ESTUDIOS SOBRE LA ECONOMIA ESPAÑOLA

The decissions of Spanish youth: A cross-section study

Maite Martínez
Javier Ruiz-Castillo

EEE 14

FEDEA
Fundación de Estudios de Economía Aplicada

http://www.fedea.es/hojas/publicado.html

THE DECISIONS OF SPANISH YOUTH: A CROSS-SECTION STUDY

Maite Martínez and Javier Ruiz-Castillo*

Departamento de EconomÌa, Universidad Carlos III de Madrid, C/. Madrid 126, 28903 Getafe, Madrid, Spain. Fax: +34 1 624 9875. e-mail address: mmartine@eco.uc3m.es and jrc@eco.uc3m.es.

Abstract

The paper presents a simultaneous model for the joint decisions of working, studying and leaving the parental household by young people in Spain. Using cross-section data from the 1990-1991 *Encuesta de Presupuestos Familiares*, the model is estimated by a two stage estimation method. Endogeneity of the three decisions proves to be important to understand the dynamics of household formation. Our results also confirm a number of plausible intuitions about the effect of demographic and economic characteristics on these decisions, illustrate important behavioural differences between men and women, and provide some new insights about the reasons for young people in Spain to remain in large numbers in the parental home.

JEL Nos.: J12, J21.

Keywords: household formation, working and studying decisions; two stage estimation.

*We are grateful to Manuel Arellano for very helpful discussions and suggestions. We would also like to acknowledge comments from Adriaan Kalwij and participants at the "Twelfth Annual Conference of the European Society for Population Economics" held in Amsterdam in June 1998. All remaining errors are our own.

I. INTRODUCTION

Contrary to Anglo-Saxon and central European countries but in line with other southern European nations, in Spain the proportion of young people living with their parents is very high[1]. Del Río and Ruiz-Castillo (1997) describes the evolution of living arrangements and living standards during the 1980's in Spain. They find that the young (defined as individuals from 16 to 30 years of age) who live as dependants in the parental home, are one of the population subgroups with the highest social welfare indices in the distribution of household expenditures adjusted for household size. However, we do not know of any work that attempts an explanation of this important demographic phenomenon.

Which forces can explain the decision by the young of remaining in the parental household until rather late in their life-cycle? In the first place, observers have pointed out that the unemployment rate among the youth in Spain is among the greatest in the EU. From this perspective, living at one's parents is supposed to provide a cushion to the young people unemployed or searching for a first job[2]. In the second place, Spain has one of the largest enrolment rates in higher education among the OCDE countries. From this perspective, parents are helping to finance their sons and daughters investment in human capital by providing them with shelter, food, tuition and possibly other goods and services. In the third place, although it is less well known, it should be emphasised that many of the young people living with their parents have already a job. From this perspective, parents are

providing means to compensate for job insecurity, low wages and/or high housing costs which act as deterrents for the young people to form their own independent household.

A general model would have to take into account the interaction between parents and their descendants decisions in a dynamic context (see, for instance, Ermish and Di Salvo, 1997, and Ermish, 1998). To test such a model one would need panel data, unavailable at present in Spain. But even with cross-section data one can address the following fundamental point: the decisions taken by young people about living arrangements, labour force participation and human capital investment are interrelated[3].

The following facts may serve to illustrate some of the interdependencies which should be disentangled. The unemployed living with their parents have a larger reservation wage than those living on their own. Being a College student decreases the probability of being employed. An individual's ability to earn some income is a necessary condition for him or her to afford leaving one's parents house.

In this paper we model the joint decision of whether to remain in the parental house, whether to work, and whether to keep on studying. A simultaneous probability model is estimated adapting a two-stage method proposed by Arellano and Bover (1997) in a related context. Our results confirm a number of plausible intuitions about the effect of demographic and economic characteristics on these decisions, illustrate important behavioural differences between men and women, and uncover some new interesting

facts about the circumstances which influence young people in Spain to remain in large numbers in the parental home.

The rest of the paper is organised in four sections. Section II presents the empirical model. Section III is devoted to the data description. Section IV contains the results, while Section V concludes.

II. EMPIRICAL MODEL SPECIFICATION

To form a household requires an economic independence that the individual achieves mainly by working in a stable job. On the other hand, an individual may postpone the formation of a new household, remaining at his/her parents while acquaring a higher educational level which may get translated into higher future earnings. Therefore, it seems reasonable to think that household formation by young people is intrinsically linked to other decisions with regard to job or educational status. In the model presented here, the young decide jointly whether to remain in the parental household, whether to work and whether to keep on studying[4].

When making such decisions, each individual analyzes the costs and benefits involved. Therefore, the following three latent equations can be defined conditional on a set of exogenous variables:

$$I_{1i}^{*} = X_{1i}\beta_1 + \delta_{12} I_{2i}^{*} + \delta_{13} I_{3i}^{*} + u_{1i} \qquad (1)$$

$$I_{2i}^{*} = X_{2i}\beta_2 + \delta_{21} I_{1i}^{*} + \delta_{23} I_{3i}^{*} + u_{2i} \qquad (2)$$

$$I_{3i}^{*} = X_{3i}\beta_3 + \delta_{31} I_{1i}^{*} + \delta_{32} I_{2i}^{*} + u_{3i}\,. \qquad (3)$$

The variable I_1^* is the underlying net utility the individual derives from leaving the parental house, I_2^* represents the net utility of working, and I_3^* is the net utility of studying. X_1, X_2 and X_3 are sets of exogenous demographic and economic variables that condition each equation. The β and δ vectors are the parameters of interest, and the error terms u_1, u_2 and u_3 are assumed to be jointly normally distributed.

In the data set we observe the outcomes of the choices, not the underlying utilities. That is, we observe whether and individual is independent from his/her parents, whether s/he is working and whether s/he is studying. The connection between our observations and the corresponding latent variables is given by the following three dichotomous variables:

$$I_{1i} = 1 \text{ if } I_{1i}^* > 0 \qquad I_{1i} = 0 \text{ otherwise} \tag{4}$$

$$I_{2i} = 1 \text{ if } I_{2i}^* > 0 \qquad I_{2i} = 0 \text{ otherwise} \tag{5}$$

$$I_{3i} = 1 \text{ if } I_{3i}^* > 0 \qquad I_{3i} = 0 \text{ otherwise} \tag{6}$$

We are interested in the estimation of the set of parameters $\Theta = \{\beta_1, \beta_2, \beta_3, \delta_{12}, \delta_{13}, \delta_{21}, \delta_{23}, \delta_{31}, \delta_{32}\}$ from the simultaneous probability model consisting of equations (1) to (3) and the observability conditions (4) to (6). Given the interdependence among the unobserved latent variables, we face a trivariate probit. Estimation by maximum-likelihood is very cumbersome: it involves computation of a triple integral and it can even be infeasible.

