

The decisions of Spanish youth: A cross-section study

Maite Martínez-Granado, Javier Ruiz-Castillo

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

Received: 18 September 1998/Accepted: 24 October 2000

Abstract. This 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 in order to understand the dynamics of household formation. Our results also confirm a number of plausible intuitions about the effect of individual characteristics and economic variables on these decisions, and provide some new insights into the reasons for young people in Spain remaining in large numbers in the parental home. Most of the results are gender independent.

JEL classification: J12, J21

Key words: Household formation, working and studying decisions, two stage estimation

1. Introduction

Contrary to Anglo-Saxon and central European countries but in line with other Southern European nations, in Spain the proportion of young people

All correspondence to Maite Martínez-Granado. Financial support from Projects 96-0118 and PB98-1058 of the Spanish DGES are gratefully acknowledged. We are grateful to Manuel Arellano for very helpful discussions and suggestions. We would also like to acknowledge comments from Cesar Alonso, Raquel Carrasco, Adriaan Kalwij, Sergi Jiménez, Jose María Labeaga, Barbara Petrongolo, María Jesús San Segundo and participants at the Universidad del País Vasco seminar and at the “Twelfth Annual Conference of the European Society for Population Economics” held in Amsterdam in June 1998. The paper benefited very much from the comments of two anonymous referees. All remaining errors are our own. *Responsible editor:* T. Paul Schultz.

living with their parents is very high. 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 dependents in the parental home, is one of the population subgroups with the highest social welfare indices in the distribution of household expenditures adjusted for household size. Jurado (1997) provides an interesting analysis of how regional variation in housing and labor market conditions, as well as education enrolment rates, are associated with regional variation in the rate of coresidence in a selected number of Spanish regions. However, we do not know of any econometric work that attempts an explanation of this important phenomenon using multivariate techniques.

A natural starting point for an economic analysis of household formation decisions by young adults is a utility comparison framework. Using a Nash bargaining model of family behavior, McElroy (1985) is the first paper to generate the indirect utility functions for this framework. The model is used to examine the joint determination of market work and family status of young men in the U.S. Among other things, the results indicate that the option to live in the parental household serves as "unemployment insurance". This provides a valuable insight into the Spanish case, where the unemployment rate among the young is one of the highest in the EU.¹

However, it should be noted that more than 50% of all young people coresiding in the parental home² have a job. It is true that the vast majority of this employment is temporary and not well paid. From this perspective, parents might be providing means to compensate for job insecurity, low wages and/or high housing costs. However, it is less well known that the family as a safety net in Spain also works the other direction. Sastre (1999) shows that, under the assumption that all household members pool their economic resources, young dependents' income contributes to a decrease in household income inequality. On the other hand, Cantó and Mercader (1999) show that the employed young dependents reduce the risk of poverty for the rest of the household members, particularly when the household head is out of work. In brief, there are many reasons why the living arrangements and the labor participation decisions in Spain should be treated simultaneously.

A general model would have to take into account the interaction between parents and their descendants' decisions in a dynamic context. The more complete dynamic model of household formation is the one by Rosenzweig and Wolpin (1993) (or RW for short). The authors formulate an altruistic, imperfect-foresight, overlapping generations model, incorporating human capital investments, interhousehold transfers, and household formation. The presence of the human capital investment dimension is important in our case, because Spain has one of the largest enrolment rates in higher education among the OECD countries. From this perspective, parents are helping to finance their sons and daughters' investment in human capital by providing them with shelter, and possibly other goods and services, while coresiding, or with direct transfers when their offspring are living apart. However, the investment in school in RW is simply modeled as an all-or-nothing option. Moreover, the labor/leisure margin is not included. Accordingly, all young people are assigned some gross potential earnings (in Mincer's sense), depending on their own accumulated human capital in school or at work. In a similar spirit but in a much simpler context, Ermish (1996) (or E for short) explicitly takes into account the public good aspect of housing, and incorporates the

human capital investment decision in a two period dynamic model in which the young adults face a binding borrowing constraint. Unfortunately, like RW, this model does not include the labor/leisure choice. Instead, in this work all young people are assigned some potential earnings, which varies with individual ability.³

Ideally, young people's decisions should be analyzed with panel data.⁴ Moreover, the theories just reviewed demand a rich data set. As RW point out, their model indicates that information on the contemporary life-cycle earnings of both generations and their components, the human capital investments of the second generation, the household living arrangements, and the parental transfers are, at a minimum, required. Unfortunately, our data are far from these standards. In this paper we use the 1990–1991 *Encuesta de Presupuestos Familiares* (EPF), a large household budget survey gathered by the *Instituto Nacional de Estadística* (INE) with the main purpose of estimating the aggregate weights of the Consumer Price Index. Like other similar surveys, the main shortcoming of the 1990–1991 EPF is that it lacks data on wealth and permanent income for all households and on parental income and parental characteristics for the young people living independently. On the bright side, our survey's information on individual and household characteristics is completed with a variety of variables that reflect potentially important regional differences in housing and labor market conditions.

Combining the residence and the parental transfer decisions, both RW and E lead to a number of mutually exclusive states, which can be analyzed by means of a multinomial empirical model. We could always attempt to extend this utility maximizing framework in order to make the two additional decisions we are interested in truly endogeneous: the labor participation decision, and the decision to pursue some education. This approach has the disadvantage of imposing a particular form of endogeneity and simultaneity on the decisions involved. Moreover, the small sample size of some of the categories does not allow the identification of all the parameters of interest with our data set.

Alternatively, in this paper we use a multivariate probit model in which it is the *propensity* to select a state, rather than its occupancy, which determines the probability that a state is actually occupied.⁵ In particular, we study the joint decision by young people aged 19–35 on whether to remain in the parental household, whether to work, and whether to keep on studying. Parental transfers and other non labor income are treated as exogenous variables conditioning these three simultaneous decisions. In principle, one could also attempt to endogenize the marriage decision. But we present evidence showing that, in the Spanish case, this is not necessary: the decision to marry or to leave the parental home almost always takes place simultaneously. Our probability model is estimated adapting a two-stage method proposed by Arellano and Bover (1997) in a related context.

Our results indicate that a rich pattern of interdependencies exists between the three decisions, and that both individual characteristics and economic variables play a significant explanatory role in the three propensities modeled in the paper. In particular, we establish that the following factors increase the propensity of young people to coreside with their parents: to be unemployed or inactive; to study; to have no income different from labor earnings or parental transfers; to have achieved a higher educational level; to live in a small village; to live in a region with high housing prices or little availability of rental hous-

ing, and to live in a region with a higher unemployment rate or a higher illiteracy rate. Most of these factors are common to male and female.

The rest of the paper is organized in four sections. Section 2 discusses in more detail the existing theory before presenting the empirical model. Section 3 is devoted to the data, Sect. 4 contains the results, and Sect. 5 concludes.

2. The empirical model

2.1. *The existing theoretical literature*

In order to structure the empirical analysis, our starting point is the single period framework of McElroy (1985) in which young people make decisions about living arrangements and labor force participation.⁶ In another strand of the literature during the late 1980s, Cox (1987) presents a utility maximizing model containing both altruistic and exchange motives for private *inter vivos* transfers. In Cox (1990), a simple two period model in which the parent has access to capital markets but the child does not, is used to analyze whether private intergenerational transfers function as loans or as subsidies that are used to help family units to overcome liquidity constraints. However, Cox's models ignore the decision concerning the coresidence of family members. But to the extent that households contain public goods that can be jointly consumed, residence sharing is cheaper for the parent than providing an almost equivalent service without coresidence.

These different strands of the literature converge in the type of full-blown dynamic theory of household formation due to RW. Combining the residence and transfer decisions, this model leads to three mutually exclusive states: (i) living apart-receiving parental transfers; (ii) coresiding, and (iii), living apart-not receiving parental transfers.⁷ As pointed out in the Introduction, we could always extend this framework to treat the decisions about labor participation and human capital investment endogenously. Consider, for example, a simple two period dynamic model in which all households consist of a one-child family. In this case, the three mutually exclusive states in RW and E become 12, depending on whether the child studies or not and, in addition, on whether s/he works or not in the first period. Even if we consider that parental transfers are predetermined variables for the young offspring, as we do in our empirical analysis, the number of states would still be 8.

