columna [1], de los 23.972 hogares, se han excluido aquellos en los que el sustentador principal no forma parte de la población activa (véase apéndice 1) o su edad es superior a 65 años. En la columna [2] se excluyen las mujeres. Finalmente, en la columna [3] se elimina adicionalmente a los menores de 25 años. Los resultados presentados en los cuadros de esta sección corresponden a las estimaciones efectuadas con la submuestra que excluye a las mujeres y a los menores de 25 años.

La razón para excluir a las mujeres y a los menores de 25 años se basa en el hecho de que son dos grupos cuya representación entre los sustentadores principales no refleja lo que ocurre en relación al paro en estos grupos en el total de la población. De hecho, el porcentaje de mujeres en paro entre los sustentadores principales (10) es del 3,3 por 100, mientras que la tasa de paro femenina para el total de la población en el tercer trimestre de 1980 es del 13,3 por 100. Por lo que se refiere a los varones menores de 25 años, el porcentaje de parados en la muestra es del 11 por 100, mientras que la tasa de paro en el total de la población es del 46 por 100. Con

CUADRO N.º 2

**ANÁLISIS DE LA INCIDENCIA DEL PARO POR GRUPOS DE EDAD**

	25-34	35-44	45-54	55-64
[1] ... ..	6,6	5,8	6,7	6,2
[2] ... ..	9,0	5,8	6,7	7,1
[3] ... ..	136,4	100,0	100,0	114,5

*Notas:* En la fila [1] se indica el porcentaje de personas en paro para cada uno de los grupos considerados en la muestra de la EPF utilizada. La probabilidad de estar en paro para un individuo tipo está indicada en la fila [2] en términos porcentuales. El individuo tipo corresponde a un sustentador principal con estudios primarios, residente en Andalucía, en un municipio con un número de habitantes comprendido entre 50.000 y 500.000, no emigrante, entrevistado en primavera y casado, con dos hijos de edades entre 9 y 13 años. El modelo utilizado para el cálculo de estas probabilidades es el correspondiente a la columna (3) del cuadro A.2 del apéndice 2. En la fila [3] se detalla el cociente entre la fila [2] y la fila [1] en términos porcentuales.

respecto a los resultados de la estimación del modelo, éstos no difieren prácticamente entre las distintas submuestras tal y como queda de manifiesto si se comparan las estimaciones que se presentan en las tres columnas del cuadro A.2 del apéndice 2.

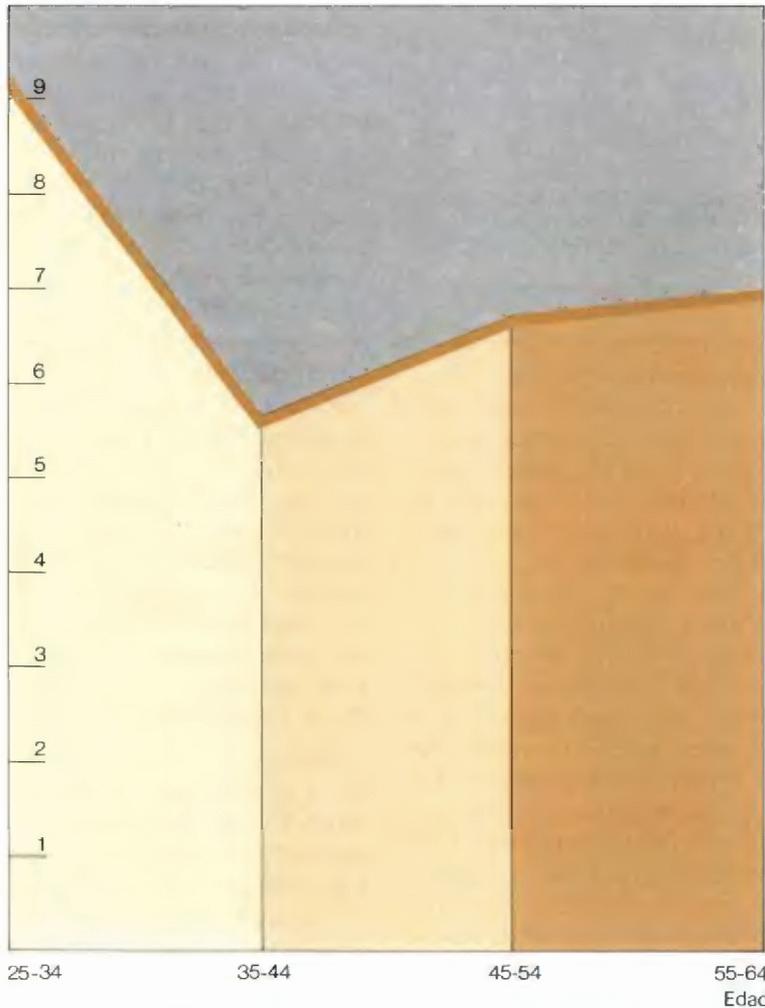
Tras estas consideraciones previas acerca de quienes se esperaba estuviesen mayormente afectados por el paro y sobre los criterios de elección de la muestra, pasaremos a discutir los resultados del análisis empírico efectuado. La descripción detallada de las variables utilizadas en el análisis se encuentra en el apéndice 1. En el apéndice 2 se presenta el modelo econométrico utilizado, así como los resultados de la estimación del mismo para distintas submuestras. Aunque la muestra de que se dispone no es la óptima para extraer conclusiones sobre los rasgos característicos del paro en España, sí que nos permitirá tener una idea aproximada de los mismos, en particular en lo que hace referencia al desempleo masculino no juvenil.

### Edad

En el cuadro n.º 2 se presentan los resultados correspondientes a las tasas de paro por grupos de edad. Los resultados de la primera fila en este cuadro (y en los que siguen en esta sección) corresponden a los porcentajes de paro en la muestra utilizada para cada uno de los grupos considerados, en este caso grupos de edad. En la segunda fila se presentan las probabilidades estimadas (11) de estar en paro para un sustentador principal con unas características determinadas. El sustentador tipo, elegido para éste y posteriores cálculos, corresponde a un individuo con las características mayoritarias para cada uno de los rasgos considerados (residente en un municipio de 50.000 a 500.000 habitantes, en Andalucía, no emigrante y con estudios primarios) con la excepción de la edad, ya que se supone que el sustentador tipo tiene entre 35 y 44 años, mientras que el mayor porcentaje corresponde al grupo cuyas edades están entre 45 y 54 años. Tal selección tra-

**GRAFICO 1  
REPRESENTACION GRAFICA DE LA INCIDENCIA  
DEL PARO POR GRUPOS DE EDAD**

Probabilidad estimada (%)



ta de hacer más acorde la edad al número de hijos de la familia tipo, que se supone es una en la que hay dos hijos con edades comprendidas entre 9 y 13 años, con lo que el análisis del efecto de la edad no es irrealista en lo que hace referencia a la edad del sustentador principal y la de sus hijos. Obviamente, la edad del sustentador tipo es lo que varía en los resultados presentados en el cuadro n.º 2.

Como cabía esperar, la distribución de las probabilidades estimadas de estar en paro tiene forma de U, como queda reflejado en el gráfico 1. Es decir, la probabilidad de estar desempleado es mayor para los más jóvenes y los mayores cuando se comparan individuos con idénticas características, pero con edades distintas.