Alternative two-stage methods can be used which provide consistent, although inefficient, estimates of the parameters of interest. These two-stage methods[5] are the equivalent to two-stage least squares for continuos variables and are easier to implement and estimate: first, the reduced form equations are derived and estimated as independent probit equations; on the second stage, the reduced form predictions are used as regressors instead of the unobservable latent variables.

Arellano and Bover (1997) proposed a two-stage estimator for limited dependent variable models from panel data. Their methodology can be readily extended to the case of simultaneous probability models using a cross-section, as the one we are dealing with. This approach is easy to implement and interpret, and leads to consistent parameter estimates. It has also an advantage over the methods that estimate also a probit on the second stage. In these methods, given the assumption of unitary variance of disturbances on the reduced form equations, the parameters of interest can only be recovered up to scale. However, the Arellano and Bover method allows us o recover the actual parameters without the scale restriction because, as we will see, OLS in the second stage do not impose such identifiability conditions.

In the first stage, we consider the reduced form equations for the three endogenous variables,

$$I_{1i}^{*} = \pi_1 X + v_{1i} \qquad (7)$$

$$I_{2i}^{*} = \pi_2 X + v_{2i} \qquad (8)$$

$$I_{3i}^{*} = \pi_3 X + v_{3i}\,, \qquad (9)$$

where X includes all variables in X_1, X_2 and X_3. The error terms v_{1i}, v_{2i} and v_{3i} are assumed to be jointly normally distributed with variance equal to 1. The parameters in equations (7) to (9) are estimated by separate probit maximum likelihood[6], and the predictions for the unobserved latent variables $\hat{I}_{1i}^{*} = \hat{\pi}_1' X_i$, $\hat{I}_{2i}^{*} = \hat{\pi}_2' X_i$ *and* $\hat{I}_{3i}^{*} = \hat{\pi}_3' X_i$, are then computed.

Following Arellano and Bover (1997), we use the predictions to replace both types of unobservable latent variables: the endogenous explanatory variables and the dependent ones. Then, the parameters can be consistently recovered by applying OLS to the following equations:

$$\hat{I}_{1i}^{*} = X_{1i}\beta_1 + \delta_{12}\hat{I}_{2i}^{*} + \delta_{13}\hat{I}_{3i}^{*} + u_{1i} \qquad (10)$$

$$\hat{I}_{2i}^{*} = X_{2i}\beta_2 + \delta_{21}\hat{I}_{1i}^{*} + \delta_{23}\hat{I}_{3i}^{*} + u_{2i} \qquad (11)$$

$$\hat{I}_{3i}^{*} = X_{3i}\beta_3 + \delta_{31}\hat{I}_{1i}^{*} + \delta_{32}\hat{I}_{2i}^{*} + u_{3i}\,. \qquad (12)$$

Given the consistency and normality of the reduced form parameters, the set of estimates $\hat{\Theta}$ is also consistent and asymptotically normal. However, the asymptotic variance matrix of the estimates is *not* the traditional for OLS estimators because the dependent and the endogenous explanatory variables have been replaced by their predicted values. The variance matrix for this particular case is presented in Appendix A.

This procedure enables us to address the principal technical issue of this study, that is, to estimate the coefficients of the endogenous variables as a mean of inferring the interdependence among the three decisions considered

here. In that sense, we will refer to the model as "structural", by opposition to the reduced form equations that are estimated in the first stage.

There are three types of exogenous variables included in the analysis. The first group refers to the observable variables influencing the individual tastes for working, studying and living on one's own, such as education, age, sex or place of residence. The second group contains economic conditions, such as the regional unemployment rate or the regional average housing price, which would influence some but not all of the tendencies providing the necessary identification conditions[7]. Finally, certain household characteristics such as household income, social group or the number of siblings would mainly affect the probability of household formation, especially if one considers that living in the parental household is the outcome of a bargain process between the parents and the young. On the other hand, household characteristics may also influence the schooling status of the individual.

Unfortunately, the data set we use lacks information about the family background of individuals living on their own, which cannot be included as explanatory variables for the probability of leaving the parental home[8]. Nevertheless, it seems interesting to describe the household type in which it is most probable to find a dependant young person.

Taking into account that young people can be participating in the labour force, studying or enjoying other forms of inactivity, we partition all households into four different types: i) Households in which there are no young[9]; ii) households in which there are some young, all of them working;

iii) households in which there are some young, all of them studying; iv) the remaining households. Given this classification, we perform a multinomial logit estimation in which the probability of household *i* to be of type *j* depends on a set of exogenous characteristics, *Z*, including socio-economic, geographic and other household characteristics,

$$P_{ij} = \exp(\alpha'_j Z_i) \Big/ \sum_k \exp(\alpha'_k Z_i) .$$

Of course, some normalisation condition, such as $\alpha_m = 0$, is necessary for identification purposes.

III. DATA

The data used in this paper comes from the *Encuesta de Presupuestos Familiares* (EPF for short) for 1990-91. This is a household budget survey collected by the Spanish *Instituto Nacional de Estadística* (INE for short) during 52 consecutive weeks, from April of 1990 to March of 1991, with the main purpose of estimating the weights of the Consumer Price Index. It contains valuable information on a variety of demographic and socio-economic household and individual characteristics which are essential to our work. The 1990-91 EPF is a representative sample consisting of 21,155 observations for a population of approximately 11 million households living in residential housing over all of Spain, including the African cities of Ceuta and Melilla. There are 72,123 individuals in the sample, representative of a population of 38.5 million people.