What in RW, or in E is a reasonable specification, becomes an intractable one with our data when, together with the residence decision, we want to cover the labor participation and the human capital decisions as well. Moreover, the multinomial logit framework precludes us from asking whether the propensity (and/or the fact) that an adult child is employed has some explanatory role in the probability that this person lives independently. Analogously, this way of setting up the estimation problem forces us to ignore the possibility that the propensity (and/or the decision) to coreside with one's parents, enjoying the corresponding material advantages, might influence the probability that one finds a job in the labor market.

Our conclusion from the above discussion is that the multinomial logit empirical model, which can be rationalized quite directly with the help of the existing optimization theory, has certain disadvantages.⁸ Alternatively, we believe that it is fruitful to take as our starting point the idea that household

formation by young people is intrinsically linked to other decisions with regard to job or education status. Given this idea, the existing literature will be systematically used in the selection of explanatory variables and the interpretation of our empirical results.

2.2. The empirical model

We assume that the young decide simultaneously whether to remain in the parental household, whether to work, and whether to keep on studying. Empirically, there are different ways of expressing the interrelation among the three decisions. Our basic assumption is that it is the propensity to select a state, rather than its occupancy, which determines the probability that a state is actually occupied. In order to analyze the mutual influences among the three propensities, we consider the following simultaneous equations model in which each of the three propensities is defined in terms of the other propensities and a set of exogenous variables:

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

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

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

The variable I_1^* is the underlying individual propensity to leave the parental house, I_2^* represents the propensity to work, and I_3^* is the propensity to study. 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. This approach has the empirical attractiveness of allowing us to test the endogeneity and simultaneity of the three decisions.⁹

In the data set we observe the outcomes of the choices, not the underlying propensities. That is, we observe whether an 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:

$$I_{1i} = 1 \quad \text{if } I_{1i}^* > 0$$

$$I_{1i} = 0 \quad \text{otherwise} \quad (4)$$

$$I_{2i} = 1 \quad \text{if } I_{2i}^* > 0$$

$$I_{2i} = 0 \quad \text{otherwise} \quad (5)$$

$$I_{3i} = 1 \quad \text{if } I_{3i}^* > 0$$

$$I_{3i} = 0 \quad \text{otherwise} \quad (6)$$

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 Eqs. (1) to (3) and the observability conditions (4) to (6). Given the interdependence among the unobserved latent variables, we face a trivariate probit. Although estimation by numerical methods in a Simulated Maximum Likelihood routine could be used to achieve full efficiency, two-stage methods provide consistent estimates of the parameters of interest and they are easy to implement in spite of some efficiency loss.¹⁰ As discussed below, the method we use minimizes this efficiency loss.

Arellano and Bover (1997) propose a two-stage estimator for limited dependent variable models from panel data. In the first stage, reduced form equations for the endogenous variables are derived and estimated as independent probit equations,¹¹ in the second stage, the reduced form linear predictions replace all the unobservable latent variables.

This methodology can be readily extended to our case in which a simultaneous probability model must be estimated using a cross-section. This approach has two advantages over other methods that also estimate a probit on the second stage. First, given the assumption of unitary variance of the disturbances on the reduced form equations in these methods, the parameters of interest can only be recovered up to scale. However, because OLS in the second stage does not impose such identification conditions, the Arellano and Bover strategy allows us to recover the actual parameters without the scale restriction. Moreover, in an additional step, the Arellano and Bover method, allows for (i) the computation of a linear GMM estimator that is asymptotically efficient (relative to the first stage estimation), and (ii) the construction of a specification test for the overidentifying restrictions.

More explicitly, in the first stage, we consider the reduced form equations for the three endogenous variables,

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

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

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

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 Eqs. (7) to (9) are estimated by separate probit maximum likelihood, 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.

In the second stage, we use these 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:

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

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

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

Given the consistency and normality of the reduced form parameters, the set of estimates $\hat{\theta}$ is also consistent and asymptotically normal. However, since the dependent and the endogenous explanatory variables have been replaced by their predicted values, the asymptotic variance matrix of the estimates is *not* the traditional one for OLS estimators.

This second stage OLS estimator $\hat{\theta}$ can be interpreted as a GMM estimator in which the weighting matrix has not been chosen optimally. An efficient

estimation relative to the first stage estimates, $\hat{\Pi} = \begin{bmatrix} \hat{\Pi}'_1 \\ \hat{\Pi}'_2 \\ \hat{\Pi}'_3 \end{bmatrix}$, can be obtained in

a third stage by choosing optimally the weighting matrix as a consistent estimate of the inverse of the covariance matrix of the orthogonality conditions. A discussion of the consistency and normality of the estimates, the variance matrix and the election of the optimal weighting matrix can be found in Appendix A.¹²

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 means of inferring the interdependence among the three decisions considered. In that sense, we will refer to the model as “structural”, as opposed to the reduced form equations that are estimated in the first stage.

3. Data

In Tables 1 to 3, we present some basic statistics illustrating the differences between some Northern and Southern European countries (including Spain), plus the United States, in the three dimensions we are concerned with.¹³ Table 1 refers to the differences in living arrangements. We observe that in all countries a significant proportion of females leave the parental home before males of the same age. However, while in the U.S. or Germany, for instance, approximately 90% of the young live on their own at the age of 29, in the three Southern countries selected about 50% of the males and 25–35% of the females at that age are still living as dependants in their parents' house. In all countries,

Table 1. Proportion of men and women still living with parents by age group in six European countries – in 1986 – and the United States, in 1987 (in % of age group total)

	Men			Women		
	15–19	20–24	25–29	15–19	20–24	25–29
Spain	95.6	88.1	53.2	93.9	76.1	35.3
Italy	97.4	87.8	49.6	95.7	70.4	25.5
Greece	94.6	76.5	53.8	89.2	52.3	23.8
France	94.8	56.9	19.3	89.8	36.4	8.4
United Kingdom	93.6	57.2	21.9	87.8	33.8	8.6
Germany	94.8	64.8	27.4	92.0	42.8	11.0
United States		Up to 24: 27.4	Up to 29: 13.0		Up to 24: 22.2	Up to 29: 8.6

Source: For the six European countries, Tables 1 and 2 in Fernández Cordón (1997); for the U.S., Table 1 in Haurin et al. (1993).

Table 2. Unemployment rates by age groups in 1989

	14–24	25–29	30–34	Total rate
Spain	34.3	23.4	14.7	17.3
Italy	31.9	16.9	9.4	11.1
Greece	24.8	7.0	4.3	7.5
France	19.6	11.3	8.4	9.6
United Kingdom	10.3	8.9	7.1	7.4
Germany	5.5	6.2	6.4	5.7
United States	10.9(*)			5.3

Source: For the six European countries, Labour Force Survey; for the U.S., Current Population Survey. (*) U.S. data is for individuals from 16 to 24 years of age.

Table 3. Net enrolment in full-time public and private university education, by age group in 1991. Selected countries

	18–21	22–25	26–29
Spain	21.3	14.2	5.3
France	18.5	10.6	3.7
United Kingdom	12.4	3.0	0.9
Germany	6.8	14.7	9.3
United States	22.8	8.5	2.5

Source: *Education at a glance*, OECD, 1993. Data for Italy and Greece, not available.

the unemployment rate is higher for young people up to 29 years of age than for the population as a whole (Table 2), but this effect is larger in the Southern countries. Finally, Table 3 shows that as far as the net enrolment rate in university education is concerned, in 1991 Spain compares very favorably with the Northern countries, all of which are at a considerably higher level of economic development.

As we said in the Introduction, the data used in this paper comes from the 1990–1991 EPF. This is a household budget survey collected 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 is a representative sample consisting of 21,155 observations for a population of approximately 11 million households living in residential housing throughout Spain. 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 elderly, 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 – for more details on the 1990–1991 EPF, see INE (1992).

In view of Table 1, we choose 35 years of age as the upper bound in our definition of the young.¹⁴ 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 beyond secondary education. This gives us a sample of 9,741 males and 9,534 females. Ceuta and Melilla residents are not considered since some regional variables are not available for them.

