HOUSING DEPRIVATION AND HEALTH STATUS: EVIDENCE FROM SPAIN

Carolina Navarro[a]

Luis Ayala[b]

José M. Labeaga[c]

IMPORTANT

This is the version accepted for publication in *Empirical Economics*.
The final published version is available in *Empirical Economics*, 38, pp. 555-582, 2010:
DOI: https://doi.org/10.1007/s00181-009-0279-5

The reference is:
Navarro, C., Ayala, L. and Labeaga, J.M. (2010): "Housing deprivation and health status: evidence from Spain". *Empirical Economics*, 38, pp. 555-582.

[a]UNED
cnavarro@cee.uned.es

[b]Universidad Rey Juan Carlos
luis.ayala@urjc.es

[c]*UNED*
jlabeaga@cee.uned.es

Abstract

Living in inadequate housing conditions not only supposes a failure of a basic functioning. It also has effects on other essential aspects of well-being such as health. This study questions to what extent living in poor housing conditions can determine individuals' health status once the possible influence of other factors is controlled for. By estimating a logistic model with individual effects and a housing deprivation index based on a latent variable model, we reach a number of relevant conclusions concerning the relationship between these two different dimensions of multidimensional well-being. We find a negative effect of housing deprivation on the individuals' health, both when housing conditions are introduced in a disaggregated manner in the model and when they are combined in a latent variable.

JEL: I10, I30.
Keywords: housing deprivation, health, latent class models.

1. Introduction[1]

In recent years, there has been a significant increase in the number of studies that make an attempt to estimate the influence of various factors on health. Among the relationships that have aroused more interest, one of the most complex is that between housing conditions and health. The publication of a growing number of studies has contributed to consolidate the evidence on the negative impact that living in poor housing conditions has on health status. Some studies have taken a cross-sectional approach (Ineichein 1993; Lowry 1991; Smith 1989) while others attempt to look into the effects exerted by housing deprivation on health from a longitudinal perspective (Marsh et al. 2000; Power and Hertzman 1997).

A consensus has been reached on identifying the housing conditions that have a negative effect on the health status of individuals. Living in dwellings with leaky roofs, damp and overcrowded living conditions have been found as some of the attributes producing more negative effects on health. These conditions are associated with health problems even after potentially confounding factors such as income, social class, and unemployment are controlled for. When comparing to other health risk factors, the negative influence exerted by multiple housing deprivation is of a similar magnitude as smoking and greater than that exerted by excessive alcohol consumption (British Medical Association Board of Science and Education 2003). Some studies even establish a strict relationship between housing conditions and mortality rates (Dale et al. 1996).

The notion of housing as a public health issue is not new. Interest in housing as a determinant of health has fluctuated in response to housing-related infectious disease outbreaks, social unrest and class conflict, industrialist interest in maintaining a wealthier workforce, and economic downturns leading to crises in housing availability and quality (Krieger and Higgins 2002). Different lines of research have been developed on the relationship between health and housing since the pioneering work done at the

[1] We thank the Inter-ministerial Commission on Science and Technology (SEC2001-0746, SEJ2004-07373-c03-03, and SEJ2005 08783-C04-04) for the funding received. We are also grateful to Rosa Urbanos, Irini Moustaki as well as to seminar participants at Institute for Social and Economic Research (ISER) University of Essex, Universidad Complutense, UNED, Encuentro de Economía Pública, Jornadas

end of the 19th century, which came about as a result of a concern for and an interest in the consequences that poor housing conditions had on health in the Victorian era in the United Kingdom. There are four main areas of analysis: 1) studying the disadvantages of individuals suffering from health problems in the housing market; 2) analysing homeless people's access to health services; 3) assessing the pathological consequences of living in bad housing conditions; and 4) examining the stress associated with not being able to gain access to adequate housing and its possible effect on mental health (Dunn 2000). Research on these relationships received an additional impetus at the end of the 1990s when it was incorporated more visibly into some government's political agenda (Department of Health 1998,1999).

The main hypothesis of this literature is that poor housing conditions should produce a poor health status. The effects of poor housing upon health, however, can not be easily isolated. Despite a rapidly expanding literature there is a need for research that provides a more complete picture of the relationship between housing and health. On the one hand, the literature examining the links between health and socio-economic circumstances might have been using measures acting as proxies for a number of different health risks. Some of these risks may have a latent and unexplored housing dimension (Marsh et al. 2000). In practice, more additional attempts are needed to overcome some of the traditional difficulties encountered in multidimensional deprivation studies to get a composite measure of this latent variable. On the other hand, after many years of research we still have relatively little insight into which framework is best to use when studying the effects of poor housing on health. The health production function framework could allow us making predictions about the impact of different housing conditions on health. A reasonable assumption could be made that living in a deficient dwelling determines health status, once other observed and unobserved factors have been controlled. We are, of course, aware of potential reverse causation and we will try to isolate causal effects in the direction of housing to health.

If the effects of these inputs could be isolated, the empirical analysis of the health production function could yield useful insights into new designs of some policies aimed at improving the health status. People receiving public support in order to get better

de Economía de la Salud, International Conference on Microeconometric Models and Simulation Tools for Fiscal Policy (Instituto de Estudios Fiscales), and 12th Panel Data Conference.

housing conditions should be more likely to report better health status. Therefore, guaranteeing adequate housing not only means covering a basic need. It also produces external effects on other dimensions of well-being. In terms of efficiency, some housing policies could reduce the health costs associated to the impacts of deficient housing conditions. Policies aimed at supporting improved housing as a means to improve health can lead to lower costs of the national health system[2].

The aim of this study is to analyse the influence that living in a dwelling suffering from certain deficiencies —such as a lack of certain facilities, overcrowding or other structural problems— has on the health status. We use a latent variable for housing deprivation and a health production function. Probability decision models are used, which are adjusted using the Spanish waves of the *European Community Household Panel* (ECHP) from 1995 to 1998. Controlling heterogeneity enables certain factors such as the pre-existing stock of health to be identified as fixed effects. A second novelty resides in the way housing deprivation is represented. The two-fold fact that the different housing conditions may be endogenous in a model that tries to explain the individuals' health status and the need to prove the validity of these factors as health proxies, both individually and as a group, have led us to construct a latent variable that summarises such characteristics. We also test some alternatives accounting for the potential problem of reverse causality.

Our results show that most housing conditions –living in a dwelling that lacks hot running water or central heating), has structural problems like a leaky roof, damp or rot in the floor or window frames, or suffers from problems of overcrowding– have effects on individuals' health status. Nonetheless, housing conditions are not the only relevant determinants of health status. The inclusion of other variables has allowed us to verify the influence exerted by other economic and social variables. These results also hold when different specifications of the reference model are conducted in order to address the potential problem of reverse causality.

[2] Evidence for Canada shows, for instance, that the cost of social and health care services for the homeless is 30 per cent higher than the costs that would arise if the homeless were re-housed and maintained in specially designed dwellings (Ministry of Social Development and Economic Security 2001).

The structure of the paper is as follows. The theoretical grounds of the relationship to be studied and the issues related to the measurement of housing deprivation are reviewed in the following section. The econometric specifications are presented in the third section. The possibilities and constraints of the data used for the empirical work as well as the results are presented and interpreted in the fourth section. An analysis of the results is contained in section 5. Section 6 concludes.

2. Background

2.1. Housing deprivation and health

The wide variety of elements that come into play when assessing the effects of housing deprivation on health makes it necessary to develop a general analytical model capable of exploring this kind of questions. If the main goal is to examine the relationship between housing inputs and health as an output, the key to this model lies in identifying a function of health production (Grossman 1972a 1972b 1999; Rosenzweig and Schultz 1983; Kenkel 1995; Contoyannis and Jones 2004). Until the first studies analysing the possible effects of a vector of factors on health within the framework of a production function, most of the empirical works focused on analysing the effects of health services. Auster *et. al* (1969) was one of the first studies that introduced an overall function of health production in which the effects of other factors were added to the health services.

The model developed by Grossman (1972a), which analyses how individuals allocate their resources to produce health, is the most relevant theoretical framework to explain an individual's health status. In this model, health represents a durable capital stock. Each individual owns an initial amount of such stock that depreciates with age and can be increased with investment. The demand for health consist of two elements: consumption commodity (which enters directly in the utility function) and investment commodity (which determines the total amount of time available for market and non-market activities).

