Health shocks, employment and income in the Spanish labour market.

by

Pilar García Gómez[a] and Angel López Nicolás[ab]*

First version April 2005

This version June 2005

[a] Departament d'Economia i Empresa and CRES, Universitat Pompeu Fabra, 08005-Barcelona, Spain

[b] Departamento de Economía. Universidad Politécnica de Cartagena. 30203-Cartagena, Spain.

Abstract

This paper investigates the relationship between health shocks and labour outcomes in the Spanish population using the European Community Household Panel. In order to control for the non-experimental nature of the data we use matching techniques. Our results suggest that there is a significant effect running from health to the probability of employment and to labour income. Moreover, while we cannot investigate the influence of childhood events and other phenomena that trigger long run causal pathways from socio-economic status to health, we are able to find a significant reduction in the probability of reporting good health in individuals who transit out of employment in comparison with individuals who are otherwise identical in terms of reported health status at the time of the transition.

JEL classification: C23, I12, J60

Keywords: health shocks; unemployment; matching; Spain.

(*) Corresponding author: Departamento de Economía. Facultad de Ciencias de la Empresa. Universidad Politécnica de Cartagena. Paseo Alfonso XIII, 50. 30203 Cartagena (Spain). Fax:+34 968325781 Tel. +34 968 325484. E-mail: angel.lopez@upct.es This paper derives from the project "La dinámica del estado de salud y los factores socieconómicos a lo largo del ciclo vital. Implicaciones para las políticas públicas", which is supported by the Fundación BBVA. The views expressed in this paper are those of the authors and not necessarily those of the funders or the authors' employers.

1. Introduction

The literature in the fields of public health and health economics contains abundant evidence on the relationship between socio-economic conditions and health status (see e.g. Acheson report, 1998, Deaton 2002, 2003) and it is often argued that income redistribution should be used as a correcting mechanism (see Navarro 2001). It is nevertheless difficult to extract policy recommendations from the existing evidence. Socio-economic conditions are a set of factors that not only include income but also education, housing, social relations, social rank among a group etc. Therefore the simple idea of redistributing income from rich to poor will only be effective if health is effectively determined by income or something that, in turn, is determined by income. That is, health and income are related through mechanisms of direct and indirect causality. On one hand income is necessary to consume basic goods and services and, on the other hand, health affects the possibility of generating income in the labour market. Deaton (2003) shows how the gradient between income and health depends on i) a plausible direct causal effect, ii) the incidence of shocks to health, iii) the ability of the health care system in helping to overcome health shocks and iv) the degree of income insurance to health shocks. The incidence of shocks depends upon public health measures such as the promotion of healthy lifestyles and the reduction of accidents. Also, the ability of the health care system depends on the volume of resources devoted as well as how efficiently are these resources used from both the technical and the allocative perspectives. Finally, the risk of income loss after a health shock depends on the system of social protection. Each one of the last three elements is policy amenable and it is important to assess their weight in each population of interest before thinking about correcting mechanisms. More fundamentally, it is important to assess whether the direct pathway from socio-economic conditions to health is less or more important than the reverse pathway. Economists have provided insightful evidence on these issues. Exploiting the fact that American children's health does not have a substantial influence on parental income, Case, Lubotsky and Paxson (2002) argue that the correlation between children's health and income reveals a causal pathway from parental socio-economic status to children's health. Moreover, these authors argue that part of the gradient found in adulthood is originated in childhood, through a lower accumulation of education capital and a greater depreciation of health capital by children coming from low socio-economic status. If this was the only story, research should focus exclusively on pinpointing the precise pathways in which low parental socio-economic status affects children's health (e.g. is it a greater incidence of shocks or a poorer management or shocks or both?) in order to find policy prescriptions. However, there is evidence pointing at the existence of a pathway from health to socio-economic status too. Contoyannis and Rice (2001) have used IV methods for panel data to show that bad psychological health affects men wages negatively while good self assessed health affects female wages positively and, more recently, Smith (2004) finds that the onset of disease affects labour supply and the ability to generate income during adulthood in a non-trivial way.

