

TAXA DE PARO EN ESPAÑA: UNHA ANÁLISE EN FUNCIÓN DAS CARACTERÍSTICAS DEMOGRÁFICAS DO INDIVIDUO

MARÍA VICTORIA VERDUGO MATÉS
Departamento de Economía Aplicada
Facultade de Ciencias Económicas e Empresariais
Universidade de Vigo

MARÍA ISABEL CAL BOUZADA
Departamento de Economía Aplicada
Facultade de Ciencias Económicas e Empresariais
Universidade de Vigo

Palabras clave: *Taxa de paro; Modelo logit; Elección discreta; Estimación.*

Key words: *Unemployment rate; Logit model; Discrete election; Estimate.*

Resumo

A forza de traballo está integrada por un conxunto de individuos moi heteroxéneos, que poden diferenciarse en virtude de cualidades tales como o sexo, o estado civil, o nivel de estudos e a idade. Parece evidente que a probabilidade de que un individuo estea desempregado diferirá, dependendo de se presenta unhas ou outras características.

O principal propósito deste artigo é cuantificar o impacto das devanditas características demográficas sobre a probabilidade individual de estar desempregado, utilizando para iso a metodoloxía econométrica dos "modelos logit binarios".

Abstract

Labour force is made up by a group of a highly heterogeneous individuals which can be differentiated by virtue of qualities such as sex, marital status, level of studies and age. It seems obvious that the probability for an individual being unemployed will differ, depending on any

of those characteristics. The aim of this paper is to quantify the impact of the demographic characteristics above on the individual probability of being unemployed, utilising the econometric methodology of "binary logit models".

1. INTRODUCCIÓN

Esta análise econométrica do desemprego baséase nos "*Modelos de elección cualitativa*", concretamente nos "*modelos Logit binarios*". Os devanditos modelos aplicáronse a unha gran variedade de problemas de elección cualitativa, e a súa utilización neste estudio está xustificada no sentido de que aínda que os individuos non teñen que realizar ningún tipo de elección, pode considerarse que se atopan nunha situación que pode ser descrita en termos probabilistas.

Coa finalidade de reduci-lo número de observacións a un tamaño razoable e evita-la utilización de algoritmos de optimización non lineal para obter a estimación máxima verosímil, optamos por traballar con datos agrupados en lugar de datos individuais. Resulta conveniente resaltar que esta aproximación tan só é razo-

ble se o número de observacións repetidas¹ é suficientemente grande. Neste sentido, a utilización de datos agrupados na presente análise está sobradamente xustificada, pois baséase nos resultados da *Encuesta de Población Activa*², na que o número de familias enquisadas (máis de 60.000) se pode considerar suficiente.

2. MODELO LOGIT BINARIO DE DESEMPREGO

Sexa Y_i unha variable binaria que toma os seguintes valores:

$Y_i = 1$, se o individuo i -ésimo está desempregado;

$Y_i = 0$, en caso contrario.

Supoñamos unha mostra de N^3 individuos clasificados en M grupos, existindo en cada un deles n_i individuos.

Sexa:

p_i , a taxa de paro mostrai⁴;

P_i , a correspondente taxa de paro poboacional.

A relación entrámbalas dúas taxas virá dada pola seguinte expresión:

$$p_i = P_i + u_i \quad (1)$$

onde u_i representa o termo de erro que fai que o valor da taxa de paro mostrai (p_i) difira da taxa de paro poboacional (P_i).

Posto que a taxa de paro poboacional (P_i) é descoñecida, para representala utilizáremo-la función de distribución loxística (F), da que a súa expresión matemática é:

$$P_i = P(Y_i = 1) = F(x'_i \beta) = \frac{e^{x'_i \beta}}{1 + e^{x'_i \beta}} \quad (2)$$

onde:

x_i , é o vector de observacións para as variables explicativas do grupo i -ésimo;

β , é o vector de parámetros común para tódolos grupos mostrais.