All the endogenous and exogenous variables are described in the Data Appendix B. As far as the endogenous variables, we observe in Table B.1 that the proportion of young females living on their own is 44% versus only 33% among the males, while the proportion of females studying is almost 5 percentage points above the males. On the other hand, the female employment rate is only 37% as opposed to 63% for the males; however, although not shown here, the female unemployment rate is 28%, which is around ten percentage points higher than the male rate. Table B.2 shows the sample distribution according to the three dependent variables. With reference only to the minority groups, around 5% of males and females work and follow some type of studies at the same time, while only 2% of females and 0.5% of males study and are independent.

The exogenous variables entering Eqs. (1) to (3) are of two types: individual characteristics, and economic variables. We have information on the following individual characteristics: education, age, whether residing in a large city or in a small village, parental transfers, and other non labor income. We should point out that, in addition to coresidence, there are two ways parents can help their offspring: by a cash transfer, and by financing all or part of the housing services consumed by their descendants when they live independently. In fact, more than 80% of those independent males and females receiving any kind of transfer live in dwellings subsidized by some family member. Both types of parental transfers are assumed to be optimally selected by the parents. Hence, they are a predetermined variable for the young individuals. Only 2.19% of young people coresiding with their parents receive a cash transfer. Therefore, we are forced to add up both types of transfers, forfeiting the possibility of identifying the role of parental cash transfers. The final individual characteristic is the remaining non labor income. Unfortunately, it has been impossible to identify the effect of both variables separately. Therefore, in the final specification non labor income is the sum of parental transfers and other non labor income. In principle, both the parental transfers and the educational attainment could be considered endogenous variables. However, their exogeneity, as well as the exogeneity of the rest of explanatory variables is tested making use of the Sargan test of overidentifying restrictions proposed in Appendix A.

Our data source lacks information on three potentially important individual characteristics: the type of contract – temporary or indefinite – of those young people holding a job, the marital status and the fertility behavior of anyone different from the household head (defined as the household member with the highest earnings) and his/her spouse. The majority of the employed young people in Spain has a temporary job.¹⁵ Our lack of data in this respect precludes a study of the interaction between the decision to leave one's parents house made by those who are employed and the type of contract they have. However, it should be noted that even having the data, we could not simply include the type of contract 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

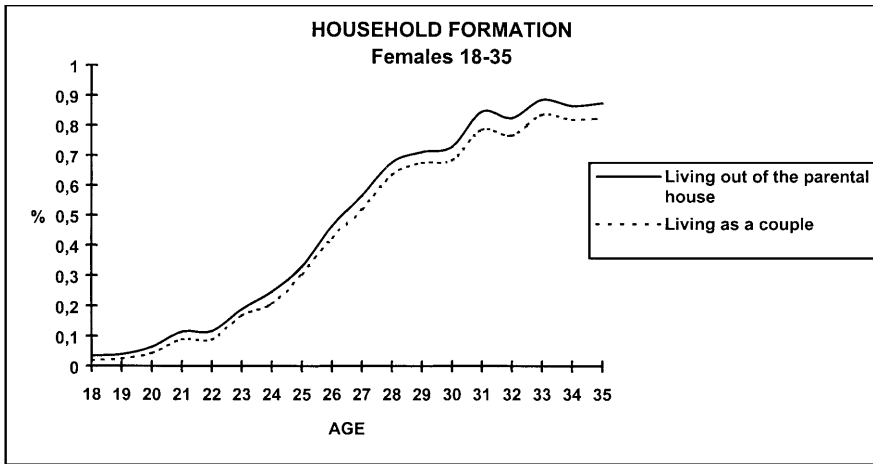


Fig. 1. Proportion of females living out of the parental house, by marital status

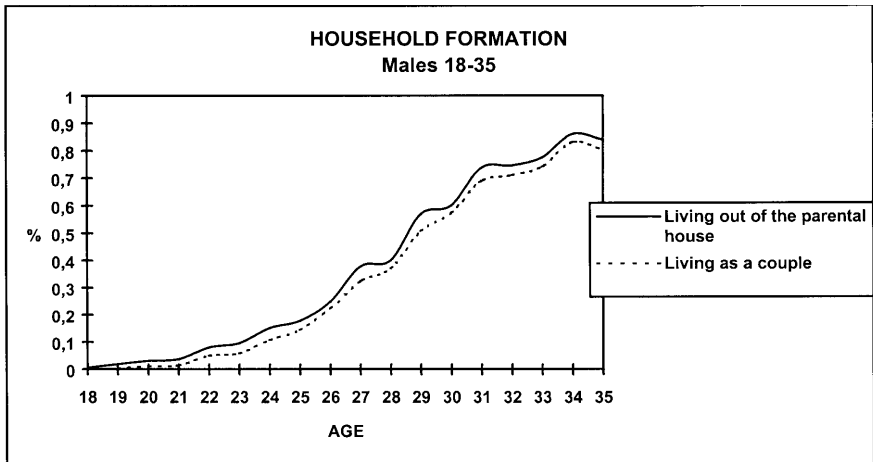


Fig. 2. Proportion of males living out of the parental house, by marital status

important explanatory variable of household formation – see, for example, Haurin et al. (1993). However, Figs. 1 and 2 indicate that in Spain almost all individuals that live outside their parents' household are married, which implies that the decision of marrying or leaving the parents' home almost always takes place simultaneously.¹⁶ Therefore, the only implication of not including marital status as an explanatory variable is that our results would refer, not only to the propensity to leave the parental house, but also to the propensity for getting married.

Finally the absence of fertility variables for those living with their parents could potentially be a shortcoming for interpreting the results, especially in the case of women's decisions. As in the case of the type of contract, fertility

decisions are likely to be endogenous and influenced by the same exogenous variables we use in our analysis.

We have two sets of economic variables, both of which are supposed to allow us to identify the equations of interest. In the first place, we have selected three variables which refer to housing conditions and capture the spatial variability of: (i) rental-equivalent values of the housing services provided by the whole stock, (ii) housing prices in the owner-occupied sector, and (iii) the relative importance of the rental housing sector, which is the one typically frequented by young people. We assume that these variables only affect directly the propensity for leaving the parental house. In the second place, we have considered the following three variables which capture the spatial variability in labor market conditions: (i) the unemployment rate for the population as a whole, (ii) the unemployment rate disaggregated by sex and age, and (iii) the illiteracy rate. We assume that these variables only affect the propensity for leaving the parental house through their effect on the propensities for working or studying. Descriptive statistics for the individual characteristics and economic variables can be found in Table B.1.

4. Results

4.1. *Simultaneity and identification*

As pointed out in Section 2.1, the existing theory suggests the interdependence among the decisions of working, leaving the parental house and studying. But one of the advantages of our approach is that we can evaluate whether a simultaneous model of the three decisions is called for. We do this by testing for pairwise independence of the equations using bivariate probits.¹⁷ Table B.3 in Appendix B reflects that for both males and females, we always find a significant correlation between the error terms of any two equations. Therefore, we conclude that the simultaneous model proposed in Section 2 has a sound empirical base.

We estimate separately each of the structural equations for males and females to reflect differences by gender. The reduced form estimates are presented in Tables B.3 and B.4 in Appendix B, but they are not discussed here because they do not provide any additional insight on the topic. Table 4 shows the optimal GMM estimates for our simultaneous equation system. In this step all endogenous variables have been replaced by their linear predicted values (see Appendix A). As previously stated, the economic variables allow us to identify the parameters of interest. In particular, the two variables on housing values and their interactions with age, as well as the variable reflecting the relative size of the rental housing market, only directly affect the propensity for living independently. Conversely, the regional unemployment, the provincial unemployment by age and sex, and the regional illiteracy rate only influence the decision of leaving the parental household through the effect of working and studying. These exclusion restrictions are not rejected by the Sargan test defined in Appendix A, whose value appears in the bottom line of Table 4. Moreover, the Sargan test does not allow us to reject the exogeneity of any of the variables since it implies that the orthogonality conditions imposed on the estimation are jointly not significantly different from zero.