This paper attempts to provide some evidence on the relative importance of the two causal pathways in the Spanish population. The study of socio-economic differences in health in the Spanish population has traditionally been associated to the realm of public health (Regidor et al. 1995, 1999, 2002), but the health economics community has contributed with recent studies (van Doorslaer and Koolman, 2004; García Gómez and López Nicolás, 2004a, 2004b) that show that the health of the Spanish population –be it self-assessed health or indicators such as the GHQ index for mental health or the Euroqol index- is relatively more concentrated in the top of the income distribution. That is, there is a significant degree of income related health inequality. These contributions from health economics also aim to decompose inequality into explanatory factors –thus advancing one step towards unveiling causal mechanisms- and reach the conclusion that, other things being equal, income is positively and significantly associated with health. Nevertheless it is difficult to identify the direction of causality, let alone the precise mechanisms of causation, behind this correlation due to the well-known limitations of cross sectional information. Our contribution is motivated, among others, by Smith's (2004) suggestion to unveil the relationship between socio-economic status and health by resorting to longitudinal information spanning several decades for representative samples of the US population. The steps taken consist in conditioning on past health shocks before evaluating current changes in labour status, income and medical outlays and vice-versa, that is, conditioning on past labour status changes and other events related to socio-economic status before evaluating current health shocks. Thus in this paper we resort to the best source of longitudinal information on health and socio-economic characteristics for the Spanish population available to researchers: the European Community Household Panel (1994-2001, hereafter ECHP). We will condition on past health (labour status) events to evaluate current changes in labour status (health). While the spirit is the same as in the study cited above, our specific methodology consists in matching individuals who experience a health shock (labour status change) with identical individuals in a control group. In this sense we follow the recent usage of the matching methods in the context of health shocks by Lechner and Vázquez Álvarez (2004), Frölich et al. (2004) and Dano (2004).

Our results suggest that there is a significant effect running from health to the probability of employment and to labour income. Moreover, while we cannot investigate the influence of childhood events and other phenomena that trigger long run causal pathways from socio-economic status to health, we are able to find a significant reduction in the probability of reporting good health in individuals who transit out of employment in comparison with individuals who are otherwise identical in terms of reported health status at the time of the transition.

In the next section we describe the methodology used to identify ways of causation between health changes to labour outcomes and viceversa. Section 4 discusses some features of the ECHP particularly relevant for this study. Section 5 presents the empirical results and section 6 concludes.

3. Methods

3.1 Outcomes of interest

In this paper we investigate, firstly, labour market outcomes potentially affected by an adverse health shock. We are particularly interested in labour market transitions that are not led by the availability of retirement –Disney et al. (2003) have recently used IV panel methods to show that adverse health shocks affect positively the probability of retirement- and therefore we restrict our population of interest to individuals below 60 years of age. The reason for this focus is that we are interested in transitions that might drive the observed gradient between income and health and, in the case of transitions to retirement, income loss is cushioned by pensions. In fact, the results by Smith for the US cited above show that individuals transiting to retirement after a health shock do not suffer income losses whereas the rest of individuals who transit out of employment suffer income losses which can cumulate up to large amounts in the course of a few years. Therefore our outcomes of interest are the probability of being in employment and the probability of being inactive or in other states, and the levels of income from labour and other sources. Secondly, we shall also investigate the effect of changes in employment status, in particular entering unemployment, on health in an attempt to capture potential causality from labour outcomes to health.

3.2 Estimating Average Treatment Effects on the Treated

As in any other evaluation exercise with non-experimental data, the problem in our setting consists in obtaining a credible counterfactual against which we may measure the impact of the health shock. Let T=1,0 indicate treatment and lack of treatment respectively and let Y_{i1} and Y_{i0} denote the outcome of interest for individual i with treatment and without treatment respectively. Since we will observe individual i either with treatment or without treatment, we cannot observe the distribution of the treatment effect $B_i=Y_{i1}-Y_{i0}$. Some features of such distribution are estimable, nevertheless. In particular, we may consider the Average Treatment Effect on the Treated

$$ATET= E(B\,|\,T=1)=E(Y_1-Y_o\,|\,T=1) \tag{1}$$

This magnitude measures how much the outcome of interest changes on average for those individuals who undergo the treatment (who suffer the health shock to be defined below . Clearly, simply computing the difference in the average outcomes of those in treatment and those out of treatment is open to bias. That is,

$$\begin{gathered}E(Y_1|T=1)-E(Y_0|T=0)= \\ E(Y_1|T=1)-E(Y_0|T=1)+E(Y_0|T=1)-E(Y_0|T=0)= \\ E(Y_1-Y_o|T=1)+E(Y_0|T=1)-E(Y_0|T=0)= \\ ATET+BIAS\end{gathered} \tag{2}$$

Only if we can guarantee that the outcomes of the control group are equal on average to what the outcomes of the treatment group would have been in the absence of treatment does this consistently estimate the ATET. With non-random sorting into treatment and control such condition is rarely met.

Now suppose that by conditioning on an appropriate set of observables, X, assignment to the treatment group becomes random (or, at least, independent of the outcomes). This is the conditional independence assumption (see Heckman et al. 1997 or Wooldridge 2002)

$$Y_o \perp T \mid X$$

(3)

This implies that

$$E(Y_0|T=1, X)-E(Y_0|T=0, X)=0$$

(4)

Therefore we could estimate the ATET from the difference in outcomes between treated and controls within each cell defined by the conditioning variables X (see Blundell and Costa Dias 2002). Using the law of iterated expectations and the conditional independence assumption, the ATET can be retrieved from observed data in the following way

$$ATET=E(Y_1 \mid T=1)- E(Y_0 \mid T=1)=E_X[(E(Y_1 \mid X, T=1)- E(Y_0 \mid X, T=1)) \mid T=1]=$$
$$E_X[(E(Y_1 \mid X, T=1)- E(Y_0 \mid X, T=0)) \mid T=1]$$

