# Is social capital good for your health?

# A European perspective

Lorenzo Rocco*, Marc Suhrcke**

This version December, 2009

Word count: 7412

**Abstract**

**Aims**: Examine the causal impact of social capital on health in 14 European countries

**Methods and data**: Using data from the European Social Survey for 14 European countries, supplemented by regional level data coming from the EUROSTAT REGIO database, we study whether individual and/or community level social capital positively affect health. Both health and social capital measures are affected by error-in-variables and the causal relationship between social capital and health is likely to be circular. We address these empirical issues by instrumenting social capital. The large variance of the error term due to measurement errors calls for strong instruments in order to get reliable estimates in finite sample.

**Results**: Our dataset is rich enough in information to allow us to find a seemingly strong causal relationship going from social capital to individual health. We also show that community social capital (defined at regional level) does not affect health when individual social capital is controlled for.

**Conclusions**: Policy interventions should be targeted at improving primarily individual social capital, because in so doing they would achieve a double effect: on the one hand they would directly improve individual health; on the other hand they would contribute to community social capital that reinforces the beneficial role of individual social capital.

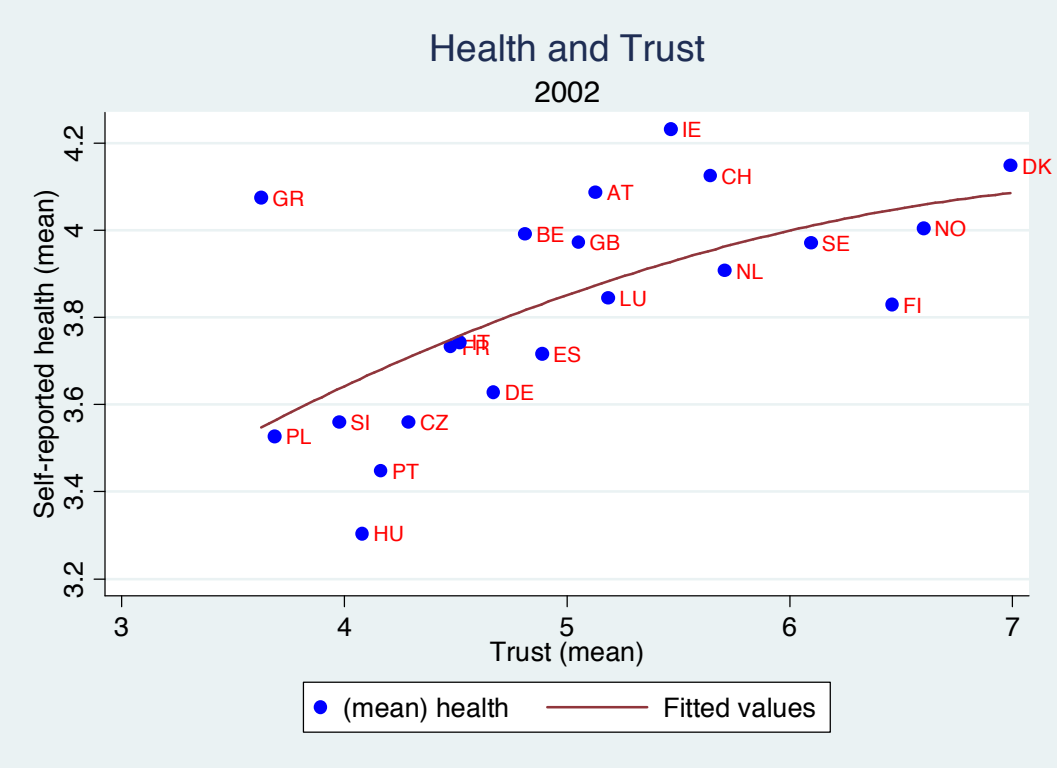
---

* University of Padova – lorenzo.rocco@unipd.it

** University of East Anglia, Norwich – m.suhrcke@uea.ac.uk

## *Introduction*

Health Economists and Public Health scientists have expended increasing effort in recent years to document what appears to be a rather close link between social capital and health (see Islam et al. 2006 for an excellent survey, and the recent contributions of Scheffler et al. 2007; Folland, 2007; D'Hombres et al. 2007). Figure 1, using data from 21 European countries included in the European Social Survey, confirms this close relationship. Here social capital is measured for each country as a weighted mean of individuals' self-reported rate of trust[1].




The positive relationship is evident and the correlation rather strong (0.51). The crucial question is whether this relationship reflects a causal impact from social capital to health, the reverse causal impact, or the influence of third factors simultaneously affecting social capital and health. A further, more specific question relates to the relevant dimension of social capital that may be responsible for this link: is it the social capital at the individual level or at community level that matters?

As for the causality question but in a different country context, Folland (2007) and D'Hombres et al. (2007) cannot reject the hypothesis of a causal impact of social capital on health after a careful analysis. Existing results are far more mixed regarding the relative importance of individual vs. community level social capital: for instance, Poortinga et al. (2006) found no independent impact of

[1] The exact wording of the survey question is: "generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?"

community social capital when individual social capital is simultaneously included in the equation to estimate. By contrast, Iversen (2007) suggests that both levels matter for individual health.

The present paper attempts to answer both questions for a set of 14 European countries, using data from the European Social Survey in 2002. Anticipating the results, first, we corroborate the hypothesis of a positive causal impact running from social capital to health and, second, we suggest that community social capital plays no autonomous effect besides individual social capital. To the best of our knowledge the present paper for the first time takes into account the empirical challenges originating from the fact that both health and social capital are measured imperfectly.

Several definitions of social capital have been proposed in the sociological literature. Bourdieu (1986) defines social capital as a capital of social connections, mutual acquaintance and social recognition. Coleman (1988) refers to social capital as all those features of the social structure which might facilitate actions of individuals within the social structure itself (for instance parental care is a social norm which facilitates children subsequent activity and success in the society; social relationships *per se* are a form of social capital as they establish obligations, expectations, trustworthiness). Putnam (1993) defines social capital as those "features of social organization, such as trust, norms, and networks that can improve the efficiency of society by facilitating coordinated actions" (Putnam 1993:167). The definition of social capital adopted by the World Bank is very close to Putnam's. Although all definitions refer directly or indirectly to social connections or social network, as elements of social capital, Putnam's definition points to the role of social capital as a catalyst of coordination/cooperation, an essential device to achieve better outcomes, either social or economic. Coordination allows to reduce transaction costs, to overcome difficulties due to incomplete or asymmetric information, to establish efficient transactions in presence of incomplete contracts (Alesina and La Ferrara, 2002). In this perspective, it seems that emphasis on trust as an indicator of social capital is well placed, as trust favours (even one-shot) cooperation, without the need of creating long standing personalized relationships and processes of reputation building. Moreover, logically, trust is a determinant of social connections, as a minimum amount of trust is required to initiate a social interaction (Gosh and Ray, 1996; Kranton, 1996).