This theoretical framework has had a great influence on numerous studies, both theoretical and empirical. Rosenzweig and Schultz (1983) differentiated two kinds of factors in the health production function. The first set of factors includes choices made by the individual, while the other set consist of exogenous elements that can be considered as given in the function of health production. This distinction between exogenous and endogenous factors was introduced by Becker (1965). Many of the studies after Rosenzweig and Schultz (Kenkel 1995; Contoyannis and Jones 2004) reflect this differentiation. Kenkel (1995) used the health production function framework to analyse the importance of lifestyles on health. The former are determined by the choices made by individuals. The stock of health is produced as a function of the production technology given by the various lifestyles (eating breakfast regularly, proper weight, number of sleeping hours, consumption of tobacco and alcohol, sport, among others), the stock of human capital and different socioeconomic characteristics (age, sex, ethnic origin, chronic diseases, among others) that can have an influence on the productivity of gross investment, the stock of pre-existing health or the determining factors of the rate of depreciation. The inclusion of these factors responds to the fact that health is considered an essential commodity including aspects of both consumption and capital (Grossman 1972a).

A possible alternative to examine the importance of housing conditions on health is to adapt this function by incorporating variables that reflect housing characteristics (lack of certain facilities, structural problems or overcrowding)[3]. The main difficulty involved in transforming the Kenkel model to empirical analyses lies in the fact that introducing variables that refer to the behaviour of individuals makes the specification of the function very complex. The need for information about lifestyles comes up against the common shortage of data on individual habits. A variation could lie in the introduction of an alternative assumption: the rest of the regressors included in the model would provide information on the individuals' lifestyles. This vector of regressors includes different variables that measure the stock of health (age), rate of depreciation, pre-existing health (chronic diseases), gross investment (educational attainment), housing deprivation (lack of certain facilities, structural problems, overcrowding) and other

[3] It must be noted that the Grossman model applies to individuals, while housing appears to be a collective choice made by households. The individual approach could be consistent under a collective model with altruistic individuals.

variables having an incidence on the investment's productivity. As Contoyannis and Jones (2004) pointed out, the latter should include information about the households' characteristics (marital status), social class (income level) or labour market conditions (employment status).

The results that can be obtained should be interpreted by taking into account that the marginal effects not only gather information about the influence of the set of regressors on the health status but also include the changes caused by lifestyles. Even in the case in which observed heterogeneity is controlled by a wide set of determining factors, unobserved effects of individuals (genetic characteristics) could directly affect health status or indirectly by not included factors in lifestyles (tastes).

The introduction of housing conditions into the health production function inevitably entails selecting the most relevant housing characteristics affecting health. The literature on housing conditions and health has been reviewed intensively (Marsh et al. 2000; Krieger and Higgins 2002; Dunn 2000). There are several avenues trough which living in a deficient dwelling could affect health. Several factors have been stressed: lack of safe drinking water, absence of hot water for washing, ineffective waste disposal or intrusion by disease vectors have been identified as determining factors of infectious diseases. Overcrowding problems could give rise to infectious and respiratory diseases. Damp, cold and mouldy housing is associated with asthma and other respiratory problems, eczemas, asthma, recurrent headaches, sore throats and rhinitis. Extreme temperatures could underpin heart diseases, bronchitis, hypothermia or respiratory diseases. More precisely, deviation of indoor temperature beyond a threshold has been associated with increased risk of cardiovascular problems.

Other studies have also emphasized the contribution of rot in floors or window frames to different diseases, including asthma. Water intrusion is a major contributor to problems with dampness. Old and dirty carpeting can also result in allergic, respiratory, neurological and haematological illnesses. Exposure to toxic substances may also lead to neuropsychiatric and chronic health disorders. Symptoms of stress, anxiety, depression, even social misconduct, may be related to indoor exposure to these substances. They have also been associated with asthma symptoms, headache, hypertension and lung cancer.

It is reasonable, therefore, to expect a poor health status when a dwelling does not cover the provision of functional and adequate physical, social and mental conditions for safety, hygiene and privacy. The nature of health as a durable capital stock makes deficient housing conditions a contributing factor of its depreciation. In this sense, including housing conditions as inputs of the health production function could yield more accurate results than other standard models. Two questions should be answered for the general understanding of the relationship. First, we should test whether or not there are different types of effects depending on specific conditions of housing deprivation. Second, it must be evaluated to what extent these effects hold when individuals suffer from different and simultaneous deficient housing conditions. We try to provide answers to both questions by considering both a disaggregated vector of housing conditions as well as a composite index of housing deprivation.

2.2. A composite index of housing deprivation

One straightforward approach for the assessment of the effects of housing deprivation on health is considering individually different housing conditions. An alternative approach consists of summarizing the different conditions into a single index. The different conditions can be an imperfect manifestation of a latent structure of housing deprivation. Furthermore, each one of the housing attributes can have a very different weight on the well-being of households. Latent variable models, and more specifically latent trait models, offer a suitable methodological framework to solve some of the usual problems in the measurement of multidimensional deprivation, such as the aggregation of different basic dimensions. These models use multivariate analysis techniques to measure an unobserved concept like housing deprivation derived from a set of observed items or attributes. Latent variable models have received a growing attention in the analysis of very different research topics (Knott and Tzamourani 1997; Marcoulides and Moustaki 2002; Moustaki and Knott 2000), including those related to the measurement of multidimensional poverty (Silber 2007).

Essentially, latent variable models allow to reduce the dimensionality of the set of observed variables and to assign the values corresponding to latent variables to each individual as a function of the responses obtained for each of the observed variables. This is done either to analyse and identify the latent variables underlying a set of observed indicators, or to prove whether or not a specific set of indicators measuring certain concrete concepts reveal a supposed *a priori* hypothetical structure. The partial indicators chosen will constitute the supposed *a priori* hypothetical structure that will be tested by means of the latent trait model. An alternative would consist of assessing the consistency of the deprivation indicators by estimating the Cronbach Alpha coefficient. Nevertheless, the use of this method suffers from some relevant constraints. Deprivation indicators have different variances and the measurement error is not random across indicators and individuals (Moisio 2004). A key advantage of the latent trait model is that it offers an empirical method to test whether the selected indicators really measure the same latent phenomenon.

The second advantage these models offer is that they allow us to synthesise a set of partial indicators of the same phenomenon into a single index. This is made taking into account the correlation among these components and their mutual dependence on the latent variable. These techniques are appropriate for the nature of a possible set of housing conditions (dichotomous variables) and allow different additive weights to be assigned to them. The latent trait model is similar to the factor analysis model but for binary observed responses (Bartholomew and Knott 1999). The linear factor analysis model for metric items is written as:

$$x_i = \alpha_{i0} + \alpha_{i1} y_1 + \ldots + \alpha_{iq} y_q + \varepsilon_i \qquad i = 1,\ldots,p \qquad (1)$$

where p denotes the total number of housing conditions, x_i denotes the *ith* metric observed item, $y = (y_1,\ldots,y_q)$ denotes the vector of latent variables and ε_i denotes the error term. We can assume that the error term follows a normal distribution with zero mean and variance σ_i^2 ($\varepsilon_i \sim N(0, \sigma_i^2)$) and that the latent variables are assumed to be independent with standard normal distributions $y_j \sim N(0, 1)$ for all j. From those two assumptions the conditional distribution of x_i given the latent variables y has a normal distribution with parameters $(\alpha_{i0} + \sum_{j=1}^{q} \alpha_{ij} y_j, \sigma_i^2)$. In the case where the observed item x_i is a binary item the model given by (1) could not work well. We are interested in modelling not the actual binary outcome but the probability that a randomly selected individual will correctly respond to an observed condition x_i given his/her position on the vector of latent variables y, $P(x_i=1|\ y)=\pi_i(y)$. In this sense, this conditional probability can be expressed as a linear function of the latent variables:

$$\pi_i(y) = \alpha_{i0} + \alpha_{i1} y_1 + ... + \alpha_{iq} y_q \qquad i = 1,..., p \tag{2}$$

There are two problems regarding the hypothetical linear relationship between the probability of a correct response and the latent variables. On the one hand, the left part of equation (2) is a probability that should take values between 0 and 1 but the right hand side is not restricted. On the other hand, the rate of change in the probability of a positive response could not be the same for the whole range of y. A curvilinear relationship might be more appropriate. To take into account those points it is necessary to introduce a link between the probability and the latent variables. That link should have an s-shape and it should also map the [0, 1] range into the (-∞, +∞) range. Link functions commonly used in practice are the logit and the probit ones.

In the latent trait model, each observed housing condition would correspond to q+1 parameters to be estimated (α_{i0} and the factor loadings $\alpha_{i1,...,}$ α_{iq}, where α_{i0} represents the probability of a median individual suffering deprivation of the observed condition, while α_{ij}, $j=1$ are discrimination parameters). The higher the value of α_{ij} is for an observed condition, the greater the difference in the probability of obtaining a positive response between two individuals situated at a certain distance apart on the latent dimension. The higher the parameter is, the easier will be to discriminate between two individuals on the basis of their level of deprivation concerning each observed condition.