(5)

This turns out to be prohibitive in terms of data, as the size of many cells will be small. An alternative is to use the results of Rosenbaum and Rubin (1983, 1984) and condition on the probability of treatment as a function of X, P(X) since the conditional independence assumption also implies that

$$E(Y_0|T=1, P(X))-E(Y_0|T=0, P(X))=0$$

(6)

Therefore we could estimate the ATET from the differences in outcomes between treated and controls within each cell defined by values of P(X).

$$ATET=E(Y_1 \mid T=1)- E(Y_0 \mid T=1)=E_{P(X)}[(E(Y_1 \mid P(X), T=1)- E(Y_0 \mid P(X), T=1)) \mid T=1]=$$
$$E_{P(X)}[(E(Y_1 \mid P(X), T=1)- E(Y_0 \mid P(X), T=0)) \mid T=1]$$

(7)

In practical terms, this requires matching treated individuals with controls on criteria based on the closeness of their P(X) score –the propensity score- as we shall see later.

The ability of this estimator to retrieve consistently the ATET relies crucially on the adequacy of the conditional independence assumption. That is, that all factors that may affect treatment and the outcomes are included in the vector of conditioning variables. Panel data –spanning periods before and after the treatment- afford the possibility to correct for the hypothetical failure of this assumption. In essence, the idea relies on the ability to first difference the outcomes of the treated and the controls in order to eliminate any unobservable fixed effects affecting selection and the outcomes. Letting the superscript A and B denote the time periods before and after treatment occurs, the conditional independence assumption would now be stated in the following terms

$$Y^A_0 - Y^B_0 \perp T \mid X \tag{8}$$

So that,

$$E(Y^A_0 - Y^B_0 \mid T=1, X) - E(Y^A_0 - Y^B_0 \mid T=0, X) = 0 \tag{9}$$

And therefore, the "differences in differences" ATET can be estimated in the following way from observed data

$$\begin{gathered} ATET_{DID} = E(Y^A_1 - Y^B_1 \mid T=1) - E(Y^A_0 - Y^B_0 \mid T=1) = \\ E_X[(Y^A_1 - Y^B_1 \mid X, T=1) - E(Y^A_0 - Y^B_0 \mid X, T=1)) \mid T=1] = \\ E_X[(E(Y^A_1 - Y^B_1 \mid X, T=1) - E(Y^A_0 - Y^B_0 \mid X, T=0)) \mid T=1] \end{gathered} \tag{10}$$

The same reasoning about the propensity score applies to the $ATET_{DID}$ estimator.

As we shall see, an alternative way to use the longitudinal perspective offered by panel data consists in using the standard ATET estimator of expression (7) including pre-treatment outcomes within the vector of conditioning variables, either by including them directly in the propensity score function or by restricting the sample of controls to individuals who are identical in terms of pre treatment outcomes.

3.3 Identifying a health shock and constructing treatment and control groups

Our measure of health shocks is based on the responses to the question on self-assessed health in the ECHP "How good is your health in general?". From the five possible responses (very good, good, fair, bad and very bad), we consider that the respondent has undergone an adverse health shock if he or she reports "fair", "bad" or "very bad" in any given period, with the timing of the shock occurring sometime between the last period when he or she recorded any of the other three alternatives.

Since we wish to evaluate whether suffering a health shock in these terms leads to any change in labour outcomes, we want to rule out the possibility that any potential anticipation of the change in labour status causes the change in self reported health, therefore we adopt the following strategy –motivated by the procedures used by Lechner and Vázquez Alvarez (2004)- in order to construct the treatment and control groups:

1) Consider a window of three years for each observed individual. This creates 6 possible sequences of three years over the time span covered by our data. To these three years, regardless of the sequence, we refer as t=1, t=2 and t=3
2) For each sequence select individuals who are healthy (SAH good or very good) at t=1, the start of the sequence, and also are employed at t=1 and t=2
3) The **treatment group** are individuals meeting selection criterion # 2 who report fair, bad or very bad health in t=2 and t=3. That is, those individuals who undergo a health shock after t=1 and for whom adverse health persists at least over t=3. The sequence of health states for these individuals is therefore GBB (Good, Bad, Bad)
4) The **control group** are individuals meeting selection criterion # 2 for whom we observe a GGG sequence of health states (Good, Good, Good)

By analogy, in order to investigate the effects of changes in employment status on health, we consider the following selection criteria when constructing treatment and control groups

5) Consider a window of three years for each observed individual. This creates 6 possible sequences of three years over the time span covered by our data. To these three years, regardless of the sequence, we refer as t=1, t=2 and t=3
6) For each sequence select individuals who are employed at t=1, the start of the sequence, and also are in good or very good health at t=1 and t=2
7) The **treatment group** are individuals meeting selection criterion # 2 who report being unemployed in t=2 and t=3. The sequence of employment states for these individuals is therefore EUU (Employed, Unemployed, Unemployed)