In this paper we embrace Putnam's perspective as it focuses on the individual origin and source of social capital. As Gleazer et al. (2002) point out, only if social capital is an individual concept can we hope to rationalize it, discuss its accumulation and its production by applying the economic

toolbox. Conversely a concept mainly defined as a communitarian phenomenon, as it is the case in the post-Coleman literature, would be hardly manageable.

As mentioned, the most widely used indicator of individual social capital is the degree of individual trust. For instance, the European Social Survey asks: "generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?". Respondents rate their agreement on a ladder ranging from 0 to 10. This is our measure of individual social capital. There is not an equally clear choice for community social capital. Our indicator of collective social capital associated to individual *i* is the average trust of all individuals living in the same region of *i*. We shall discuss this choice more in depth in the following sections.

Before proceeding with the estimation of the relationship between social capital and health let us clarify the mechanisms likely to be responsible for the positive role of social capital on individual health. The most intuitive is the easier access to health relevant information, which is related to the intensity of social interaction (frequent meeting of friends and relatives, participation to social events and meetings, membership to formal and informal organization, etc.). The more involved is one agent in continuous social interaction, the easier and cheaper his access to information on diseases, remedies, past experiences with hospitals, health personnel, doctors or drugs. In an environment such health care, where asymmetric information between health suppliers and consumers is widespread and represents a central issue in any tentative of reform and regulation of the sector, having better and more complete information is certainly important, as it allows to choose carefully the most suited hospital or doctor.

A second mechanism is that related to the provision of informal health care and support in case of illness. Even in developed countries, where formal health care is ubiquitous, there still exists a substantial demand for informal assistance[2], housing services and babysitting in case of temporary illness (sometimes even financial support is required to cover the out-of-pocket costs of health care). Usually the market or the public health system are unable to provide these kind of services, either because of the short length of the illness periods which makes organization difficult or because the costs of provision which might exceed people's budget. For these reasons, people tend to agree on informal and tacit rules, such as reciprocal assistance between neighbours or between friends, which act as risk sharing devices that supplement formal health insurance. In less developed countries these agreements are even more important and frequent given that the sphere of action of the formal

[2] A recent report valued the the cost of informal care in the UK at £87bn – almost the same as the NHS budget as a whole. See http://www.carersuk.org/Policyandpractice/Research/Profileofcaring/1201108437 (last accessed 04/12/2009).

system is much narrower. Reciprocal support and assistance are possible only in a context of reciprocal trust, as there is no enforceable contract guaranteeing obligations. Repeated interactions facilitate such forms of cooperation: in LDC repetition is further guaranteed by the fact that households or kin groups, i.e. long-lasting social institutions which stand beyond individuals, are the true subjects of the reciprocal obligations; instead in developed countries obligations are taken by individuals (who possibly change residence rather often) and do not usually transfer to the heirs, making cooperation much more dependent on reciprocal trust.

As informal assistance can be formalized as a prisoner dilemma, it is well known that the cooperative Pareto efficient equilibrium is achievable only if each player trusts that his partner is cooperative. Actually reciprocal assistance is not a Nash equilibrium and cannot be the outcome of the game if players do not trust each other. Moreover, if only one player trust the other, the predictable outcome is that the trusting player provides assistance to his partner but he does not receive assistance from the partner when in need. This outcome is hardly positive to the trusting player's health.

At a more aggregate level, social capital may also serve to coordinate people's lobbying efforts and power towards public authorities to obtain health infrastructure, traffic regulations, sport facilities and so on. Note however that the level of aggregation cannot be too high, as different communities can have opposite views on the proper location for and the right cost of these facilities. Moreover, these investments are essentially non excludable public goods that will indistinctly benefit both citizens who participate in and who abstained from lobbying. Therefore, it can be possible that such coordination can be realized at village or borough level and not at higher level, such as regional or city where problems of coordination are more acute. Due to only very aggregate information about the respondents' residence place, here we are constrained to define community as the region of residence at the level of aggregation given in the survey. This meant that what we call "communities" corresponds to local units classed as NUTS2 by the European NUTS classification system.

The paper is organized as follows. Section 2 describes data and provides some summary statistics. Section 3 introduces the model of social capital and health we estimate. Section 4 extensively discuss the instruments used in the analysis, their relevance and their exogeneity. Section 5 presents estimates corresponding to four quasi-nested specification of the general model and Section 6 discusses the results and concludes.

## *Data*

We use the first and second round of the European Social Survey, a repeated cross-country survey which covers 21 European Countries. The ESS provides information about individual social behaviour and perception, such as political opinions, political participation, exposition to media and news, social relationships, trust of other people and institutions. In addition, ESS is particularly valuable because it provides detailed information about respondents socio-economic characteristics and parental background. Unfortunately health has not been a major focus in the survey design: respondents are only asked to self-report their current health status and whether they are hampered in daily activities by illness or disability. Nevertheless, despite the generality and potential bias in those questions, the advantages are that a) they have been widely used in many surveys and, more importantly, b) self reported as been shown to be a good predictor of mortality at the individual level.

Over 80 thousand people answered the questionnaire, equally shared between the two rounds dated 2002/03 and 2004/05. In each round the survey defines a representative sample for each of the twenty-one European Countries covered, so that about two thousand residents of each country provide information in each round.

Information on the region of residence, at NUTS2 in most cases, is also available. This feature has allowed us to supplement additional data about regional characteristics coming from the Eurostat REGIO dataset. In particular we have added regional indicators of development (GDP per capita, GDP growth and employment) of health supply (number of beds in hospitals and number of health personnel per 100,000 residents) as well as population density, length of road network and number of beds in hotels and residences. Such data is not available for every country at regional level, so that the dataset resulting from the merging between ESS and REGIO covers only fourteen countries.

The variables of primary interest are health on the one hand and individual and community social capital on the other. Regarding health, people are asked to rate their current health on a five-steps ladder ranging from very bad (1) to very good (5). Our health indicator (*goodhealth*) is a binary variable which takes 1 (good health) if respondents judge their health fair, good and very good. Instead very bad and bad health are coded as 0. Although self-reported health is of course a noisy measure of true health, it has shown to be a rather reliable indicator and it has been largely used in the literature. In particular, it has been shown that self-reported health is highly correlated with mortality.