A household is defined as "the person or set of persons who jointly occupy

a residential family dwelling, or part of it, and consume or share food and other commodities under a common budget." Therefore, people living in collective housing -residences for College students or the old, hospitals, hotels, prisons and the like- are not directly interviewed. However, expenditures and characteristics of household members who are entirely dependant on household resources but who live elsewhere at the time of the interview, are recorded in our data[10].

We concentrate on a subsample of young individuals aged between 18 and 35 years old at the interview date, which gives us a sample of 9,844 males and 9,652 females. Table 1 illustrates the differences in living arrangements between the Spanish and, for instance, the US situation. We observe that in both countries a significant proportion of females leave the parental home before the males of the same age. However, while in the US approximately 90 per cent of the young live on their own at 29 years of age, in Spain about 43 per cent of the males and 29 per cent of the females at that age are still living as dependants in their parents house. This explains why we choose 35 years of age as the upper bound in our definition of the young. On the other hand, since we are interested in the education decision, we choose 18 years of age as the lower bound, the earliest age at which people in Spain are supposed to decide whether to continue their studies at the College level.

[Table 1 around here]

All the endogenous and exogenous variables are described in the Data Appendix B. As we can see in Table B.1, the proportion of young females

living on their own is 44 per cent versus only 33 per cent among the males, while the proportion of females studying is also almost 4 percentage points above the males. On the other hand, the female participation rate is only 51.5 per cent against 77.3 for the males; however, the female unemployment rate is about ten percentage points higher.

Table B.2 shows the sample distribution according to the three dependent variables. Around 5 per cent of males and females work and follow some type of studies at the same time. Only 2 per cent of females and 0.5 per cent of males study and are independent. Most of these women (47 per cent) declare to follow “other studies” different from primary, secondary, or College studies.

The EPF provides information on education, job status and income for each household member, but it lacks information on two potentially important aspects: the type of contract, temporary or fixed, of those holding a job, and the marital status of any member apart from the household head (defined as the household member with the highest earnings) and his/her spouse. The first limitation precludes an study of the interaction between the decision to leave one’s parents house on the part of those employed and the type of contract they have. Notice, however, that had we have the data we could not simply include the contract type as an exogenous variable in the present framework. The recognition of its endogenous nature would possibly call for an independent analysis among the employed.

On the other hand, there is some empirical evidence that marriage is an important explanatory variable of household formation (see, for example, Haurin *et al.* ,1993). However, looking at Figures 1 and 2 it can be seen that in Spain almost all individuals that live out of their parents household are married, which implies that the decision of marrying or leaving the parents' home takes place almost always simultaneously[11]. Therefore, the only implication of not including marital status as an explanatory variable is that the results would not only show the propensity to leave the parental house but also the propensity of getting married.

[Figures 1 and 2 around here]

The exogenous variables entering equations 10 to 12 consist of individual characteristics (education, age, age squared, whether residing in a large city or in a small village and non labour income), and economic variables (the regional unemployment rate for the whole population and disaggregated by sex and age group, and the regional average prices of renting and owning a house). These economic variables are the ones that allow us to handle the identification problem. We need at least two exclusion restrictions for each equation. We assume that housing prices would only affect directly the probability of leaving the parental house, while the regional unemployment rates only affect this probability through its effect on the probability of working or studying. The education variables do not enter the "Studying" equation, because the level of education attained is precisely the result of the

decision to keep on studying. On the other hand, we allow the effect of the endogenous variables to vary with age in every equation[12].

For the analysis of the characteristics of households with young dependants, we select a subsample of 8,910 observations all of which include at least a couple in which the household head is between 40 and 65 years old. We include as explanatory variables those referring to the social and economic status of the parental household (the educational level attained by both the household head and his/her spouse, the household head occupation, the couple's total income, the household composition, and the region of residence), as well as the regional unemployment rate to control for general economic conditions.

IV. RESULTS

IV.1. Individual Decisions: "Structural" OLS Estimates

We will start by discussing the results for the three structural equations[13]. Each of them was estimated separately for males and females to capture differences by gender.

Table 2 shows estimates for the propensity to live out of the parental household. Age is a significant factor influencing positively this propensity for men and, especially, for women. For the latter group, there is a non-linear positive effect at a decreasing rate: they tend to leave the parents house sooner than men, but this differential effect is smaller the older the woman gets.

As expected, working increases the probability of leaving the parental household for males as well as females, because individuals who work have access to the necessary funds that allow them to be independent of their family. The probability of working has a smaller effect for women although, contrary to men, for them this effect increases with age.

[Table 2 around here]

For males, education has a negative but not well defined independent effect on the probability of forming a household. The effect of the educational variables is mainly transmitted through the higher probability of finding a job. For women, the effect is clearer: on the one hand, higher education increases the probability of having a job, and therefore the probability of leaving the parents household; but on the other hand, it has a strong negative independent effect. As pointed out in Section III, most individuals living on their own are married; therefore, the negative effect of education could just reflect the postponement of marriage decisions by more educated individuals, specially women.

Although the student condition indirectly reduces the probability of living the parents household via the reduction of the probability of working, the variable “Studying” has no significant direct effect for men. This is not the case for women, for whom this variable has a somehow puzzling positive effect. However, in the first place remember that the variable "Studying" does not only include what can be called official education, but also other type of studies (followed by 47 per cent of independent women). In the second place,

it is interesting to point out that while almost 20 per cent of the independent women that are studying are classified as the household head, only 9 per cent of those who are not studying behave as such. In the third place, the proportion of independent women who are working is higher if they are studying (45.4 per cent) than if they are not (36.1 per cent). Therefore, after controlling for the level of education attained, it looks like women prone to follow some type of study are also prone to live independently[14].

Non labour income has a positive effect on the probability of leaving the parental house but it is only well defined for women: access to more economic resources helps individuals to afford the expenses of living on their own and therefore increases the probability of leaving the parental house. Housing costs, whether owning or renting, provides a significant disincentive effect to move out of the parents household for both men and women. Finally, to live in a large city increases the probability of forming a household for men, while to live in a small village decreases this probability for both genders. This effect probably reflects the fact that the traditional pattern of “extended families”, where several generations cohabit under the same roof and cooperate on the same productive activities, is more prevalent in the countryside.