8) The **control group** are individuals meeting selection criterion # 2 for whom we observe a EEE sequence of employment states (Employment, Employment, Employment)

We shall match individuals in the treatment and control groups on the basis of the propensity score. Thus we do not resort to first differences, but from the discussion in 3.2 it follows that we nevertheless exploit the longitudinal perspective of our data by conditioning on the labour status at times t=1 and t=2 –in the estimation of the ATET of a health shock on labour outcomes- and on self-assessed health at times t=1 and t=2 –in the estimation of the ATET of a transit into unemployment on health. A similar strategy has also been adopted by Dano (2004) when evaluating the effects of road accidents on labour outcomes.

4. Data

Table 1 shows that, in the ECHP 1994-2001, we can observe a total of 34830 GGG sequences, that is sequences of three consecutive years of good health, and 3080 GBB sequences, that is sequences of an adverse health shock lasting at least two years. Similarly we observe 2181 Employment, Unemployment, Unemployment sequences and 22724 sequences of three years in employment.

When we apply the age selection criterion, the number of sequences is reduced substantially. When we condition on being employed (or being in good health) during the first two years of the sequences, sample sizes further reduced according to the figures in the table.

Table1. Sample sizes in treatment and control groups

	Treated (GBB)	Control (GGG)	Treated (EUU)	Control (EEE)
Initial sample	3080	34830	2181	22724
Individuals aged <=60	1557	31543	1530	21831
Without missing values in the propensity score	1528	31000	1495	21467
Conditional to have been working in periods 1 and 2	691 (45% of previous sample of treated)	15015 (48% of previous sample of control)		
Conditional to have been in Good or excellent health in periods 1 and 2			876 (59% of previous sample of treated)	15686 (73% of previous sample of control)

Note: GBB: Good/excellent health- Fair/Bad/Very bad health - Fair/Bad/Very bad health

GGG: Good/excellent health - Good/excellent health - Good/excellent health
EUU: Employed – Unemployed – Unemployed
EEE: Employed – Employed – Employed

5. Empirical results

We estimate ATET effects by means of the Stata procedures written by Becker and Ichino (2002). First we estimate the probability (probit specification) of being in the treatment group (the propensity score) as flexible functions of age and gender, educational attainment, the logarithm of equivalent household income at the start of the sequence, regional indicators and wave indicators. These specifications pass the "balancing hypothesis". That is, there are no systematic differences in observable characteristics between treated and controls once we condition on the propensity score. Subsequently we match treated individuals with controls using three alternative methods: i) nearest neighbour matching, ii) radius matching and iii) kernel matching (see Becker and Ichino for the relevant technical details). There are no a priori grounds to expect any of these matching methods to be preferable, so reporting the three estimates allows us to assess the robustness of the results.

5.1 Effects of a health shock on labour outcomes

i) Effects on the probability of employment

Table 2 presents the estimates of the ATET on the probability of employment. The three estimates are remarkably close at nearly minus 5%. These effects are statistically significant at conventional levels (we do not report standard errors for the kernel matching estimates in this version since they are currently being bootstrapped).

Table 2 ATET on the probability of being employed versus any other labour status outcome (student, military service, unemployed, retired, housework or inactive).

	#Treated	#Control	ATET	Std	t
Neighbour Matching	691	656	-0.046	0.015	-3.127
Radius Matching (r=0.1)	691	14854	-0.047	0.012	-4.078
Kernel Matching	691	14854	-0.048	.	.

So, if those affected by an adverse health shock are less likely to be employed, where do they end up?

Table 3 ATET on the probability of several activity status

		Number of treated	Number of control	Coefficient	Std Error	t
Student	N. Neighb.	691	656	0,003	0,002	1,415
	Radius	691	14854	0,001	0,002	0,31
	Kernel	691	14854	0,001	.	.
Militar	N. Neighb.	691	656	0,001	0,001	1
	Radius	691	14854	0	0,001	0,245
	Kernel	691	14854	0	.	.
Unemployed	N. Neighb.	691	656	0,001	0,01	0,146
	Radius	691	14854	-0,005	0,007	-0,727
	Kernel	691	14854	-0,004	.	.
Retired	N. Neighb.	691	656	0,003	0,003	0,981
	Radius	691	14854	0,003	0,003	1,35
	Kernel	691	14854	0,003	.	.
Housework	N. Neighb.	691	656	-0,003	0,006	-0,454
	Radius	691	14854	0,002	0,004	0,589
	Kernel	691	14854	0,002	.	.
Inactive	N. Neighb.	691	656	0,041	0,009	4,588
	Radius	691	14854	0,046	0,008	5,625
	Kernel	691	14854	0,045	.	.

Note: These are ATETs on the probability of being in each of the status versus all other possible status.