Individual social capital is captured by an indicator extensively used in the literature since Putnam (1993), i.e. the individual degree of trust. Precisely, the question posed in the ESS (and very similar to that included in many other surveys covering social aspects) is "generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?". Respondents rate their trust on a ladder ranging from 0 to 10. We maintain this codification in our individual social capital variable (*trust*). Community social capital associated to individual *i* is measured as the mean trust of the residents in the same region of individual *i*. Average measures of individual social capital are frequently used in the literature (see Islam et al., 2006). However the definition of what is to be intended community is debated. Here community is defined as the region of residence (coded as NUTS2). Unfortunately, ESS do not provide more refined information about respondents' residence. As we have already mentioned in the introduction it is possible that the relevant level of aggregation is lower, such as village or borough.

Simple summary statistics of the variables included in the model are reported in Table 1.

[TABLE 1]

## *The model*

We want to test whether individual and community social capital – proxied by trust and average trust in the region – affect individual health.

Moreover, we want to test whether the impact of individual social capital is strengthen or weakened by community social capital. For instance, it is likely that an individual endowed with a sufficiently high level of trust and living in a community rich in social capital can easily develop an intense and wide social network. In contrast the same individual placed in a community lacking social capital, might find it difficult to set up social links as his potential partners are rather diffident.

We estimate the following regression model:

$$H^*_{irc} = \alpha_0 + \alpha_1 S^*_{irc} + \alpha_2 S^*_{irc}\overline{S}^*_{rc} + \alpha_3 \overline{S}^*_{rc} + X_{irc}\alpha_4 + R_{rc}\alpha_5 + u_c + \varepsilon_{irc} \qquad (1)$$

where $H^*_{irc}$ is an objective indicator of the health of an individual $i$, living in region $r$ of country $c$, and $S^*_{irc}$ is an objective measure of micro/individual social capital. The objective indicator of community social capital is $\overline{S}^*_{rc}$, defined in region $r$ of country $c$. We allow for an effect of $S^*_{irc}$ varying with $\overline{S}^*_{rc}$ by introducing the product of the two variables. In addition to community social capital, we control for individual and regional specific characteristics, $X_{irc}$ and $R_{rc}$ respectively. Finally, the error term is composed of country fixed effects $u_c$ and an individual specific error component $\varepsilon_{irc}$. We allow $E(S^*_{irc}\varepsilon_{irc})$ to be non zero while we assume that $\varepsilon_{irc}$ is uncorrelated with all other variables and in particular with objective community social capital. Reverse causality and

the possibility that some unobservable individual characteristics are omitted from the model justify the assumption $E(S^*_{irc}\varepsilon_{irc}) \neq 0$. Instead objective community social capital is claimed to be exogenous because it seems unlikely that individual health might have a feedback on community social capital. Moreover, this assumption is supported by the fact that a series of regional indicators ( $R_{rc}$ ) and country fixed effects are included in the model in order to limit the possibility of misspecification.

Unfortunately all $H^*_{irc}$ , $S^*_{irc}$ and $\overline{S}^*_{rc}$ are unobservable and only self-reported measures of health and social capital are available. Specifically, we observe self-reported health $H_{irc}$ and self-reported social capital $S_{irc}$ . We assume that respondent's self-reported variables are noisy measures of the true indicators $H^*_{irc}$ and $S^*_{irc}$ , and we define them as follows

$$H_{irc} = H^*_{irc} + \eta_{irc} \tag{2}$$

$$S_{irc} = S^*_{irc} + \lambda\overline{S}_{rc} + \mu_{irc} \tag{3}$$

where $\eta_{irc}$ and $\mu_{irc}$ are zero-mean noises such that $E(H^*_{irc}\eta_{irc}) = 0$ , $E\left(S^*_{irc}\mu_{irc}\right) = 0$ and $E\left(\overline{S}_{rc}\mu_{irc}\right) = 0$ . A central feature of our model is that we allow for the measurement error in reporting individual social capital to be correlated with *reported* community social capital. Finally, our observed measure of community social capital is the sample average of reported individual social capital in the region *r* of country *c* and we assume the following relationship with the objective (and unobserved) community social capital

$$\overline{S}_{rc} = \overline{S}^*_{rc} + \theta_{rc} \tag{4}$$

with the usual assumption that $E\left(\overline{S}^*_{rc}\theta_{rc}\right) = 0$ . Looking at equations equation (3) and (4), it is worthwhile noticing that we do not assume that objective community social capital is simply the sum (or the mean) of objective individual social capital, because, as it is suggested by the World Bank definition[3], community social capital is more than this[4].

---

[3] "Social capital is not just the sum of the institutions which underpin a society – it is the glue that holds them together" (quoted from The WorldBank's social capital website).

[4] Anyway, our equations (3) and (4) imply a relationship between objective individual and community social capital. By taking the sample mean at regional level of equation (3) we get $(1-\lambda)\overline{S}_{rc} = \frac{1}{N_{rc}}\sum_i S^*_{irc} + \frac{1}{N_{rc}}\sum_i \mu_{irc}$ so that

$$\overline{S}^*_{rc} = \frac{\sum_i S^*_{irc} + \sum_i \mu_{irc}}{N_{rc}(1-\lambda)} - \theta_{rc} \tag{5}$$

By substituting (2), (3) and (4) in (1) we obtain an expression relating observed self-reported individual health, reported individual social capital and our indicator of community social capital:

$$H_{irc} = \alpha_0 + \alpha_1 S_{irc} + \alpha_2 S_{irc}\overline{S}_{rc} + (\alpha_3 - \alpha_1\lambda)\overline{S}_{rc} - \alpha_2\lambda\overline{S}_{rc}^2 + X_{irc}\alpha_4 + R_{rc}\alpha_5 + u_c + \\ + \varepsilon_{irc} + \eta_{irc} - \alpha_1\mu_{irc} + \alpha_2\theta_{rc}(\lambda\overline{S}_{rc} - S_{irc}) - \alpha_2\mu_{irc}\overline{S}_{rc} + \alpha_2\mu_{irc}\theta_{rc} - \alpha_3\theta_{rc} \tag{6}$$

which can be rewritten as

$$H_{irc} = \gamma_0 + \gamma_1 S_{irc} + \gamma_2 S_{irc}\overline{S}_{rc} + \gamma_3\overline{S}_{rc} + \gamma_4\overline{S}_{rc}^2 + X_{irc}\gamma_5 + R_{rc}\gamma_6 + u_c + \tau_{irc} \tag{7}$$

where

$$\tau_{irc} = \varepsilon_{irc} + \eta_{irc} - \alpha_1\mu_{irc} + \alpha_2\theta_{rc}(\lambda\overline{S}_{rc} - S_{irc}) - \alpha_2\mu_{irc}\overline{S}_{rc} + \alpha_2\mu_{irc}\theta_{rc} - \alpha_3\theta_{rc} \tag{8}$$