Table 4. GMM optimal estimates

	Males		Females	
	Coeff	<i>t</i> -ratio	Coeff	<i>t</i> -ratio
<i>Work equation</i>				
Independence (P)	0.383	1.491	0.395	1.772
Studying (P)	-1.464	-6.443	-1.454	-8.598
Age	-1.479	-2.213	-2.576	-4.764
Age ²	0.426	2.128	1.087	4.836
Primary School	0.276	1.236	0.756	3.192
Secondary School	1.835	5.393	2.714	7.797
Higher Education	2.390	5.839	3.490	9.150
City	0.267	2.965	0.390	5.048
Village	-0.189	-1.657	-0.062	-0.605
Non labor Income	-0.039	-1.685	-0.064	-1.754
General UR	-2.524	-2.869	-1.739	-2.087
UR by sex and age group	0.132	0.321	0.435	0.883
Illiteracy rate	-4.544	-2.260	-3.379	-2.242
Intercept	-1.468	-2.806	-3.243	-7.412
<i>Independence equation</i>				
Working (P)	0.462	2.558	0.507	2.564
Studying (P)	-1.589	-3.084	-0.998	-2.453
Age	-0.492	-0.588	0.451	0.764
Age ²	1.203	2.669	0.545	1.444
Primary School	0.002	0.006	0.363	1.432
Secondary School	2.119	3.183	1.163	1.656
Higher Education	2.596	3.030	0.818	1.016
City	0.637	3.302	0.322	2.609
Village	-0.427	-3.590	-0.268	-3.153
Non labor Income	0.141	3.769	0.108	1.775
Owning Costs	-0.385	-2.134	-0.406	-3.312
Age x Owning Costs	0.355	1.190	0.903	3.655
Rental Values	-0.093	-0.341	0.004	0.015
Age x Rental Values	-1.218	-1.861	-2.058	-3.425
Rental Accommodation	2.016	2.999	1.056	1.620
Intercept	-4.133	-4.114	-1.819	-1.841
<i>Studying equation</i>				
Working (P)	-0.686	-7.486	-0.679	-11.039
Independence (P)	0.259	1.444	0.274	1.876
Age	-1.012	-2.494	-1.783	-5.854
Age ²	0.296	2.926	0.754	7.017
Primary School	0.189	1.160	0.514	3.173
Secondary School	1.253	8.287	1.861	11.463
Higher Education	1.631	10.897	2.391	12.955
City	0.181	4.235	0.268	6.731
Village	-0.130	-1.690	-0.042	-0.615
Non labor Income	-0.027	-1.661	-0.044	-1.977
General UR	-1.640	-2.522	-1.109	-2.050
UR by sex and age group	0.018	0.073	0.296	0.958
Illiteracy Rate	-3.172	-2.367	-2.508	-2.669
Intercept	-1.003	-3.876	-2.227	-12.941
Number Observations	9741		9535	
Sargan Test χ^2 (7)				
(Overidentifying restrictions)	5.839		6.892	
(<i>p</i> -value)	(0.559)		(0.440)	

(P): Predicted value; Age = (age-25)/10; UR: Unemployment Rate;

4.2. *Interdependence*

As expected, working increases the propensity to leave on one's own since individuals who work have access to the necessary funds that allow them to be independent from their family. By the same token, not working – i.e., being unemployed or inactive – increases the propensity to coreside. This is the important “unemployment insurance” effect discovered by McElroy (1985) for the 19–24, never married, out of school, U.S. males in 1971. On the other hand, living independently has a positive effect, although only marginally significant, on the propensity to work; that is to say, the increased costs that the individuals face when living on their own act as an incentive to work. Equivalently, it appears that coresiding is weakly associated with a greater propensity for being unemployed or out of the labor force.

Studying strongly increases the propensity to coreside.¹⁸ This result reflects the conditions of the publicly dominated Spanish university system: as in Greece, Italy, France or Portugal, tuition costs in public Universities are nominal – less than 500 Ecus per year; moreover, the spatial dissemination of College centers in Spain reduces the costs of studying through coresidence. The opposite result in Ermish (1996) reflects the British University system, which strongly encourages people to study in a different locality from their parents.¹⁹ On the other hand, although not very significant, the positive effect of living apart on the propensity to study is the only puzzling result in this subsection. We have to bear in mind that we are considering every type of course not only official studies. In particular, among the women living independently, 47% report following “other studies” different from primary, secondary, or College studies.

Unsurprisingly, the propensity to work strongly reduces the propensity to study. Conversely, to follow any type of studies significantly reduces the propensity to work.

It is important to emphasize that the nature of the interaction between the three pairs of decisions is similar for both men and women.

4.3. *Individual characteristics*

Age, although significant in most equations, does not have a clear interpretation since it enters the equations through many channels (age, age squared, the group specific unemployment rate, and interactions with housing prices).

The role of the education variables is important. In the first place, the higher the level of education attained by the individual, the higher the propensity to work – a plausible result. In the second place, given that the level of education can be considered a predetermined variable in the equation for studying, it is also plausible that the higher the level of education attained the higher the propensity to continue studying. In the third place, the effect of education on the propensity to leave the parental house has a differentiated effect for males and females. For males, education has a positive independent effect on the probability of forming a household, which reflects the fact that the more educated individuals are also those with higher earnings and, presumably, the more attractive partners in the marriage market. There is also a strong indirect effect through the increase on the propensity of working, offset by the increase in the

propensity to study. For women, the indirect effects work in the same direction and are even stronger, but education does not have an independent effect on its own: only if the woman has finished secondary school is she more likely to be living on her own. Since most individuals living apart from their parents are married, this decreasing effect of education may merely reflect the postponement of marriage decisions by more educated women.

Parental transfers and other non labor income have a positive effect on the probability of leaving the parental home: 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. On the other hand, non labor income has the usual negative effect on the propensity to work (and to study). This effect is stronger for females, possibly because a larger non labor income allows them to spend their time in other activities such as taking care of the family or having children.

Living in a large city increases the propensity to form a household, while living in a small village decreases it. This effect is especially strong for males. It probably reflects the fact that the traditional pattern of “extended families”, where several generations cohabit under the same roof and cooperate in the same productive activities, is more prevalent in rural areas. On the other hand, the effect of the municipal size on the propensity to work and to study reveals that there are more jobs and opportunities to study as the municipal size increases. However, the effect pattern differs by gender. For males, relative to medium sizes municipalities, there is a negative effect in small villages, and a positive effect in big cities. For females, there is a clear discontinuity: there is only a significant strong positive effect in big cities.

4.4. *Economic variables*

All the variables reflecting the spatial variation in housing conditions have the expected influence on the propensity to live apart. In the first place, the higher the rental equivalence value of the whole stock and, above all, the higher the owner-occupied housing prices are, the lower the propensity to live independently. The difference between these two variables is that the effect of the second one decreases with age. In the second place, the variable reflecting the availability of rental housing has a strong and significant positive effect on the probability of living independently. These results are in agreement with those obtained in Börsch-Supan (1985), Haurin et al. (1993, 1994), Ermish and Di Salvo (1997) and Ermish (1999), emphasizing the role of housing costs – and not only own income – as a basic determinant of the household formation and related demographic decisions.

The gender and age group specific unemployment rate does not have any significant effect on the propensity to work: it appears that individuals perceive the general unemployment rate as the relevant variable. Of course, the general unemployment rate has a significant negative effect on the propensity to work, revealing the lower probability of receiving a job offer, as well as the discouraging worker effect that reduces the effort of looking for a job. Note that the indirect effect of unemployment on coresidence by lowering the propensity to work reinforces the role of the family in Spain as a cushion in the face of unfavorable labor market conditions.²⁰ On the other hand, given

job status, the unemployment rate has a negative effect on the probability of studying for both men and women. Therefore, unemployment works in two ways: it reduces the probability of finding a job, favoring the studying option (opportunity cost effect), but it *directly* reduces the probability of studying (discouraging effect due to the poorer job perspectives).

The illiteracy rate has a strong negative effect on the propensity to study and to work. The first effect is consistent with different explanations. It can reflect a peer effect, so that young individuals in areas where not many people have studied in the past also tend not to study. It could also be reflecting how regions with a more educated stock of human capital have developed a wider net of possibilities for the young people to continue their studies after the compulsory age. Finally, the negative effect on the propensity to work can reflect the fact that areas with a more educated workforce may have better job opportunities through higher investment (and job creation) by the firms settled there – for a theoretical exposition of this idea, see Acemoglu (1996).