Table 3 reveals that the exit from employment as a result of a health shock leads to inactivity rather than unemployment. In fact, the ATET effect on the probability of being inactive ranges between 4% and 5% and, according to the nearest neighbour and radius estimates, it is statistically significant. In contrast, none of the ATET effects for the rest of activity status is statistically significant. An interesting question for further research is how long does this period of inactivity last and whether it is a waiting room for retirement.

ii) Effects on income

Incomes may suffer after a health shock even without transiting out of employment. One immediate way is through productivity losses -which in some cases may be absorbed by the employer-, or the inability to work extra time. When the health shock leads to transitory inability to work (incapacidad laboral transitoria), the worker's labour income is reduced approximately either 25% or 40% depending on the length of the inability period. Transitions out of employment into permanent inability to work will lead generally to reductions to labour income of up to 45%. Some health shocks may lead to unemployment

–e.g. contracts expiring after the health shock may not be renewed-, in which case the worker will qualify for unemployment benefits. The duration of unemployment benefits ("prestaciones de desempleo") is currently set at around 9 days per month of work after having worked a minimum of 12 months, and it lasts for a maximum of 24 months. Unemployment benefits are linked to the previous wage within a lower and an upper bound. In 2005, the size of the benefit varies between a minimum of 375, 84 €/month (gross) to a maximum of 822,3 €/month gross for a worker without dependents or 1057 €/month (gross) for a worker with two or more dependent children. After this –or in cases who are not eligible for unemployment benefit-, unemployed workers may be eligible for unemployment subsidy ("subsidio de desempleo") depending on their age and their family responsibilities (e.g. no person below 45 without dependents can receive this subsidy unless special cases such as ex-convicts, returned migrants etc.). This subsidy lasts up to 30 months for workers above 45 years of age and it may be extended until the worker qualifies for pension receipt if he/she is above 52. For 1995, the subsidy ranges between 375,84 €/month (gross) for workers without dependents to 624 375,84 €/month (gross) for workers with two or more dependent children. To give an idea of the degree of income insurance afforded by the unemployment legislation, we may contrast these figures with the latest statistic on wages (gross of income tax and worker contributions to the social security) in the Spanish economy, whose mean for the last quarter of 2004 is 1641,14 €/month.

For these reasons we may expect reductions in personal labour income and increases in personal social security transfers following a health shock. It is possible that other components of households income are affected by the health shocks, this would be the case when another worker in the households must adjust his/her labour supply to provide care. In this paper we calculate the effect of a health shock on all sources of household and personal income reported in the ECHP. The treatment and control groups are defined in the same terms as in the previous estimation of the ATET on the probability of employment. That is, we match individuals with a GBB sequence of health states with individuals with a GGG sequence on the basis of the propensity score and the extra condition that they are employed at t=1 and t=2.

The following tables present the ATET of a health shock effects on different sources of income. It should be noted that in the ECHP income data refers to the year prior to the date of the survey, so we report the effects on income at t=2, that is the year in which the respondent reports a health shock, and at t=3. Note that the ATET estimates suggest a significant reduction in total household income at both t=2 and t=3. These are driven by the reductions in labour income, which are not compensated by the parallel increases in social security transfers. For instance, in the year when the shock is reported, the reduction in personal labour income is estimated at around 2154 € (all money figures expressed in constant terms at 2001 prices) –this is the average of the three different estimates. In contrast, the increase in social security transfers is estimated at around 544 €. In the second year after having suffered the shock, the estimates suggest a reduction of 2322 € in personal labour income and an increase of 600 € in personal social security transfers. It would be interesting to observe whether the reduction in income continues beyond the second year after the shock, but, as suggested by the smaller sample sizes for treatment and control groups in the bottom panel of the table, extending the time span prohibitively reduces the number of valid observations. In any case the results lend support to the idea that there is a direct causal effect from health to income and that this has to do with the fact that the system of social security provisions does not fully insure labour income against illness.

Table 4. ATET on different income measures on t=2 (year coinciding with the health shock.

Income measure	Method	# treated	# controls	ATET	S. Error	t
Household total Income	N. Neighb.	691	658	-3041	926	-3,29
	Radius	.	.	.	.	.
	Kernel	691	14854	-2314	.	.
Household total Labour income	N. Neighb.	691	658	-2134	757	-2,82
	Radius	.	.	.	.	.
	Kernel	691	14854	-3155	.	
Household total private transfers	N. Neighb.	691	658	-11	33	-0,33
	Radius	.	.	.	.	.
	Kernel	691	14854	-12	.	
Household total social security transfers	N. Neighb.	691	660	408	249	1,64
	Radius	691	14854	454	181	2,51
	Kernel	691	14854	448	.	.
Personal total Income	N. Neighb.	691	660	-1310	625	-2,10
	Radius	691	14854	-1214	424	-2,86
	Kernel	691	14854	-1226	.	
Personal total Labour income	N. Neighb.	691	660	-1689	501	-3,37
	Radius	691	14854	-2434	325	-7,48
	Kernel	691	14854	-2338	.	.
Personal total private transfers	N. Neighb.	691	660	-32	30	-1,05
	Radius	691	14854	-15	8	-1,95

	Kernel	691	14854	-14	.	.
Personal total social security transfers	N. Neighb.	691	660	583	99	5,92
	Radius	691	14854	525	98	5,38
	Kernel	691	14854	524	.	.