Se na expresión (1) substituímos (2), obtemos:

$$p_i = \frac{e^{x'_i \beta}}{1 + e^{x'_i \beta}} + u_i \quad (3)$$

Para solucionar-lo problema da non linealidade da ecuación (3), tanto nas variables como nos parámetros, recórrase á aproximación lineal en serie de Taylor:

$$\ln \frac{p_i}{1 - p_i} = x'_i \beta + \varepsilon_i \quad (4)$$

Na ecuación (4) recóllese a especificación do modelo logit, co que se obtén unha especificación lineal non só nas variables senón tamén nos parámetros. Polo tanto, trátase dunha especificación facilmente manexable de cara á estimación, coa única salvidade de que o termo de erro presenta problemas de heterocedasticidade.

A estrutura da varianza do termo de erro vén dada pola seguinte expresión:

$$\text{Var}(\varepsilon_i) = \frac{1}{n_i P_i (1 - P_i)}$$

Esta varianza depende da probabilidade poboacional (P_i) e, dado que a devandita probabilidade é descoñecida, como aproximación para a súa estimación se utiliza a probabilidade mostrai (p_i). Esta varianza estimada pódese utilizar para construí-la variable de ponderación adecuada e poder obter uns estimadores máis eficientes cós estimadores de mínimos cuadrados ordinarios.

Antes de continuarmos, resulta conveniente facer unha breve referencia acerca dos coeficientes da ecuación (4). Como xa sinalamos, no lado dereito aparecen só variables explicativas cualitativas, polo que os coeficientes da devandita ecuación poden interpretarse en termos de análise da varianza.

Para simplificármola exposición supoñemos que a variable dependente está caracterizada só por dous atributos; onde o atributo A presenta m modalidades e o atributo B , k .

Polo tanto, o modelo pode reformularse como:

$$\ln \frac{p_i}{1-p_i} = \sum_{r=1}^m \alpha_r A_r + \sum_{s=1}^k \beta_s B_s + u_i \quad (5)$$

onde:

$A_r=1$, se a observación pertence á r -ésima clasificación do atributo A ;

$A_r=0$, nos casos restantes.

$B_s=1$, se a observación pertence á s -ésima clasificación do atributo B ;

$B_s=0$, nos casos restantes.

Sen embargo, coa fin de facilita-la interpretación dos coeficientes, utilízase a expresión equivalente:

$$\ln \frac{p_i}{1-p_i} = \mu + \sum_{r=1}^{m-1} \alpha_r AT_r + \sum_{s=1}^{k-1} \beta_s BT_s + u_i \quad (6)$$

onde as novas variables se definen como:

$$AT_r = A_r - A_m \quad \text{para } r = 1, 2, \dots, m-1$$

$$BT_s = B_s - B_k \quad \text{para } s = 1, 2, \dots, k-1$$

Na expresión (5), a interpretación dos coeficientes é clara e directa. Cada conxunto de coeficientes das variables ficticias ten un significado intuitivo independente da categoría á que pertencen as variables ficticias do modelo.

μ : "media xeral": é constante para tódolos grupos de individuos. Este coeficiente representa a transformación logística da probabilidade de estar desempregado para o "individuo medio" na mostra, sen consideración diferenciada dos atributos A e B.

α_r para $r = 1, 2, \dots, m$: cuantifica a diferenza da media xeral para o grupo r -ésimo do atributo A.

β_s para $s = 1, 2, \dots, k$: cuantifica a diferenza da media xeral para o grupo s -ésimo do atributo B.

Os coeficientes α_m e β_k omitidos na ecuación (6) poden ser calculados a través das seguintes expresións:

$$\sum_{r=1}^m \alpha_r = 0 \quad \text{y} \quad \sum_{s=1}^k \beta_s = 0$$

3. ESTIMACIÓN DO MODELO

Na táboa 1 recóllense os resultados da estimación dunha primeira especificación do modelo, na que se inclúen as catro características (sexo, estado civil, nivel de educación e idade), sen considera-la posible interacción entre elas. O máis relevante da análise é o seguinte:

VARIABLE	COEFICIENTE	ERRO ESTÁNDAR
C	-1.0131761	0.0272826 **
SEXFT	0.2573805	0.0334722 **
SEXMT	-0.2573805	0.0334722 **
CASAT	-0.4399452	0.0467793 **
NONCASAT	0.4399452	0.0467793 **
E1T	0.5347199	0.1369827 **
E2T	0.1036285	0.0843944
E3T	-0.0784272	0.0762175
E4T	-0.1171896	0.1171607
E5T	-0.4427316	0.1346842 **
IDADE1T	1.1044487	0.0901598 **
IDADE2T	0.7374207	0.0611366 **
IDADE3T	0.2644419	0.0610185 **
IDADE4T	-0.3470429	0.0483045 **
IDADE5T	-0.7510092	0.0702767 **
IDADE6T	-1.0082594	0.0903461 **
U	0.115450	$U_M=0.000405$
		$U_D=0.206864$
		$U_C=0.792731$
NOTA: T indica que se trata de variables transformadas; ** indica que o coeficiente é significativamente distinto de cero para un nivel de confianza do 95%.		
FONTE: Elaboración propia.		

Táboa 1.- Resultados da estimación. Especificación I

1) A variable sexo ten un efecto estatisticamente significativo, sendo o devandito efecto negativo para os homes e positivo para as mulleres.

Este resultado reflicte a distribución das taxas de desemprego por sexo e idade presentadas na táboa 2, na que se pon de manifesto que os ratios

de desemprego dos homes son inferiores ós das mulleres.

IDADE	MULLER	HOME
De 16-19	58.0685	47.7234
De 20-24	46.8169	39.0806
De 25-29	37.8140	26.3018
De 30-44	27.5104	15.3814
De 45-54	20.7400	12.1350
De 55 e máis	8.8230	13.1737

FONTE: Encuesta de Población Activa e elaboración propia.

Táboa 2.- Ratios de desemprego por grupos de idade e sexo

2) Os coeficientes da variable estado civil son negativos para os individuos casados e positivos para os non casados.

As diferencias substanciais (sobre todo para os homes) existentes entre as taxas de paro dos individuos co mesmo sexo pero diferente estado civil (táboa 3) suxiren a existencia de interaccións entrámbolos dous atributos.

	CASADO	NON CASADO
Muller	26.5990	36.9221
Home	12.7298	33.1664

FONTE: Encuesta de Población Activa.

Táboa 3.- Ratios de desemprego por sexo e estado civil

3) A priori é de esperar que, a medida que o nivel de estudos do individuo se incrementa, a probabilidade de estar desempregado diminúa.

Os individuos analfabetos ou sen estudos (E1) e os individuos con estudos primarios (E2) teñen unha probabilidade de estar desempregados por riba da media, mentres que aqueles con niveis medios (E3⁵ e E4⁶) ou superiores (E5⁷) teñen unha probabilidade de estar desempregados menor cá media; isto resulta significativo só para os individuos analfabetos ou sen estudos e para os individuos con estudos superiores.

4) A idade parece ser un factor moi importante na determinación da probabilidade de desemprego. Os individuos entre 16 e 29 anos de idade (IDADE1, IDADE2 e IDADE3) teñen unha

probabilidade de estar desempregados por riba da media, mentres que aqueles por riba dos trinta anos de idade (IDADE4, IDADE5 e IDADE6) teñen unha probabilidade de estar desempregados inferior á media. En calquera caso, o efecto é tanto maior en valor absoluto canto máis nos achegamos ós extremos, é dicir, ós individuos máis novos e ós individuos máis vellos. Así, por exemplo, o ratio de desemprego dos individuos entre 16 e 19 anos é do 52,26%, mentres só o 11,94% dos individuos por riba dos 55 anos están desempregados (táboa 4).

IDADE	TOTAL
De 16-19	52.2619
De 20-24	42.5644
De 25-29	31.3045
De 30-44	19.9921
De 45-54	14.7308
De 55 e máis	11.9375

FONTE: Encuesta de Población Activa e elaboración propia.

Táboa 4.- Ratios de desemprego por grupos de idade

Na táboa 5 preséntanse os resultados da estimación dunha segunda especificación do modelo na que se recolle a interacción entre sexo e estado civil e entre idade e o nivel de estudos.

Como xa se mencionou, a existencia de diferencias substanciais entre os ratios de desemprego dos individuos do mesmo sexo e distinto estado civil suxiren a conveniencia de ter en conta a posible interacción entrámbolos dous atributos.