5. Conclusions

As far as we know, this paper constitutes the first econometric attempt to explain why parents and young descendants between 18 and 35 years of age decide to live together in Spain in rather large proportions. Lacking longitudinal data, we have found it interesting to work with a sufficiently rich, large and readily available household budget survey – the 1990–1991 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 of whether to continue 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-section data. Our results indicate that the behavior of Spanish youth is amenable to careful empirical analysis with standard tools.

From a methodological point of view, we have shown that the endogenous and simultaneous treatment of the three decisions should occupy the core of any attempt to understand the issues involved. Otherwise, seriously misleading results could be obtained, a possibility we have illustrated when the propensities to work or to study are treated as mere exogenous variables influencing the decision to abandon the parental household.

As pointed out in the Introduction, we have confirmed that there exists a rich pattern of interdependencies between the three decisions, and that both individual characteristics and economic variables have a significant explanatory role in the three propensities modeled in the paper. More specifically, our results show that the pattern of interdependencies among the three decisions, as well as the effect of the economic variables, are qualitatively the same for males and females. In particular, we have found support for the following regularities: (i) parents help their young offspring through coresidence when the latter do not have a job or are studying; (ii) living independently has a positive effect on the propensity to work, (iii) housing conditions significantly affect

the living arrangements of the young in a direct way, while (iv) unemployment exerts its influence indirectly through its negative effect on the propensities to work and to study.

However, there are also subtle differences in gender behavior. (i) More educated women tend to postpone longer than men the decision to marry and form a new household; (ii) increases in non labor income have a stronger negative effect on the propensities to work and to study for women, and (iii) the decrease in the propensities to work, to study and to form a new household in a small village is stronger for men, while the increase in the first two propensities in a large city is greater for women.

This paper has several obvious shortcomings. In the first place, as we have pointed out from the beginning, even in a static framework the data we have used is rather incomplete. It lacks information on potentially important individual characteristics of young people, as well as on a host of fundamental variables reflecting the family background of the individuals living apart from their parents. More importantly, perhaps, is that all of our cross-section results are rather suspect because they reflect the combined impacts of both inflows and outflows from a certain work, study or living arrangement situation. In this respect, the present study can only be taken as a first step towards an understanding of the decisions of the Spanish youth that we have focused on here.

In the second place, we have been unable to provide any explanation for the role of the Spanish family, mentioned in the Introduction, as a safety net in the opposite direction of the one usually stressed. We refer to the evidence indicating that young people with a job in poorer households are making a decisive contribution in raising the standard of living of the remaining household members. To study this question, one would certainly need data on such "reverse transfers" – as well as data on the parental socioeconomic characteristics of all the young people in the population. Furthermore, to rationalize these transfers in a theoretical model, one could follow one of two routes: to postulate strong altruistic motives on the part of the young offspring; or to extend the existing exchange models substituting transfers from the young for parental loans. As a matter of fact, this line of thought might lead to a better understanding of the fact revealed by Rosenzweig and Wolpin (1993) and Ermish (1996), that, contrary to existing theory, higher parental income increases the probability of their offspring living independently.

In the third place, only an appropriate comparative study including non Southern European countries, would allow us to estimate the effect that institutions and public policies may have on the behavior of the young. Meanwhile, we have already mentioned that public policy in Spain favors high enrolment rates in public universities. To this we may add that, as pointed out in Cantó and Mercader (1999), social protection in Spain has developed during the 1980s, maintaining the pension and the unemployment subsidy programs at the center of the system. Thus, public protection for unemployed youth is either not available at all in the case of the "first-job-seekers" (35 and 28% of the unemployed males and females in our sample, respectively), or very limited indeed for "early-age-unemployed" holding mere temporary jobs. Moreover, general family support systems and child-care policies in particular are underdeveloped in the Spanish welfare state. Finally, housing policies tend to favor owner-occupied housing, which is the most inaccessible tenure choice for the young. One may assume, of course, that all these policies tend to reinforce the importance of coresidence in Spain.

Appendix A: The Arellano and Bover (1997) estimator

Consider Eqs. (1) to (3) from Section 2:

$$I_{1i}^* = X_{1i}'\beta_1 + \delta_{12}I_{2i}^* + \delta_{13}I_{3i}^* + u_{1i} \tag{1}$$

$$I_{2i}^* = X_{2i}'\beta_2 + \delta_{21}I_{1i}^* + \delta_{23}I_{3i}^* + u_{2i} \tag{2}$$

$$I_{3i}^* = X_{3i}'\beta_3 + \delta_{31}I_{1i}^* + \delta_{32}I_{2i}^* + u_{3i}. \tag{3}$$

They can be rewritten as

$$\begin{bmatrix} I_{1i}^* \\ I_{2i}^* \\ I_{3i}^* \end{bmatrix} = \begin{bmatrix} W_{1i} & O & O \\ O & W_{2i} & O \\ O & O & W_{3i} \end{bmatrix} \times \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} + \begin{bmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \end{bmatrix}, \tag{4}$$

where $W_{1i} = (X_{1i}' I_{2i}^* I_{3i}^*)$ is a $1 \times (k_1 + 2)$ vector, $W_{2i} = (X_{2i}' I_{1i}^* I_{3i}^*)$ is a $1 \times (k_2 + 2)$ vector and $W_{3i} = (X_{3i}' I_{1i}^* I_{2i}^*)$ is a $1 \times (k_3 + 2)$; O are vectors of zeros that conform the W_{ji} . More compactly, (4) can be written as

$$I_i^* = \mathbf{W}_i\delta + u_i. \tag{5}$$

Let $\mathbf{X}_i = (I_3 \otimes X_i)$, where I_3 is a (3×3) identity matrix, and X is the set of all different exogenous variables in X_{1i} , X_{2i} and X_3 . Since the error term in expression (5) is uncorrelated with the exogenous variables, we can write

$$E\{\mathbf{X}_i'(I_i^* - \mathbf{W}_i\delta)\} = 0. \tag{6}$$

Using the law of iterated expectations,

$$E\{\mathbf{X}_i'[E(I_i^*|\mathbf{X}_i) - E(\mathbf{W}_i|\mathbf{X}_i)\delta]\} = 0, \tag{7}$$

where

$$E(\mathbf{W}_i|\mathbf{X}_i) = \begin{bmatrix} X_{1i}' & \pi_2 X_i & \pi_3 X_i & O & 0 & 0 & O & 0 & 0 \\ O & 0 & 0 & X_{2i}' & \pi_1 X_i & \pi_3 X_i & O & 0 & 0 \\ O & 0 & 0 & O & 0 & 0 & X_{3i}' & \pi_1 X_i & \pi_2 X_i \end{bmatrix},$$

is a $(3 \times (k_1 + k_2 + k_3 + 6))$ matrix, and $E(I_i^*|\mathbf{X}_i) = \begin{bmatrix} \pi_1' X_i \\ \pi_2' X_i \\ \pi_3' X_i \end{bmatrix} = \Pi X_i$.

This suggests to consider GMM estimators of δ based on the following sample orthogonality conditions:

$$b_N(\delta) = 1/N \sum_{i=1}^N \mathbf{X}_i'(\hat{I}_i - \hat{\mathbf{W}}_i\delta), \tag{8}$$

where I_{1i}^* , I_{2i}^* and I_{3i}^* are replaced by their linear predictions from the first stage

independent probit estimates. A GMM estimator of δ based on (8) takes the form

$$\tilde{\delta}_A = \left[\left(\sum_i \hat{\mathbf{W}}_i' \mathbf{X}_i \right) A_N \left(\sum_i \mathbf{X}_i' \hat{\mathbf{W}}_i \right) \right]^{-1} \left[\left(\sum_i \hat{\mathbf{W}}_i' \mathbf{X}_i \right) A_N \left(\sum_i \mathbf{X}_i' \hat{I}_i \right) \right], \quad (9)$$

where A_N is a weighting matrix. When $A_N = (\sum_i \mathbf{X}_i' \mathbf{X}_i)^{-1}$, $\tilde{\delta}_A$ coincides with an OLS estimator, $\hat{\delta}$, applied to Eq. (5), where the endogenous variables have been replaced by their linear predictions from the independent probit estimates, that is,

$$\hat{\delta} = \left(\sum_i \hat{\mathbf{W}}_i' \hat{\mathbf{W}}_i \right)^{-1} \left(\sum_i \hat{\mathbf{W}}_i' \hat{I}_i \right). \quad (10)$$

