Health Shocks and the Hazard Rate of Early Retirement in the ECHP

Ronald Hagan[a], Andrew M. Jones[b], and Nigel Rice[c]

JEL-Classification: H55 I12 J26
Keywords: Retirement; Health; ECHP; European Union

1. Introduction

In this paper we use discrete-time proportional hazard models with internationally comparable longitudinal data from the full eight waves of the European Community Household Panel (ECHP) to study the relationship between retirement, health levels and health shocks across nine European Union countries. We use a stock sample approach that conditions on individuals within a set age range and in employment at the first wave of the ECHP and follows them over subsequent waves. A variety of health measures, including self assessed health (SAH), a latent health stock and graduated acute health shocks are used, as are alternative definitions of retirement. Self-reported health shocks are recorded prior to the individual's date of exit from the labour market and this should help to mitigate the influence of justification bias, although anticipation effects may still be a factor.

a School of Women's and Infants' Health, University of Western Australia, Perth, Western Australia.
b Department of Economics and Related Studies, University of York, United Kingdom.
c Centre for Health Economics, University of York, United Kingdom.
Acknowledgements: The *European Community Household Panel Users' Database*, version of December 2003, was supplied by Eurostat. We are grateful to Cristina Hernández Quevedo for her assistance in preparing the version of the ECHP-UDB that was used in this paper. Material from this paper was presented at the Swiss Society of Economics and Statistics (SSES), Lausanne, March 2008.

2. Data

2.1 The Stock Sample

Data drawn from the full eight waves (1994–2001) of the European Community Household Panel (ECHP) dataset are used. The ECHP is a standardized annual longitudinal survey carried out among the pre-enlargement member countries of the European Union (EC-15). The ECHP was designed and coordinated by the Statistical Office of the European Communities (Eurostat) covering, at the level of households and individuals, information on demographics, participation in the labour market, income, health, education and training, housing etc. With some minor exceptions, the same sampling methodology and an identical questionnaire was used in each participating member state (see e.g., PERACCHI, 2002, NICOLETTI AND PERACCHI, 2005, JONES et al., 2006). The first wave was conducted in 1994 and the User's Data Base (ECHP-UDB), which is an anonymised and user-friendly version of the data, has been available since December 1998. The version containing the complete eight waves (December 2003) is used in this study. Nine countries were selected for study, listed in alphabetical order: Belgium, Denmark, France, Greece, Ireland, Italy, Portugal, Spain and the UK. These particular countries were selected as they all had data from the first wave of the panel in relation to health variables and employment activity[1].

To define an appropriate risk set for the hazard of retirement we selected those individuals who were present at wave 1 in the ECHP, were aged 50–64 years in wave 1, were employed or self-employed in the first wave and had measures of health and employment activity recorded for all available waves. Once an individual retired or was missing then their data after that point was excluded. This stock sample consists of all individuals in work in the first wave of the ECHP. The transitions to other employment states are summarised for each wave of the data, pooled across countries. This confirms the selection of only individuals who are in work, either employed or self-employed, in wave 1 (Table 1).

The nine countries selected for study have a total of 105,613 participants in the ECHP. Of these 22,456 met the age selection criteria for the stock sample with 9,631 (43% of age group) meeting the employment and complete data requirements for the stock sample. Those analysed regarding retirement totalled 8,629 (38% of age group). The mean age of individuals in each country's stock sample is similar though there are differences in the distribution across the age groups as shown in Table 2. There are differences in employment status and in

1 For the UK we use the comparable sample of BHPS data that is supplied with the ECHP-UDB, rather than the original ECHP sample which is only available for 3 waves.