La tercera fila de los cuadros

corresponde al cociente (en términos porcentuales) de la segunda fila por la primera, siendo un indicador de las discrepancias entre las probabilidades estimadas y los porcentajes observados. Aunque las probabilidades estimadas dependen de las características del sustentador principal considerado, el *ratio* que representa la tercera fila debería ser homogéneo para todos los grupos si el análisis basado en el modelo de elección discreta no tuviese nada que añadir al simple cálculo de los porcentajes por grupos de edad. Caso de no ser homogéneo el *ratio*, existe evidencia de que los porcentajes por grupos de edad (o cualquier otro tipo de clasificación) están recogiendo aspectos que son propios de otros rasgos característicos del individuo. En el caso de la división por grupos de edad, el *ratio* es homogéneo con la excepción que hace referencia al grupo de sustentadores principales con edades comprendidas entre 25 y 34 años y, en menor grado, al grupo con edades comprendidas entre 55 y 64 años. La probabilidad estimada es notoriamente superior al porcentaje observado para el grupo de 25 a 34 años, por lo que el porcentaje observado para este grupo puede estar recogiendo los efectos de otras variables que reducen la probabilidad de estar en paro, por ejemplo, el nivel de instrucción. Los niveles de instrucción de los sustentadores principales con edades de 25 a 34 años deben ser mayores que los del resto de la muestra, por lo que el porcentaje observado para dicho grupo de edad está recogiendo parte del efecto debido a poseer niveles de instrucción más elevados. Este caso ilustra la discusión realizada en la sección introductoria acerca de las ven-

CUADRO N.º 3

ANÁLISIS DE LA INCIDENCIA DEL PARO POR NIVELES DE INSTRUCCIÓN

	Analfabetos	Sin estudios	Primarios	Secundarios	Superiores
[1] .....	18,2	11,3	5,2	2,7	1,5
[2] .....	18,3	11,7	6,1	2,6	1,3
[3] .....	100,5	103,5	117,3	96,3	86,7

Notas: Las mismas que en el cuadro n.º 2, con la diferencia de que el sustentador principal tipo no tiene fijado el nivel de instrucción, pero su edad está comprendida entre 35 y 44 años.

tajas de utilizar modelos de elección discreta sobre el simple uso de los porcentajes de paro observados, a la hora de determinar los efectos de distintas características sobre la probabilidad de estar en paro (o empleado).

Además de las consideraciones anteriores, debe mencionarse que la única tasa de paro significativamente distinta a la del grupo de 35 a 44 años es la correspondiente al grupo de 25 a 34 años.

### Nivel de instrucción

Los resultados correspondientes a los porcentajes y probabilidades de estar en paro según el nivel de instrucción se presentan en el cuadro n.º 3. En cuanto a los resultados de la estimación, los cuatro parámetros de las cuatro variables ficticias correspondientes al nivel de instrucción son muy significativos, con estadísticos 't' superiores a 6 y con los signos y magnitudes esperados. Cuanto mayor es el nivel de instrucción menor es la probabilidad de estar en paro, hecho que queda reflejado en el gráfico 2.

La relación entre los porcentajes observados y las probabi-

lidades estimadas es homogénea para los analfabetos y sin estudios. La probabilidad estimada sobreestima el porcentaje observado en el caso de poseer estudios primarios, mientras que lo subestima en el caso de poseer estudios medios o superiores. Ello puede ser debido a que los porcentajes observados están recogiendo el efecto de otras variables, por ejemplo el tamaño municipal. Así, uno esperaría que en los municipios con menos de 2.000 habitantes la población tuviera en su mayoría estudios primarios, por lo que el porcentaje observado puede estar recogiendo el efecto negativo sobre la probabilidad de estar en paro de vivir en un municipio con menos de 2.000 personas, en donde la mayoría de sus habitantes están dedicados a las labores del campo y en donde la actividad económica tiene un carácter más familiar, con lo que es más probable que, a la hora de indicar cuál es su situación laboral, declaren estar trabajando. De hecho, el 62,9 por 100 de los no parados en poblaciones con menos de 2.000 habitantes tiene estudios primarios, frente a un 42,4 por 100 de toda la muestra de no parados. Asimismo, el 54,5 por 100 de los que trabajan en estos municipios pequeños se

dedican a tareas agrícolas, frente a un 17,3 por 100 de toda la muestra de los ocupados. De forma similar, la gente con estudios secundarios y superiores reside mayormente en poblaciones con un elevado número de habitantes, por ejemplo en núcleos con más de medio millón de habitantes. El hecho de que la probabilidad estimada sea inferior al porcentaje observado para estos niveles de instrucción puede deberse a que en el porcentaje observado se recoge el efecto negativo sobre la probabilidad de estar empleado de vivir en núcleos urbanos muy poblados y en los que, consecuentemente, el paro en el sector industrial tiene una incidencia notoria. De hecho, en los centros urbanos con más de medio millón de habitantes, el 41,6 por 100 de los que trabajan tienen estudios secundarios o superiores, mientras que ese porcentaje se reduce al 4,2 por 100 cuando se considera toda la muestra.

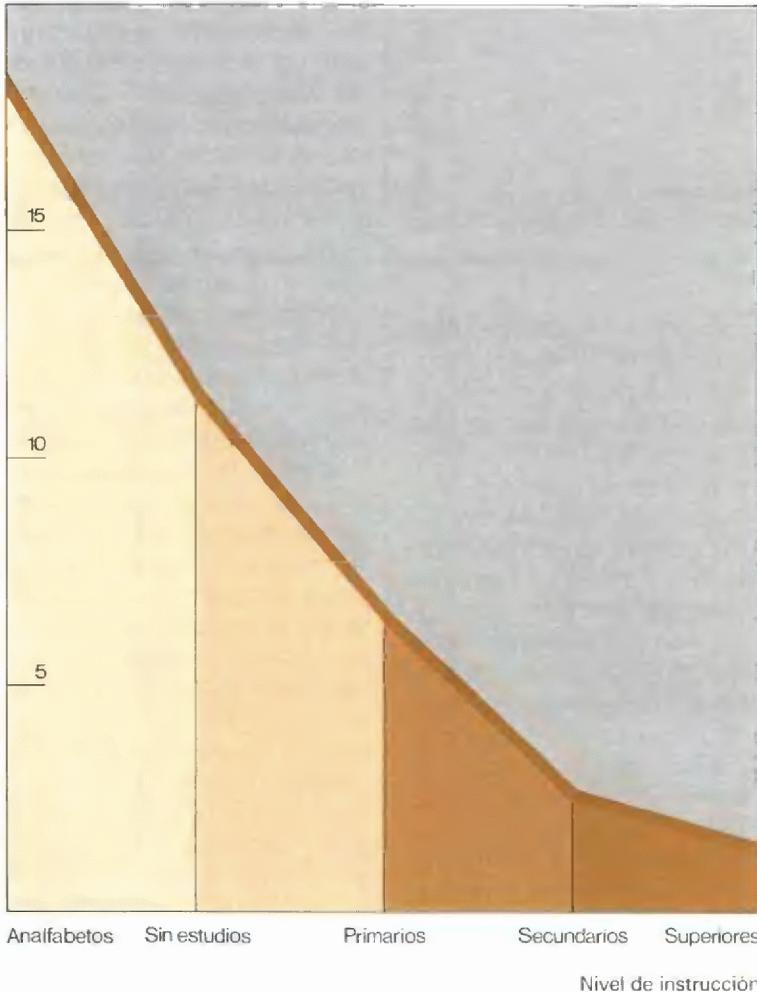
Nótese que las variaciones en las probabilidades estimadas son mucho más notorias por niveles de instrucción que por grupos de edad. Para individuos con idénticas características, la probabilidad de estar en paro es 0,17 puntos mayor para un analfabeto que para alguien con estudios superiores.

### Condición de emigrante

El fenómeno migratorio interno en un país puede responder a las diferencias en los niveles de actividad y los niveles de vida entre las distintas regiones. Una cuestión de interés es conocer si los emigrantes, particularmente los recién emigrados, sufren una situación discriminatoria en lo que hace referencia

**GRAFICO 2  
REPRESENTACION GRAFICA DE LA INCIDENCIA  
DEL PARO POR NIVELES DE INSTRUCCION**

Probabilidad estimada (%)



probabilidades estimadas, según se trate de un emigrante o no, para un sustentador principal con determinadas características.