Do mesmo modo, parece conveniente ter en conta a posible interacción entre a idade e o nivel de estudos. Como se recolle na táboa 6, a probabilidade de estar en paro de dous individuos cun mesmo nivel de estudos difire se a súa idade é diferente. Co fin de simplificala exposición e para evita-la inclusión no modelo dun número excesivo de variables ficticias, lévase a cabo unha reclasificación dos individuos de acordo co seu grao de educación. Só se distinguen tres categorías: individuos analfabetos ou sen estudos (E1), individuos con estudos intermedios (EI) e individuos con estudos superiores (E5). Se nos fixamos de novo na táboa 6, a agrupación de indivi-

duos con estudos primarios (E2) e de individuos con estudos medios (E3 e E4) nunha soa categoría (EI) é defendible no sentido de que entre eles as diferencias son menores.

VARIABLE	COEFICIENTE	ERRO ESTÁNDAR
C	-1.0435749	0.0254358 **
(SEXF*CASA)T	0.0290775	0.0457052
(SEXF*NONCASA)T	0.5085571	10.0450971 **
(SEXM*CASA)T	-0.8809887	0.0437108 **
(SEXM*NONCASA)T	0.3433542	0.0422297 **
(IDADE1*EI)T	1.7193686	0.8278918 **
(IDADE1*EI)T	1.1249058	0.1017021 **
(IDADE1*E5)T	0.5929075	1.3545519
(IDADE2*EI)T	1.0401980	0.4718871 **
(IDADE2*EI)T	0.7095268	0.0669520 **
(IDADE2*E5)T	1.1545062	0.2322319 **
(IDADE3*EI)T	1.0956998	0.4157884 **
(IDADE3*EI)T	0.2465171	0.0710261 **
(IDADE3*E5)T	0.2263003	0.1398497
(IDADE4*EI)T	0.3714559	0.1678001 **
(IDADE4*EI)T	-0.2607387	0.0540175 **
(IDADE4*E5)T	-1.1973580	0.1468317 **
(IDADE5*EI)T	-0.0522167	0.1501306
(IDADE5*EI)T	-0.7504367	0.0875040 **
(IDADE5*E5)T	-1.9198346	0.3406186 **
(IDADE6*EI)T	-0.6348531	0.1648295 **
(IDADE6*EI)T	-0.9828686	0.1252009 **
(IDADE6*E5)T	-2.4830797	0.5953944 **
U	0.077468	$U_A=0.024370$
		$U_B=0.080642$
		$U_C=0.894987$

NOTA: T indica que se trata de variables transformadas; ** indica que o coeficiente é significativamente distinto de cero para un nivel de confianza do 95%.

FONTE: Elaboración propia.

Táboa 5.- Resultados da estimación. Especificación II

IDADE	E1	E2	E3	E4	E5
De 16-19	66.9643	54.8571	50.5783	55.3702	66.6667
De 20-24	50.1608	45.9147	41.2656	40.2847	52.8287
De 25-29	51.4925	35.9423	30.4641	28.8453	30.6519
De 30-44	33.8293	23.2127	20.0734	19.6920	9.6214
De 45-54	25.0710	15.4773	10.8638	11.9330	4.9344
De 55 e máis	15.7458	11.6678	11.1819	12.8253	2.9002

FONTE: Encuesta de Población Activa e elaboración propia.

Táboa 6.- Ratios de desemprego por grupos de idade e nivel de estudos

Dos resultados da táboa 5, cabe sinalalo seguinte:

1) Respecto á interacción entre sexo e estado civil cómpre subliñar que, mentres para os homes ámbolos dous coeficientes son significativos —negativos para os casados e positivos para os non casados—, para as mulleres tan só resulta significativo se non están casadas, aínda que nos dous casos o efecto é positivo. Do anterior e máis da análise da primeira especificación, pódese deducir que para os homes hai un predominio da variable estado civil sobre a variable sexo mentres que para as mulleres ocorre todo o contrario.

Na táboa 3, vemos que os individuos casados, xa sexan homes ou mulleres, presentan ratios de desemprego inferiores ós dos individuos non casados, pero as diverxencias son maiores para os homes.

2) Respecto á interacción entre idade e nivel de educación, hai que sinalar que os individuos de 16 a 29 anos con calquera nivel de educación teñen un efecto positivo, e significativo na maioría dos casos, na probabilidade de estar desempregados. Para os restantes grupos de idade, sexa cal sexa o seu nivel de educación, o efecto diferencial é case sempre negativo, resultando significativo na maior parte dos casos.