Turning to the work equation, Table 3 presents the results of the "structural" OLS estimation. For males, a prime determinant of working is the level of education attained, with an expected strong and positive effect. The general regional unemployment rate has also a significant negative effect,

reflecting both the lower probability of receiving a job offer and the discouraging worker effect that reduces the effort of looking for a job. The gender-age group specific unemployment rate, however, does not have any significant effect: it appears that males perceive the general unemployment rate as the one influencing their probability of receiving a job offer. After controlling by the demand side factors, the place of residence does not affect the probability of working.

The fact that an individual is studying reduces greatly his probability of working although this effect becomes smaller the older the individual is, as the sign of the variable “Studying x Age” reflects. As expected, income from other sources reduces the marginal benefits from working.

The decision to form an independent household has a positive effect on the probability of working: the increased costs the individual has to face in this case acts as an incentive to work. Finally, age has no significant effect on its own. This is understandable if we take into account that we are dealing with prime age males after controlling by whether they are studying or not.

[Table 3 around here]

For women, most of the previous effects prevail but also some new emerge. Age and living out of the parents house decrease the probability of working. That could just be reflecting that a great number of women that leave the parental house do it to get married and decide to take care of the family instead of working. Fertility decisions, linked with marriage, would enhance such an effect. On the other hand, the negative effect of the non

labour income variable can be seen again as a pure income effect on the probability of working: the higher the income from sources other than wages, the lower the probability that women work and the greater the time they can spend in other activities like taking care of the family or having children. As we will see below, this effect is reinforced through a positive impact of non labour income on the probability of studying.

With respect to the demand side variables, it is noticeable that both the general and the gender-age specific unemployment rates influence negatively the probability of working. Living in a village decreases the probability of working while living in a city increases it, reflecting the higher job opportunities and the life style in the city compared with the countryside.

To finish this part of the analysis, Table 4 shows the estimates for the studying equation. Results are quite similar for both men and women. Age increases the probability of studying but at a decreasing rate, specially for men. Living independently and working reduces the probability of studying, although the effect diminishes the older the individual is. The negative effect of living on their own is particularly high for women, reflecting again that family care and fertility decisions, which are intrinsically linked to the abandonment of the parental household, make women less likely to work or to study.

[Table 4 around here]

Given job status, the regional unemployment rate has a negative effect on the probability of studying for both men and women. As for the working

equation, the relevant unemployment rates are the general regional rate for men, and the gender and age group specific rate for women. Therefore, unemployment works through two channels: it reduces the probability of finding a job, favouring the studying option (opportunity cost effect), but it reduces the probability of studying (discouraging effect, due to the poorer job perspectives). Non labour income has a different influence by gender, being only relevant for women: it functions as an incentive to keep on studying. Finally, estimates reflect the fact that to live in a large city provides more opportunities for studying than to live in a small village.

IV. 2. Household Characteristics: Multinomial Logit Estimates

The last part of the empirical analysis consists on a multinomial logit to describe the characteristics of a sample of households headed by someone between 40 and 65 years of age, classified by their living arrangements. The reference group is formed by households without young dependants (40.7 per cent of the total sample). Household type 2 consist of those households in which all young dependants are working (24 per cent), and household type 3 are those in which all young dependants are studying (15.5 per cent). Although the complete results appear in Table 5, we concentrate on the effect of the explanatory variables in these two groups.

[Table 5 around here]

Social and economic variables influence mainly the probability of having young people studying in the household: the higher the household head is in

the social ladder (effect of the variables manager, self-employed, and non manual worker outside the agricultural sector), the higher the probability of having young dependants studying. The positive effect of the income variable leads to the same conclusion. The parents educational background works in the same direction: the higher the education level attained by the parents, the lower the probability of having dependants working and the higher the probability of having them studying. However, if both the father and the mother have a College degree, it seems that young individuals are more likely to leave the parental home.

It is interesting to note that the dummy variable for when the mother is not working but taking care of the family has a positive effect on having dependants at home. Therefore, it appears that the utility that the young enjoy from staying at the parental house increases with the availability of the mother to take care of them. In the same vein, we find a negative effect of household size (especially in the presence of children) on the probability of having young dependants: the greater the number of siblings competing for goods and services, the smaller the probability that the young remain at the parental home.

The regional unemployment rate has a positive effect on having dependants, especially if they are not working, confirming that the family works as a cushion when the young face adverse economic conditions. The parents age has a positive effect on the probability of having any dependent

at home but at a decreasing rate, and the effect is lower if, keeping constant the age of one of the parents, the spouse's age increases.

V. CONCLUSIONS

As far as we know, this paper constitutes the first systematic attempt to explain why parents and young descendants between 18 and 35 years of age in Spain decide to live together in rather large proportions.

Lacking longitudinal data, we have found it interesting to work with a rather rich, large and readily available household budget survey -the 1990-91 EPF- collected by the Spanish INE with completely different aims in mind. The reason is that, even with cross-section data one can start addressing the issues involved in joint decision making. The major novelty in the paper is that, in addition to the joint decision of whether to remain in the parental house and whether to work, we have been able to add the decision on whether to keep on studying.

The analysis is implemented through a two-stage method developed by Arellano and Bover (1997) for limited dependent variable models from panel data, which has been adapted here to the case of simultaneous probability models using cross-sectional data. Finally, some insight is gained through a multinomial logit analysis on the type of households in which it is more probable to find young dependants.

As far as our results are concerned, the following four points deserve to be emphasised. In the first place, like in previous studies for other countries, we

have shown that the endogenous treatment of the three decisions should occupy the core of any attempt to understand the issues involved. Otherwise, seriously misleading results can be obtained, a possibility we have illustrated when the decisions to work or study are treated as mere exogenous variables influencing the decision to abandon the parental household. As an example of the advantages of a simultaneous approach, we refer only to the role of the education variables on the probability of living independently: although these variables exercise an indirect influence through the probability of working for both genders, they exert an independent and positive effect in the women sample.

In the second place, our results confirm a number of plausible intuitions, such as i) the role of having a job, the age, or the cost of housing on the probability of leaving the parental home; ii) the role of education and the regional unemployment rate on the probability of working; iii) the role of living in a large city or in the countryside on the three decisions; or iv) the fact that the socio-economic and educational status of parents influence mainly the probability of having young dependants studying at home.