Es de destacar que la probabilidad estimada de estar en paro para los emigrantes es casi un 29 por 100 mayor que el porcentaje observado, mientras que para los no emigrantes son parecidas. Ello es indicativo de que el porcentaje observado está recogiendo el efecto de rasgos que, siendo predominantes entre los emigrantes comparados con el total de la muestra, tienen un efecto positivo sobre la probabilidad de estar empleado. De hecho, el que el concepto de emigrante se refiera a emigración en los últimos cinco años hace que sea de esperar que la mayor parte de ellos no sean personas que estén al final de su vida laboral (12). Asimismo, el nivel de instrucción de los emigrantes es superior a la media (13). El hecho de ser en su mayoría menores de 35 años tendería a reducir la probabilidad de estar empleado, mientras que el mayor nivel de instrucción actúa en el sentido contrario, aunque el efecto del nivel de instrucción es mayor, por lo que el porcentaje observado de paro entre los emigrantes es menor que el obtenido una vez que se controlan los efectos de otros factores.

a su situación laboral. Esta discriminación o bien puede manifestarse en una distinta incidencia sobre la probabilidad de estar empleado de las variables consideradas en el análisis, o bien puede haber una discriminación que actúe a través de otros factores no considerados en el análisis, lo cual queda reflejado por diferencias en el término constante.

Observando los resultados del cuadro A.2 del apéndice 2, se deduce que, independientemente de la submuestra que se utilice, el hecho de ser emigrante tiene un efecto negativo y significativo sobre la probabilidad de estar empleado de parecida magnitud en las tres submuestras. En el cuadro n.º 4 se presenta el análisis comparativo entre los porcentajes observados y las

En la columna [1] del cuadro A.3 del apéndice 2 se presentan los resultados de estimar el modelo que explica la situación laboral del sustentador principal para la submuestra de no emigrantes. Comparando estas estimaciones con las del total de la muestra (columna [3] del cuadro A.2), se observa que las diferencias son mínimas. Por ello, la posible discriminación, reflejada por el hecho de que el pa-

CUADRO N.º 4

ANÁLISIS DE LA INCIDENCIA DEL PARO SEGUN LA CONDICION DE EMIGRANTE

	Emigrante	No emigrante
[1] .....	6,6	6,3
[2] .....	8,5	6,1
[3] .....	128,8	96,8

Notas: Las mismas que en el cuadro n.º 2, con la diferencia de que el sustentador principal puede ser o no ser emigrante, pero su edad está comprendida entre 35 y 44 años.

rámetro de la variable que representa la condición de emigrante sea significativo y negativo, no actúa a través de los factores incluidos en el modelo sino a través de otros mecanismos no considerados en el análisis. En este sentido, hubiese sido de interés utilizar la información referente al lugar de procedencia, información contenida en la EPF pero no en la base de datos utilizada en este trabajo.

**Comunidad autónoma de residencia**

Las variables ficticias correspondientes a las comunidades autónomas (14) están recogiendo efectos de distinto signo relacionados con el lugar de residencia. Entre estos efectos podemos incluir las condiciones de la demanda de trabajo (puestos vacantes, tasas de paro) o el nivel de industrialización de la comunidad autónoma a la que la variable ficticia está referida.

De la simple observación de los resultados de la estimación del modelo presentados en cualquiera de las columnas del cuadro A.2 del apéndice 2, se desprende que, tomando como dadas las restantes características,

la probabilidad de estar empleado es menor en Andalucía (comunidad autónoma de referencia) que en las restantes, ya que los signos de las variables ficticias regionales son todos positivos. Pese a ello, la diferencia no es significativa (15) en relación a Canarias, Cataluña, Extremadura, Madrid, Navarra, La Rioja y País Vasco. De hecho, todas ellas, excepto La Rioja y Navarra, tienen porcentajes de paro en la muestra utilizada superiores al 5 por 100, mientras que en Navarra el porcentaje es del 3,7 por 100 y en La Rioja del 3,3 por 100. Esto último puede deberse a que el porcentaje de paro observado esté recogiendo el efecto positivo de otros factores sobre el empleo, como puede ser el que los municipios sean de tamaño pequeño, con lo que la población estará empleada en su mayoría, debido a que está dedicada a tareas agrícolas o a negocios de tipo familiar.

El mismo modelo de elección discreta se ha estimado para dos submuestras de sustentadores principales: aquellos que residen en Andalucía y Cataluña. Los resultados de las estimaciones se presentan en las columnas [2] y [3] del cuadro A.3 del apéndice 2. Comparando dichos re-

sultados con los correspondientes a toda la muestra (columna [3] del cuadro A.2), se observan algunos rasgos distintivos. Así, por ejemplo, en Andalucía el efecto positivo del tamaño municipal inferior a 2.000 habitantes desaparece, e incluso, aunque no sea significativo, adopta un valor negativo. Ello puede ser consecuencia de que el paro en el sector agrícola está mucho más extendido entre los jornaleros de Andalucía. De hecho, el 82,2 por 100 de los trabajadores en el sector agrícola están empleados, frente a un 88,9 por 100 para el total de la muestra. Asimismo, la condición de emigrante sigue siendo relevante a la hora de determinar la probabilidad de estar empleado en Andalucía, aunque ello puede estar recogiendo más el efecto de la emigración interna dentro de cada comunidad autónoma que el de la emigración proveniente de otras comunidades. Por lo que hace referencia a Cataluña, el hecho de ser analfabeto no parece tener ningún efecto significativo sobre la situación laboral del sustentador principal. Ello puede ser debido al escaso porcentaje de analfabetos entre los sustentadores principales (1,6 por 100) en Cataluña (16), en relación al porcentaje de todo el país (3,3 por 100) según la muestra que se utiliza.

En el cuadro n.º 5 se presentan las probabilidades estimadas de estar desempleado para Andalucía y Cataluña según el nivel de instrucción (17). En general, las probabilidades obtenidas con las estimaciones correspondientes a las submuestras de cada comunidad son algo inferiores a las obtenidas con las estimaciones del modelo para toda la muestra, en donde las diferencias entre las distintas co-

comunidades autónomas quedan únicamente reflejadas en un distinto valor del término constante, con la excepción, en Cataluña, para aquellos que no tienen estudios. También en lo referente a Cataluña, la discrepancia para los analfabetos en las probabilidades estimadas para cada modelo es notoria, siendo consecuencia lógica de la no significación del parámetro perteneciente a este grupo cuando se utiliza la submuestra de residentes en Cataluña. Por ello, se ha optado por no consignar esta probabilidad estimada en el cuadro n.º 5.

Debe destacarse que las probabilidades estimadas para Cataluña son siempre inferiores a las correspondientes a Andalucía, lo cual está en consonancia con los porcentajes de paro observados para nuestra muestra (10,6 por 100 en Andalucía y 5,7 por 100 en Cataluña).

El último extremo a considerar es si las diferencias observadas al analizar los resultados de las dos submuestras regionales, en relación a los de toda la

muestra, están reflejando que el efecto de los distintos rasgos considerados es diferente para cada comunidad, o si simplemente se trata de diferencias que responden a aspectos no incluidos en el modelo. Con un nivel de significación del 5 por 100 se puede afirmar que, en el caso de Cataluña, las diferencias en los efectos de los distintos rasgos característicos sí que son significativas, no pudiéndose decir lo mismo en el caso de Andalucía (18). En este sentido, dado el reducido tamaño muestral para algunas de las comunidades, no es posible realizar el mismo tipo de análisis que el efectuado en este apartado. Por ello, sería conveniente poder utilizar la información de la EPA para estudiar estas cuestiones.