Before discussing the consistency and asymptotic normality of $\tilde{\delta}_A$, we need an expression for the asymptotic variance of $b_N(\delta)$. Notice that (8) can be written as

$$\begin{aligned} b_N(\delta) &= 1/N \sum_{i=1}^N \mathbf{X}_i' (\hat{\Pi} X_i - \hat{\mathbf{W}}_i' \delta) \\ &= 1/N \sum_{i=1}^N \mathbf{X}_i' (\Gamma \hat{\Pi} X_i - X_i^* \beta)' \end{aligned} \quad (11)$$

with $\Gamma = \begin{bmatrix} 1 & -\delta_{12} & -\delta_{13} \\ -\delta_{21} & 1 & -\delta_{23} \\ -\delta_{31} & -\delta_{32} & 1 \end{bmatrix}$, $X_i^* = \begin{bmatrix} X_{1i}' \\ X_{2i}' \\ X_{3i}' \end{bmatrix}$, and $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}$. Equations (1) to (3) can be conveniently rewritten as

$$\Gamma I_i^* = X_i^* \beta + u_i. \quad (12)$$

A comparison of expression (12) with the reduced form $I_i^* = \Pi X_i + \varepsilon_i$, pre-multiplied by Γ , proves that the following restriction holds:

$$X_i^* \beta = \Gamma \Pi X_i. \quad (13)$$

Using this restriction in expression (11), we obtain

$$\begin{aligned} b_N(\delta) &= 1/N \sum_{i=1}^N (I_3 \otimes X_i) (\Gamma \hat{\Pi} X_i - \Gamma \Pi X_i) \\ &= 1/N \left(I_3 \otimes \sum_{i=1}^N X_i \right) (\Gamma \otimes I_m) \text{vec}(\hat{\Pi} - \Pi), \end{aligned} \quad (14)$$

¹ The numerical equivalence follows from the fact that the columns in $\hat{\mathbf{W}}_i$ are linear combinations of those in \mathbf{X}_i

and consequently

$$b_N(\delta) \sim_a N(0, E(\mathbf{X}'_i \mathbf{X}_i) V^* E(\mathbf{X}'_i \mathbf{X}_i)), \tag{15}$$

with $V^* = (\Gamma \otimes I_m) \text{var}(\text{vec}(\hat{\Pi}))(\Gamma' \otimes I_m)$. Hence, it can be proven that

$$\sqrt{N}(\tilde{\delta}_A - \delta) \sim_a N(0, (M'_{\mathbf{XW}} A_N M_{\mathbf{XW}})^{-1} M'_{\mathbf{XW}} A_N V_b A_N M_{\mathbf{XW}} (M'_{\mathbf{XW}} A_N M_{\mathbf{XW}})^{-1}), \tag{16}$$

where $M_{\mathbf{XW}} = p \lim \left(\frac{\mathbf{X}'_i \mathbf{W}_i}{N} \right)$, and $V_b = \text{var}(b_N)$. A consistent estimate of the asymptotic variance of $\tilde{\delta}_A$ is given by

$$\begin{aligned} AVAR(\tilde{\delta}_A) &= (\hat{M}'_{\mathbf{XW}} A_N \hat{M}_{\mathbf{XW}})^{-1} \hat{M}'_{\mathbf{XW}} A_N \hat{V}_b A_N \hat{M}_{\mathbf{XW}} (\hat{M}'_{\mathbf{XW}} A_N \hat{M}_{\mathbf{XW}})^{-1}, \end{aligned} \tag{17}$$

where $\hat{M}_{\mathbf{XW}}$ and \hat{V}_b are consistent estimates of $M_{\mathbf{XW}}$ and V_b respectively. The most efficient estimator relative to $\hat{\Pi}$ is obtained by choosing optimally A_N as a consistent estimate of the inverse of the covariance matrix of the orthogonality conditions

$$A_n = \hat{V}_b^{-1} = (M_{\mathbf{XX}} \text{var}(\text{vec}(\hat{\Pi})) M_{\mathbf{XX}})^{-1}. \tag{18}$$

Now we only need an estimate for $\text{var}(\text{vec}(\hat{\Pi}))$. Let us consider

$$\text{vec}(\hat{\Pi}) = \begin{pmatrix} \hat{\pi}_1 \\ \hat{\pi}_2 \\ \hat{\pi}_3 \end{pmatrix} = \arg \max(L) = \arg \max(L_1 + L_2 + L_3), \tag{19}$$

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

$$\begin{aligned} &\left(-1/N \text{diag} \left(\frac{\partial^2 L_j}{\partial \pi_j \partial \pi'_j} \right) \right) \sqrt{N} \text{vec}(\hat{\Pi} - \Pi) \\ &= 1/\sqrt{N} \sum_i \begin{pmatrix} \partial L_{i1}/\partial \pi_1 \\ \partial L_{i2}/\partial \pi_2 \\ \partial L_{i3}/\partial \pi_3 \end{pmatrix} + O_p(1), \end{aligned} \tag{20}$$

which suggests an estimate for the variance of the form

$$\text{var}(\text{vec}(\hat{\Pi})) = \hat{H}^{-1} \hat{\Psi} \hat{H}^{-1}, \tag{21}$$

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

Finally, under the null hypothesis of lack of misspecification, a test statistic of the overidentifying restrictions can be derived for the optimal GMM estimator,

$$S = N \left(\sum_i^N \hat{u}_i' \mathbf{X}_i \right) \hat{V}_b^{-1} \left(\sum_i^N \mathbf{X}_i' u_i \right) \sim \chi_{3m - (k_1 + k_2 + K_3 + 6)}^2, \quad (22)$$

where $\hat{u}_i = \hat{I}_i^* - \hat{\mathbf{W}}_i \tilde{\delta}_A$.

In brief, the method discussed above works as follows:

1. First, the reduced form estimates are obtained.
2. Second, using these reduced form estimates, a sub-optimal GMM estimator is implemented that allows us to compute a consistent estimate of V_b .
3. The consistent estimate of V_b and the reduced form estimates are used to obtain the most efficient estimates relative to the first-stage reduced form ones.

Appendix B: Data

Endogenous variables

Independent: Dummy variable that equals one if the individual does not live in the parental home.

Work: Dummy variable that equals one if the individual is working (full or part time) at the interview date.

Studying: Dummy variable that equals one if the individual is carrying on any type of education. It is worth noting that 21.9% of female students are said to attend “other type of education”, different from primary, secondary, and College education, while only 14.9% of male students declare to do so.

Exogenous variables

- *Individual characteristics*:

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).

Parental transfers: Regular or occasional cash transfers, plus housing subsidies received by the young offspring living apart.

Other non labor income: This is the summation of all types of current income sources, different from parental transfers and labor earnings or public sub-

sities related to the economic activity (like the unemployment compensation). It includes lotteries, returns on capital (positive and negative when borrowing), school grants, and other public transfers.

Non labor income: Parental transfers plus other non labor income.

• *Economic variables:*

Rental values: Regional average across the 50 Spanish provinces (excluding Ceuta and Melilla) of: (i) annual rents by square meter actually paid in rental housing, and (ii) self-imputed annual rents by square meter for owner-occupied and other non-rental housing. Source: 1990–1991 EPF.

Owning costs: Regional average across the 17 Spanish *Comunidades Autónomas* (excluding Ceuta y Melilla) of house prices by square meter. Source: “*Precio medio del m² de las viviendas*”, Ministerio de Fomento.

Rental accommodation: Regional percentage (across the 50 Spanish provinces) of rental accommodation. Source: 1990–1991 EPF

Unemployment: Regional unemployment rates (across the 17 Spanish *Comunidades Autónomas*) for the population as a whole and disaggregated by sex and age. Source: *Encuesta de Población Activa*.

Illiteracy rate: Number of illiterate individuals older than 10 per thousand inhabitants in every province. Source: *Encuesta de Población Activa*.