First of all, note that model (1) is identified, as parameters $\alpha_3$ and $\lambda$ can be obtained from the reduced form estimates of (7) thanks to the relationships

$$\lambda = -\frac{\gamma_4}{\gamma_2} \quad \text{and} \quad \alpha_3 = \frac{\gamma_2\gamma_3 - \gamma_1\gamma_4}{\gamma_2} \tag{9}$$

Next, note that $E(S_{irc}\tau_{irc})$, and $E(\overline{S}_{rc}\tau_{irc})$ are all nonzero by construction. Therefore, although in model (1) we assumed that (objective) community social capital was exogenous, and there were good theoretical reasons to claim so, this is not the case anymore when we acknowledge that only a proxy of community social capital is actually observed. In other words the empirical version of model (1), the one that can be estimated in practice, introduces further complications. Precisely, endogeneity of $\overline{S}_{rc}$ depends on the joint role of 1) measurement error in $S_{irc}$ and 2) the presence of the interaction $S^*_{irc}\overline{S}^*_{rc}$. Thus, we end up with a model where all social capital variables (four!) are endogenous. Hence OLS estimates of (5) will be biased and inconsistent.

In addition, $\sigma^2_\tau = Var(\tau_{irc})$ can be expected to be large, as it is a combination of three error terms. Not only, it also depends on both $S_{irc}$ and $\overline{S}_{rc}$ so that homoskedasticity cannot be assumed. Finally, the presence of the regional-level error term $\theta_{rc}$ and of $\overline{S}_{rc}$ generates spatial correlation among people living in the same region.

Being OLS unavailable, we rely on IV estimators with standard errors robust to both heteroskedasticy and spatial correlation. As usual the choice of proper instruments is not easy, but in our setting is even more difficult, because of the large variance of $\tau_{irc}$ due to the measurement

---

i.e. objective community social capital is a linear transformation of average individual objective social capital within the region.

errors in self-reported social capital and health. A large variance of the error term has two negative impacts on the reliability of IV estimates.

First, it is well known that IV are unbiased only asymptotically: in finite samples, even if rather large, IV estimates are biased and the bias increases with the standard error of $\tau_{irc}$ , that in our case we have just shown to be large[5]. Second, IV estimate's variance increases with the variance of $\tau_{irc}$ . Both problems can be alleviated by adopting strong instruments, i.e. highly correlated with the endogenous variables, as both IV bias and IV variance decrease with the correlation between instruments and instrumented variables (see Wooldridge, 2002 – equations 5.37 and 5.39).

Fortunately our dataset is rather rich and several variables are suitable to be used as instruments for $SC_{irc}$.

### *Instruments.*

Overall we have four variables to instrument, individual and community social capital with their interaction and squared community social capital. Fortunately, once we have proper instruments for individual and community social capital, it is easy to produce additional instruments for the transformed variables by applying appropriate operators.

Given that community social capital is computed as the average of individual social capital reported by the residents of a region, each instrument suitable for individual social capital is potentially relevant also for community social capital.

Our set of instruments is as follows:

1) whether the birth-country of respondent's father and mother is the same as respondent's country of residence (*Z1* and *Z2* respectively)
2) whether the respondent or a member of his household has been victim of a burglary or an assault in the past 5 years (*Z3*).
3) population density at regional level (*Z4*)
4) length of roads network at regional level (*Z5*)
5) the percentage of residents without internet connection living in respondent's region (*Z6*)
6) the percentage of residents with the status of citizen in respondent's region (*Z7*)

Instruments *Z1*, *Z2* and *Z3* are defined at individual level, the remaining four at regional level.

All instruments must satisfy two requirements, they must be relevant, i.e. correlated with the endogenous variables and they must be exogenous, i.e. they must affect individual health only through the instrumented variables, without independent and autonomous role. We first discuss the

[5] Of course the bias is zero if instruments and the error term have exactly zero correlation, but this is the case only asymptotically

relevance of each instrument and next their exogeneity.

As regards *Z1* and *Z2* and *Z7*, there is an extensive literature pointing out that ethnic/national minorities in a country tend to remain less socially integrated (Alesina and La Ferrara, 2002; La Porta et al. 1999, Easterly and Levine, 1997). Social marginalization experienced early in the individual life might also shape subsequent social behaviour and ability of creating and managing social ties.

Having been recently a victim of a burglary or an assault (*Z3*) is certainly related with the degree of trust towards other people. A similar shock is likely to induce a widespread feeling of fear and distrusts against people outside a relatively narrow circle of close friends and relatives. Results suggest that the negative correlation is very strong.

Next *Z4*, *Z5* and *Z6* are correlated with the costs of social interaction. In more sparsely populated regions with a less developed read network the cost of social interaction, i.e. the cost of meeting together is higher. When internet connection is widespread, opportunities of creating relationships are larger and social interaction cheaper (think for instance to the role of e-mail and chat in the daily life). There exists a reciprocal relationship between trust and social interaction: while social interaction is certainly easier in a climate of reciprocal trust, people learn to trust or distrust their partners after a substantial amount of repeated interactions. Therefore *Z4*, *Z5* and *Z6* by favouring contacts and meetings among people are indirectly related with people's trust.

Exogeneity (i.e. the excludability condition of the proposed instruments) requires a more extensive and careful discussion.

Birthplace of respondent's father and mother, *Z1* and *Z2*, could have an independent effect on respondent's individual health, as they are likely to be related to parents' education, income and working conditions which are shown to be important inputs in the health production function (Rosenzweig and Shultz, 1982). To capture this effect of *Z1* and *Z2*, we have included in the set $X_{irc}$ rather detailed information on parental background, including father's and mother's education, employment status when respondent was 14, distinguishing between employed, self-employed and unemployed, and a couple of dummies indicating whether the father or the mother were dead when respondent was 14. We also control for whether respondent's birth country coincides with that of current residence to capture possible impediments or discrimination in accessing education and health care services. Moreover cultural aspects such as respondent's religion – generally transmitted by parents – are controlled for. Indeed, we are fairly sure that parents' birthplace has no effect per se on respondent's health except from the impact on opportunities of social relationship and

interaction (in other words parent's birthplace can reasonably be assumed to influence health only through social capital, once parental background is controlled for).