### Otros aspectos

Los dos últimos rasgos que consideramos son el tamaño del municipio de residencia y la composición familiar. En relación al primero, la residencia en municipios con menos de 2.000 ha-

bitantes o en grandes núcleos urbanos (más de medio millón de habitantes) parece tener incidencia sobre la probabilidad de estar empleado, aunque con signo contrario. Así, la residencia en municipios pequeños tiene un efecto positivo, probablemente como consecuencia de que un gran número de personas están empleadas en labores agrícolas o bien el tipo de actividades económicas es, en gran parte, de carácter familiar. Por otro lado, el residir en una de las grandes ciudades está asociado a una mayor probabilidad de estar desempleado, posiblemente como consecuencia de la mayor incidencia del paro industrial en estos núcleos urbanos.

En lo que hace referencia al efecto de la composición familiar, no parecen existir razones teóricas muy claras por las que su incidencia deba tener un signo u otro. Por ello, en el presente análisis dichas variables se han introducido a modo de variables de control, sin que sea intención del trabajo identificar la relación entre la situación laboral y la composición del hogar. Pese a ello, debe destacarse el efecto significativo y negativo de la presencia de hijos entre 5 y 13 años, así como el efecto positivo, aunque no significativo, de los hijos menores de 4 años, factor este último que sí es un determinante importante de la participación de la mujer en el mercado de trabajo. Una posible explicación para el efecto negativo sobre la probabilidad de estar empleado de la presencia de hijos de mediana edad está basada en la teoría de la búsqueda. Si se reciben prestaciones cuya cuantía depende del número de hijos, la situación de un desempleado podría ser mejor que si no los tuviese, por lo que su

CUADRO N.º 5

#### ANÁLISIS DE LA INCIDENCIA DEL PARO EN ANDALUCÍA Y CATALUÑA SEGUN EL NIVEL DE INSTRUCCIÓN

	Analfabetos	Sin estudios	Primarios	Secundarios	Superiores
Andalucía (1) ...	15,2	9,5	4,7	2,0	1,0
Andalucía (2) ...	14,7	8,5	3,9	1,0	1,0
Cataluña (1) ...	12,5	7,6	3,7	1,5	0,7
Cataluña (2) ...	—	8,1	3,2	0,9	0,2

Notas: El sustentador principal típico está definido como en el cuadro n.º 2, con la diferencia de que la edad del sustentador principal está comprendida entre 35 y 44 años y la familia no tiene hijos. El modelo utilizado difiere según los casos.

(1) Estas probabilidades corresponden al modelo estimado para la totalidad de la muestra, suponiendo que el sustentador principal reside en Andalucía o Cataluña según el caso.

(2) Estas probabilidades corresponden a los modelos estimados con las submuestras de las comunidades autónomas correspondientes.

salario de reserva sería más alto. De hecho, este tipo de resultados también se obtuvieron para Gran Bretaña, y Nickell (1980) apunta esta explicación entre las posibles razones (19). Pese a ello, este tipo de explicación no puede utilizarse en el caso español, ya que no existen prestaciones que dependan del número de hijos.

Por otro lado, es de destacar el efecto positivo de la presencia de adultos mayores de 25 años en el hogar. De hecho, estas dos variables a las que hacemos referencia (25-60, >60), están recogiendo, entre otros, el efecto de estar o no estar casado, con lo que su efecto positivo sobre la probabilidad de estar empleado es una indicación de que el estar casado está asociado a una mayor probabilidad de estar empleado. Esto último es acorde con lo apuntado en la introducción sobre la relación positiva entre las obligaciones económicas del individuo y el estar casado, en lo que hace referencia al desempleo voluntario. Lancaster y Nickell (1980) apuntan otra razón para el efecto positivo sobre la probabilidad de estar empleado de la circunstancia de estar casado, en el sentido de que para el empresario es una cierta garantía de la seriedad del individuo.

### 3. POLITICAS PARA COMPENSAR LA DESIGUAL DISTRIBUCION DEL PARO

En principio, la desigual distribución del paro no implica que estar en paro durante más tiempo o más frecuentemente signifique estar peor. Los costes ori-

ginados por el desempleo pueden quedar compensados por los salarios recibidos cuando se está empleado o por el mayor ocio del que se disfruta o por las prestaciones percibidas en concepto de seguro de desempleo cuando se está en paro. De todas formas, la creciente dificultad para obtener un empleo, así como el hecho de que la duración de los periodos de desempleo parece crecer y exceder, para un elevado porcentaje de la población afectada, al tiempo durante el que se recibe el seguro de desempleo, hace que aquellos que se hallan en paro, por lo general, estén peor que los que no pertenecen a tal categoría. Estos aspectos relacionados con la duración de los periodos de desempleo son particularmente relevantes en el caso de España, tal y como queda reflejado en el cuadro n.º 6.

Del anterior cuadro se desprende que en la actualidad más del 50 por 100 de los parados lo están por un periodo superior al año. La larga duración de los periodos de desempleo tiende a reducir el capital humano específico del individuo, con lo que, en principio, reduciría la probabilidad de encontrar un nuevo empleo. Este aspecto es particularmente grave para los jóvenes,

pues a ello hay que añadir que no poseen mucha experiencia laboral.

A fin de hacer más equitativa la situación de los parados, y en particular la de los que llevan largo tiempo en esta situación, con el resto de la población activa, dos tipos genéricos de políticas se pueden arbitrar: aquellas que persiguen el cambio en la situación laboral del individuo, es decir, que obtenga un empleo, y las que persiguen mejorar su situación en el paro. De entre las segundas, la más inmediata es incrementar las prestaciones recibidas del seguro de desempleo a aquellos que llevan más de un determinado periodo de tiempo desempleados. Esta política, por sí sola, tiene el inconveniente de, por una parte, poder reducir los esfuerzos del parado para buscar empleo y, por otro lado, tender a subir su salario de reserva. Por ello, este tipo de política debería ir acompañada de alguna medida para tratar de facilitar la obtención de puestos de trabajo para los parados, lo que configuraba el primer grupo de políticas antes mencionado. De entre estas políticas, merecen un breve comentario las siguientes: fijar cuotas de contratación de parados a las empresas, dar subsidios a las empresas para la

CUADRO N.º 6

#### DISTRIBUCION DE LOS VARONES QUE ESTAN EN PARO SEGUN EL TIEMPO DE BUSQUEDA DE EMPLEO

	< 1 m.	1-3 m.	3-6 m.	6 m.-1 a.	1-2 a.	> 2 a.
1980-III ... ..	3,8	17,6	18,9	24,6	23,9	10,9
1984-III ... ..	2,2	11,7	12,2	20,3	22,3	30,5

Notas: Las abreviaturas m. y a. corresponden respectivamente a meses y años. El 0,9 por 100 de los parados en el tercer trimestre de 1980 y el 0,7 por 100 en el tercer trimestre de 1984 están considerados como no clasificables en función del tiempo de búsqueda de empleo.

Fuente: Encuesta de Población Activa.

contratación de parados o la reducción de la contribución a la Seguridad Social de las empresas por cada trabajador, en base a distintos criterios.

La fijación de cuotas de contratación de parados cuando se contrata nuevo personal es una política que incide sobre los beneficios de las empresas, puesto que los desempleados son, por lo general, menos productivos, por lo que reducirían dichos beneficios. Esto podría reducir la demanda de trabajo, aunque, en cualquier caso, la política de fijación de cuotas no incrementa el empleo. Esta política estaría fundamentalmente pensada para aquellos que han estado parados durante largo tiempo.