Given that only the observed variables $x_1,...,x_p$ can be known, the estimation of the unknown parameters is based on their joint distribution function:

$$f(x_1,..., x_p) = \int ... \int g(x_1,..., x_p \mid y) h(y) dy \tag{3}$$

where we assume the conditions of conditional independence, a Bernoulli distribution for each x_i and independent latent variables:

$$g(x_1,\dots,x_p \mid y) = g(x_1 \mid y)\dots g(x_p \mid y) = \prod_{i=1}^{p} g(x_i \mid y),\ g(x_i \mid y) = \{\pi_i(y)\}^{x_i}\{1-\pi_i(y)\}^{(1-x_i)},$$
$$h(y) = h(y_1)\times\dots\times h(y_q). \qquad (4)$$

The parameters α_{i0} and $\alpha_{i1},\dots,\ \alpha_{iq}$, included in $\pi_i(y)$ can be estimated by maximum likelihood. An EM algorithm (Bartholomew and Knott 1999; Bock and Aitkin 1981) will be employed to estimate the model with unobserved variables, using the TWOMISS program (Albanese and Knott 1990). Estimating the parameters allows us to give a response to the above mentioned objective of assigning the latent variable values to each individual or household as a function of the presence or lack of the observed conditions. All the information about latent variables is contained in the posterior distribution of such variables given a set of observed responses ($h(y\ /x_1,\dots,\ x_p)$), which we will call the response pattern $[x = (x_1,\dots,\ x_p)]$. Using the logit link function shows that the posterior distribution depends on the observed variables through q components. These components, called 'sufficient statistics', are given by:

$$X_j = \sum_{i=1}^{q} \alpha_{ij} x_i\ , \qquad j = 1,\dots,q, \text{ with } q<p \qquad (5)$$

The components, which are a weighted sum of the observed responses using as weights the discrimination coefficients (α_{ij}), are used to score the individuals on the latent dimensions. The mean of that distribution, E ($y_j \mid x_1,\dots,\ x_p$),$j = 1,\dots,q$, can also be used to scale individuals. Therefore, an index of housing deprivation can be defined for every individual synthesising a set of partial indicators based on the correlation of such dimensions and their mutual dependence.

3. A discrete choice panel data model for the relationship between housing deprivation and health

The potential relationship between housing deprivation and health can be tested by transforming the health production model into an empirical one including the different dimensions previously outlined. This model could include the effects that the different dimensions of housing deprivation have on the health status expressed as a binary variable, namely being in good or bad health. To do so, it would be necessary to ensure that each of the housing deprivation components is a suitable measurement of such latent dimension and constitutes a good proxy of the health status, both individually and as a group (Marsh et al. 2000). For these reasons, we will include a disaggregated vector of household characteristics in a preliminary approximation to the relationship between housing and health. In a subsequent approximation, we will set out a model that sums up all the housing dimensions in a single variable representing deprivation in order to compare its relevance as a proxy for individuals' health.

A simple way to approach the relationship between the two variables mentioned above consists of estimating a binary logistic regression model. Although the reported health status is usually a multinomial variable in which the different characteristics have an order, for the sake of simplicity we choose to transform the dependent variable into a dichotomous indicator (Contoyannis and Jones 2004; Wagstaff et al. 1989; Kunst et al. 1995). A way of constructing this variable is based on using each individual's own perceptions. Despite the fact that incorporating subjective variables can be open to criticism, there are arguments that back the notion of subjectivity as one of the main advantages provided by this kind of assessments (Manning et al. 1981; Kemna 1987; Taubman and Rosen 1981). Some studies also reveal that there is a significant relationship between this variable and other health measures (Borg and Kristensen 2000; Burstrom and Fredlund 2001).

Because data on the same individuals are available for various years, the binary choice model has to take into consideration both the temporal and individual dimensions:

$$H_{it}^{*} = z_{it}^{'}\beta + v_{it}\text{, i = 1,…, n;   t = 1,…,T} \qquad (6)$$

where $z = (L, V, S, X)$, L represents the various lifestyles, V represents housing conditions, S represents the stock of human capital and X includes, as aforementioned, different variables that can have an influence on the productivity of gross investment, the stock of pre-existing health or the determining factors of the rate of depreciation.

In order to provide an overall simplified version, we begin by adopting the assumptions of common idiosyncratic shocks and uncorrelated mixed errors. In other words, $v_{it} = \varepsilon_i + u_{it}$, in which $\varepsilon_i = \varepsilon$, $\forall\ i$ with u_{it} serially uncorrelated. Despite the limitations of this approach, its purpose is to obtain some *naïve* estimators in order to compare them with results corresponding to more complex models subsequently.

Some of the previous assumptions like the use of data on the same individuals for different time units are then reviewed in a second specification and the results are compared. There are two related aims. First, we like to relax the hypothesis of homogeneity in the response of the individuals; second, we want to check whether unobserved individuals' lifestyles or their genetics are correlated with the observable explanatory factors included and to what extent their exclusion bias the results. In doing this we exploit the most important advantages of using panel instead of cross-section data (Hsiao 1985; Baltagi 2001; Matyas and Sevestre 1992). In terms of the specification, the structural logistic model can now be formulated as in equation (6), but we relax the assumption that $\varepsilon_i = \varepsilon = 0$, $\forall i$ although maintaining that u_{it} is serially uncorrelated. The probability of being unhealthy when homogeneity is imposed is:

$$\Pr[H_{it} = 1] = \Pr[H_{it}^* > 0] = \Pr[v_{it} > -z_{it}^{'}\beta] = F(z_{it}^{'}\beta) \qquad (7)$$

where F represents the cumulative probability function of the logistic distribution. When relaxing the homogeneity assumption, this probability becomes:

$$\Pr[H_{it} = 1] = \Pr[H_{it}^* > 0] = \Pr[u_{it} > -\beta z_{it}^{'} - \varepsilon_i] = F(z_{it}^{'}\beta + \varepsilon_i) \qquad (8)$$

Since the time dimension of the data is small, we have in (8) a problem of *nuisance parameters*, which aggravates when the heterogeneous effects and the regressors are allowed to be correlated. The relaxation of the assumption of uncorrelated effects is due to the fact that unobserved factors such as genetics are expected to be correlated with some demographic variables, so that:

$$E\left(z_{it}\varepsilon_i\right) \neq 0 \tag{9}$$

In order to incorporate this possibility, we assume that $E\left(\varepsilon_i / z_{it}\right) = \bar{z}_i^{'}\lambda$ and then,

$$H_{it}^{*} = z_{it}^{'}\beta + \bar{z}_i^{'}\lambda + \eta_i + \nu_{it} \tag{10}$$

where the individual effects are given by $\varepsilon_i = \lambda \bar{z}_i + \eta_i$ and we now assume that $E(\eta_i | z_i)=0$. At the end, in equation (10) we account for correlated effects and we end up with an specification that can be estimated under the assumption that η_i are not correlated with the z's. We are aware of two restrictions incorporated in the specification of the effects, assuming a linear function for them as well as a lack of correlation between the η's and z's. In any event, these two assumptions can be relaxed and tested.

Estimating the logistic model with individual effects enables us to confirm or refute, should it be the case, the results obtained by the model that does not take into consideration the panel structure of the data. In order to fulfil the two previously mentioned aims, it is necessary to conduct a prior comparison that would allow us to know whether or not unobserved heterogeneity actually exists. A likelihood ratio (LR) test compares the two specifications. In our case, we have under the null a homogenous model and we relax this assumption under the alternative hypothesis by allowing for the presence of individual heterogeneity. Thus, the LR can be written as:

$$-2Ln\lambda = -2Ln\left(\frac{\hat{L}_R}{\hat{L}_U}\right) \tag{11}$$

where $\hat{L}_R$ and $\hat{L}_U$ are the values of the log-likelihood function at the optimum evaluated for the restricted and unrestricted models, respectively. It follows under the null a χ^2 distribution with degrees of freedom equal to the number of restrictions imposed when moving from the unrestricted to the restricted specification.

An important limitation which has been already mentioned both in the theoretical and empirical literature in this field can affect the results of this exercise: potential reverse causality. It implies that the direction of the causation could be happening from housing to health or vice versa. This is normally a very hard problem to solve using any kind of data because of the difficulty of finding instruments for the potential endogenous explanatory variables. However, panel data offers another advantage to deal with this limitation. We have instruments inside the dataset because we have repeated observations for each individual in the sample. Moreover, the ECHP is a sufficiently rich dataset to include variables related to housing which are not potentially related to health. So, we have two alternatives. One consists of using lags of the variables as instruments, at the cost of losing some observation for each individual. Another one is a two step procedure. In the first step we estimate an auxiliary regression for the potential endogenous variables and then, at the second we estimate the structural form using the predictions obtained at the first stage as instruments.