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

	Males		Females	
	Mean	Standard error	Mean	Standard error
Endogenous variables				
Independent	0.334	0.472	0.446	0.497
Work	0.635	0.481	0.372	0.483
Studying	0.224	0.417	0.271	0.444
Exogenous variables				
<i>Individual characteristics</i>				
Age	25.8	5.2	26.0	5.2
Educ2	0.499	0.500	0.474	0.499
Educ3	0.355	0.479	0.350	0.477
Educ4	0.105	0.306	0.133	0.339
City	0.532	0.499	0.557	0.497
Village	0.145	0.352	0.129	0.335
Non labor income (N.L.I.)	103,607	282,892	168,636	344,843
• Parental transfers only	300,538	257,160	334,997	423,849
(Observations %)	(4.54%)		(5.86%)	
• Other N.L.I. only	29,723	242,582	57,822	217,738
(Observations %)	(17.16%)		(12.42%)	
• Both	328,842	353,161	445,495	441,016
(Observations %)	(1.66%)		(1.45%)	
<i>Economic variables</i>				
Rental values	3,287	1,344	3,345	1,369
Owning costs	86,987	22,426	87,476	22,987
Rental accommodation	0.117	0.041	0.117	0.041
Regional unemployment (RU)	16.927	5.468	16.796	5.459
RU by age-males	17.533	11.187		
RU by age-females			30.918	12.027
Illiteracy rate	0.041	0.026	0.041	0.026
Sample size	9,741		9,535	

Table B.2. Individual distribution according to dependent variables

	Males		Females	
	Number of observations	%	Number of observations	%
Working = 1, studying = 0, independent = 0	2984	30.63	1,675	17.57
Working = 0, studying = 1, independent = 0	1,631	16.74	1,931	20.25
Working = 0, studying = 0, independent = 0	1,609	16.52	1,385	14.53
Working = 1, studying = 0, independent = 1	2,695	27.67	1,417	14.86
Working = 0, studying = 0, independent = 1	268	2.75	2,475	25.96
Working = 1, studying = 1, independent = 0	268	2.75	295	3.09
Working = 1, studying = 1, independent = 1	242	2.48	161	1.69
Working = 0, studying = 1, independent = 1	44	0.45	196	2.06

Table B.3. Bivariate probit correlation coefficients

	Males		Females	
	ρ	LR test(*)	ρ	LR test(*)
Work – independence	0.448	390.239	-0.157	59.324
Work – studying	-0.575	726.757	-0.466	524.355
Independence – studying	-0.099	11.329	-0.287	128.212

(*) The likelihood ratio test of $H_0 : \rho = 0$ follows a $\chi^2(1)$

Table B.4. Reduced form first stage estimates: females

	Work		Independence		Study	
	Coeff	t-ratio	Coeff	t-ratio	Coeff	t-ratio
Age	-0.135	-1.188	1.660	10.469	-1.256	-8.823
Age ²	-0.341	-5.534	-0.498	-5.768	0.845	11.087
Primary school	0.367	5.028	0.231	2.970	0.338	2.443
Secondary school	0.411	5.446	-0.154	-1.905	1.552	11.268
Higher education	0.907	11.293	-0.378	-4.410	1.681	11.860
Regional UR	-0.695	-1.516	-0.171	-0.324	-1.128	-1.965
Regional UR – by age (females)	-1.216	-5.358	-1.311	-4.716	0.942	3.675
City	0.003	0.097	0.049	1.153	0.264	5.956
Village	-0.064	-1.382	-0.241	-4.334	-0.048	-0.784
Non labor income	-0.046	-4.457	0.073	5.744	0.008	0.747
Owning costs	-0.037	-0.414	-0.384	-3.341	-0.063	-0.596
Age \times owning costs	-0.178	-1.203	0.602	2.855	0.212	1.100
Rental values	-0.122	-0.652	-0.072	-0.312	0.208	0.952
Age \times rental values	1.443	5.788	-0.468	-1.342	-0.773	-2.376
Rental accommodation	0.647	1.921	1.186	2.954	0.698	1.665
Illiteracy rate	-0.423	-0.535	2.420	2.600	-0.495	-0.500
Intercept	-0.177	-1.389	0.203	1.349	-2.241	-12.12
Log-likelihood	-5,956		-4,050		-3,728	
Number observations			9535			

Table B.5. Reduced form first stage estimates: males

	Work		Independence		Study	
	Coeff	<i>t</i> -ratio	Coeff	<i>t</i> -ratio	Coeff	<i>t</i> -ratio
Age	1.024	8.006	2.049	11.485	-1.296	-8.855
Age ²	-0.685	-9.957	-0.269	-2.626	0.670	8.241
Primary school	0.591	8.311	0.441	5.390	-0.087	-0.733
Secondary school	0.259	3.554	0.378	4.445	1.192	10.136
Higher education	0.203	2.489	0.231	2.522	1.571	12.646
Regional UR	-3.909	-7.710	-2.694	-5.012	-0.313	-0.516
Regional UR – by age (males)	-0.288	-1.051	-0.062	-0.164	0.448	1.431
City	-0.206	-5.378	0.025	0.581	0.318	7.025
Village	-0.107	-2.227	-0.272	-4.942	-0.122	-1.934
Non labor income	-0.077	-7.510	0.046	4.121	0.038	3.324
Owning costs	-0.163	-1.729	-0.394	-3.097	-0.080	-0.730
Age <i>X</i> owning costs	0.047	0.277	0.176	0.781	0.209	1.055
Rental values	0.430	2.126	0.358	1.430	-0.061	-0.270
Age <i>X</i> rental values	0.923	3.142	0.239	0.625	-0.627	-1.819
Rental accommodation	-0.147	-0.408	1.480	3.674	-0.017	-0.041
Illiteracy rate	2.176	2.604	6.162	6.531	-2.484	-2.525
Intercept	0.872	6.929	-0.915	-6.086	-1.714	-10.20
Log-likelihood		-5,086		-3,859		-3,571
Number observations				9741		

Endnotes

- ¹ For an early exposition of this idea in Spanish literature, see Revenga (1991). For more recent analysis, see Robinson (1998) and Toharia et al. (1998).
- ² Among the young people studied in this paper who reside in a household headed by someone else, almost 7% live in a household headed by a relative, other than a parent. However, to simplify the terminology, in what follows we will refer to all such households as the “parental” household.
- ³ On the other hand, Ermish and Di Salvo (1997) and Ermish (1999) develop a model of household formation in which parents are altruistic about their children, and housing is a local public good in the sense that housing services per person are not affected by household size. The theoretical model derives predictions about the impact of the price of housing, young adults’ income and parental income on the probability that a young adult lives apart from his/her parents.
- ⁴ Recent empirical work in this area is of this type. RW use the kinship-linked cohorts of 9 National Longitudinal Surveys from 1967 to 1981 in the U.S. Their sample consists of more than 5,000 young men who were ages 14–17 at the time of the 1966 interview. E uses people aged 16–29 in the first four waves of the 1991–1994 British Household Panel Survey (BHPS). Ermish and di Salvo (1997) use a sample of 10,500 British people born in 1958 and interviewed in 1991 when they were 33 years old. Ermish (1999) focus on persons aged 16–30 from the first five waves of the 1991–1995 BHPS.
- ⁵ This is the Ashford and Sowden (1970), Amemiya (1975) and Zellner and Lee (1965) model, discussed as *Case 3* in Heckman (1978). For the panel data case, see Heckman (1981).
- ⁶ Within the empirical literature, Haurin et al. (1993, 1994) also insist that the child’s income is endogenous because labor supply is jointly decided along household formation.
- ⁷ As far as the fourth possible state, in RW transfers are always positive whenever the two generations live together, while in E there is no data on non-housing transfers to the coresiding child.
- ⁸ In a multinomial logit framework, data limitations would force us to ignore the human capital investment decision. At any rate, estimates of a 4 state multinomial logit arising from the labor force participation and the household formation decisions are available upon request.