Los subsidios por la contratación de desempleados representan para el Estado un coste a la hora de generar empleo, quedando abierta la cuestión de si otro uso de esos fondos hubiera tenido un mayor impacto sobre la demanda de trabajo. Otra cuestión es la cuantía del subsidio, ya que éste debe ser lo suficientemente atractivo para que las empresas lo soliciten, pero sin que sea tan cuantioso que las empresas puedan hacer un mal uso de él, en el sentido de contratar nuevo personal que sustituya al ya existente. Jackman y Layard (1985) demuestran cómo una política de subsidios financiada por un impuesto sobre el total de salarios pagados por la empresa puede reducir, bajo ciertas hipótesis, la tasa natural de paro. Esto sería debido a dos razones: por un lado, afectaría al proceso de fijación de salarios y, por otro, desplazaría la demanda de empleo hacia aquellos sectores con tasas de paro más elevadas.

La exención de pago de con-

tribuciones a la Seguridad Social por parte de las empresas por trabajadores no cualificados, junto con una política de garantizar el puesto de trabajo para los mismos, desplazaría la demanda de trabajo hacia aquellos sectores en los que existe un exceso de oferta. Layard (1985) propone este tipo de exenciones para aquellas empresas que contraten parados. El período de exención debería ser fijado de acuerdo a como se financia este tipo de política. La financiación podría efectuarse mediante una reestructuración en el sistema de contribuciones a la Seguridad Social por parte de las empresas.

Por lo general, este tipo de políticas no son efectivas por sí solas, debiéndose considerar la aplicación simultánea de algunas de ellas conjuntamente con políticas específicas según los rasgos característicos de los afectados. Las políticas de reciclaje y de jubilación anticipada son ejemplos de este tipo de políticas complementarias.

#### 4. CONCLUSIONES

En el presente trabajo se ha pretendido caracterizar los rasgos más relevantes del paro masculino en España, haciendo uso de la información contenida en la EPF 1980-81. Debido al tipo de datos disponibles, hemos debido centrar el análisis en la muestra de sustentadores principales de los hogares que fueron entrevistados. Aunque los sustentadores principales no son representativos de la totalidad de la población activa, una vez eliminados aquellos grupos peor representados por dicha muestra, se pueden extraer algunas conclusiones sobre quienes son

los mayormente afectados por el paro en parte de la población activa española. Es por ello por lo que las mujeres y los menores de 25 años, que son grupos mal representados en la muestra de sustentadores principales, dada la definición de este concepto, no han sido considerados en la mayoría de los análisis efectuados. El paro juvenil y la participación de la mujer en el mercado de trabajo son temas que requieren un tratamiento especial. Las mujeres y los varones menores de 25 años representaban el 41,6 por 100 de la población activa en el tercer trimestre de 1980.

Son aquellos que se encuentran al final de su vida laboral, con un bajo nivel de instrucción, que viven en los grandes núcleos urbanos, que son recién llegados a la zona en donde residen, con hijos, y que pertenecen a determinadas comunidades autónomas, los que mayormente sufren la incidencia del paro y a quienes han de ir dirigidas las políticas tendentes a hacer más equitativa la situación entre parados y no parados, y, en particular, aquellas que faciliten la obtención de nuevos empleos a los parados.

Los resultados obtenidos son los que corresponden a un primer análisis de los rasgos característicos del paro en España, aunque a lo largo de la exposición se han destacado aspectos que, con la información disponible para este trabajo, deberían ser objeto de un análisis más detallado (20). Asimismo, se han mencionado otros cuyo estudio podría ser abordado, caso de disponer de información proveniente de otras fuentes. De entre los primeros cabría destacar: un análisis más detallado de la posible discriminación de los emigrantes, incorporando la información

relativa al lugar de donde provienen; una caracterización más detallada del efecto de la composición familiar sobre la probabilidad de estar empleado, utilizando información sobre el número de perceptores de ingresos; o la descomposición de los efectos «regionales» en diferentes componentes, así como la incorporación de efectos «sectoriales» en el modelo. Por último, el análisis aquí presentado corresponde a un modelo de elección binaria (estar empleado o en paro), pero la información disponible permite caracterizar la situación laboral como el resultado de una elección con tres alternativas: estar en paro o estar empleado, bien menos de un tercio de la jornada normal o bien más de un tercio. Por lo tanto, sería interesante estudiar en forma más detallada los rasgos característicos de aquellos que se encuentran en estas situaciones (21). Asimismo, es de interés el análisis de los rasgos del paro según se trate de paro agrícola o no agrícola. De hecho, podría esperarse que el efecto de los distintos aspectos considerados fuese también distinto para cada uno de los sectores.

Caso de que se dispusiese de información proveniente de otras fuentes, se podría abordar la descomposición de la tasa de paro en sus dos componentes: duración y probabilidad de entrar a formar parte del grupo de parados. Un análisis inicial podría efectuarse haciendo uso de la información contenida en la EPA. Esto también permitiría analizar la evolución temporal del efecto de los rasgos característicos, evolución que no tiene por qué ser la misma.

Finalmente, querríamos concluir el presente trabajo con la

petición de que se incluyan en las nuevas encuestas efectuadas por el INE ciertas preguntas ya reseñadas y que permitirían efectuar análisis más completos del comportamiento del individuo, con lo que sería posible llevar a cabo estudios de simulación del impacto de distintas políticas, fundamentalmente fiscales, en el bienestar individual y global.

## NOTAS

(\*) Agradecemos a J. Ruiz-Castillo su ayuda para la utilización de la base de datos de la Encuesta de Presupuestos Familiares empleada en este estudio. Asimismo, agradecemos la colaboración como ayudante de investigación de Y. Sorolla en las primeras etapas de este trabajo.

(1) La información publicada en los boletines trimestrales de la Encuesta de Población Activa no permite saber el número de nuevos desempleados y personas que han abandonado el desempleo.

(2) La tasa de paro anual ( $u$ ) puede definirse como:

$$u = \frac{D}{52} \times \frac{S}{L} \times 100$$

en donde  $D$  es el tiempo medio que está un individuo en paro durante ese año,  $S$  es el número en individuos que durante ese año, en uno u otro momento, han estado en paro y  $L$  es el total de la población activa. El número de años que tardaría un individuo en estar desempleado vendría dado por  $(L/S)$ .

(3) Nótese que el estudio sólo dispondría de información referida a períodos incompletos de desempleo, al igual que en Nickell (1980), por lo que sería mucho más relevante para este tipo de análisis disponer de los datos de una encuesta en que se conociese la situación laboral de los individuos no sólo en un momento en el tiempo, sino a lo largo de un período determinado. Un ejemplo de este tipo de encuestas es el «Cohort Study of the Unemployed (Autumn, 1978)», realizado por el Department of Health and Social Security de Gran Bretaña.

(4) En la EPF se entiende por sustentador principal: «aquél miembro del hogar cuya aportación periódica —no ocasional— al presupuesto común sirva actualmente para atender a la mayor parte de los gastos del mismo». (INE [1983], pág. 12.)

(5) El concepto de persona principal en la EPA es más parecido al concepto tradicional de «cabeza de familia».

(6) Si, como es de esperar, en un futuro próximo se dispone de la información por hogares de la EPF 1973-74, así como de las nuevas versiones de la EPF, será posible hacer un estudio comparativo de los cambios estructurales en los rasgos característicos del paro entre los sustentadores principales. Este estudio podría llevarse a cabo para la

totalidad de la población caso de disponer de la información individual contenida en la EPA.

(7) En las publicaciones trimestrales de la EPA no existe información alguna sobre las causas por las que el individuo está en paro.

(8) El tipo de ocupación es una información que está contenida en la EPF, pero de la que no se disponía en el banco de datos utilizado para el presente trabajo.