Por último, na táboas 7, 8, 9 e 10 presentámoslos ratios de desemprego estimados a partir da segunda especificación do modelo⁸. Estes ratios están referidos a diferentes "individuos medios", concretamente preséntanse os ratios de desem-

prego de homes e mulleres desagregados por estado civil, idade e nivel de educación.

IDADE	E1	EI	E5
De 16-19	66.9267	52.7574	39.6136
De 20-24	50.6425	42.4343	53.4943
De 25-29	52.0289	31.6916	31.2556
De 30-44	34.4559	21.8362	9.8691
De 45-54	25.6028	14.6163	5.0482
De 55 e máis	16.1197	11.9480	2.9381

FONTE: Elaboración propia.

Táboa 7.- Ratios de desemprego estimados para as mulleres casadas por grupos de idade e nivel de educación

IDADE	E1	EI	E5
De 16-19	76.5729	64.3339	51.4468
De 20-24	62.3676	54.3517	65.0102
De 25-29	63.6610	42.8371	42.3428
De 30-44	45.9200	31.0934	15.0284
De 45-54	35.7270	21.6623	7.9084
De 55 e máis	23.6878	17.9773	4.6615

FONTE: Elaboración propia.

Táboa 8.- Ratios de desemprego estimados para as mulleres non casadas por grupos de idade e nivel de educación

IDADE	E1	EI	E5
De 16-19	44.8880	31.0099	20.8886
De 20-24	29.2274	22.8841	31.6467
De 25-29	30.3885	18.9733	17.4638
De 30-44	10.1078	4.2212	12.1662
De 45-54	6.4465	2.0951	7.1796
De 55 e máis	5.1787	1.2037	12.7354

FONTE: Elaboración propia.

Táboa 9.- Ratios de desemprego estimados para os homes casados por grupos de idade e nivel de educación

IDADE	E1	EI	E5
De 16-19	73.4807	60.4604	47.3197
De 20-24	58.4185	50.2327	61.1658
De 25-29	59.7601	38.8481	38.3689
De 30-44	41.8541	27.6686	13.0383
De 45-54	32.0290	18.9900	6.7859
De 55 e máis	20.8321	15.6687	3.9799

FONTE: Elaboración propia.

Táboa 10.- Ratios de desemprego estimados para os homes non casados por grupos de idade e nivel de educación

Do resultado de estas táboas, cabe sinalalo seguinte:

1) Para o mesmo sexo, idade e nivel de educación, os ratios de desemprego dos individuos casados son menores cós dos individuos non casados, sendo esta diferenza máis considerable nos homes.

2) Para a mesma idade, nivel de educación e estado civil, os ratios de desemprego dos varóns son menores cós das mulleres.

3) Os individuos que posúen unha probabilidade estimada maior de estar desempregados son os individuos non casados, entre 16 e 19 anos e sen estudos. Estes ratios son do 73,48% para os homes e do 76,57% para as mulleres.

4. CONCLUSIONES

Este artigo presenta unha sinxela visión dalgun dos aspectos do desemprego en España. As altas taxas de desemprego que se veñen presentando e as discrepancias substanciais entre grupos poboacionais diferentes motiváronnos a cuantificalo impacto desas características diferenciais sobre a probabilidade de que un individuo estea en paro.

Os resultados indican que a medida que a idade se incrementa a probabilidade de que un individuo estea desempregado diminúe. Sen embargo, o resultado esperado de que a medida que o nivel de educación é maior a probabilidade de estar desempregado diminúe non é correcto. Así, os individuos entre 20 e 24 anos con estudos

universitarios teñen moita máis probabilidade de estar desempregados cós individuos coa mesma idade e menor educación. Os individuos casados teñen unha menor probabilidade de estar desempregados cós individuos non casados. Ademais, constátase que a probabilidade de que un individuo estea desempregado, é menor se se trata dun home.

De todo o anterior, parece evidente afirmar como conclusión final que o desemprego en España se nutre fundamentalmente da poboación nova e dos individuos sen estudos.