- ⁹ Similar simultaneous equation models have been used in other contexts in order to analyze the relationship between disability and labor force participation (Stern 1989), or the relationship between employment and care giving to elderly parents (Wolf and Soldo 1994).
- ¹⁰ Mallar (1977) or Heckman (1978), among others, propose two-stage methods in this type of situation. For a revision of these methods, see Maddala (1983).
- ¹¹ Independent equation estimation at the first-stage will in general yield inefficient estimates because it ignores the possible dependence among the equations through the error terms, u_{ji} . Taking into account such dependence would typically require the use of simulation based estimators (see Hajivassiliou and Ruud 1994).
- ¹² If the three decisions in Eqs. (1) to (3) were not simultaneous and we were only interested in the study of the decision of leaving the parental household, a single equation with two endogenous dummies could also be estimated. However, the two-stage estimators we have discussed could not be implemented, and some simulation technique of estimation had to be applied. Finally, note that a multivariate probit with endogenous dummies will also complicate the estimation.
- ¹³ The six European countries are the ones studied in Fernández Cordón (1997) using the Labor Force Survey. For another comparative study between Northern and Southern European countries, using the European Household Panel, see Iacovou (1998).
- ¹⁴ Results for an upper bound of 30 years are available on request. They are not very sensitive to the choice of the upper bound, especially for males. For females, although in general the sign and size of the estimated parameters are comparable for both age groups, the overidentifying restrictions are marginally rejected. Moreover, the independence variable, that is marginally significant for women between 18–35 in the equations of working and studying losses completely its significance when considering only women between 18 and 30 years of age.
- ¹⁵ Table 2 of Cantó and Mercader (1999) show the entry-level jobs in 1996 by new school leavers (aged 16 to 29) one year after leaving education in different selected European countries. Spain is an outlier: 80% of this group hold a temporary job. On the other hand, according to Güell and Petrongolo (2000), in 1990 less than 15% of all temporary contracts become indefinite. Finally, Jimeno and Toharia (1993) and De la Rica and Felgueroso (1999) find that the temporary workers earn approximately 10% less than permanent ones, after controlling for observable personal and job characteristics.
- ¹⁶ This is confirmed by all demographic studies in the subject. See, for instance, Vergés (1997) or Jurado (1997).
- ¹⁷ We thank an anonymous referee for this suggestion.
- ¹⁸ The results are substantially different when ignoring the endogeneity of the working and studying decisions. For women, we find a positive effect of the propensity to work and to study on the probability of coresidence. For males, working provides incentives to leave the parental house but studying has no effect on the household formation decision.
- ¹⁹ According to Ermish (1996), the positive effect found by Rosenzweig and Wolpin (1993) may reflect the different definition of coresidence among students: young men who are at college and away from home are classified as coresiding with their parents if they report themselves as attached to their parents' household. Nevertheless, in their fixed effects multinomial logit model, Rosenzweig and Wolpin (1993) also find a positive effect of attending college on the probability of receiving a parental transfer.
- ²⁰ Rosenzweig and Wolpin (1993) find a positive but small direct effect of the weeks spent in unemployment on coresidence, as well as on the parental transfers when living apart. Thus, in the U.S., the family acts as a cushion in bad times independent of living arrangements.

References

- Acemoglu D (1996) A Microfoundation for Social Increasing Returns in Human Capital Accumulation. *Quarterly Journal of Economics* 111(3):779–804
- Amemiya T (1975) Qualitative Response Models. *Annals of Economic and Social Measurement* 4:363–372
- Arellano M, Bover O (1997) Estimating Dynamic Limited Dependent Variable Models from Panel Data. *Investigaciones Económicas XXI*(2):141–165

- Ashford JR, Sowden RR (1970) Multivariate Probit Analysis. *Biometrics* 26:535–546
- Börsch-Supan A (1986) Household Formation, Housing Prices, and Public Policy Impacts. *Journal of Public Economics* 30:145–164
- Cantó-Sánchez O, Mercader-Prats M (1999) Poverty among Children and Youth in Spain: The Role of Parents and Youth Employment Status. Document de Treball 99-07, Departament d'Economia Aplicada, Universitat Autònoma de Barcelona
- Cox D (1987) Motives for Private Income Transfers. *Journal of Political Economy* 95:508–546
- Cox D (1990) Intergenerational Transfers and Liquidity Constraints. *Quarterly Journal of Economics* 105:187–217
- De la Rica S, Felgueroso F (1999) Wage Differentials Between Permanent and Temporary Workers: Further Evidence. mimeo
- Del Río C, Ruiz-Castillo J (1997) Demographic Trends and Living Standards. Working Paper 97-67, Economic Series 33, Universidad Carlos III de Madrid
- Ermish J (1996) Parental Support for Human Capital Investment by Young Adults. Discussion Paper No. 1536, CEPR
- Ermish J (1999) Prices, Parents and Young People's Household Formation. *Journal of Urban Economics* 45:47–71
- Ermisch J, Di Salvo P (1997) The Economic Determinants of Young People's Household Formation. *Economica* 64:627–44
- Fernández Córdón JA (1997) Youth Residential Independence and Autonomy: A Comparative Study. *Journal of Family Issues* 18:576–607
- Güell M, Petrongolo B (2000) Workers' Transitions from Temporary to Permanent Employment: the Spanish Case. mimeo, Universidad Carlos III
- Hajivassiliou VA, Ruud PA (1994) Classical Estimation Methods for LDV Models Using Simulation. In: Engle RF, McFadden DL (eds) *Handbook of Econometrics*, vol. 4. North-Holland, Amsterdam
- Haurin DR, Hendershott PH, Dongwook K (1993) The Impact of Real Rents and Wages on Household Formation. *The Review of Economics and Statistics* 75:284–293
- Haurin DR, Hendershott PH, Dongwook K (1994) Housing Decisions of American Youth. *Journal of Urban Economics* 35:28–45
- Heckman JJ (1978) Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica* 46(6):931–957
- Heckman JJ (1981) Statistical Models for Discrete Panel Data. In: Manski CF, McFadden D *Structural Analysis of Discrete Data with Econometric Applications* (ed) The MIT Press, Cambridge, Massachusetts
- Iacovou M (1998) Young People in Europe. mimeo, Institute for Social and Economic Research, University of Essex
- INE (1992) *Encuesta de Presupuestos Familiares 1990–1991*. Metodología, Madrid
- Jimeno JF, Toharia L (1993) The Effects of Fixed-term Employment on Wages: Theory and Evidence from Spain. *Investigaciones Económicas* 27:475–494
- Jurado Guerrero T (1997) Un análisis regional de los modelos de convivencia de los jóvenes españoles. Las cuatro Españas de la emancipación familiar. *Estudios de Juventud* 39:17–35
- Maddala GS (1983) Limited Dependent and Qualitative Variables in Econometrics. Econometric Society Monographs, Cambridge University Press
- Mallar CD (1977) The Estimation of Simultaneous Probability Models. *Econometrica* 45(7):1717–1722
- McElroy MB (1985) The Joint Determination of Household Membership and Market Work: The Case of Young Men. *Journal of Labor Economics* 3(3):293–316
- Revenga A (1991) La liberalización económica y la distribución de la renta: la experiencia española. *Moneda y Crédito* 193:179–224
- Robinson A (1998) Spanish Family Values: Spain's Welfare system is let off the hook by the family. *New Economy* 4:188–92
- Rosenzweig M, Wolpin K (1993) Intergenerational Support and the Life Cycle Incomes of Young Men and Their Parents: Human Capital Investments, Coresidence, and Intergenerational Financial Transfers. *Journal of Labor Economics* 11:84–112
- Sastre M (1999) *Los ingresos y los gastos en las EPF. Ensayos sobre desigualdad y bienestar*. Unpublished Ph. D. Dissertation, Universidad Complutense de Madrid
- Stern S (1989) Measuring the Effect of Disability on Labor Force Participation. *Journal of Human Resources* 24(3):361–395

- Toharia LA, Cebrián C, García Serrano I, García Mainar C, Malo I, Moreno MA, Villagómez E (1998) *El mercado de trabajo en España*. McGraw-Hill/Interamericana de España, S.A.U.
- Vergés R (ed) (1997) *La edad de emancipación de los jóvenes*. Centre de Cultura Contemporània de Barcelona
- Wolf DA, Soldo BJ (1994) Married Women's Allocation of Time to Employment and Care of Elderly Parents. *Journal of Human Resources* 29(4):1259–1276
- Zellner A, Lee T (1965) Joint Estimation of Relationships Involving Discrete Random Variables. *Econometrica* 33:382–394