(9) En Mortensen (1970) y Lippman y McCall (1976a y 1976b) se encuentran excelentes exposiciones sobre la teoría de la búsqueda.

(10) El porcentaje de mujeres en paro entre los sustentadores principales en nuestra muestra (3,3 por 100) es inferior al porcentaje de varones (6,4 por 100), lo cual se refleja en el signo positivo de la variable que indica el sexo del sustentador principal. Así, las probabilidades de estar empleados para un hombre y una mujer con edades entre 35 y 44 años, y con características idénticas a las del sustentador tipo definido en la nota del cuadro n.º 2, serían de 0,938 y 0,971, respectivamente.

(11) La probabilidad estimada de estar en paro es función de los parámetros estimados del modelo del apéndice 2 que corresponda. Por lo tanto, pueden definirse las desviaciones *standard* de dichas probabilidades estimadas a partir de la matriz de covarianzas de los parámetros estimados del modelo que corresponda. A fin de simplificar la exposición, se ha preferido omitir dichas desviaciones *standard* de los cuadros de la sección 2.

(12) En la muestra 3 (varones mayores de 25 años) el 52,4 por 100 de los que son emigrantes tienen edades de 25 a 34 años, el 28 por 100 de 35 a 44, el 15,1 por 100 de 45 a 54 y el 4,5 de 55 a 64 años. Por otra parte, el 19,4 por 100 de los no emigrantes tienen edades de 25 a 34 años, el 28,5 por 100 de 35 a 44, el 32,2 por 100 de 45 a 54 y el 19,9 por 100 de 55 a 64 años.

(13) Entre los emigrantes, el 1 por 100 son analfabetos, el 11,1 por 100 no poseen estudios, el 40,8 por 100 poseen estudios primarios, el 26,2 por 100 poseen estudios secundarios y el 20,9 por 100, estudios superiores. Entre los no emigrantes el 3,4 por 100 son analfabetos, el 22,9 por 100 no poseen estudios, el 51,7 por 100 tienen estudios primarios, el 14,8 por 100, estudios secundarios y el 7,3 por 100, estudios superiores.

(14) La distribución de la muestra por comunidades autónomas no es muy acorde con lo que ocurre a nivel de toda la población. Los porcentajes son los siguientes: Andalucía (18,3 por 100 de la muestra y 17,2 por 100 de la población), Aragón (5,5

y 3,2), Asturias (2,8 y 3,0), Baleares (1,8 y 1,7), Canarias (3,7 y 3,6), Cantabria (2,2 y 1,4), Castilla-León (13,3 y 6,9), Cataluña (10,8 y 16,1), Extremadura (3,9 y 2,8), Galicia (6,5 y 7,5), Madrid (5,5 y 12,4), Castilla-La Mancha (7,6 y 4,4), Murcia (1,9 y 2,5), Navarra (1,6 y 1,3), La Rioja (1,6 y 0,6), Valencia (7,5 y 9,7) y País Vasco (5,5 y 5,7). Ello aconsejaría utilizar el análisis *probit* con las variables ponderadas, haciendo uso de las ponderaciones que se incluyen en la EPF.

(15) Los estadísticos «t» de dichos coeficientes son inferiores a 1,96, que es el valor correspondiente al elevado número de grados de libertad de que disponemos y a un nivel de significación del 5 por 100.

(16) En realidad, son sólo 26 entre los 1.656 sustentadores principales residentes en Cataluña, lo cual es un número pequeño para poder identificar efectos significativos sobre la situación laboral del hecho de ser analfabeto.

(17) Los resultados del cuadro n.º 5 hacen referencia a una familia sin hijos, ya que los parámetros de la composición familiar (-0,20 y -0,18, respectivamente) son grandes en valor absoluto y, en particular, el que hace referencia al número de adultos no es significativo, por lo que las probabilidades calculadas haciendo uso de este parámetro tendrían una desviación típica elevada.

(18) Estas conclusiones están basadas en un test de razón de verosimilitudes. En el caso de Cataluña, dicho estadístico toma el valor 37,47 y en el de Andalucía 27,93. El valor tabulado de una  $\chi^2$  con 22 grados de libertad, para un nivel de confianza del 95 por 100, es 33,92.

(19) En Nickell (1980) se obtienen resultados similares, y se apunta como posible explicación el que el efecto positivo del número de hijos sobre la probabilidad de estar desempleado se deba a que existe una relación positiva entre fertilidad y paro, ya que los «débiles» tienen un menor control sobre la fertilidad y, a su vez, los débiles tienen un mayor temor a perder su empleo, lo cual puede estar relacionado con su experiencia como parado. Véase Nickell (1980), páginas 786-787, para una discusión más detallada de estos aspectos.

(20) El equipo de investigadores que firma el presente trabajo está abordando en la actualidad los aspectos mencionados en este apartado de conclusiones, cuya realización puede ser llevada a cabo con la información del banco de datos proveniente de la EPF 1980-81 a nuestra disposición.

(21) Este análisis podría efectuarse como si se tratase de una «elección» con tres alternativas o bien una «elección» secuencial: estar empleado o en paro, y si está empleado, a tiempo completo o a tiempo parcial.

## REFERENCIAS BIBLIOGRAFICAS

- AMEMIYA, T. (1981), «Qualitative Response Models: A Survey», *Journal of Economic Literature*, 19, págs. 1483-1536.
- GARCIA, J., y STERN, J. (1985), «Real Earnings Gains and Losses from Unemployment», Centre for Labour Economics, London School of Economics, *Discussion Paper*, No. 213.
- HERNÁNDEZ-IGLESIAS, F., y RIBOUD, M. (1985), «Trends in Labour Force Participation of Spanish Women: An Interpretive Essay», *Journal of Labor Economics*, 3, páginas 201-217.
- INSTITUTO NACIONAL DE ESTADÍSTICA (1983), *Encuesta de Presupuestos Familiares 1980-81. Tomo I. El Gasto y el Ingreso de los hogares. Conjunto Nacional*.
- (1985), *El paro. Sus repercusiones en el nivel de vida*.
- JACKMAN, R., y LAYARD, R. (1985), «A Wage-Tax Worker-Subsidy Policy for Reducing the "Natural" Rate of Unemployment», Centre for Labour Economics, London School of Economics, *Discussion Paper*, No. 228.
- LANCASTER, T., y NICKELL, S. J. (1980), «The Analysis of Re-Employment Probabilities of the Unemployed», *Journal of the Royal Statistical Society, Series A*, 143, páginas 141-165.
- LAYARD, R. (1985), «How to Reduce Unemployment by Changing National Insurance and Providing Job-Guarantee», Centre for Labour Economics, London School of Economics, *Discussion Paper*, No. 218.
- LIPPMAN, S. A., y MCCALL, J. J. (1976a), «The Economics of Job Search: A Survey. Part I: Optimal Job Search Policies», *Economic Inquiry*, 14, págs. 155-189.
- (1976b), «The Economics of Job Search: A Survey. Part II: Empirical and Policy Implications of Job Search», *Economic Inquiry*, 14, págs. 347-368.
- MADDALA, G. S. (1983), *Qualitative and Limited-Dependent Variables in Econometrics*, Econometric Society Monographs, Cambridge University Press.
- MORTENSEN, D. T. (1970), «Job Search, the Duration of Unemployment and the Phillips Curve», *American Economic Review*, 60, págs. 847-862.
- NICKELL, S. J. (1980), «A Picture of Male Unemployment in Britain», *Economic Journal*, 90, págs. 776-794.