Having been victim of a crime or knowing a close person that have been in the recent past is certainly not an individual decision or under individual control. However we cannot claim that being victim of a crime is a truly random event, i.e. a completely exogenous accident, as people are able to modify the probability of these events, by avoiding risky borough of their town, by installing security devices in their houses etc. Moreover it is likely that the risk of burglary increases with people age and gender (elderly and women are easier targets), income and with crime intensity in the place of residence. Hence *Z3* is likely to be correlated with age, gender, household income and with crime intensity in the region of residence. The latter variables have an impact on individual health, age and gender obviously, income by determining the opportunities of investment in health, criminality by reducing people's mobility. Therefore we need to include them among individual and regional controls to being allowed to claim that *Z3* has no independent effects on health.

One might be concerned that having been victim of a crime has a direct impact on health. A more suitable instrument would have been whether a household member was victim of a crime, excluded the respondent himself. Data does not allow this distinction. However we are considering small events of thefts in victims' houses, unlikely to harm people physical health directly and permanently. There exist studies pointing at the emotional negative consequences of rapes and violence (WHO, 2002) but this type of crimes are not considered here.

The remaining four instruments, *Z4*, *Z5*, *Z6* and Z7, defined at regional level, could have an impact on individual health only trough their correlation with the economic development of the region of residence. For instance, in Europe, more densely populated regions and those more endowed of transport infrastructures are also the richer and those with higher supply of health care. Therefore we have controlled for regional GDP per capita, nominal GDP growth rate, employment and variables of health supply, such as number of bed place in the hospitals and number of health personnel.

Overall, all our instruments could have an impact on individual health that is not mediated exclusively through social capital. Such impact however would not be an autonomous, independent one, but would instead be due to instruments' correlation with third variables (parental background, income, regional development...). Therefore, once we control for the latter, we claim that the

exogeneity requirement is met. In the following estimates we report the Sargan or the Hansen J test which do not reject the hypothesis of exogeneity.

## *Estimates*

We proceed by estimating progressively more general models, by applying progressively less stringent restriction to the model (7). We are going to show that only the most general specification produces results which are interpretable. Constrained models appear to be mis-specified. The interest of showing constrained models is that they have been largely used in the literature. Note that by applying the restriction we are discussing below, constrained specifications of model (7) are not proper nested models, as the set of endogenous variables to be instrumented depends on the restriction imposed. Therefore also the set of instruments and their functional form must vary accordingly, in order to maximize their correlation to the instrumented variables (i.e. their relevance).

**Model 1**. Consider the simplest model. Assume first that $\alpha_3 = \alpha_2 = \lambda = 0$ . Indeed we estimate

$$H_{irc} = \gamma_0 + \gamma_1 S_{irc} + X_{irc}\gamma_5 + R_{rc}\gamma_6 + u_c + \tau'_{irc} \tag{10}$$

with

$$\tau'_{irc} = +\varepsilon_{irc} - \eta_{irc} - \alpha_1 \mu_{irc} \tag{11}$$

In this case homoskedasticity could be retained. The only variable to instrument is $SC_{irc}$. The set of excluded instruments is *I1={Z1,...,Z7}*. We report in Table 2 probit estimates (marginal effects) of equation (8), with and without region dummies (column 1 and 2), without accounting for endogeneity of individual social capital. It appears that our regional controls $R_{rc}$ are sufficient to capture regional heterogeneity as estimates of the social capital coefficient are very close in column 1 and 2. In column 3, OLS linear probability estimators of equation (8) are reported and their size is comparable to the probit estimates. Therefore linear probability model can be safely adopted. Finally in column 4 IV linear probability estimates are reported. Note that the magnitude of $SC_{irc}$ is more than ten times higher than in column 3. Table 2 also reports the p-value associated to the Anderson canonical correlation LR test which tests the strength of the instruments, the p-value of the Hansen overidentification test (which test instruments exogeneity) and the F test that all excluded instruments *I1* are jointly zero. While Hansen test does not reject the hypothesis of exogeneity of the instruments, both Anderson and the F tests prove the strength of our instruments. Overall, these results suggest that OLS are largely biased due to measurement errors and reverse causality. Looking at column 4, increasing *trust* by one will increase the probability of being in good health by about 9 percent.

[TABLE 2]

**Model 2**.Next assume that $\alpha_2 = \lambda = 0$. The model to estimate is then

$$H_{irc} = \gamma_0 + \gamma_1 S_{irc} + \gamma_3 \overline{S}_{rc} + X_{irc}\gamma_5 + R_{rc}\gamma_6 + u_c + \tau''_{irc} \tag{12}$$

with $\tau''_{irc} = +\varepsilon_{irc} - \eta_{irc} - \alpha_1 \mu_{irc} - \alpha_3 \theta_{rc}$ and $\gamma_3 = \alpha_3$. The assumption of spherical errors cannot be retained as there might be special correlation within each region. This specification controls for possibly independent effects of community social capital. By construction, community social capital is endogenous as well as $SC_{irc}$ : we instrument them by the set *I1* as in Model 1. Results are reported in Table 3, first and second columns. In the former, OLS estimates of (10) are reported as a benchmark and in the latter we present IV estimates. In both cases we opt for a linear probability model, and so we do in the following. While individual social capital is largely significant and positive, community social capital is not significant. However, instruments, although likely to be exogenous, seem to be poorly correlated with community social capital, a fact which might bias our estimate. This concern pushes us to adopt the procedure described in the following model 3 and 4 to improve the fit.

**Model 3**. Next consider restriction $\lambda = 0$. The model to estimate is

$$H_{irc} = \gamma_0 + \gamma_1 S_{irc} + \gamma_2 S_{irc}\overline{S}_{rc} + \gamma_3 \overline{S}_{rc} + X_{irc}\gamma_5 + R_{rc}\gamma_6 + u_c + \tau_{irc} \tag{13}$$

with the most general error term. In this case both individual and community social capital are endogenous and need to be instrumented. The error term is heteroskedastic and spatially correlated. Instruments are functionally transformed to maximize the correlation with the endogenous variables and maximize their strength. Precisely, the set of instruments is

$$I3 = I1 \cup \{Z3/Z4, Z3/Z6, Z3/Z7, \sqrt{Z5}, \sqrt{Z5}\ln(Z5), \sqrt{Z4/1000}, (Z4/1000)^3, Z6^2, Z7^3\}$$

which is obtained by a multiple fractional polynomial fit of each endogenous variable with included and excluded instruments. Relevance and Hansen test are clearly passed by *I3*. Note the improvement in the F test corresponding to $\overline{S}_{rc}$ with respect to model 2. Results are reported in column 3 and 4 of Table 3. In the former OLS are reported as benchmark and in the latter IV. Now community social capital seems to have a significantly negative independent effect on individual health. Individual social capital marginal effect, given by $\gamma_1 + \gamma_2 \overline{SC}_{rc}$, is not significantly different from zero. Note however that it is positive when computed at the mean of $\overline{S}_{rc}$. Mis-specification of the model due to the assumption λ=0 is likely to be responsible for these results as we shall see below.