In the third place, our results illuminate important behavioural differences between males and females. In particular, women tend to leave their parents home sooner than men, but less so the older they get. This is consistent with the fact that, contrary to men, the more educated a woman is, the more she prefers to postpone marriage staying longer in the parental home. Also, women living independently have a lower probability of working than men.

These facts can be understood in a context in which many women leave the parental household to get married, take care of the home and/or have children. But this is more so the lower her educational level and the younger a woman is. It is worth pointing out also that young males seem to form expectations about the state of the labor market based only on the regional unemployment rate, while young females take into account the disaggregated rate by gender and age.

In the fourth place, we have uncovered some interesting facts on individual decisions, such as the positive effect of living independently on the probability of working for men, and a negative direct effect of regional unemployment on the probability of studying together with the more predictable positive indirect effect through the decrease in the probability of working. As far as having young dependants at the parental home, we would single out the positive effect that appears when the mother is not working and when there are fewer minors to compete for household resources, as well as the negative effect arising when both the father and the mother have a College degree.

The availability of a longitudinal data set, such as the forthcoming European Panel, would allow us to analyse more adequately the transition pattern towards forming a household, considering jointly the parents and the young decisions. Given the nature of the data available at present, this is beyond this paper's scope and is left to future research.

NOTES

[1] For a comparative study from a demographic perspective of the situation of the young in three southern European countries -Spain, Greece and Italy- and three central ones - France, Germany and the UK- see Fernández Cordón (1997).

[2] For an early exposition of this idea, see Revenga (1991).

[3] For previous work in this area, see McElroy (1985) on market work and family status, and Börsch-Supan (1986) and Haurin *et al* (1993, 1994) on housing demand and household formation.

[4] McElroy (1985) suggests that these three decisions, as well as marriage, should be treated jointly. However, she only analyzed the first two.

[5] Mallar (1977) or Heckman (1978), among others, proposed two-stage methods in this type of situation. For a revision of these methods see Maddala (1983).

[6] Estimates would be consistent and asymptotically normal, but not efficient: we estimate separately the three equations in spite of the fact that, in general, the error terms will be correlated.

[7] Note that, as we will see below, identification of each of the equations (10) to (12) requires at least two exclusion restrictions.

[8] As we said in the Introduction, there is evidence for Spain that, due to economies of scale and other factors, young people living in their parents household enjoy higher consumption than living alone (for the US, see McElroy, 1985). However, the lack of information about the standard of living of the parents of the young living by themselves forces us to abstract from these considerations.

[9] This type includes families that never had a child and families in which all the children had already left. Ideally we should separate these two groups, but the data set described in next section only informs about the number of children that remain at home rather than the number of children an individual has had in his/her life.

[10] For more details on the 1990-91 EPFs, see INE (1992).

[11] This is confirmed by all demographic studies in the subject. See, for instance, Vergés (1997).

[12] Other specifications, including the interactions with education, were tried but proved to fit worse the data.

[13] The reduced form equations, which are available from the authors on request, are not presented here because they do not provide any additional insight on the topic.

[14] It is important to point out that if we do not correct by the endogeneity of working and studying, the fact that the individual is studying has a positive and marginally significant effect on the probability of leaving the parental household for men and a negative and strongly significant effect for women. Estimates with exogenous working and studying decisions are available from the authors on request.

REFERENCES

Arellano, M. and Bover, O. (1997), "Estimating dynamic limited dependent variable models from panel data". *Investigaciones Económicas*, vol. XXI(2), pp. 141-165.

Börsch-Supan, A. (1986), "Household formation, housing prices, and public policy impacts". *Journal of Public Economics*, vol. 30, pp. 145-164.

Del Río, C. and J. Ruiz-Castillo (1997), "Demographic Trends and Living Standards", Working Paper 97-67, Economic Series 33, Universidad Carlos III de Madrid.

Ermisch, J. (1998), "Prices, Parents and Young People's Household Formation". *Mimeo.*

Ermisch, J. and Di Salvo, P. (1997), "The economic determinants of young peopleís household formation". *Economica*, vol. 64, pp. 627-44.

Fernández Cordón, J. a. (1997), "Youth Residential Independence and Autonomy. A Comparative Study", *Journal of Family Issues*, vol. 6, pp.563-575.

Haurin, D. R., Hendershott, P. H. and Dongwook, K. (1993), "The impact of real rents and wages on household formation". *The Review of Economics and Statistics.*

Haurin, D. R., Hendershott, P. H. and Dongwook, K. (1994), "Housing decisions of American Youth". *Journal of Urban Economics*, vol. 35, pp. 28-45.

INE (1992), *Encuesta de Presupuestos Familiares 1990-91. Metodología*, Madrid.

Maddala, G. S. (1983), "Limited dependent and qualitative variables in econometrics". Econometric Society Monographs. Cambridge University Press.

Mallar, C. D. (1977), "The estimation of simultaneous probability models". *Econometrica*, vol. 45(7), pp. 1717-22.

McElroy, M. B. (1985), "The joint determination of household membership and market work: the case of young men". *Journal of Labor Economics*, vol. 3 (3), pp. 293-316.

Revenga, A. (1991), "La liberalización económica y la distribución de la renta: la experiencia española," *Moneda y Crédito*, **193**: 179-224.

Vergés, R., ed. (1997), *La edad de emancipación de los jóvenes*, Centre de Cultura Contemporània de Barcelona.

APPENDIX A: ASYMPTOTIC VARIANCE MATRIX

In this section we derive the asymptotic variance matrix for equation (10) following Arellano and Bover (1997). The same procedure can be applied to equations (11) and (12).