## APENDICE 1: DATOS Y VARIABLES

La información utilizada en el presente trabajo proviene de la EPF 1980-81. Se han excluido inicialmente de la muestra aquellos sustentadores principales mayores de 65 años (edad de jubilación) y los que, a la pregunta correspondiente a su situación laboral responden estar jubilados o retirados, vivir de rentas, estar cumpliendo el servicio militar, ser estudiantes o dedicarse a las labores del hogar (muestra 1). Es decir, centramos nuestro análisis en los sustentadores principales que forman parte de la población activa, tal y como la define el INE. Posteriormente, y por las razones mencionadas en el texto, se han excluido las mujeres que son sustentadores principales (muestra 2) y, finalmente, los menores de 25 años (muestra 3).

La variable endógena del modelo es una variable binaria que toma el valor 1 si el individuo trabaja, siendo indistinto si es más o es menos de un tercio de la jornada normal, y el valor 0 si en la encuesta declara estar parado, habiendo trabajado antes o buscando el primer empleo. Las variables explicativas, que recogen los rasgos característicos que se analizan en este estudio, se definen del siguiente modo:

- *Edad*: Cinco variables ficticias para los siguientes grupos de edad: menores de 20 años, 20-24, 25-34, 45-54 y 55-64. El grupo omitido para evitar problemas de multicolinealidad perfecta es 35-44.
- *Nivel de instrucción*: Cuatro variables ficticias para los siguientes grupos: analfabetos, sin estudios (alfabetos, no habiendo cursado más de cinco cursos primarios), estudios secundarios (Bachillerato Elemental o Superior, o equivalentes, y Formación Profesional) y estudios superiores (enseñanzas de tipo superior universitarias y enseñanzas de tercer grado no universitarias). La variable ficticia omitida corresponde a los estudios primarios.
- *Composición familiar*: Siete variables que indican el número de miembros del hogar en los siguientes grupos de edad: de 0 a 4 años, 5-8, 9-13, 14-17, 17-24, 25-60 y más de 60 años.
- *Tamaño municipal*: El número de habitantes del municipio en que reside el sustentador principal sirve para definir cuatro variables ficticias para los siguientes grupos: menos de 2.000 habitantes, 2.000-10.000, 10.000-50.000 y más de 50.000 habitantes. La variable ficticia omitida corresponde al grupo 50.000-500.000.
- *Condición de emigrante*: Una variable ficticia que toma el valor 1 si el sustentador principal no residía en su actual municipio a finales de 1975.

- *Variables estacionales*: Dado que las entrevistas se efectúan a lo largo de un año, se han incluido tres variables ficticias, correspondientes a invierno, verano y otoño, para recoger la posible estacionalidad existente. La variable omitida corresponde a primavera.
- *Sexo*: Variable ficticia que toma el valor 1 si el sustentador principal es mujer y 0 en caso contrario.

Las medias de estas variables para la submuestra de varones no menores de 25 años (columna [1]), así como para su correspondiente submuestra de parados (columna [2]), se indican a continuación en el cuadro A.1. Téngase en cuenta que, con la excepción de las variables que hacen referencia a la composición familiar, todas las variables son binarias (0,1), por lo que las medias se representan en términos de porcentaje.

CUADRO A.1  
MEDIAS DE LAS VARIABLES UTILIZADAS

	[1]	[2]
Variable endógena	93,7	0,0
Edad:		
25-34	21,1	22,0
45-54	31,3	26,2
55-64	19,1	33,2
Nivel de instrucción:		
Analfabetos	3,3	9,4
Sin estudios	22,3	39,7
Estudios secundarios	15,4	6,6
Estudios superiores	8,0	1,9
Composición familiar:		
0-4	0,3	0,3
5-8	0,5	0,6
9-13	0,5	0,5
14-17	0,4	0,4
17-24	0,4	0,5
25-60	2,0	2,0
> 60	0,2	0,2
Tamaño municipal:		
< 2.000	10,8	5,6
2.000-10.000	16,1	20,1
10.000-50.000	18,9	26,2
> 50.000	12,0	11,8
Emigrante	5,3	21,1
Comunidad autónoma:		
Aragón	5,5	3,0
Asturias	2,8	1,0
Balears	1,8	0,7
Canarias	3,7	5,1
Cantabria	2,2	1,1
Castilla-León	13,3	7,9
Cataluña	10,8	9,8
Extremadura	3,9	5,8
Galicia	6,5	5,7
Madrid	5,5	5,7
La Mancha	7,6	8,2
Murcia	1,9	1,6
Navarra	1,6	0,9
La Rioja	1,6	0,8
Valencia	7,5	7,4
País Vasco	5,5	4,4
Variables estacionales:		
Invierno	25,0	29,4
Verano	23,7	22,3
Otoño	25,4	25,9
Tamaño muestral	15.297	967

Nota: Los valores de la columna [1] corresponden a las medias de las variables para la totalidad de la muestra. En la columna [2] se reportan las de la submuestra de parados.

## APENDICE 2: MODELOS DE ELECCION DISCRETA. RESULTADOS DE LA ESTIMACION

El modelo econométrico utilizado en el presente trabajo es un modelo en el que la variable endógena es de tipo cualitativo, ya que se trata de explicar si el individuo está empleado o no lo está. Estos aspectos cualitativos se pueden representar mediante una variable ficticia que, en nuestro caso, toma el valor 1 si el individuo está empleado y 0 si no lo está. Las variables explicativas que se han utilizado en el modelo han sido detalladas en el Apéndice 1. La especificación del modelo, basada en la del modelo de regresión clásico, sería

$$Y_i = X_i' \beta + u_i \quad [1]$$

en donde el subíndice «i» indica el sustentador principal de que se trata.  $Y_i$  es la variable endógena, que toma el valor 1 si el individuo trabaja y 0 si está en paro.  $X_i'$  es un vector de características del sustentador principal, que están definidas en el Apéndice 1, y  $u_i$  es el término de perturbación.

La especificación [1] es lo que se conoce en la literatura como modelo de probabilidad lineal. La estimación del mismo por mínimos cuadrados ordinarios es consistente, aunque en dicho modelo las perturbaciones son heteroscedásticas y las predicciones

de la variable endógena, que no son más que las probabilidades de estar empleado, no están restringidas a hallarse en el intervalo 0,1. Estas limitaciones del modelo de probabilidad lineal se solventan haciendo uso de modelos probabilísticos. Así, podemos definir la probabilidad de estar empleado del siguiente modo

$$\text{prob} (Y_i = 1) = F (X_i' \beta) \quad [2]$$

en donde  $F$  es una función de distribución determinada. Si se trata de una distribución logística, tenemos lo que se conoce en la literatura como el Modelo Logit, mientras que si se trata de una distribución normal tenemos el Modelo Probit, que es el que se ha utilizado en el presente estudio. (Véanse Amemiya (1981) y Maddala (1983) para una detallada descripción y discusión de este tipo de modelos.)

La formulación del modelo sería

$$Y_i = \begin{cases} 1 & \text{si } X_i' \beta + u_i > 0 \\ 0 & \text{si } X_i' \beta + u_i \leq 0 \end{cases} \quad [3]$$

en donde  $u_i$  se distribuye según una normal.

Los resultados de la estimación del modelo [3] para distintas submuestras se presentan en los cuadros A.2 y A.3 de este apéndice. Dichos modelos se han estimado por máxima verosimilitud. La función de verosimilitud (en logaritmos) ( $L^*$ ) del Modelo Probit es

$$L^* = \sum_{i=1}^M \ln F (X_i' \beta) + \sum_{i=M+1}^N \ln |1 - F (X_i' \beta)| \quad [4]$$

en donde  $N$  es el tamaño de la muestra, que está ordenada de tal modo que los  $M$  primeros individuos están empleados y los  $N - M$  restantes en paro.