Although we propose to use a latent index as an instrument for the housing deprivation indicators, we are aware that it really does not solve the problem of simultaneity between health and housing variables. Moreover, the use of lagged variables has the problem of correlation (by construction) with the heterogeneous component of the error and it does not solve the problem[4]. So, we propose the alternative approach consisting of running an auxiliary regression for each of the housing deprivation variables. In particular, we propose to use in the auxiliary equations household variables (family composition, family size and number of children and their squared), variables related to the house (age and age squared of the house and tenure status), whether the family owns vehicle, education indicators and time dummies.

This procedure has two main problems. First, it is necessary to impose exclusion restrictions and we solve it by including some variables not in use in the health equations (i.e., composition and size of the family or age of the house). Second, the inclusion of the same set of variables in all equations modelling the indicators of housing deprivation can generate a problem of high collinearity in the posterior estimation of the structural equation. We deal with this second limitation by means of excluding some non-significant regressors from the auxiliary equations. The results do not change and at the same time exclusion of some non-significant variables helps in identifying the parameters of the model. We prove to predict the indicators of housing deprivation by two different auxiliary specifications: a standard binary logit and a random effects binary logit for each indicator.

Once we have parameter estimates for each auxiliary equation, we get the predicted indicator for each housing deprivation variable and then we use these predicted variables as instruments for the original ones. We solve the reverse causality problem whenever we are willing to assume that all the conditionings in each auxiliary regression are strictly exogenous, which constitutes our identification restrictions.

4. Data

The interest in the influence of housing characteristics on health implies that the choice of a dataset is conditioned by the available information on both variables. The ECHP contains valuable information on housing facilities and specific housing problems. In addition, it includes enough details on other variables such as economic, employment and living conditions of the household. The data used come from waves for 1995, 1996, 1997 and 1998[5]. As usual in this kind of studies, the unit of analysis is the individual, although the family context is sometimes relevant in health economics (Jacobson 2000; Bolin et al. 2001, 2002a, 2002b).

The dependent variable is the health status. It is defined in dichotomous terms: "being unhealthy" or "being healthy". It takes on the value of 1 whenever individuals state that

---

[4] Discrete variables are difficult to instrument using lags when there are changes in time in the regime of the variable (i.e., observations on a given year taken value 1 and in the subsequent year taken value 0).

[5] The ECHP was conducted until 2001. We opted for using only the first four waves due to the high attrition rates in the Spanish sample (over 50 per cent in subsequent years). Attrition is especially higher in the case of people stating that they are in bad or very bad health, renters and several housing conditions. The 1994 survey is not used because all values for the variable providing information on the existence of chronic diseases are missing.

they are in “very bad”, “bad” or “fair” health and the value of 0 when they state they are in “good” or “very good” health. Although this constitutes our final choice, we have proved with grouping the fair alternative with the “good-very good” ones without finding significant changes in the results[6]. The explanatory variables are dictated both by the economic model for the production of health and the available information in the ECHP.

Housing conditions are defined by a set of dummies that take the value of 1 when an individual is deprived on some specific housing condition and 0 otherwise (see Table A.3. in the Appendix for details). More specifically, the variables representing a lack of separate kitchen, bath, indoor flushing toilet, hot running water, central heating, garden and natural light indicate if the dwelling has this kind of facilities. After conducting a detailed study on the relationship between a lack of central heating, income and weather conditions for different Spanish regions, we consider that the lack of central heating in Southern Spain (Andalucia, Murcia, Ceuta and Melilla) and the Canary Islands does not necessarily mean being in a state of deprivation. We considered households that do not have central heating in these areas as non-deprived. Nevertheless, in order to account for territorial singularities we also include a variable representing regional residence. Another group of variables, which includes the presence of a leaky roof, damp and rot in floors and window frames, indicates if the dwelling is suffering from some kind of structural problem. The noise, pollution and vandalism variables represent the presence of environmental problems or crime and vandalism in the dwelling’s surrounding area. Lastly, the overcrowding variable indicates whether or not the dwelling has space problems. Overcrowding is defined as a dwelling having a number of rooms less than the number of adults in the household. Establishing the space a person needs to live is necessarily a subjective matter. This is why alternative indicators have been used to take into account household composition (Chiappero 2000). Sensitivity analyses do not show relevant effects on the estimates.

In order to isolate the effects of housing deprivation from neighbourhood influence we have included, in addition to the prior variables of vandalism or pollution in the

[6] In order to test the equivalence of the categories chosen with those in the observed data, we have also proved ordered logit models (results are available upon request). We have also considered the

residence area, the neighbourhood unemployment rate. Beyond the condition of the housing unit itself, the site of the home may be a determinant of health. Results for this variable could yield relevant policy connotations. This variable is defined as the proportion of individuals registered as unemployed in each one of the census tracks. A set of dummies reflecting the tenure status are also used as control variables. More precisely, we include three dummies reflecting whether or not the household owns the dwelling, it is rented or accommodation is provided rent-free.

We have also considered suitable to include a subjective variable representing satisfaction with the dwelling's conditions. It aims to complement the information on housing deprivation by incorporating a measure of what could be called an individual's "psychological welfare". This variable varies in a scale ranging from "not satisfied at all" to "fully satisfied". Since it is possible to argue that satisfaction with housing can have an endogenous nature, when we exclude it from the specification the results are unaffected.

Other explanatory variables that have an influence on the stock of health are included, such as variables that explain individuals' economic, employment and social situations. Educational attainment is included as one of the determining factors for gross investment. It includes three dummies that reflect the different categories that refer to education. A dummy that provides information on the existence of chronic diseases is included as a variable representing the stock of pre-existing health. The stock of pre-existing health contains, by construction, the heterogeneous effects. This is another reason to acknowledge correlated effects and to model them as function of explanatory variables as we do in our estimates. A set of dummies reflecting the different age and sex categories of individuals (see Appendix) is used as determining factors for the depreciation rate of the stock of health.

There are more explanatory variables that could have an influence on the productivity of health investment and, hence, on the stock of health. The variable representing equivalent income gathers information on a household's total net real income per equivalent adult (the modified OECD equivalence scale recommended by EUROSTAT

---

dichotomous variables but changing the classification *fair health* from the bad to the good alternative and results do not change.

is used, which takes a single-person household as a reference and gives a weighting of 0.5 to the rest of the household's adults and 0.3 to children). In order to reflect different profiles at different income levels, we use a disaggregation by income deciles.

The social behaviour of individuals is represented by variables that reflect their level of social integration based on a set of five dummies (see again the Appendix). A set of dummies that gather information of the type of employment contract and the main source of income is included as an indicator of the employment situation. Dummies on the marital status of individuals are also added. Lastly, we include annual dummies in an attempt to control common time shocks.

5. Results

5.1. Results with different housing conditions

The results of the model with individual effects and disaggregated housing conditions (*Model 1*) are reported in the first column of Table 1 as odds ratios. The likelihood ratio (LR) tests show the need to estimate the model using individual effects. Although we focus our attention on these models, we present in Table A.1 of the Appendix the results of the logistic model without individual effects. We must note that housing conditions are not observed but predicted variables. Of the whole set of variables representing individuals effects, we only report significant variables. Most of the explanatory variables contribute significantly to explaining health status considered as a dichotomous variable.

Concerning housing conditions, the results corroborate the expected negative effect on health of lacking some of these conditions with some exceptions that we comment on later on. Individuals who live in a dwelling that lacks hot running water or central heating have respectively a 21 and 17 percent higher probability of being unhealthy than individuals whose dwellings do not lack these conditions. Likewise, individuals belonging to households whose dwellings suffer from structural problems have a greater chance of being unhealthy than the rest of individuals. The overcrowding variable shows that living in an overcrowded dwelling increases the probability of being unhealthy too. Some of these probabilities are slightly lower than the ones estimated without controlling heterogeneity and reverse causation. As regards to the satisfaction expressed by individuals with their housing conditions, there is also an inverse relationship between being satisfied with the dwelling and being unhealthy.

The results for the variables representing dwelling's facilities coincide with the findings obtained by studies carried out for other countries, showing a negative effect on health. In almost all cases, to a high degree of statistical confidence the effects of these variables variable are well defined. These results therefore confirm the negative implications that the various symptoms of housing deprivation could have on health, an essential dimension of individual well-being.

Results regarding neighbourhood effects show that people living in areas with vandalism problems have a higher probability of suffering from bad health. On the other hand, the variable representing the neighbourhood unemployment rate is not significant upon health. Once we control for these proxies for poor local condition, the effects of housing conditions still hold. We also include tenure status as a control factor. Results show that renters have a lower probability of being unhealthy than owners. This result, however, must be considered taking into account the singularities of the Spanish housing market. A huge difference as compared to other European countries is the extremely high proportion of owners (85 percent). In practice, descriptive statistics show that poor owners are poorer than poor renters.