Table 1: Labour Market Status by Wave in ECHP Sample

	1	2	3	4	5	6	7	8
Employed	6119	4795	3985	3169	2539	2103	1705	1398
self-employed	3512	2772	2358	1968	1664	1385	1183	1035
unemployed		190	235	250	207	163	125	102
retired		478	938	1371	1699	1939	2180	2344
other		394	421	454	457	474	447	428
attrition or death		1002	1694	2419	3065	3567	3991	4324
total	9631	9631	9631	9631	9631	9631	9631	9631
in-work	9631	7576	6343	5137	4203	3488	2888	2433

Table 2: Comparisons of Characteristics of Individuals in Wave 1 of Stock Samples by Country

Country	Be	Dk	Fr	Gr	It	Ire	Pt	Sp	UK
n	440	651	965	1226	1504	755	1354	1180	799
self employed (%)	25	20	20	68	38	44	45	35	20
mean age (yrs)	55	55	54	56	55	56	56	56	55
age 50–54 (%)	55	51	58	42	51	45	40	46	53
age 55–59 (%)	30	33	35	34	35	33	34	33	33
age 60–64 (%)	15	16	7	24	14	22	27	21	15
males (%)	66	56	59	70	73	78	63	74	55
SAH									
very good (%)	20	48	14	32	10	39	3	12	25
good (%)	54	34	46	38	41	43	42	46	53
fair (%)	23	14	35	23	39	16	37	30	17
bad (%)	2	3	3	6	9	1	16	11	5
very bad (%)	0	0	2	1	2	0	2	2	1
severe limitation (%)	3	1	8	4	5	1	6	4	8[1]
some limitation (%)	18	11	10	13	19	13	21	12	–
no limitation (%)	79	87	82	83	76	86	73	84	92

the proportion of males between countries in the sample. There are also differences in the reported self-assessed health status and the proportion having some degree of limitation because of health problems (Table 2) similar to that reported by HERNANDEZ-QUEVEDO et al. (2006).

2.2 Health Shocks

It was emphasised earlier that using measures of self-reported health that are recorded prior to an individual's exit from the labour market should help to mitigate problems of justification bias. However there may be antipcipation effects and measurement errors may still occur. To address this possibility the method of estimating a model of SAH as a function of a set of health indicators is used to define a latent "health stock". This follows the approach of STERN (1989) and BOUND (1991) as implemented in BOUND et al. (1999) and more recently used by DISNEY et al. (2006) and RICE et al. (2006). There are differences in how the latter two studies have created their latent health variable. DISNEY *et al* (2006) use personal characteristics as well as health indicators while RICE et al. (2006) only include the health indicators when constructing the health stock. In this study both forms of the latent health variable are used.

The main decisions in computing the health stock variable are in the choice of health indicators and whether to include demographic variables in the model. The ECHP does not have the same extensive range of objective health variables as the BHPS, as used by DISNEY et al. (2006) and RICE et al. (2006), but does have a set of measures relating to limitation in daily activities, recent illness or mental problem and a history of in-patient stay in hospital. Two latent health stock measures were created; one using only the health indicators and one using demographic variables in addition. The first model regressed the measure of self-reported self-assessed health (SAH) on the health indicators.[2] The coefficients from the ordered probit models were used to obtain the latent health stock variable for each country. In all countries the estimated coefficients displayed the expected negative sign demonstrating that indicators of health problems were associated with lower reporting of self-assessed health. We refer to this measure as the non-normalised latent health stock.

2 The UK data has some limitation recorded only in wave 6. In the other waves it is collapsed into severe limitation. The UK data also does not contain the mental problem data. The other countries in the ECHP have all variables available in each wave except France which codes illness and mental problem in waves 1, 3, 4 only. The ordered probit models were adjusted accordingly.

The same approach was taken in estimating the health stock variable using the method of Disney et al. (2006). They estimated the individual's health stock for each wave of the data using the wave specific values of the objective health variables but also included some demographic variables. These estimated values are then normalised as a deviation of the individual's health index from the cohort mean for each year. This predicted individual "health stock" thus creates a health stock for each individual in relation to the year on year average for the sample. The normalised variable has a mean of 0 and a standard deviation of 1 for each wave of the sample and we refer to this measure as the normalised latent health stock. This normalisation is carried out separately for each county in order to address concerns about cross-country differences in self-reporting of health.