CUADRO A.2  
RESULTADOS DE LA ESTIMACION

	(1)	(2)	(3)
Constante	1,40 (13,60)	1,43 (13,40)	1,47 (13,40)
Edad:			
16-19	- 0,25 (1,02)	- 0,19 (0,65)	
20-24	- 0,33 (3,02)	- 0,32 (2,73)	
25-34	- 0,23 (4,22)	- 0,23 (4,16)	- 0,23 (4,14)
45-54	- 0,06 (1,15)	- 0,07 (1,32)	- 0,07 (1,29)
55-64	- 0,11 (1,90)	- 0,11 (1,74)	- 0,10 (1,70)
Nivel de instrucción:			
Analfabetos	- 0,58 (8,07)	- 0,64 (8,66)	- 0,64 (8,60)
Sin estudios	- 0,34 (8,45)	- 0,35 (8,54)	- 0,36 (8,59)
Estudios secundarios	0,40 (6,92)	0,39 (6,50)	0,39 (6,26)
Estudios superiores	0,64 (6,79)	0,66 (6,62)	0,67 (6,62)
Composición familiar:			
0-4	0,03 (0,85)	-0,03 (0,81)	0,02 (0,68)
5-8	- 0,07 (2,68)	- 0,07 (2,84)	- 0,07 (2,82)
9-13	- 0,06 (2,66)	- 0,06 (2,69)	- 0,06 (2,65)
14-17	- 0,04 (1,32)	- 0,03 (1,09)	- 0,03 (1,00)
17-24	- 0,02 (1,92)	- 0,03 (1,19)	- 0,03 (1,17)
25-60	0,13 (4,34)	0,12 (3,80)	0,10 (3,10)
> 60	0,13 (3,29)	0,12 (2,92)	0,11 (2,70)
Tamaño municipal:			
< 2.000	0,29 (4,29)	0,30 (4,41)	0,30 (4,31)
2.000-10.000	- 0,00 (0,07)	- 0,00 (0,04)	- 0,01 (0,13)
10.000-50.000	- 0,07 (1,61)	- 0,06 (1,38)	- 0,08 (1,65)
> 500.000	- 0,11 (1,69)	- 0,11 (1,72)	- 0,14 (2,06)
Emigrante	- 0,18 (2,56)	- 0,17 (2,39)	- 0,18 (2,34)
Comunidad autónoma:			
Aragón	0,35 (3,71)	0,39 (3,93)	0,38 (3,79)
Asturias	0,45 (3,45)	0,48 (3,51)	0,48 (3,33)
Balears	0,34 (2,37)	0,44 (2,80)	0,52 (3,05)
Canarias	0,11 (1,39)	0,12 (1,50)	0,13 (1,50)
Cantabria	0,34 (2,38)	0,34 (2,32)	0,32 (2,21)
Castilla-León	0,26 (3,97)	0,28 (4,21)	0,26 (3,86)
Cataluña	0,12 (1,94)	0,12 (1,95)	0,12 (1,88)
Extremadura	0,05 (0,67)	0,04 (0,55)	0,03 (0,37)
Galicia	0,21 (2,81)	0,23 (3,05)	0,22 (2,80)
Madrid	0,07 (0,87)	0,09 (1,10)	0,08 (0,92)
La Mancha	0,13 (2,01)	0,14 (2,10)	0,16 (2,35)
Murcia	0,30 (2,50)	0,36 (2,83)	0,35 (2,72)
Navarra	0,29 (1,87)	0,29 (1,84)	0,26 (1,62)
La Rioja	0,31 (1,93)	0,31 (1,88)	0,29 (1,76)
Valencia	0,18 (2,62)	0,22 (3,15)	0,23 (3,22)
País Vasco	0,15 (1,82)	0,16 (1,97)	0,16 (1,89)
VARIABLES ESTACIONALES:			
Invierno	- 0,15 (3,26)	- 0,15 (3,28)	- 0,15 (3,12)
Verano	- 0,03 (0,59)	- 0,03 (0,68)	- 0,03 (0,63)
Otoño	- 0,08 (1,81)	- 0,09 (1,81)	- 0,08 (1,67)
Sexo (mujer = 1)	0,36 (4,39)		
Tamaño muestral	16.759	15.699	15.297

Notas:

1. La variable endógena toma el valor 1 si el sustentador principal trabaja, tomando el valor 0 si el individuo está en paro o busca su primer empleo.
2. La muestra utilizada corresponde a sustentadores principales, menores de 65 años y que no sean rentistas, jubilados, amas de casa o estudiantes.
3. Los resultados de las distintas columnas corresponden a las siguientes submuestras:
  - (1) Muestra definida en la nota 2.
  - (2) Submuestra de varones.
  - (3) Submuestra de varones mayores de 25 años.
4. Entre paréntesis, los valores del estadístico «t».

CUADRO A.3  
RESULTADOS DE LA ESTIMACION

	(1)	(2)	(3)
Constante	1,47 (15,44)	1,88 (7,55)	2,54 (6,08)
Edad:			
25-34	- 0,24 (4,10)	- 0,17 (1,49)	- 0,36 (1,84)
45-54	- 0,06 (1,18)	- 0,02 (0,16)	- 0,22 (1,29)
55-64	- 0,09 (1,51)	0,02 (0,15)	- 0,26 (1,11)
Nivel de instrucción:			
Analfabetos	- 0,64 (8,53)	- 0,71 (5,51)	- 0,06 (0,15)
Sin estudios	- 0,36 (8,49)	- 0,39 (4,90)	- 0,45 (2,94)
Estudios secundarios	0,41 (6,22)	0,58 (3,25)	0,50 (2,45)
Estudios superiores	0,66 (5,98)	0,57 (2,88)	1,10 (2,27)
Composición familiar:			
0-4	0,03 (0,71)	0,05 (0,71)	0,02 (0,14)
5-8	- 0,07 (2,64)	- 0,09 (1,96)	- 0,17 (2,01)
9-13	- 0,06 (2,50)	- 0,07 (1,60)	- 0,20 (1,98)
14-17	- 0,04 (1,34)	- 0,06 (1,03)	0,01 (0,08)
17-24	- 0,03 (0,96)	- 0,11 (2,18)	0,04 (0,35)
25-60	0,10 (3,12)	0,11 (1,70)	- 0,18 (1,36)
> 60	0,10 (2,46)	0,11 (1,31)	- 0,02 (0,15)
Tamaño municipal:			
< 2.000	0,32 (4,44)	- 0,14 (0,63)	0,45 (1,98)
2.000-10.000	- 0,00 (0,03)	- 0,16 (1,65)	0,69 (2,54)
10.000-50.000	- 0,07 (1,52)	- 0,10 (1,21)	- 0,11 (0,71)
> 500.000	- 0,13 (1,90)	- 0,03 (0,20)	- 0,03 (0,19)
Emigrante		- 0,34 (2,07)	- 0,33 (1,30)
Variables estacionales:			
Invierno	- 0,14 (2,76)	- 0,21 (2,20)	- 0,09 (0,60)
Verano	- 0,03 (0,60)	0,05 (0,50)	0,26 (1,41)
Otoño	- 0,09 (1,76)	- 0,12 (1,25)	- 0,12 (0,79)
Tamaño muestral	14.480	2.804	1.656

Notas:

1. La variable endógena toma el valor 1 si el sustentador principal trabaja, tomando el valor 0 si el individuo está en paro o busca su primer empleo.
2. La muestra utilizada corresponde a sustentadores principales, mayores de 25 años y menores de 65 años y que no sean rentistas, jubilados, amas de casa o estudiantes.
3. Los resultados de las distintas columnas corresponden a las siguientes submuestras:
  - (1) Submuestra de varones no emigrantes.
  - (2) Submuestra de varones residentes en Andalucía.
  - (3) Submuestra de varones residentes en Cataluña.
4. Entre paréntesis, los valores del estadístico «t».