Note: Money figures in 2001€
Standard errors for the kernel matching estimates are not available for this version of the paper since they require bootstrapping. In some instances we were not able to obtain ATET estimates with the radius method.

Table 5. ATET on different income measures on t=3 (one year after the health shock)

Income measure	Method	# treated	# controls	ATET	S. Err.	t
Household total Income	N. Neighb.	504	476	-3798	1094	-3,47
	Radius	504	10777	-2735	613	-4,46
	Kernel	504	10777	-2650	.	.
Household total Labour income	N. Neighb.	504	476	-2843	889	-3,20
	Radius	504	10777	-3744	571	-6,55
	Kernel	504	10777	-3558	.	.
Household total private transfers	N. Neighb.	504	476	74	61	1,21
	Radius	504	10777	57	54	1,05
	Kernel	504	10777	58	.	.
Household total social security transfers	N. Neighb.	504	476	365	320	1,14
	Radius	504	10777	727	232	3,14
	Kernel	504	10777	715	.	.
Personal total Income	N. Neighb.	504	476	-1527	679	-2,25
	Radius	504	10777	-1725	418	-4,12
	Kernel	504	10777	-1725	.	.
Personal total Labour income	N. Neighb.	504	476	-1923	631	-3,05
	Radius	504	10777	-2560	398	-6,43
	Kernel	504	10777	-2482	.	.
Personal total private transfers	N. Neighb.	504	476	9	19	0,48
	Radius	504	10777	-7	14	-0,49
	Kernel	504	10777	-6	.	.
Personal total social security	N. Neighb.	504	476	585	158	3,70

transfers						
	Radius	504	10777	613	143	4,28
	Kernel	504	10777	601	.	.

Note: Money figures in 2001€
Standard errors for the kernel matching estimates are not available for this version of the paper since they require bootstrapping.

5.3 Effects of labour transitions on health

Without aiming to establish the precise causal pathway from transitions into unemployment to health –which could range from psychological effects to economic hardship induced disease-, we now turn to the estimation of the ATET of a transition into unemployment on the probability of reporting good health. In this version we shall only report the estimate that uses the nearest neighbor matching algorithm for our sample of 867 individuals in the treatment group. This estimate is –0.031 with a standard error of 0.016 (t-value=-1.905). This preliminary evidence thus suggests that individuals with good health who experience a transition into unemployment are 3% more likely to suffer a health shock after being unemployed for some time.

6. Discussion and conclusion

In this paper we have provided evidence for the Spanish population suggesting that adverse health shocks have an important causal effect on the probability of being in employment. While this is relatively unsurprising, these shocks also seem to lead to non-negligible reductions in income. This lends support to the idea that, as shown by Smith for the USA population, socio-economic changes occurring in adulthood are able to affect the relationship between income and health. Therefore social security policies aimed at insuring income against adverse health event may have some room for weakening the relationship between income and health documented for the Spanish population. Of course, this does not diminish the potential importance of events occurring early in life and which may be caused by the socio-economic conditions of parents, neither the potential importance of channels of causation from socio-economic status to health later on in life. In fact, even within a short time span of three years we are able to find that a transition into unemployment increases significantly the chances of reporting bad health thereafter.

While these results are not directly translatable into policy, they are able to suggest further avenues of research. For instance, which are the precise ways in which the health shock leads to income reductions (on the job productivity losses borne by the worker, less extra time, exit and entry into sick leave etc..). Another important question is how different institutional arrangements in social and worker protection might influence the regional differences in income related health inequalities reported elsewhere.

References

Acheson Donald (1998). Independent Inquiry into Inequalities in Health.

Becker Sasha and Ichino Andrea (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal* 2(4): 358-377.

Blundell Richard and Costa Dias Monica (2002). Alternative Approaches to Evaluation in Emprical Microeconomics. *Portuguese Economic Journal* 1: 91-115.

Case Anne, Lubotsky Darren and Paxson Chris (2002). Economic Status and Health in Childhood: the Origins of the Gradient. *American Economic Review* 92(5): 1308-1334.

Contoyannis, P. and N. Rice (2001). The impact of health on wages: Evidence from the British Household Panel Survey. *Empirical Economics* 26(4): 599-622.

Dano (2004). Road injuries and long-run effects on income and employment. Paper presented at the XXIII European Workshop on Econometrics and Health Economics. Venafro, Italy. September 2004.

Deaton Angus (2002). Policy Implications of the Gradient of Health and Wealth. *Health Affairs* 21(2): 13-30.

Deaton Angus (2003). Health, inequality and economic development. *Journal of Economic Literature* 41: 113-158.