Both health stock variables can be used to demonstrate gradual deterioration in an individual's health. However the decision to retire may be more related to a "health shock", i.e. a more acute deterioration. To identify a health shock we have conditioned on the initial levels of health stock and used the lagged value of the health stock variable to capture deviations or "shocks". Conditioning on the initial level of health stock allows the estimated coefficients on lagged health to represent the deviation from the underlying health stock and has the advantage of helping to control for individual-specific unobserved health-related heterogeneity.

In addition to this approach we have used a more direct measure of a health shock based on the differences between consecutive waves in an individual's normalised health stock value. This is similar to the concept of an acute health shock described by Riphahn (1999), although she uses reported health rather than a purged measure. Three mutually exclusive binary dummy variables were created for a "small" acute health shock representing a decrement of $\geq 0.5 - < 1.0$ standard deviations; a "medium" acute health shock representing a decrement of $\geq 1.0 - < 2.0$ standard deviations and a "large" acute health shock representing a decrement of ≥ 2.0 standard deviations. These three binary dummy variables thus represent graduated health shocks and are used in a model that also conditions on the initial normalised health stock value.

A clear gradient of increasing severity of health challenge is apparent, with increasing severity of health shock associated with both an increase in the number and in the severity of adverse health events. This is illustrated for wave 2 of the data in Table 3. This is the first wave where an acute health shock can be measured. What this can mean at an individual level is shown by the two examples of the reported experiences of two individual respondents from the Irish data shown in Table 4.

For all countries there was a general decline in the health status of the stock sample as individuals aged during the study period. This deterioration in health

Table 3: Correlation of Health Shocks with Objective Health Measures

	No shock	Small shock	Medium shock	Large shock
n	5963	320	335	307
Illness (%)	3.2	37	39	50
Inpatient (%)	3.4	37	32	38
Some or severe limitation (%)	6.3	34	68	99
Number of health events[a] (%)				
0	89	15	1	0
1	7	56	57	27
2	2	19	28	34
3	1	6	9	22
4–5	0.2	4	6	16

a sum of binary dummy variables illness, mental problem, inpatient, some limitation and 2* severe limitation; range 0–5

Table 4: Health Shocks – Individual Examples

	SAH	health stock	health shock	health events	outcome
52 year old male – employed					
W1	v gd	.525		no problems, no limitations	
W2	v gd	.540	–	no problems, no limitations	
W3	Fair	–3.6	large	illness, mental problem, severe limitation	retires
W4	Fair	–3.4	–	illness, severe limitation	retired
W5	Fair	–3.5	–	illness, mental problem, severe limitation	retired
53 year old female, unpaid work in family enterprise					
W1	Fair	.452		no problems, no limitation	
W2	Fair	–1.7	large	illness, mental prob, in-patient, no limitation	
W3	Fair	–2.9	medium	in-patient, some limitation	house work
W4	Fair	–2.9	–	illness, in-patient, some limitation	other – inactive
W5	Bad	–4.1	medium	illness, mental prob, in-patient, severe limitation	other – inactive

was accompanied by the occurrence of acute health shocks though these did not increase in prevalence across the waves occurring in 13.9% of individual's in wave 2 and 13.8% in wave 8 (Table 5). The large acute health shocks are the least frequent and in the individual country data involve only small numbers of individuals.