[Insert Table 1 around here]

However, there are other determinants of health status. The stock of pre-existing health is a very relevant factor. Hence, individuals who suffer from chronic illnesses have a greater probability of being unhealthy at a specific time than the rest. In quantitative terms, this factor has the greatest impact, which to some extent reflects state dependence while at the other picks up unobserved heterogeneity. Income also contributes significantly to explaining the health status (Van Doorslaer et al. 1997). Individuals with incomes in the upper deciles of the distribution have a lower probability of being unhealthy than individuals located in the lowest decile[7]. Marital status also has a significant effect. In particular, being separated is the status with the greatest probability of being unhealthy. The regional residence has also an important impact on health. People living in North East, Madrid, East and South have a lower probability of suffering from bad health than individuals located in North West, Centre and Canary Islands.

[7] Income is included as a relative measure. The implicit hypothesis is that income might have some effects that are not always linear. There is some prior empirical evidence for Spain on this issue (Ayala and Navarro 2007).

Other factors that appear to have a positive influence on health are educational attainment and social relationships[8]. Individuals whose highest level of education is secondary education are 13 percent less likely to be unhealthy than those with only primary school. Concerning social relationships, the results show that those individuals who are less socially integrated —those who meet people less often— have a greater likelihood of being unhealthy than those who meet people most days.

The last set of results deals with the relationship between the employment situation and the health status. Pensioners or individuals perceiving other social benefits show a greater chance of being in bad health than wage earners. As in the case of the pre-existing stock of health, the effect of this variable diminishes considerably when individual effects are taken into consideration, thus pointing towards potential correlation among them. One explanation could be that most of these individuals receive disability or sickness benefits. Lastly, it can be observed that the fact of having a permanent or even a fixed/short-term employment contract, as opposed to not having one, increases the probability of being healthy.

5.2. Results with the latent variable of housing deprivation

In order to assess the impact of different housing characteristics on health, it is necessary to demonstrate that each of the components or indicators making up the housing deprivation index serves as a proxy for the health status, both individually and as a group. The previous model has allowed us to make a preliminary approximation to the relationship between housing and health based on the different housing characteristics. The next step consists of summing up these dimensions in a single variable representing housing deprivation to confirm or refute the previous results. The interest of this second model is also due to the fact that the variables representing housing may be endogenous in a model that attempts to explain health status. The latent variable of housing deprivation can also be interpreted as an instrument for them. A

[8] The decision to include in the models both current levels and within-individual means of near time-invariant variables such as education or marital status can make identification of the coefficients more complex but it is still possible. The purpose is i) control for fixed effects; ii) control for their direct effects. Education as well as other variables show very small time variation but the changes are enough to identify both effects. Transitions into states (marriage or divorce) affect equivalent income and trough it could be important for deprivation and accumulation of human capital is a key issue to avoid deprivation.

composite index can therefore yield a clearer picture of the relationship between these two very different dimensions of multidimensional well-being.

The latent variable can be constructed from the housing variables which show a negative impact on health in the previous regression model. These variables are the same that turned out to be determining factors for social deprivation in a previous study that analysed the basic conditions needed to define housing deprivation in accordance with the criteria commonly used in the poverty literature (Navarro and Ayala 2008). The indicators that make up the latent variable are hot running water, central heating, leaky roof, damp, rot in floor or window frames, and overcrowding. We should also emphasize that the index presents correct correlations with most of the regressors in the model. It is positively correlated with age and negatively correlated with income, for instance.

The results of the latent trait model for each of the sub-periods show that the vector of observed variables [$x = (x_1,..., x_p)$] consisting of *hot running water, central heating, leaky roof, rot in floor or window frames* and *overcrowding* variables, can be explained by a small number of latent variables. In this case, a single latent variable allows us to identify the variable underlying housing deprivation, and confirm the assumed *a priori* structure. The estimated parameters, their standard errors as well as goodness-of-fit measures are reported in Table 2. Results show that the latent variable model can explain between 85 and 90 percent of the association among the *x* variables[9].

[Insert Table 2 around here]

The last column in Table 2 shows the probabilities of a median individual responding positively to each of the six indicators. The indicators corresponding to the hot running water and rot in floor or window frames variables are the least likely to receive positive responses. The estimates of the discrimination parameters α_{ij} representing the weight of each of the observed variables are shown in the fourth column. The values of these parameters show that the central heating and overcrowding indicators have less weight

[9] Another alternative consists of calculating the Pearson χ^2 for combinations of two or three responses. The residuals offer information about the predictions the model makes on the response patterns composed of two or three elements (see Table A.2 in the Appendix).

than the other observed variables. At the same time, they are the ones that have the greatest probability of receiving a positive response. Additionally, the values corresponding to the hot running water and rot indicators show that they have greater weight than the previous ones, as they have a lower probability of receiving a positive response. To a certain extent, we can thus say that the latent variable constructed takes into account the housing conditions that only a very small percentage of the population lacks and assigns them greater weights.

We now re-estimate the logistic model with individual effects incorporating as a new regressor the latent variable representing housing deprivation. This procedure is aimed to check the robutness of the results reached by the specification including disaggregated housing characteristics as well as to control the possible endogeneity of these variables. As before, the values of the LR test provide clear evidence for the presence of unobserved heterogeneity, thus pointing towards the necessity of estimating the model with individual effects (Table 1). The results presented under the heading *Model 2* confirm the findings of the previous specification. Most of the variables make a significant contribution to explain health status. The probability changes evolve as expected in the face of marginal changes in the explanatory variables. Moreover, the goodness-of-fit measures continue to show an adequate fit.

Once the existence of specific individual effects is controlled, we again find evidence of negative influence of housing deprivation on the health status. More precisely, individuals living in houses having structural problems, lacking certain facilities or suffering from overcrowding have a greater probability than the rest of the population of stating they suffer from bad health. Moving up to the next score of the housing deprivation scale increases the probability of bad health by 80 percent. The relationship that exists between suffering housing deprivation, both individually and as a group, and the health status is therefore confirmed even when unobserved heterogeneity is controlled for.

Moreover, we confirm the results for those variables representing neighbourhood characteristics. As before, people living in areas with vandalism problems have a higher probability of suffering form bad health. On the other hand, the neighbourhood unemployment rate doesn't have any significant effect on health. So it seems that once we have controlled for these neighbourhood effects, housing conditions still have an important effect on health. The results about housing tenure are also similar. Owners have a higher probability of being unhealthy than renters, for the reasons explained before.

Previous results are also confirmed in the case of most of the (other) control variables. The presence of chronic illnesses diminishes the probability of being in good health although the effect is attenuated when heterogeneity is taken into account. Individuals in the upper deciles of the income distribution –particularly the last decile- have a lower probability of suffering from bad health than those whose income levels lie in the lowest income decile. Living in some regions, like North East, Madrid, East or South, reduces the probability of being unhealthy. The positive influence exerted on health by factors like educational attainment, personal satisfaction with housing conditions, marital status and the level of social integration are also confirmed. Lastly, the results obtained for the employment situation are also similar to those of the previous model. More precisely, the fact of not having a contract or having a pension or other social benefits as the main source of income increases the chances of being in fair, bad or very bad health. As before, we observe that the quantitative impact of social benefits diminishes considerably when individual effects are taken into consideration. This is why this factor is also considered as a fixed component representing the stock of pre-existing health.

5. Conclusions

The study of the possible correlation between different dimensions of well-being, like health and housing, can improve the interpretation of multidimensional deprivation. A better understanding of the effects of housing on health can be helpful for an adequate weighting structure of the different dimensions synthesised by composite indices. In this sense, the introduction of housing conditions into the health production function can yield relevant results in terms of a better explanation of health determinants. The nature of health as a durable capital stock makes deficient housing conditions a contributing factor of its depreciation. In practice, there are several avenues trough which living in a deficient dwelling could affect health.

The aim of this study has been to show the influence of housing deprivation on health status, as well as to assess the extent to which the existence of unobserved individual characteristics can condition this relationship. More precisely, we have tested the hypothesis that living in a deficient dwelling determines health status, once other observed and unobserved factors have been controlled. Our interest in obtaining answers to these questions is the relevant implications they have both in the assessment of multidimensional deprivation as well as in the design of social policies. The reduced form health production, estimated as a discrete choice model with individual effects, has confirmed the negative relationship between housing deprivation and health. We have tested the robustness of the results using both a set of housing characteristics and an index of housing deprivation.

Our results also confirm the importance of controlling both observed as well as unobserved heterogeneity among individuals. We must emphasize the importance that living in a dwelling that lacks hot running water or central heating (note that deprivation due to lack of central heating has been restricted to very specific regions), has structural problems like a leaky roof, damp or rot in the floor or window frames, or suffers from problems of overcrowding, has on individuals' health status.