**Model 4**. Finally, we estimate equation (5) without restrictions. IV estimates are reported in column 5 of Table 3. Now we have four endogenous variables $\left\{SC_{irc}, SC_{irc}\overline{SC}_{rc}, \overline{SC}_{rc}, \overline{SC}^2_{rc}\right\}$. The set of

instruments we adopt is $I4 = I3 \cup \{\sqrt{Z6}\}$ again obtained by multiple fractional polynomial fit. Hansen J test is passed, F tests for each of the endogenous variables are satisfactory. However, *I4* performs badly in the Anderson canonical correlation test: yet, this should not be interpreted as an indication that the excluded instruments are not relevant, as they have proved to be rather strong in each specification and as the F tests indicate that singularly each endogenous variable is well correlated with *I4*. The negative result is likely due to the large degree of collinearity between $\overline{SC}_{rc}$ and $\overline{SC}_{rc}^2$. Estimates show that individual social capital marginal effect $\gamma_1 + \gamma_2 \overline{SC}_{rc}$ is significant and positive for individuals living in communities with sufficiently high social capital (higher than 4.655, to be compared with an average value of 4.846). Moreover, by using equation (7), we obtain that $\alpha_3$ is not significantly different from zero while $\lambda$ is highly significant ($\alpha_3$=-0.0263 (s.e. 0.0597), $\lambda$=0.9428 (s.e. 0.1343)), i.e. community social capital plays no autonomous role in determining individual health, but the bias in reporting individual social capital does depend on community social capital.

[TABLE 3]

## *Results*

Overall we found that individual social capital seems to be a significant, true determinant of individual health only if trusting individuals live in regions with sufficiently high community social capital. In particular in regions with average community social capital, the marginal effect of individual social capital on health is 0.028, i.e. increasing individual trust by one unit (on a scale from 0 to 10) will increase the probability of being in good health of 2.8%. Compared to OLS/PROBIT, IV estimates are much larger: this is because we control for error-in-variable issues and reverse causality. There is no evidence that community social capital plays an autonomous role ($\alpha_3$=0), but instead it contributes to the effect of individual social capital. In other words, social capital is effective if diffused at community level. Namely, trusting the others has a positive impact on health only if trust is reciprocal: trusting when not trusted will lead to the worst outcome to the player in a prisoner dilemma-like setting as discussed in the introduction.

Finally, acknowledging that community social capital alters the reporting of individual social capital proved to be crucial in producing reliable and sensible results. Actually, this effect is highly significant and large enough to determine the apparently negative independent impact of community social capital that we obtained in Model 2 and 3, while otherwise community social capital plays no autonomous role.

## *Conclusions*

This paper improves on the previous literature that explores the relationship between social capital and health in that it explicitly accounts for measurement error in self-reported variables. With few exceptions (see e.g. Folland, 2007; D'Hombres et al. 2007), also endogeneity of social capital has been overlooked so far in the widely used multilevel analysis. Here, instead, much emphasis is put on the choice of adequate instruments to break down the circular relationship relating social capital and health, in order to identify the causal impact of social capital.

Overall we have obtained that individual social capital increases the probability of being in good health if the community has sufficiently high social capital. However, community social capital does not affect health directly.

The latter result should taken with care. It is possible that the communities should be smaller than those we consider, as cooperation is usually easier among a limited number of agents. Other limitations of the present analysis are related to the fact that in spite of the many individual and regional controls included in the regression, some unobservables could remain (e.g. risk aversion, time preferences, predisposition to relating with other people). However a proper account of unobserved heterogeneity can be achieved only by disposing of panel data: unfortunately most existing dataset which pay attention to social issues tend to be designed as repeated cross sections (with the notable exception of SHARE).

As in previous research, we get that social capital cannot be overlooked by policy makers when they seek at improving health conditions. The novelty of this study in terms of policy implications is that policy interventions should be targeted at improving primarily individual social capital, because in so doing they would achieve a double effect: on the one hand they would directly improve individual health; on the other hand they would contribute to community social capital that reinforces the beneficial role of individual social capital. Exploiting such reinforcing mechanism could improve the (cost-)effectiveness of policies: an intervention that succeeds in improving the social capital of a large number of individuals in one community would produce a larger health benefit than an intervention that targets the same number of individuals located in different communities..

Further research should be devoted to model theoretically and formally the role social capital in influencing individual health: so far empirical literature advanced without a clear theoretical guideline. This papers partly goes in this direction, but mainly from an empirical perspective.

## *References*


1. Alesina, A., and La Ferrara E. (2002). “Who trusts others?” Journal of Public Economics, 85(2): 207-234.
2. Bourdieu, P. (1986), “The Forms of Capital”, in John G. Richardson (edt), Handbook of Theory and Research in the Sociology of Education, New York, Greenwald Press.
3. Coleman, J. (1988) "Social Capital in the Creation of Human Capital", American Journal of Sociology. 94 Supplement: (pp. S95-S-120)
4. D’Hombres B., Rocco L., Suhrcke M. and McKee M. (2007) “Does Social Capital Determine Health: Evidence from Eight Transition Countries”, EU Commission - JRC working paper.
5. Durlauf, S., and Fafchamps, M. (2005). “Social Capital” in Handbook of Economic Growth, Steven Durlauf and Philippe Aghion (eds.), Amsterdam, New Holland
6. Easterly,W., Levine, R., (1997). “Africa’s growth tragedy: policies and ethnic divisions.” Quarterly Journal of Economics, 1203–1250.
7. Ghosh, P. and Ray, D. (1996) "Cooperation in Community Interaction without Information Flows", Review of Economic Studies, vol. 63, pp. 491-519.
8. Glaeser, E.L., Laibson, D., and Sacerdote B. (2002). “An economic approach to social capital.” The Economic Journal, 112: F437-F458
9. Islam, K., Merlo, J., Kawachi, I., Lindstrom, M., and U. Gerdtham (2006). “Social capital and health: Does egalitarianism matter? A literature review” International Journal for Equity in Health, 5(3):
10. Iversen, T. (2007). “An exploratory study of associations between social capital and self-assessed health in Norway” University of Oslo, HERO Working paper 2007: 9
11. Kranton R.E (1996), “The Formation of Cooperative Relationships”, Journal of Law Economics and Organization 12 : 214-233.
12. La Porta, R., Lopez de Silanes, F., Shleifer, A., Vishny, R., (1999). “The quality of government”. Journal of Law and Economics January, 222–228.
13. Poortinga, W. (2006). “Social capital: An individual or collective resource for health?” Social Science and Medicine 62(2): 292–302.
14. Putnam, R. D., Leonardi R., and R. Nanenetti (1993). “Making democracy work: civic traditions in modern Italy”. Princeton, N.J.: Princeton University Press.
15. Rosenzweig M.R. and Schultz P.T. (1982) “The Behavior of Mothers as Inputs to Child Health: The Determinants of Birth Weight, Gestation, and Rate of Fetal Growth”, in Economic Aspects of Health, edited by V.R. Fuchs, University of Chicago Press
16. Scheffler, R., Brown, T.T. and J. Rice (2007). “The Role of Social Capital in Reducing Non-specific Psychological Distress: The Importance of Controlling for Omitted Variable Bias.” Social Science and Medicine, 65: 842-854.
17. WHO (2002) “World report on violence and health”, WHO Geneva
18. Wooldridge J.M. (2002) “Econometric Analysis of Cross Section and Panel Data”, The MIT Press