Disney, R.., C. Emmerson and M. Wakefield (2003). Ill health and retirement: A panel data based analysis. WP 03/02. The Institute for Fiscal Studies. London.

Frölich Markus, Heshmati Almas and Lechner Michael (2004). A Microeconometric Evaluation of Rehabilitation of Long-Term Sickness in Sweden. *Journal of Applied Econometrics* 19: 375-396.

García Gómez Pilar and López Nicolás Ángel (2004a). Regional Differences in Socio-Economic Health Inequalities in Spain. University Pompeu Fabra. Department of Economics and Business. Working paper #757.

García Gómez Pilar and López Nicolás Ángel (2004b) Socio-Economic Inequalities in Health in Catalonia Autores: University Pompeu Fabra. Department of Economics and Business. Working paper #758.

Heckman J, Ichimura H and Todd P (1997). Matching as an Econometric Evaluation Estimator. *Review of Economic Studies* 64: 605-654.

Lechner Michael and Vázquez Álvarez Rosalia (2004). The effect of disability on labour outcomes in Germany: evidence from matching. Centre for Economic Policy Research. London. Discussion Paper #4223.

Navarro Vicente (2001). The Political Economy of Social Inequalities: Consequences for Health and Quality of Life. New York. Baywood

Regidor E, Gutiérrez-Fisac J L, Rodríguez C, de Mateo S and Alonso I (1995). Las desigualdades sociales y la salud en España. La salud y el sistema sanitario en España. Informe SESPAS 1995: 19-43.

Regidor E, Gregorio B, de la Fuente L, Domingo A, Rodriguez C and Alonso J (1999). Association between educational level and health related quality of life in Spanish adults. Journal of Epidemiology community Health 53: 75-82

Regidor E, Gutiérrez-Fisac JL, Domínguez V, Calle M E and Navarro P (2002). Comparing social inequalities in health in Spain: 1987 and 1995/98. Social Science and Medicine 54: 1323-1332.

Rosenbaum P and Rubin DB (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70: 41-55.

Rosenbaum P and Rubin DB (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association* 79: 516-524

Smith James P (2004). Unravealing the SES Health Connection. The Institute for Fiscal Studies. Working Paper WP04/02.

Van Doorslaer Eddy and Koolman Xander (2004). Explaining the differences in income-related health inequalities across European countries. *Health Economics* 13(7): 609-628.

Wooldridge Jeffrey (2002). Simple Solutions to the Initial Conditions Problem in Dynamic, Non Linear Panel Data Models with Unobserved Heterogeneity. The Institute for Fiscal Studies. Deparment of Economics, UCL. Working Paper CWP18/02.

Table A1. Descriptive statistics of control and treatment groups

	GGG		GBB		EEE		EUU	
	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error	Mean	Standard Error
Age	38.12	9.74	44.57	9.47	38.66	9.68	35	11.21
Male	0.67	0.47	0.66	0.47	0.68	0.47	0.42	0.49
Married	0.68	0.47	0.78	0.41	0.69	0.46	0.59	0.49
Never married	0.28	0.45	0.16	0.37	0.27	0.44	0.37	0.48
Separated/Divorced	0.027	0.16	0.039	0.19	0.028	0.16	0.029	0.17
Widowed	0.010	0.099	0.017	0.13	0.01	0.10	0.0091	0.095
Children aged less than 16	0.73	0.88	0.69	0.91	0.74	0.89	0.73	0.91
Household size	3.65	1.32	3.80	1.39	3.66	1.33	3.93	1.420379
Total household equivalent income (Adjusted OECD Scale)	11931.44	7174.23	10413.41	6892.27	11973.38	7223.01	8005.71	6844.46
Total income Household	25411.76	16094.49	22983.84	14868.17	25543.76	16215.87	18330.09	18098.36
Total income from labour in the Household	18684.15	14587.45	15277.75	12655.45	18724.66	14640.33	10563.47	10844.47
Private transfers in the household	52.51	829.49	39.35	363.42	52.64	831.65	130.7	852.85
Total social transfers in the Household	2314.66	5791.25	2755.38	4579.04	2259.29	5686.29	4913.74	12500.92
Total income Individual	13819.02	10316.83	12716.98	10916.33	14131.65	10470.59	4650.41	14105.51
Total income from labour in the Individual	11258.43	9540.38	8896.76	8293.44	11403.84	9543.19	2058.41	3874.83
Private transfers in the Individual	27.17	597.09	12.17	158.7	26.51	599.76	52.24	492.99
Total social transfers in the Individual	162.01	2327.19	688.62	2517.23	141.36	2241.2	2252.97	11703.41
Less than secondary school	0.45	0.50	0.66	0.47	0.45	0.50	0.63	0.48
Secondary school	0.22	0.41	0.14	0.34	0.22	0.41	0.22	0.41
Third level education	0.33	0.47	0.20	0.40	0.33	0.47	0.15	0.36
Work	0.94	0.23	0.90	0.30				
Employ	0.76	0.43	0.65	0.48				
Self-employ	0.18	0.39	0.24	0.43				
Student	0.0023	0.048	0.0029	0.054				
Militar	0.0014	0.037	0.0014	0.038				
Unemployed	0.039	0.19	0.033	0.18				
Retired	0.0009	0.030	0.0043	0.066				
Housework	0.0090	0.094	0.012	0.11				
Inactive	0.0020	0.045	0.048	0.21				
Sah very good					0.24	0.43	0.24	0.43
Sah good					0.67	0.47	0.64	0.48
Sah fair					0.08	0.28	0.11	0.31
Sah poor					0.0067	0.081	0.015	0.12