Nonetheless, housing conditions are not the only relevant determinants of health status. The inclusion of other variables has allowed us to verify the influence exerted by other economic and social variables. The presence of chronic illnesses, age, the fact of being a woman, belonging to the lower deciles of the income distribution, being unemployed, or having a pension or other social security benefits as the main source of income all increase the probability of individuals stating they are in fair, bad or very bad health. On the other hand, the higher the educational attainment and the social integration the higher the probability of being healthy. All these effects still hold when different specifications of the reference model are conducted in order to address the potential problem of reverse causality.

In any case, the evidence that housing conditions have a clear influence on health –confirmed when the existence of a great degree of heterogeneity is taken into account– allows us to underline some possible implications on health care policies. With all due caution, it seems that people receiving public support in order to improve their housing conditions should be more likely to report better health status.

REFERENCES

Albanese MT, Knott M (1990) TWOMISS: A Computer Program for Fitting a One -or Two- Factor Logit-Probit Latent Variable Model to Binary Data When Observations May be Missing. Technical Report, Statistics Department, London School of Economics and Political Science.

Auster R, Leveson I, Sarachek D (1969) The Production of Health, an Exploratory Study. Journal of Human Resources 4: 411-36.

Ayala L, Navarro C (2007) The dynamics of housing deprivation. Journal of Housing Economics 16: 72-97.

Baltagi BH (2001) Econometric Analysis of Panel Data. John Wiley & Sons, Chichester, United Kingdom.

Bartholomew D J, Knott M (1999) Latent Variable Models & Factor Analysis. Kendall's Library of Statistics, Vol. 7, Oxford University Press, New York.

Becker GS (1965) A Theory of the Allocation of Time. Economic Journal 75: 493-517.

Bock R.D, Aitkin M (1981) Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM Algorithm. Applied Psychological Measurement 37: 29-51.

Bolin K, Jacobsson L, Lindgren B (2002) Employer Investments in Employee Health. Implications for the Family as Health Producer. Journal of Health Economics 21: 563-83.

Bolin K, Jacobsson L, Lindgren B (2001) The Family as Health Producer-The Case when Family Members are Nash-Bargainers. Journal of Health Economics 20: 349-62.

Bolin K, Jacobsson L, Lindgren B (2002) The Family as Health Producer - When Spouses Act Strategically. Journal of Health Economics 21: 475-95.

Borg V, Kristensen T (2000) Social Class and Self-Rated Health: Can the Gradient be Explained by Differences in Lifestyle or Work Environment? Social Science and Medicine 51: 1019-30.

British Medical Association Board of Science and Education (2003) Housing and Health: Building for the Future. United Kingdom: British Medical Association.

Burstrom B, Fredlund P (2001) Self-Rated Health: Is it a Good Predictor of Subsequent Mortality among Adults in Lower as well as in Higher Social Classes? Epidemiology and Community Health 55: 836-40.

Chiappero EA (2000) Multidimensional Assessment of Well-being Based on Sen's Functioning Approach. Rivista Internazionale di Scienze Sociali 2: 207-39.

Contoyannis P, Jones AM (2004) Socioeconomic Status, Health and Lifestyle. Journal of Health Economics 23: 965-995.

Dale A, Williams M, Dodgeon B (1996) Housing Deprivation and Social Change : A report based on the analysis of individual level census data for 1971, 1981 and 1991 drawn from the Longitudinal study and the Samples of anonymised records. Office For National Statistics, Series LS No. 8. London: HMSO.

Department of Health (1998) Green Paper: Our Healthier Nation. London: The Stationery Office.

Department of Health (1999) White Paper: Our Healthier Nation. London: The Stationery Office.

Dunn JR (2000) Housing and Health Inequalities: Review and Prospects for Research. Housing Studies 15: 341-66.

Grossman M (1972a) On the Concept of Health Capital and the Demand for Health. Journal of Political Economy 80: 223-55.

Grossman M (1972b) The Demand for Health: A Theoretical and Empirical Investigation. Occasional Paper National Bureau of Economic Research 119.

Grossman M (1999), The Human Capital Model of the Demand of Health. *NBER* Working Paper 7078.

Hsiao C (1985) Benefits and Limitations of Panel Data. Econometric Reviews 4: 121-74.

Ineichein B (1993) Homes and Health. London: E & FN Spon.

Jacobson L (2000) The Family as a Producer of Health – An Extended Grossman Model. Journal of Health Economics 19: 611-37.

Kemna H (1987) Working Conditions and the Relationship between Schooling and Health. Journal of Health Economics 6: 189-210.

Kenkel DS (1995) Should you Eat Breakfast? Estimates from Health Production Functions. Health Economics 4: 15-29.

Knott M, Tzamourani P (1997) Fitting a Latent Trait Model for Missing Observations to Racial Prejudice Data. In: Rost J, Langeheine R (eds.), Applications of latent trait and latent class models in the social sciences. Waxmann, Münster.

Krieger J, Higgins DL (2002) Housing and Health: Time Again for Public Health Action. American Journal of Public Health 92: 758-768.

Kunst AE, Geurts J, Van Den Berg J (1995) International Variation in Socioeconomic Inequalities in Self-Reported Health. Journal of Epidemiology and Community Health 49: 117-123.

Lowry S (1991) Housing and Health. London: BMJ.

Manning W, Newhouse, JP, Ware JE (1981) The Status of Health in Demand Estimation; or, Beyond Excellent, Good, Fair, Poor. In: Fuchs V (ed), Economic Aspects of Health. Chicago: University of Chicago Press.

Marcoulides GA, Moustaki I (eds.) (2002) Latent Variable and Latent Structure Models. Lawrence Earlbaum Publishers, NJ, United States.

Marsh A, Gordon D, Heslop P, Pantazis C (2000) Housing Deprivation and Health: A Longitudinal Analysis. Housing Studies 15: 411-28.

Matyas L, Sevestre P (eds) (1992) The Econometrics of Panel Data. Kluwer Academic Publishers: The Netherlands.

Ministry of Social Development and Economic Security (2001) Homelessness-Causes & Effects: The Cost of Homelessness in British Columbia. British Columbia Ministry of Social Development and Economic Security and BC Housing Management Commission, Volume 3, Toronto.

Moisio P (2004) Latent Class Application to the Multidimensional Measurement of Poverty. Quality and Quantity 38: 703-717.

Moustaki I, Knott M (2000) Weighting for Item non-response in attitude scales using latent variable models with covariates. Journal of the Royal Statistical Society (Series A) 163: 445-459.

Navarro C, Ayala L (2008) Multidimensional Housing Deprivation Indices with application to Spain. Applied Economics 40, 597-611.

Navarro C (2002) Privación en vivienda y estado de salud: una estimación para el caso español. In: Cabases JM et al (eds): La organización territorial de la sanidad. Navarra: Universidad Pública de Navarra, Asociación de Economía de la Salud: 377-404.

Power C, Hertzman C (1997) Social and Biological Pathways Linking Early Life and Adult Disease. In: Marmot M, Wadsworth M (eds), Fetal and Early Childhood Environment: Long-Term Health Implications. British Medical Bulletin 53: 210-23.

Rosenzweig M, Schultz TP (1983) Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and their Effects on Birth Weight. Journal of Political Economy 91: 723-46.

Silber J (2007) Measuring Poverty: Taking a Multidimensional Perspective. FEDEA Working Papers 2007-14.

Smith SJ (1989) Housing and Health: A Review and Research Agenda. Discussion Paper 27, Centre for Housing Research, Glasgow University.

Taubman P, Rosen S (1981) Healthiness, Education and Marital Status. In: Fuchs, V. (ed)., Economic Aspects of Health. Chicago: University of Chicago Press.

Van Doorslaer E, Wagstaff A, Bleichrodt H, Calonge S, Gerdtham, U-G, Gerfin M, Geurts J, Gross L, Hakkinen U, Leu RE (1997) Income-Related Inequalities in Health: Some International Comparisons. Journal of Health Economics 16: 93-112.

Wagstaff A, Van Doorslaer E, Paci P (1989) Equity in the Finance and Delivery of Health Care: Some Tentative Cross-Country Comparisons. Oxford Review of Economic Policy 5: 89-112.