Table 1

| Variable | Description | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| goodhealth | good health | .9186977 | .2733031 | 0 | 1 |
| trust | trust | 4.905488 | 2.410518 | 0 | 10 |
| trust * mean trust | interaction of trust and mean trust | 24.51799 | 13.91545 | 0 | 67.18062 |
| mean trust | average trust of people other than i living i's region | 4.846362 | .8793804 | 1.419355 | 6.851852 |
| mean trust ^ 2 | squared mean trust | 24.26051 | 8.702694 | 2.014568 | 46.94788 |
| Z4 | population density at regional level | 443.4313 | 830.8 | 3.3 | 6231.8 |
| Z5 | roadway network length | 194.5568 | 360.3147 | -2.3 | 1903 |
| Z1 | birthplace of respondent's father | .9103256 | .2857192 | 0 | 1 |
| Z2 | birthplace of respondent's mother | .9084031 | .2884605 | 0 | 1 |
| Z3 | victim of a burglary/assault in the last 5 years | .2063876 | .4047182 | 0 | 1 |
| Z6 | percentage of people without internet access at regional level | .3884515 | .1578319 | 0 | .8987342 |
| Z7 | percentage of residents with citizenship at regional level | .9721319 | .0281157 | .8089887 | 1 |
| Z3*Z7 | interaction of Z3 and Z7 | .2004388 | .3932965 | 0 | 1 |
| Z3*Z4 | interaction of Z3 and Z4 | 99.84439 | 475.8085 | 0 | 6231.8 |
| Z3*Z6 | interaction of Z3 and Z6 | .0780283 | .1701172 | 0 | .8987342 |
| sqrt(Z5) | squared root of Z5 | 9.716104 | 19.81376 | .724466 | 622.3228 |
| sqrt(Z5)*ln(Z5) | squared root of Z5 * natural log of Z5 | -53.37346 | 246.0712 | -8007.377 | .4670204 |
| sqrt(Z4/1000) | squared root of Z4/1000 | 14.13051 | 34.89007 | .1604673 | 303.0303 |
| (Z4/1000)^3 | cubed Z4/1000 | 3.371096 | 20.12082 | 3.59e-08 | 242.014 |
| Z6^2 | squared Z6 | .1758387 | .1395285 | 1.91e-09 | .8078017 |
| Z7^3 | cubed Z7 | .920975 | .0763897 | .529453 | 1 |
| sqrt(Z6) | squared root of Z6 | 17.96063 | 583.011 | 1.112622 | 22857.24 |
| Mcrime | percentage of residents which reported to have been victims of burglary/assault in last 5 years | .2050076 | .0909894 | 0 | .6 |
| fatheredu | respondent's father education (level) | 1.788403 | 1.604824 | 0 | 6 |
| motheredu | respondent's mother education (level) | 1.555845 | 1.38593 | 0 | 6 |
| father_employed | respondent's father employed (dependent) when respondent was 14 | .6619225 | .4730624 | 0 | 1 |
| mother_employed | respondent's mother employed (dependent) when respondent was 14 | .3825736 | .486023 | 0 | 1 |
| father_selfemployed | respondent's father self-employed when respondent was 14 | .232155 | .4222139 | 0 | 1 |
| mother_selfemployed | respondent's mother self-employed when respondent was 14 | .1088992 | .3115176 | 0 | 1 |
| fatherdied | respondent's father died before respondent's was 14 | .0602171 | .2378922 | 0 | 1 |
| motherdied | respondent's father died before respondent's was 14 | .0208682 | .1429453 | 0 | 1 |
| male | gender (male = 1) | .4752248 | .4993936 | 0 | 1 |
| age | age in years | 47.71191 | 17.69981 | 14 | 99 |
| age2 | squared age | 2589.7 | 1777.702 | 196 | 9801 |
| brncntr | respondent born in the country of residence | .9366202 | .2436487 | 0 | 1 |
| urban | urban residence | .6489922 | .4772928 | 0 | 1 |
| married | marital status | .5511938 | .49738 | 0 | 1 |
| eduyrs | respondent's education years | 11.69479 | 4.122879 | 0 | 40 |
| HHincome | respondent's household income | 5.861364 | 2.311042 | 1 | 12 |
| HHmembers | number of household members | 2.694822 | 1.369369 | 1 | 13 |
| religious | self-reported rate of religiosity | 4.720992 | 3.027331 | 0 | 10 |
| bed | number of bed places in the region hospitals per 100000 inhabitants | 628.651 | 241.0924 | 154.6 | 1822.6 |
| hp | number of heath personnel at regional level per 100000 inhabitants | 346.7084 | 107.4007 | 154.8 | 830.3 |
| gdp | GDP per capita at regional level | 21.64352 | 7.121882 | 7.5681 | 53.3811 |
| gro | growth rate at regional level | 1.995475 | 2.019083 | -2.3 | 15 |
| emp | employment in thousands at regional level | 978.3604 | 689.8418 | 27.69268 | 5345.4 |
| essround | ESS round | 1.533147 | .4989078 | 1 | 2 |
| Religion dummies | Roman Catholic, Protestant, Eastern Orthodox, Other Christian denomination, Islam, Eastern religions, other non Christian Religion (reference: non religious) | | | | |
| Occupation dummies | one digit ISCO codes 1-9 (reference: unemployed) | | | | |
| Country dummies | Austria (reference), Belgium, Czech Republic, Germany, Spain, Finland, France, Greece, Hungary, Italy, Netherlands, Poland, Portugal, Sweden | | | | |