Sah very poor	0.00032	0.018	0.0023	0.048

Note: GBB: Good/excellent health- Fair/Bad/Very bad health - Fair/Bad/Very bad health
GGG: Good/excellent health - Good/excellent health - Good/excellent health
EUU: Employed – Unemployed – Unemployed
EEE: Employed – Employed – Employed
All the monetary outcomes are in 2001€

Table A2. Estimates of the Propensity Score in Health (GBB versus GGG)

	Coeficient	Standard Error
Age	0.0404605	0.0224025
Age^2	-0.0001567	0.0002754
Male	-0.1331892	0.0671929
Secondary school	-0.2341249	0.0735729
Third level education	-0.3153846	0.0649699
Noroeste	0.0263392	0.1050098
Noreste	-0.3560797	0.1076573
Centro	-0.3742696	0.1045942
Este	-0.3795403	0.0976885
Sur	-0.3064672	0.1063649
Canarias	-0.2509086	0.1323385
Log household equivalent income (Modified OECD scale)	-0.0630104	0.0327256
Children aged less than 16 in the household	0.025881	0.0299213
Married	-0.004039	0.1245213
Never married	0.1453231	0.1470275
Widowed	-0.1633508	0.246201
Wave 4	0.0793927	0.0745428
Wave 5	0.0872208	0.0728775
Wave 6	0.0104539	0.0802536
Wave 7	0.0287288	0.0792103
Wave 8	0.1649996	0.1050717
Constant	-1.765172	0.6962369
Number of observations	15706	
Log-likelihood	-2760.8366	

Note: GBB: Good/excellent health- Fair/Bad/Very bad health - Fair/Bad/Very bad health
GGG: Good/excellent health - Good/excellent health - Good/excellent health
All the variables are evaluated at t=1.
Wave variables are referred to the year t=3. By construction, the two first waves are not directly included in the analysis, although their information is used for the lag values.

Table A3. Estimates of the Propensity Score in Employment (EEE versus EUU)

	Coefficient	Standard Error
Woman aged 16-19	0.656191	0.2604122
Woman aged 20-24	0.5265783	0.2262388
Woman aged 25-29	0.4653507	0.2036933
Woman aged 30-34	0.1258579	0.2048409
Woman aged 35-39	0.0335524	0.2101807
Woman aged 40-44	0.0827569	0.2069122
Woman aged 45 or more	0.0662092	0.2032388
Man aged 16-19	0.5660509	0.1649883
Man aged 20-24	0.1267446	0.1301004
Man aged 25-29	-0.2174067	0.1177734
Man aged 30-34	-0.2317837	0.1203927
Man aged 35-39	-0.4692337	0.0881645
Less than secondary school	0.4507885	0.0668708
Secondary school	0.3480276	0.0741121
Sur	0.1939044	0.0728665
Canarias	0.142289	0.0963944
Este	-0.115545	0.0718511
Comunidad de Madrid	-0.0923488	0.0869824
Noreste	-0.0720323	0.0792742
Noroeste	-0.0088516	0.0914846
Log (equivalent household income)* Woman	-0.2309412	0.0331956
Log (equivalent household income)*Man	-0.1134973	0.0265706
Household size	0.0000508	0.0161712
A woman that never married	1.384037	0.4180948
A married woman	1.804111	0.4108958
A woman separated / divorced	1.547118	0.4368643
Widowed	0.6426849	0.2644814
Married	-0.1699486	0.0969535
Man that worked full time in period 1	-0.6478095	0.1188452
Woman that worked full time in period 1	-0.5246831	0.071709
Age when start working	0.0038265	0.0045466
Wave 4	-0.0757816	0.0691541
Wave 5	-0.2891586	0.0793017
Wave 6	-0.3167993	0.0743257
Wave 7	-0.2365832	0.0752364
Wave 8	-0.2825054	0.0849349
Constant	0.4620678	0.4224887
Number of observations	16368	
Log-likelihood	-2923,6516	

Note: EUU: Employed – Unemployed – Unemployed ; EEE: Employed – Employed – Employed
All the variables are evaluated at t=1.
Wave variables are referred to the year t=3. By construction, the two first waves are not directly included in the analysis, although their information is used for the lag values
Equivalent household income is calculated using the Modified-OECD scale.