Table 5: Occurrence of Acute Health Shocks by Wave in the Pooled Data

Health shock	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8
None	7355	6865	6164	5599	5192	4864	4541
Small	411	327	360	287	305	198	265
%	4.8	4.2	5.0	4.4	5.1	3.5	5.0
Medium	421	401	356	414	362	365	338
%	4.9	5.1	5.0	6.4	6.0	6.5	6.4
Large	357	268	260	214	156	162	127
%	4.2	3.4	3.6	3.3	2.6	2.9	2.4
Total	8544	7861	7140	6514	6015	5589	5271

3. Econometric Models and Results

We estimate hazard functions for the first transition to self-reported retirement. This is done using the stock-sampling approach of Jenkins (1995) which represents the transition to retirement as a discrete-time hazard specification, based on the Prentice and Gloekler (1978) model. For this analysis the data are organised so that the unit of analysis is the time at risk of the event i.e., "retirement". This organisation of the data and conditioning on the stock sampling – such that time periods prior to selection into the sample can be ignored – means that the estimation of a discrete-time hazard model is simplified, to the extent that any method suitable for the estimation of a binary responses may be used. We use a complementary log-log specification which is consistent with an underlying continuous time proportional hazard model.

Separate models were estimated using the following measures of health: the observed data on the absence of health limitations; the subjective measure of SAH; the estimated latent health stock variables; and the graduated health shock variables. Each model includes the same set of demographic and financial variables. All models were estimated using two alternative definitions of retirement:

self-reported retirement and an expanded definition based on inactivity in the labour market.

The models pool all of the data across countries. In the hazard models the latent health measures, the subjective SAH and the absence of health limitations are entered as both a lagged and the initial (wave 1) value. In the models with acute health shocks the shocks are not lagged but are also conditioned on the initial value of the normalised latent health stock. The complete regression results are available from the authors on request, for brevity we summarize the key results in terms of hazard ratios. The hazard ratio measures the proportional effect on the underlying (instantaneous) hazard of retiring of a one unit change in the value of the variable in question.

For example, in the models with acute health shocks the hazard ratios for the three health shock variables are 1 or greater indicating that, once conditioned on initial latent health, the occurrence of an acute health shock is associated with an increase in the probability of retiring during that time period. Hence, the occurrence of a medium acute health shock would (ceteris paribus) increase the probability of retiring (as measured by self-report) by approximately 44% and for a large acute health shock this increases slightly to 47%. These effects are consistently greater across all three acute health shocks in the expanded retirement definition model and the gradient between the three graduated acute shocks is more pronounced demonstrating a clear "dose-response" relationship in the gradient of the effects.

All measures of a change in health are associated with a change in the hazard of retirement. This is detailed for each measure and both definitions of retirement in Table 6 with the change in the hazard of retiring for a one unit change in each health measure shown as an absolute value. The results are consistent in direction within and across the definitions of retirement. Higher values of the latent health stock measures and SAH, and the absence of limitations due to health, are all associated with a decrease in the likelihood of retiring whereas the medium and large acute health shocks are associated with a significant increase in the likelihood of retiring. All measures consistently show a greater strength of effect in the expanded retirement models and this is most marked with the largest of the acute health shocks. There is evidence of a gradient in the response to the acute health shocks and this is clearer in the expanded retirement models. The effect size in absolute terms is also greater for the acute health shocks suggesting that acute changes have more impact than gradual changes as represented by the lagged latent health stock variables. The absolute effect size for the subjective lagged SAH – which is increasing in good health – is closer to the effect size of the normalised rather than the non-normalised latent health stock.

Table 6: Hazard Ratios for Models with Different Health Measures by Retirement Definition – Pooled Data (95% Confidence Intervals in Parentheses)

Health measure	Self-reported retirement		Expanded retirement	
SAH	0.85	(.80, .90)	0.82	(.78, .86)
No limitation due to health	0.75	(.67, .84)	0.70	(.63, .78)
Latent health variables:				
non-normalised	0.82	(.77, .87)	0.77	(.73, .82)
normalised	0.87	(.83, .91)	0.83	(.80, .87)
Acute health shocks:				
small	1.00	(.82, 1.20)	1.14	(.97, 1.35)
medium	1.44	(1.24, 1.67)	1.50	(1.31, 1.72)
large	1.47	(1.21, 1.78)	2.06	(1.76, 2.40)

The models include country-specific intercepts and there is significant variation in these country effects. This is demonstrated in Table 7 using the United Kingdom as the base country. Belgium, France and Italy have an increased hazard of retiring, ceteris paribus, compared to the UK and this is consistent in direction across retirement definitions and predominantly consistent in size of effect within each of the retirement definitions.