Sample size: 31914

Table 2

| | Model 1 | | | |
|---|---|---|---|---|
| | Probit (ME) | Probit (ME) | OLS | IV |
| | goodhealth | goodhealth | goodhealth | goodhealth |
| trust | 0.0049 | 0.0051 | 0.0077 | 0.0905 |
| | (10.84)*** | (10.71)*** | (11.73)*** | (5.05)*** |
| male | 0.0033 | 0.0029 | 0.0044 | -0.0020 |
| | (1.48) | (1.23) | (1.41) | (0.49) |
| age | -0.0046 | -0.0048 | -0.0038 | -0.0025 |
| | (11.88)*** | (11.91)*** | (7.34)*** | (3.57)*** |
| age2 | 0.0000 | 0.0000 | 0.0000 | -0.0000 |
| | (8.51)*** | (8.56)*** | (2.39)** | (0.13) |
| brncntr | 0.0031 | 0.0034 | 0.0055 | -0.0049 |
| | (0.63) | (0.67) | (0.87) | (0.60) |
| urban | 0.0066 | 0.0054 | 0.0090 | 0.0186 |
| | (2.68)*** | (2.15)** | (2.74)*** | (4.11)*** |
| married | 0.0159 | 0.0167 | 0.0303 | 0.0237 |
| | (6.07)*** | (6.11)*** | (8.42)*** | (5.11)*** |
| eduyrs | 0.0030 | 0.0032 | 0.0055 | 0.0006 |
| | (8.14)*** | (8.15)*** | (11.05)*** | (0.48) |
| HHincome | 0.0036 | 0.0038 | 0.0048 | -0.0000 |
| | (5.29)*** | (5.45)*** | (5.46)*** | (0.03) |
| HHmembers | 0.0015 | 0.0016 | 0.0021 | 0.0010 |
| | (1.38) | (1.45) | (1.54) | (0.61) |
| religious | -0.0003 | -0.0004 | -0.0009 | -0.0043 |
| | (0.60) | (0.78) | (1.40) | (3.99)*** |
| fatheredu | -0.0007 | -0.0007 | -0.0011 | -0.0068 |
| | (0.54) | (0.57) | (0.74) | (3.11)*** |
| motheredu | 0.0032 | 0.0037 | 0.0019 | -0.0012 |
| | (2.15)** | (2.32)** | (1.12) | (0.52) |
| father_employed | 0.0104 | 0.0114 | 0.0160 | -0.0100 |
| | (2.12)** | (2.22)** | (2.24)** | (0.96) |
| mother_employed | -0.0040 | -0.0041 | -0.0077 | -0.0042 |
| | (1.49) | (1.48) | (2.12)** | (0.95) |
| father_selfemp | 0.0185 | 0.0197 | 0.0320 | 0.0035 |
| | (3.92)*** | (3.97)*** | (4.14)*** | (0.31) |
| mother_selfemp | -0.0075 | -0.0075 | -0.0124 | -0.0088 |
| | (1.84)* | (1.80)* | (2.32)** | (1.33) |
| fatherdied | 0.0086 | 0.0085 | 0.0080 | 0.0007 |
| | (1.58) | (1.49) | (0.88) | (0.06) |
| motherdied | 0.0038 | 0.0040 | 0.0070 | 0.0092 |
| | (0.60) | (0.59) | (0.67) | (0.72) |
| essround | -0.0008 | -0.0020 | -0.0018 | 0.0046 |
| | (0.35) | (0.71) | (0.47) | (0.96) |
| bed | | -0.0000 | -0.0000 | -0.0001 |
| | | (0.36) | (1.00) | (2.77)*** |
| hp | | -0.0000 | -0.0000 | 0.0000 |
| | | (0.78) | (0.54) | (0.40) |
| gdp | | 0.0008 | 0.0011 | -0.0001 |
| | | (2.34)** | (2.62)*** | (0.25) |
| gro | | 0.0002 | -0.0001 | -0.0014 |
| | | (0.20) | (0.09) | (1.08) |
| emp | | -0.0000 | -0.0000 | -0.0000 |
| | | (1.24) | (1.49) | (3.43)*** |
| Mcrime | | -0.0312 | -0.0363 | 0.0662 |
| | | (1.58) | (1.39) | (1.71)* |
| religion dummies | yes | yes | yes | yes |
| occupation dummies | yes | yes | yes | yes |
| region dummies | yes | no | no | no |
| country dummies | yes | yes | yes | yes |
| Constant | | | 0.8928 | 0.6297 |
| | | | (39.23)*** | (9.95)*** |
| Observations | 31914 | 31914 | 31914 | 31914 |
| Anderson LR (p) | | | | 0.00 |
| Sargan (p) | | | | 0.25 |
| F trust | | | | 9.16 |
| R-squared | | | 0.11 | |

Absolute value of z statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3

| | Model 2 | | Model 3 | | Model 4 |
|---|---|---|---|---|---|
| | OLS | IV | OLS | IV | IV |
| | goodhealth | goodhealth | goodhealth | goodhealth | goodhealth |
| trust | 0.0078 | 0.0936 | 0.0177 | -0.0972 | -0.6889 |
| | (11.82)*** | (4.43)*** | (3.64)*** | (1.05) | (2.83)*** |
| mean trust | -0.0086 | -0.0152 | 0.0004 | -0.2335 | 0.6231 |
| | (1.73)* | (0.35) | (0.05) | (2.76)*** | (2.14)** |
| trust*mean trust | | | -0.0021 | 0.0343 | 0.1480 |
| | | | (2.25)** | (1.93)* | (3.10)*** |
| mean trust ^ 2 | | | | | -0.1395 |
| | | | | | (2.85)*** |
| Observations | 31914 | 31914 | 31914 | 31914 | 31914 |
| R-squared | 0.11 | | 0.11 | | |
| Anderson LR (p) | | 0.00 | | 0.00 | 0.57 |
| Sargan / Hansen J (p) | | 0.60 | | 0.15 | 0.66 |
| F trust | | 8.45 | | 7.23 | 7.41 |
| F trust*mean trust | | | | 8.74 | 8.46 |
| F mean trust | | 2.24 | | 6.36 | 6.00 |
| F mean trust^2 | | | | | 5.65 |

Absolute value of t statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%