ANEXO DEFINICIÓN DAS VARIABLES

SEXF=1 Se o individuo é de sexo feminino; e
SEXF=0 nos casos restantes.
SEXM=1 Se o individuo é de sexo masculino; e
SEXM=0 nos casos restantes.
CASA=1 Se o individuo está casado; e
CASA=0 nos casos restantes.
NOCASA=1 Se o individuo non está casado; e
NOCASA=0 nos casos restantes.
E1=1 Se o individuo é analfabeto ou sen estudos; e
E1=0 nos casos restantes.
E2=1 Se o individuo posúe estudos primarios; e
E2=0 nos casos restantes.
E3=1 Se o individuo posúe estudos secundarios ou medios (agás técnicos-profesionais medios); e
E3=0 nos casos restantes.
E4=1 Se o individuo posúe estudos de técnico-profesional (graos medio e superior).
E4=0 nos casos restantes.
E5=1 Se o individuo posúe estudos universitarios ou outros estudos superiores; e
E5=0 nos casos restantes.
IDADE1=1 Se o individuo posúe de 16 a 19 anos de idade; e
IDADE1=0 nos casos restantes.
IDADE2=1 Se o individuo posúe de 20 a 24 anos de idade; e
IDADE2=0 nos casos restantes.
IDADE3=1 Se o individuo posúe de 25 a 29 anos de idade; e
IDADE3=0 nos casos restantes.
IDADE4=1 Se o individuo posúe de 30 a 44 anos de idade; e

IDADE4=0 nos casos restantes.

IDADE5=1 Se o individuo posúe de 45 a 54 anos de idade; e

IDADE5=0 nos casos restantes.

IDADE6=1 Se o individuo posúe 55 anos de idade ou máis; e

IDADE6=0 nos casos restantes.

NOTAS

1. Na nosa análise, aínda que non hai observacións repetidas dos mesmos individuos, aqueles que posúen un mesmo vector de características son indistinguíbles para os efectos do modelo.
2. Concretamente nos resultados do segundo trimestre de 1994.
3. No noso caso, a mostra estaría formada por todos aqueles individuos que integran a poboación economicamente activa. De acordo coa *Encuesta de Población Activa*, a mostra estaría integrada por todos aqueles individuos de 16 ou máis anos que durante a semana de referencia (a anterior a aquela na que se realiza a enquisa) subministran man de obra para a produción de bens e servizos económicos ou que están dispoñibles e fan xestións para se incorporar á devandita produción. E dentro deste grupo de individuos, podemos atoparnos con dúas situacións: individuos ocupados e individuos parados.
4. A taxa de paro mostrál defínese como o ratio entre o número de individuos desempregados e o número de individuos total, é dicir, a suma dos individuos desempregados e os individuos ocupados.
5. Na categoría E3 inclúense os estudos de secundaria xeral obrigatoria e de bacharelato.
6. Na categoría E4 inclúense os estudos de formación profesional I e II, idiomas, deseño e outros estudos regrados de igual nivel.
7. En estudos superiores, ademais de incluírse os títulos universitarios, inclúense os tres primeiros cursos de estudos universitarios de dous ciclos que non dan lugar a título, titulacións equivalentes ás universitarias e estudos postsecundarios non conducentes a titulacións equivalentes ás universitarias.
8. Escollémo-la segunda especificación do modelo dado que, se nos fixamos no coeficiente de desigualdade de Theil (U), se trata da especificación que presenta unha maior habilidade para estimar os ratios de desemprego. Trátase dun coeficiente de desigualdade relativamente baixo (0,077468), ade-

mais a descomposición do erro indica que a proporción debido ó nesgo e á diferente variación son moi reducidas, co que case a totalidade corresponde á compoñente de correlación, que é síntoma dun bo comportamento estimativo do modelo.

BIBLIOGRAFÍA

CRAMER, J.S. (1991): *An Introduction the Logit Model for Economists*. London: Edward Arnold.

DHRYMES, P.J. (1984): *Econometría*, pp. 205-215. Madrid: AC.

MADDALA, G.S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

NOVALES, A. (1993): *Econometría*, pp. 529-548. Madrid: McGraw-Hill.

PULIDO, A. (1989): *Modelos econométricos*, pp. 185-215. Madrid: Pirámide.