Table 1. Binary logistic model with individual effects (odds ratios)

	Model 1		Model 2	
Health	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.
Housing conditions				
Have separate kitchen	1,048	0,124		
Have separate bath	1,099	0,148		
Have indoor flushing toilet	0,960	0,157		
Hot running water	1,214	0,035		
Heating	1,167	0,101		
Place to sit outside	1,056	0,029		
Noise problems	1,036	0,029		
Too dark	1,106	0,036		
Leaky roof	1,170	0,090		
Damp	0,939	0,049		
Rot in window frames or floor	0,911	0,065		
Overcrowding	1,019	0,004		
Pollution	1,045	0,038	1,076	0,037
Vandalism	1,199	0,038	1,213	0,037
Housing deprivation			1,798	0,166
Marital Status				
Separated	1,462	0,175	1,432	0,157
Divorced	1,055	0,177	1,103	0,169
Widowed	1,264	0,175	1,197	0,160
Married	1,052	0,178	1,007	0,166
Main source of income				
Self-employment	0,961	0,061	0,966	0,060
Pensions	1,295	0,083	1,327	0,083
Unemployment benefits	1,104	0,083	1,099	0,080
Other social benefits	1,359	0,102	1,402	0,103
Private income	1,008	0,080	1,025	0,080
Person has no income	0,933	0,068	0,948	0,068
Education				
2nd stage (isced 3)	0,874	0,067	0,862	0,065
3rd level (isced 5-7)	1,164	0,166	1,166	0,163
Social Relationship				
Once/twice a week	1,202	0,039	1,204	0,039
Once/twice a month	1,196	0,070	1,221	0,070
Less often	1,410	0,137	1,413	0,135
Never	1,306	0,310	1,231	0,284
Age				
< 25 years	0,974	0,094	0,968	0,093
25-65 years	0,659	0,121	0,652	0,119
Chronic condition	3,199	0,130	3,140	0,126
Sex	0,802	0,024	0,810	0,023
dt95	0,897	0,040	1,062	0,037
dt96	0,840	0,031	0,954	0,033
dt97	0,904	0,035	0,973	0,034

Income				
Decile 2	0,989	0,052	1,025	0,053
Decile 3	0,946	0,051	0,948	0,050
Decile 4	0,946	0,051	0,956	0,050
Decile 5	0,974	0,053	0,995	0,053
Decile 6	0,920	0,051	0,953	0,051
Decile 7	0,862	0,049	0,898	0,049
Decile 8	0,906	0,052	0,943	0,052
Decile 9	0,858	0,051	0,912	0,052
Decile 10	0,820	0,052	0,848	0,052
Satisfaction with housing	0,922	0,014	0,916	0,013
Type of employment contract				
Fixed/short-term contract	0,690	0,080	0,673	0,076
Other arrangement	0,811	0,148	0,785	0,140
Permanent employment	0,794	0,089	0,799	0,088
Not working	0,837	0,093	0,827	0,090
Tenure status				
Rent	0,785	0,059	0,862	0,056
Accommodation provided free	0,989	0,097	0,992	0,095
Regions	0,799	0,020	0,829	0,021
Neighbourhood Unemployment Rate	1,116	0,153	1,101	0,147
$\bar{z}_{mstatus}$	1,118	0,048	1,104	0,046
$\bar{z}_{sourceincome}$	1,039	0,015	1,042	0,015
$\bar{z}_{education}$	0,653	0,050	0,672	0,049
$\bar{z}_{relationship}$	1,185	0,040	1,192	0,040
$\bar{z}_{satisfactionwithhou\sin g}$	0,874	0,018	0,909	0,019
$\bar{z}_{age}$	0,525	0,051	0,496	0,048
$\bar{z}_{chronic}$	8,946	0,492	9,123	0,495
$\bar{z}_{tenurestatus}$	0,669	0,043	0,770	0,043
$\bar{z}_{hou\sin gdeprivation}$			1,199	0,030
Number of observations	56622		58043	
Log Likelihood	-22436837		-23103685	
Chi2	27107.51(63)		27576.27(53)	
Prob>Chi2	0.0000		0.0000	
LR test				
Prob>Chi2 (LR)	0.0000		0,0000	

Notes.

1. Baseline: not suffering housing deprivation, never married, earnings as the main source of income, less than 2nd stage as highest level of education attained, first decile of equivalent income, not at all satisfied with housing, meet people most days, not chronic condition, casual work, woman, more than 65 years old, owner, household situated in North West, Centre of Spain and Canary.
2. Degrees of freedom are in parenthesis after the test figures.

Table 2. Maximum Likelihood Estimates of Item Parameters

Items (year 1995)	α_{0i}	Standard error	α_{1i}	Standard error	Standardised α_{1i}	P(X=1/Z=0)
Hot running water	-3.952	0.122	1.255	0.101	0.782	0.019
Heating	-0.084	0.026	0.477	0.041	0.431	0.479
Leaky roof	-4.027	0.191	2.585	0.171	0.933	0.018
Damp	-2.879	0.161	2.794	0.199	0.942	0.053
Rot in window/frames/floor	-4.056	0.165	2.231	0.138	0.913	0.017
Overcrowding	-2.679	0.053	0.274	0.068	0.264	0.064
% explained				88.31		
LR test				88.89		
$\chi^2(29)$ for observed response patt.				46.85		
$\chi^2(29)$ for all response patterns				68.25		
Items (year 1996)	α_{0i}	Standard error	α_{1i}	Standard error	Standardised α_{1i}	P(X=1/Z=0)
Hot running water	-4.377	0.154	1.432	0.116	0.82	0.012
Heating	-0.246	0.027	0.503	0.041	0.449	0.439
Leaky roof	-3.823	0.182	2.608	0.171	0.934	0.021
Damp	-3.352	0.257	3.645	0.322	0.964	0.034
Rot in window/frames/floor	-4.014	0.151	1.957	0.122	0.891	0.018
Overcrowding	-2.59	0.051	0.252	0.066	0.244	0.07
% explained				89.42		
LR test				98.59		
χ^2 (26) for observed response patt.				61.57		
χ^2 (26) for all response patterns				79.88		
Items (year 1997)	α_{0i}	Standard error	α_{1i}	Standard error	Standardised α_{1i}	P(X=1/Z=0)
Hot running water	-4.512	0.164	1.483	0.121	0.829	0.011
Heating	-0.382	0.028	0.431	0.041	0.396	0.406
Leaky roof	-3.738	0.185	2.615	0.175	0.934	0.023
Damp	-3.07	0.213	3.332	0.273	0.958	0.044
Rot in window/frames/floor	-4.5	0.205	2.393	0.161	0.923	0.011
Overcrowding	-2.547	0.054	0.333	0.069	0.316	0.073
% explained				84.86		
LR test				154.56		
χ^2 (26) for observed response patt.				104.41		
χ^2 (26) for all response patterns				133.50		
Items (year 1998)	α_{0i}	Standard error	α_{1i}	Standard error	Standardised α_{1i}	P(X=1/Z=0)
Hot running water	-4.997	0.209	1.612	0.142	0.85	0.007
Heating	-0.396	0.029	0.406	0.044	0.376	0.402
Leaky roof	-4.521	0.284	2.86	0.246	0.944	0.011
Damp	-3.531	0.352	3.431	0.415	0.96	0.028
Rot in window/frames/floor	-4.365	0.181	2.037	0.139	0.898	0.013
Overcrowding	-2.639	0.056	0.205	0.076	0.201	0.067
% explained				90.40		
LR test				89.12		
χ^2 (20) for observed response patt.				53.69		
χ^2 (20) for all response patterns				71.52		

Appendix

Table A.1 Standard binary logistic model (odds ratios)

	Model 1		Model 2	
Health	Odds Ratio	Std. Err.	Odds Ratio	Std. Err.
Housing conditions				
Have separate kitchen	1,036	0,119		
Have separate bath	1,216	0,159		
Have indoor flushing toilet	0,924	0,147		
Hot running water	1,208	0,034		
Heating	0,932	0,071		
Place to sit outside	1,071	0,029		
Noise problems	1,026	0,028		
Too dark	1,116	0,035		
Leaky roof	1,326	0,099		
Damp	1,128	0,053		
Rot in window frames or floor	0,885	0,061		
Overcrowding	1,015	0,004		
Pollution	1,055	0,038	1,083	0,036
Vandalism	1,231	0,038	1,245	0,037
Housing deprivation			2,291	0,202
Marital Status				
Separated	1,802	0,189	1,702	0,160
Divorced	1,431	0,205	1,409	0,181
Widowed	1,899	0,115	1,688	0,096
Married	1,732	0,069	1,532	0,057
Main source of income				
Self-employment	0,952	0,059	0,945	0,058
Pensions	1,514	0,093	1,583	0,095
Unemployment benefits	1,123	0,081	1,160	0,081
Other social benefits	2,104	0,135	2,236	0,140
Private income	1,188	0,077	1,196	0,076
Person has no income	1,106	0,054	1,116	0,054
Education				
2nd stage (isced 3)	0,596	0,026	0,542	0,020
3rd level (isced 5-7)	0,547	0,028	0,507	0,021
Social Relationship				
Once/twice a week	1,290	0,037	1,286	0,036
Once/twice a month	1,414	0,070	1,446	0,070
Less often	1,839	0,156	1,861	0,154
Never	2,008	0,437	1,972	0,418
Age				
< 25 years	0,487	0,022	0,465	0,020
25-65 years	0,161	0,011	0,146	0,009
Chronic condition	11,060	0,312	11,087	0,307
Sex	0,797	0,022	0,800	0,022
dt95	1,054	0,044	1,172	0,039
dt96	0,865	0,031	1,008	0,034
dt97	0,859	0,032	0,983	0,033

Table A.1 (cont.)