Replacing the predictions for their expression in equation (10), we have that

$$\hat{\pi}_1' X_i = X_{1i}\beta_1 + \delta_{12}\hat{\pi}_2' X_i + \delta_{13}\hat{\pi}_3' X_i + \varepsilon_{1i} \quad .$$

Or in more compact notation,

$$\hat{\pi}_1' X_i = \hat{W}_i \delta_1 + \varepsilon_{1i}$$

where $\hat{\hat{W}}_i = \left[X_{1i}', \hat{\pi}_2' X_i, \hat{\pi}_3' X_i\right]$ and $\delta_1 = \begin{bmatrix} \beta_1 \\ \delta_{12} \\ \delta_{13} \end{bmatrix}$. The OLS estimator for δ_1 is:

$$\hat{\delta}_1 = \left(\sum_i \hat{W}_i' \hat{W}_i\right)^{-1} \left(\sum_i \hat{W}_i' \hat{W}_i\right) =$$

$$\delta_1 + \left(\sum_i \hat{W}_i' \hat{W}_i\right)^{-1} \left(\sum_i \hat{W}_i' \left(\hat{\pi}_{1i}' X_i - \hat{W}_i \delta_1\right)\right)$$

Rearranging terms, we have

$$\left(\sum_i \hat{W}_i' \hat{W}_i\right)\left(\hat{\delta}_1 - \delta_1\right) = \left(\sum_i \hat{W}_i' \left(\hat{\pi}_{1i}' X_i - \hat{W}_i \delta_1\right)\right) \qquad \text{(A.1)}$$

Note that the expected value of the variable I_1^* has the following form:

$$E\left(I_{1i}^* / X_i\right) = \pi_1' X_i = X_{1i}'\beta_1 + \delta_{12}\pi_2' X_i + \delta_{13}\pi_3' X_i$$

and therefore

$$\pi_1' X_i - \delta_{12}\pi_2' X_i - \delta_{13}\pi_3' X_i = X_{1i}'\beta_1 \qquad \text{(A.2)}$$

Replacing (A.2) in (A.1) we get:

$$\left(\sum_i \hat{W}_i' \hat{W}_i\right)\left(\hat{\delta}_1 - \delta_1\right) = \sum_i \hat{W}_i' \left(X_i', -X_i'\delta_{12}, -X_i'\delta_{13}\right) \begin{pmatrix} \hat{\pi}_1 - \pi_1 \\ \hat{\pi}_2 - \pi_2 \\ \hat{\pi}_3 - \pi_3 \end{pmatrix}$$

Therefore, the variance of $\hat{\delta}_1$ can be written as

$$\text{var}(\hat{\delta}_1) = \left(\sum_i \hat{W}_i' \hat{W}_i\right)^{-1} \left(\sum_i \hat{W}_i' \hat{M}_i\right) \text{var}(\hat{\Pi}) \left(\sum_i \hat{M}_i' \hat{W}_i\right) \left(\sum_i \hat{W}_i' \hat{W}_i\right)^{-1}$$

with $\hat{M}_i = \left(X_i', -X_i'\hat{\delta}_{12}, -X_i'\hat{\delta}_{13}\right)$. Now we only need an estimate for var($\hat{\Pi}$). Let us consider

$$vec(\hat{\Pi}) = \begin{pmatrix} \hat{\pi}_1 \\ \hat{\pi}_2 \\ \hat{\pi}_3 \end{pmatrix} = \arg max(L) = \arg max\left(L_1 + L_2 + L_3\right)$$

where L_j, j=1,2,3, is the corresponding likelihood function. Subject to suitable regularity conditions, a first order expansion of $\partial L(\hat{\Pi})/\partial\Pi$ around the true value of Π gives

$$\left(-1/N\, diag\left(\frac{\partial^2 L_j}{\partial\pi_j \partial\pi_j'}\right)\right)\sqrt{N}\left(\hat{\Pi} - \Pi\right) = 1/\sqrt{N}\sum_i \begin{pmatrix} \partial L_{i1}/\partial\pi_1 \\ \partial L_{i2}/\partial\pi_2 \\ \partial L_{i3}/\partial\pi_3 \end{pmatrix} + O_p(1)$$

which suggest an estimate for the variance, $\hat{V}$, of the form

$$\hat{V} = \hat{H}^{-1}\hat{\Psi}\hat{H}^{-1}$$

where $\hat{H} = diag\left(N^{-1}\partial^2 \hat{L}_j / \partial\pi_j \partial\pi_j'\right)$ *and* $\hat{\Psi} = N^{-1}\sum_i \left\{\partial\hat{L}_{ij}/\partial\pi_j\right\}.\left\{\partial\hat{L}_{ih}/\partial\pi_h'\right\}$

APPENDIX B: DATA

Endogenous variables

***Independent*:** dummy variable that equals one if the individual does not live in their parents house and zero otherwise.
Work: dummy variable that equals one if the individual is working (full or part time) at the interview date and zero otherwise.
Studying: dummy variable that equals one if the individual is carrying on any type of education and zero otherwise. It is worthy to note that 21.9 per cent of female students are said to attend "other type of education", different from primary, secondary, and College education, while only 14.9 per cent of male students declare to do so.

Exogenous variables

Educational: we define three dummy variables reflecting the highest degree completed by the individual. ***Educ2*** equals one if the individual has finished primary school, ***Educ3*** equals one if s/he has finished secondary school and ***Educ4*** equals one if some College degree has been attained.
***City*:** dummy variable that equals one if the individual lives in a large city (more than 500,000 inhabitants). ***Village*:** dummy variable that equals one if the individual lives in a small village (less than 2,000 inhabitants).
Non labour income: this is the summation of incomes different from labour earnings or subsidies related to the economic activity like the unemployment compensation. It includes returns to capital, private and other public transfers, grants, lotteries, etc. The proportion of young people who declare to have some non labor income is 19.7 per cent of males and 14.7 per cent of females.
Rent costs: regional average of rents by square metre paid and imputed rent for owner-occupying and other non-rental housing, by square metre. Imputed rents are estimated by the owner or the occupying household. Source: EPF.
***Owning costs*:** regional average of house prices by square metre. Source: *"Precio medio del m^2 de las viviendas",* Ministerio de Fomento.
***Unemployment*:** regional unemployment rates for the population as a whole and disaggregated by sex and age. Source: *Encuesta de Población Activa.*
Economic status: dummy variables for following socio-economic categories of the household head: ***agricultural, self-employed, manager, and manual or non manual worker outside of agricultural.***
***Income*:** income from all sources for the household head and his/her spouse.
***Mother*:** dummy variable that equals one if the mother is devoted to family care.
***Old*:** number of individuals in the household older than 35 years of age.
Children between 0-17: number of children in the household between 0 and 17 years old.