There are differences between these countries in the incentives in their social security and tax systems in relation to retirement, in the age at which early retirement is permitted and in the manner that benefits are accrued (Gruber and Wise, 1998, 1999; Duval, 2003). Factors that may influence the decision to retire include the standard and, in particular, early age of entitlement to public pension benefits, the generosity of pension benefits as reflected in the replacement rate and levels of pension wealth, and the implicit marginal tax on continued work. Not surprisingly then, the results reveal major differences between countries in the probability of retiring when compared to the UK. In summary, the group of countries (Belgium, France, Italy) where the hazard of retirement is significantly greater than in the UK are those whose public pension systems create higher incentives to retire in comparison to the UK. These findings are robust across the different health measures as well as across the two definitions of retirement.

Table 7: Hazard Ratios for Country Effects Relative to UK by Health Model and Retirement Definition – Pooled Data (95% Confidence Intervals in Parentheses)

	Self-reported retirement					
	Model with acute health shock			Model with normalised latent health		
Belgium	0.95	(.67,	1.35)	2.01	(1.57,	2.56)
Denmark	0.68	(.51,	.91)	1.11	(.89,	1.39)
France	1.38	(1.08,	1.77)	2.12	(1.74,	2.57)
Greece	0.35	(.24,	.51)	1.04	(.85,	1.28)
Ireland	0.52	(.40,	.66)	0.51	(.40,	.65)
Italy	1.56	(1.30,	1.88)	1.78	(1.48,	2.16)
Portugal	0.24	(.17,	.33)	0.64	(.52,	.78)
Spain	0.35	(.24,	.50)	0.92	(.75,	1.12)

	Inactive in labour market					
	Model with acute health shock			Model with normalised latent health stock		
Belgium	1.43	(1.15,	1.77)	1.43	(1.15,	1.77)
Denmark	0.78	(.64,	.96)	0.78	(.64,	.96)
France	1.42	(1.20,	1.68)	1.41	(1.19,	1.67)
Greece	0.83	(.69,	.99)	0.82	(.69,	.98)
Ireland	0.56	(.46,	.69)	0.55	(.44,	.68)
Italy	1.24	1.06,	1.46)	1.21	(1.03,	1.42)
Portugal	0.48	(.40,	.57)	0.48	(.40,	.57)
Spain	0.96	(.81,	1.14)	0.95	(.80,	1.13)

4. Conclusions

The results show a consistency of effect with only small variation in magnitude both for these latent health stock variables as well as the subjective SAH measures. This is compatible with previous findings that declining health is associated with an increased probability of retirement (see Currie and Madrian, 1999; Deschryvere, 2004). Additionally a three-level graduated acute health shock variable was created by adapting the method of Riphahn (1999) and using the normalised latent health stock variable. The effects of an acute $\geq 0.5 - < 1.0$ standard deviation, $\geq 1.0 < 2.0$ standard deviation or a ≥ 2.0 standard deviation decrement showed evidence of a gradient, with increasing hazard of retiring with increasing magnitude of the acute health shock.

There are differences between models based on the definition of retirement. This is most noticeable in relation to financial factors. A higher mean personal income up to the time of retirement is associated (ceteris paribus) with an increased hazard of self-reporting retirement but a decreased hazard of being inactive in the labour market, including being on disability benefits. This is compatible with individuals with higher personal incomes choosing at an earlier age to retire officially rather than use a non-retirement or disability route. This could relate to better pension entitlement and the implicit tax on continuing to work (Gruber and Wise, 1998; Duval, 2003). Higher personal incomes will also more commonly be associated with sedentary occupations where an acute health shock may have less impact on the ability of the individual to perform their work tasks and permit the individual to focus on the single decision of permanent retirement.