Income				
Decile 2	1,058	0,054	1,067	0,053
Decile 3	1,004	0,053	0,974	0,050
Decile 4	1,011	0,053	0,971	0,050
Decile 5	1,045	0,056	1,008	0,052
Decile 6	1,006	0,054	0,974	0,050
Decile 7	0,918	0,050	0,882	0,047
Decile 8	0,984	0,055	0,935	0,050
Decile 9	0,927	0,054	0,894	0,050
Decile 10	0,873	0,054	0,819	0,048
Satisfaction with housing	0,865	0,009	0,853	0,008
Type of employment contract				
Fixed/short-term contract	0,696	0,080	0,680	0,076
Other arrangement	0,809	0,147	0,803	0,142
Permanent employment	0,791	0,088	0,793	0,085
Not working	0,910	0,100	0,916	0,098
Tenure status				
Rent	0,669	0,044	0,675	0,030
Accommodation provided free	0,614	0,042	0,633	0,036
Regions	0,797	0,020	0,794	0,019
Neighbourhood Unemployment Rate	1,060	0,141	1,079	0,140
Number of observations	56.622		58.043	
Log Likelihood	25253.80		25546.31	
Chi2	0.0000		0.0000	
Prob>Chi2	0.3508		0.3462	

Table A.2. Summary of Second and Third Order Observed and Expected Margins

Year	Response	Observed	Expected	Observed-Expected	Max.(O-E)^2/E	Items
1995	(1.1)	88	67.5241	20.4759	6.2091	(5.1)
	(1.0)	167	181.1074	-14.1074	1.0989	(5.2)
	(0.1)	150	170.5435	-20.5435	2.4746	(5.1)
	(0.0)	5844	5823.3717	20.6283	0.0731	(5.1)
	(1.1.1)	14	7.2054	6.7946	6.4073	(1.5.6)
1996	(1.1)	65	49.7372	15.2628	4.6837	(5.1)
	(1.0)	125	157.6978	-32.6978	6.7797	(5.2)
	(0.1)	80	67.3970	12.6030	2.3567	(4.1)
	(0.0)	3335	3301.6362	33.3638	0.3371	(5.2)
	(1.1.1)	47	34.0366	12.9634	4.9374	(1.2.5)
1997	(1.1)	105	87.4148	17.5852	3.5376	(2.1)
	(1.0)	137	164.8155	-27.8155	4.6943	(5.2)
	(0.1)	55	73.3834	-18.3834	4.6053	(2.1)
	(0.0)	2813	2845.0000	-32.7030	0.3758	(4.2)
	(1.1.1)	17	27.0595	-10.0595	3.7397	(3.5.6)
1998	(1.1)	184	161.7401	22.2599	3.0636	(5.2)
	(1.0)	111	133.0863	-22.0863	3.6653	(5.2)
	(0.1)	40	52.2466	-12.2466	2.8706	(2.1)
	(0.0)	3140	3118.2857	21.7143	0.1512	(5.2)
	(1.1.1)	3	5.6697	-2.6697	1.2571	(1.2.6)

Note: An alternative goodness-of fit is to check how well the model fits not the whole response pattern but the one-two and three way margins. It will indicate pair or triples of items where the model does not fit. As we said previously, it consists of calculating the Pearson χ^2 for combinations of two or three responses. The residuals offer information about the predictions the model makes on the response patterns composed of two or three elements. The sixth column of this table shows the highest residuals of each of the second and third order combinations. Hence, for instance, the second row shows that there are 167 individuals that responded positively to indicators 5 and 2 and the highest residual registered for this pattern of responses in 1.0989. Only values higher than 4 or 5 can indicate a bad fit for the model. Despite the fact that there are some residuals with a value above five, we can therefore conclude that the fit of the model with the unidimensional latent variable is satisfactory.

Table A. 3. Variable definition and descriptives

Variables	Definition	μ 1995	1996	1997	1998
Dependent variable					
Health	Unhealthy: (yes=1) (not=0).	0.34	0.33	0.32	0.32
Health in general (original variable)	Health in general: very good(=1) good(=2) fair(=3) bad(=4) and very bad(=5).	2.29	2.26	2.26	2.28
Housing conditions					
Have separate kitchen	deprived (yes=1) (not=0)	0.01	0.01	0.01	0.01
Have separate bath	deprived (yes=1) (not=0)	0.01	0.01	0.01	0.01
Have indoor flushing toilet	deprived (yes=1) (not=0)	0.01	0.01	0.01	0.00
Hot running water	deprived (yes=1) (not=0)	0.03	0.02	0.02	0.02
Heating	deprived (yes=1) (not=0)	0.46	0.43	0.39	0.40
Place to sit outside	deprived (yes=1) (not=0)	0.27	0.24	0.27	0.24
Noise problems	Noise problems (yes=1) (not=0)	0.18	0.21	0.17	0.15
Too dark	Not enough light (yes=1) (not=0)	0.10	0.12	0.11	0.09
Leaky roof	Leaky roof (yes=1) (not=0)	0.20	0.21	0.23	0.18
Damp	Damp (yes=1) (not=0)	0.08	0.07	0.06	0.05
Rot in window frames or floor	Rot in window/frame/floor (yes=1) (not=0)	0.31	0.35	0.33	0.29
Pollution	Pollution problems (yes=1) (not=0)	0.19	0.14	0.13	0.13
Vandalism	Vandalism problems (yes=1) (not=0)	0.24	0.21	0.19	0.18
Overcrowding	Less rooms than number of adults (yes=1)(not=0)	0.14	0.15	0.17	0.15
Neighbourhood Unemployment Rate	Neighbourhood Unemployment Rate	0.10	0.11	0.09	0.08
Housing Deprivation	Index of housing deprivation	0.01	0.02	0.02	-0.01
Satisfaction with housing	Level of satisfaction: scale 1-6. Not at all satisfied(=1). fully satisfied(=6).	4.39	4.39	4.35	4.49
Marital Status	separated(yes=1) (not=0)	0.01	0.01	0.01	0.01
	divorced(yes=1) (not=0)	0.01	0.01	0.01	0.01
	widowed(yes=1) (not=0)	0.08	0.08	0.08	0.08
	married(yes=1) (not=0)	0.60	0.59	0.57	0.56
Sex	male(=1), female(=0)	0.48	0.48	0.49	0.48
Age	25-65 years old(yes=1) (not=0)	0.64	0.64	0.63	0.63
	16-25 years old(yes=1) (not=0)	0.19	0.19	0.20	0.19
Regions (aggregate)	Situated in North West, Centre and Canary(=0), in North East, Madrid, East and South(=1)	0.63	0.62	0.62	0.62
Chronic condition	Chronic condition: (yes=1) (not=0).	0.23	0.23	0.23	0.23
Type of employment contract	Fixed/short-term contract(yes=1) (not=0)	0.09	0.08	0.08	0.08
	Other arrangement(yes=1) (not=0)	0.01	0.01	0.01	0.01
	Permanent employment(yes=1) (not=0)	0.18	0.18	0.19	0.20
	Not working(yes=1) (not=0)	0.72	0.72	0.71	0.70
Tenure Status	Rent: (yes=1) (not=0).	0.12	0.11	0.10	0.09
	Accommodation provided free: (yes=1) (not=0).	0.05	0.06	0.05	0.05
Education	2nd stage (isced 3) (yes=1) (not=0)	0.18	0.19	0.19	0.19
	3rd level (isced 5-7)(yes=1) (not=0)	0.14	0.14	0.14	0.19
Social relationship	Once/twice a week(yes=1) (not=0)	0.22	0.21	0.25	0.24
	Once/twice a month(yes=1) (not=0)	0.06	0.06	0.06	0.05
	Less often(yes=1) (not=0)	0.02	0.02	0.02	0.02
	Never(yes=1) (not=0)	0.00	0.00	0.01	0.00
Main source of income	Self-employed (yes=1) (not=0)	0.06	0.06	0.06	0.07
	Pensions (yes=1) (not=0)	0.17	0.17	0.17	0.17
	Unemployment. benefits (yes=1) (not=0)	0.04	0.03	0.03	0.03
	Other social benefits (yes=1) (not=0)	0.07	0.07	0.07	0.06
	Private income (yes=1) (not=0)	0.06	0.05	0.07	0.09
	Person has no income (yes=1) (not=0)	0.27	0.30	0.28	0.26

Notes:

1. Weighted data based on the variable representing cross-sectional weighting for each year.
2. Baseline: to be healthy, not suffering housing deprivation, not at all satisfied with housing, never married, woman, more than 65 years old, not chronic condition, casual work (type of employment contract), less than 2nd stage as highest level of education attained, first decile of equivalent income, meet people most days, earnings as the main source of income, owner, household situated in North West, Centre of Spain and Canary.