Table B.1: Characteristics for young individuals between 18 and 35 years old

	Males (9,844)		Females (9,652)	
	Mean	Standard error	Mean	Standard error
Independent	.333	.471	.445	.497
Participation	.773	.419	.515	.500
Work	.633	.482	.370	.483
Unemployed	.181	.385	.282	.450
Studying	.225	.418	.271	.444
Age	25.8	5.2	26.0	5.2
Educ2	.498	.500	.475	.499
Educ3	.355	.479	.350	.477
Educ4	.104	.305	.132	.338
City	.537	.499	.562	.496
Village	.144	.351	.127	.333
Non labour income	8,958	120,689	14,393	112,931
Rent Costs	3,217	1,270	3,271	1,292
Owning Costs	86,952	22,321	87,555	22,858
Regional Unemployment	17.700	11.649	31.160	12.695

Table B.2: Individual distribution according to dependent variables

	MALES		**FEMALES**	
	Number of observations	**%**	**Number of observations**	**%**
Working=0, Studying=0, Independent=0	1,639	16.65	1,412	14.63
Working=1, Studying=0, Independent=0	3,003	30.51	1,687	17.48
Working=0, Studying=1, Independent=0	1,655	16.81	1,958	20.29
Working=0, Studying=0, Independent=1	271	2.75	2,515	26.05
Working=1, Studying=1, Independent=0	270	2.74	297	3.08%
Working=0, Studying=1, Independent=1	46	0.47	196	2.03
Working=1, Studying=0, Independent=1	2,716	27.59	1,424	14.75
Working=1, Studying=1, Independent=1	244	2.48	163	1.69

Table B.3: Household Characteristics

	Mean	Standard Error
Agricultural	.058	.233
Manual	.290	.454
Non manual	.157	.364
Self-employed	.195	.396
Manager	.081	.272
Educ2 H.H.	.578	.494
Educ3 H.H.	.094	.292
Educ4 H.H.	.083	.276
Educ2 S.	.618	.486
Educ3 S.	.055	.228
Educ4 S.	.046	.209
Both Educ4	.031	.174
Age H.H.	52.8	7.3
Age S.	49.9	8.2
City	.523	.499
Village	.165	.371
Income	1,625,490	1,249,673
Mother	.751	.432
Old	.162	.435
Children bt 0-17	1.0	1.2
Observations	**8,910**	

Table 1. Percentage of youth residing out of the parental home

	SPAIN*		US*	
Age	**Male**	**Female**	**Male**	**Female**
22	7.8	11.4	54	71
29	56.8	71.1	87	91
35	83.8	87.6	-	-

*Spanish data, from the 1990-91 EPF. US data for 1987, from TABLE 1A in Haurin *et al.* (1994)

FIGURE 1

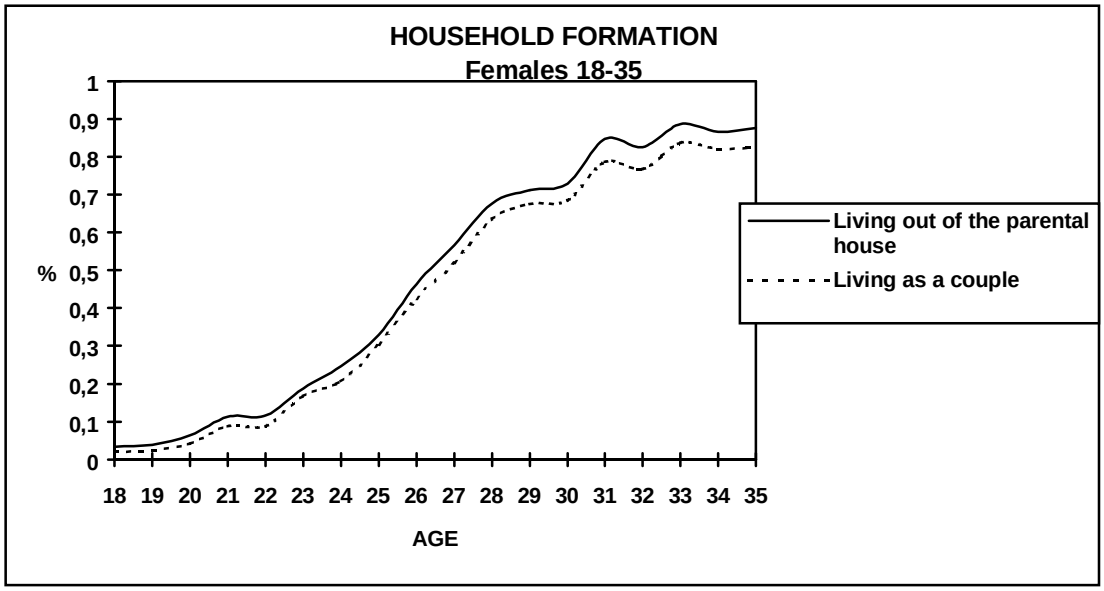

FIGURE 2

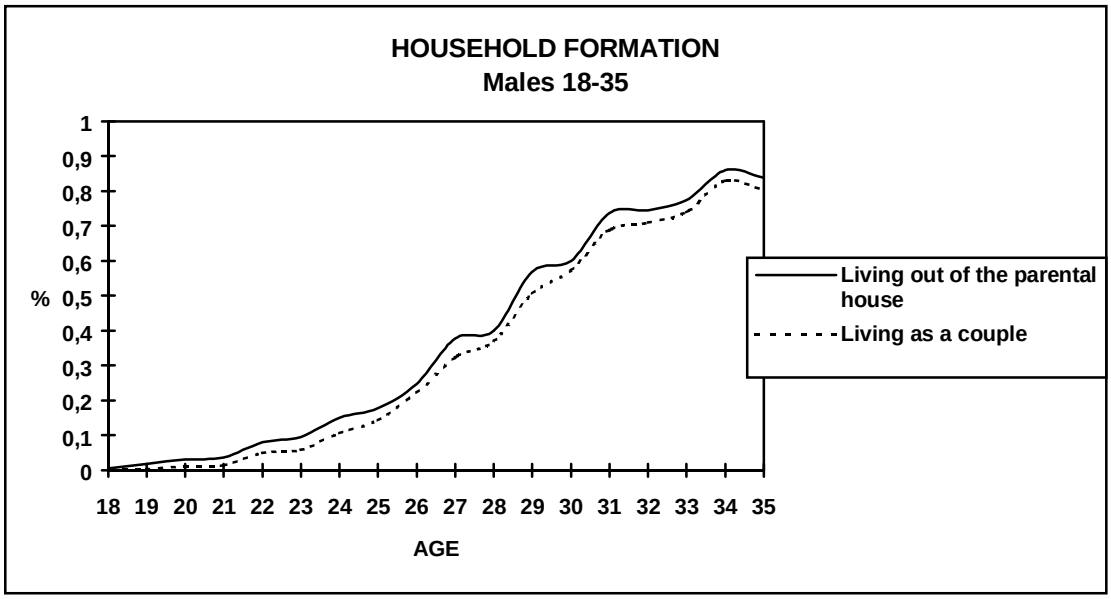

Table 2: Leaving the Parental House Equation*

(OLS "Structural" Estimates)