References

Bound, J. (1990) "Self-Reported versus Objective Measures of Health in Retirement Models", *Journal of Human Resources,* 26, pp. 106–138.

Bound, J., Schoenbaum, M., Stinebrickner, T. R. and Waidmann, T. (1999) "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers", *Labour Economics,* 6, pp. 179–202.

Currie, J. and Madrian, B. (1999) "Health, Health Insurance and the Labor Market", in: Ashenfelter, O. and Card, D., *Handbook of Labor Economics*. Vol 3C. Amsterdam: Elsevier.

Deschryvere, M. (2004) "Health and Retirement Decisions: An Update of the Literature", Discussion Paper No. 932. *The Research Institute of the Finnish Economy*, Helsinki, Finland.

Disney, R., Emmerson, C. and Wakefield, M. (2006) "Ill Health and Retirement in Britain: A Panel Data-Based Analysis", *Journal of Health Economics,* 25, pp. 621–649.

Duval, R. (2003) "Retirement Behaviour in OECD Countries: Impact of Old-Age Pension Schemes and other Social Transfers", *OECD Economic Studies* No. 37, 2003/2.

Gruber, J. and Wise, D. (1998) "Social Security Programs and Retirement: An International Comparison", *American Economic Review,* 88, pp. 158–163.

Gruber, J. and Wise, D. (1999) *Social Security Programs and Retirement around the World*. Chicago: University of Chicago Press.

Hernandez-Quevedo, C., Jones, A. M., Lopez-Nicolas, A. and Rice, N. (2006) "Socioeconomic Inequalities in Health: A Comparative Longitudinal Analysis using the European Community Household Panel", *Social Science and Medicine,* 63, pp. 1246–1261.

Jenkins, S. P. (1995) "Easy Estimation Methods for Discrete-Time Duration Models", *Oxford Bulletin of Economics and Statistics,* 57, pp. 129–138.

Jones, A. M., Koolman, X. and Rice, N. (2006) "Health-Related Non-Response in the British Household Panel Survey and the European Community Household Panel: Using Inverse Proability Weighted Estimators in Non-Linear Models", *Journal of the Royal Statistical Society Series A (Statistics in Society),* 169, pp. 543–569.

Nicoletti, C. and Peracchi, F. (2005) "Survey Response and Survey Characteristics: Microlevel Evidence from the European Community Household Panel", *Journal of the Royal Statistical Society Series A (Statistics in Society),* 168, pp. 763–781.

Peracchi, F. (2002) "The European Community Household Panel: A Review", *Empirical Economics,* 27, pp. 63–90.

Prentice, R. and Gloekler, L. (1978) "Regression Analysis of grouped Survival Data with Application to Breast Cancer Data", *Biometrics,* 34, pp. 57–67.

Riphahn, R. T. (1999) "Income and Employment Effects of Health Shocks. A Test Case for the German Welfare State", *Journal of Population Economics*, 12, pp. 363–389.

Rice, N., Roberts, J. and Jones, A. M. (2006) "Health and Retirement in the UK: Duration Analysis of the BHPS", HEDG Working Paper #06/13, *University of York*.

Stern, S. (1989) "Measuring the Effect of Disability on Labor Force Participation", *Journal of Human Resources,* 24, pp. 361–395.

SUMMARY

In this paper we use discrete-time proportional hazard models with internationally comparable longitudinal data from the full eight waves of the European Community Household Panel (ECHP) to study the relationship between retirement, health levels and health shocks in nine European countries. The results show a consistency of effect with only small variation in magnitude both for latent health stock variables as well as the subjective SAH measures. This is compatible with previous findings that declining health is associated with an increased probability of retirement. Additionally a three-level graduated acute health shock variable was created, that shows evidence of a gradient in the effect.