DEP: Independent**	Males		Females	
	Coefficient	t-statistic***	Coefficient	t-statistic
Work	0.404	5.678	0.279	2.599
Work x Age	0.073	0.981	0.375	4.985
Studying	0.127	1.038	0.335	2.418
Age	1.755	14.322	2.751	15.829
Age Squared	-0.095	-0.713	-0.956	-6.735
Educ2	0.112	1.194	-0.068	-0.521
Educ3	-0.097	-0.630	-0.975	-3.711
Educ4	-0.324	-1.646	-1.439	-4.447
City	0.085	1.611	-0.022	-0.359
Village	-0.241	-3.498	-0.236	-3.389
Non labour income	0.022	1.195	0.046	2.965
Owning Costs	-0.240	-2.123	-0.129	-1.167
Rent Costs	-0.126	-0.525	-0.608	-2.524
Intercept	-0.631	-2.796	0.994	2.654
Observations	9,844		9,652	

* For a definition of the variables see Data Appendix. Age=(Age-26)/10. Non labour income and the price of housing are divided by 10^{-5} and the price of renting by 10^{-4}.

** The variables *Independent*, *Work*, *Studying* and their interactions are predictions from their reduced form equations.

*** Standard errors for the parameters computed as in Appendix A.

Table 3: Work Equation*

(OLS "Structural" Estimates)

DEP: Work	Males		Females	
	Coefficient	t-statistic	Coefficient	t-statistic
Independent	0.629	4.299	0.014	0.128
Independent x Age	0.088	0.565	-0.303	-2.416
Studying	-0.625	-3.426	-0.750	-5.646
Studying x Age	0.351	3.119	0.239	2.962
Age	-0.308	-0.643	-0.785	-3.080
Age Squared	0.122	0.372	1.183	5.770
Educ2	0.294	2.545	0.597	4.659
Educ3	0.908	4.438	1.520	7.163
Educ4	1.147	4.059	2.024	8.373
City	0.004	0.060	0.197	3.636
Village	-0.021	-0.243	-0.156	-2.291
Non labour income	-0.078	-2.557	-0.056	-1.976
Regional Unemployment	-3.374	-6.772	-1.775	-4.046
Reg. Unemp. by age and sex	0.099	0.264	-0.718	-2.441
Intercept	0.271	0.740	-1.603	-5.048
Observations	9,844		9,652	

* See Notes for Table 2.

Table 4: Studying Equation*

(OLS "Structural" Estimates)

DEP: Studying	Males		Females	
	Coefficient	t-statistic	Coefficient	t-statistic
Independent	-1.094	-4.997	-1.377	-13.174
Independent x Age	0.546	1.665	0.505	2.296
Work	-0.531	-3.961	-0.592	-4374
Work x Age	1.080	7.003	1.057	8.835
Age	2.091	3.759	2.321	9.340
Age^2	-2.197	-4.116	-1.308	-3.551
City	0.317	4.672	0.302	4.892
Village	-0.548	-4.779	-0.518	-5.007
Non labour income	0.023	0.911	0.052	1.769
Regional Unemployment	-2.746	-3.553	-0.862	-1.193
Reg. Unemp. by age and sex	-0.420	-0.647	-1.634	-3.026
Intercept	-1.393	-4.634	-1.060	-7.634
Observations	9,844		9,652	

* See Notes for Table 2.

Table 5: Multinomial logit for type of household*

	Household type 2		Household type 3		Household type 4	
	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio
Agricultural	.158	1.180	-.387	-1.731	-.049	-0.307
Manual	.176	1.851	.238	1.866	.296	2.879
Non manual	.045	0.404	.516	3.820	.198	1.695
Self-employed	.101	1.011	.402	3.042	.242	2.234
Manager	-.549	-2.567	.327	1.907	.089	0.529
Educ2 H.H.	-.086	-0.755	.293	1.834	.002	0.013
Educ3 H.H.	-.588	-3.321	.521	2.713	.032	0.183
Educ4 H.H.	-.876	-3.379	.853	3.768	.255	1.186
Both Educ4	-.346	-0.574	-.635	-2.028	-.450	-1.245
Educ2 S.	-.195	-1.779	.388	2.640	.108	0.893
Educ3 S.	-.363	-1.690	.346	1.718	-.279	-1.352
Educ4 S.	-1.056	-2.705	.673	2.381	.081	0.269
Age H.H.	1.212	12.104	1.184	10.623	1.137	10.872
Age^2 H.H.	-.010	-6.292	-.009	-5.226	-.008	-5.239
Age S.	.677	9.343	.898	10.725	1.019	12.367
Age^2. S.	-.005	-5.172	-.006	-5.178	-.008	-6.548
Age H.H. x Age S.	-.003	-1.414	-.006	-2.690	-.004	-1.959
City	-.053	-0.749	.150	1.802	.071	0.948
Village	-.254	-2.904	-.206	-1.774	-.444	-4.380
Reg. Unemp	.019	3.225	.036	5.344	.037	6.093
Income	3.01e-08	0.803	1.29e-07	3.932	6.18e-08	1.756
Mother	.137	1.855	.234	2.744	.362	4.423
Old	-.039	-0.565	-.109	-1.347	-.075	-1.009
Children bt 0-17	-.106	-2.989	-.389	-9.728	-.1303	-3.516
Intercept	-50.915	-23.403	-54.202	-21.524	-58.339	-24.640
Observations	2,138		1,380		1,763	
Pseudo R^2	0.140					

* For a definition of the variables see Data Appendix. Variables followed by HH refer to the household head and those followed by S refer to the spouse or partner.