HEDG
HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 11/21

From social capital to health - and back

Lorenzo Rocco
Fumagalli Elena
Marc Suhrcke

July 2011

york.ac.uk/res/herc/hedgwp

From social capital to health - and back

Lorenzo Rocco — University of Padova
Fumagalli Elena — University of East Anglia
Marc Suhrcke — University of East Anglia

July 28th, 2011

Abstract

We assess the causal relationship between health and social capital, measured by generalized trust, both at the individual and the community level. The paper contributes to the literature in two ways: it tackles the problems of endogeneity and reverse causation between social capital and health by estimating a simultaneous equation model, and it explicitly accounts for mis-reporting in self reported trust.

The relationship is tested using data from the first four waves of the European Social Survey for 26 European countries, supplemented by regional data from the Eurostat. Our estimates show that a causal and positive relationship between self-perceived health and social capital does exist and that it acts in both directions. In addition, the magnitude of the structural coefficients suggests that individual social capital is a strong determinant of health, while community level social capital plays a considerably smaller role in determining health.

Introduction

Social capital, despite some ambiguity as to how it should best be defined, has increasingly been recognised among economists as an important concept that matters for a range of key economic outcomes, including financial development (Guiso et al, 2004), innovation (Akçomak and ter Weel, 2009), the spread of secondary education (Goldin and Katz, 2001) and economic growth (Zak and Knack, 2001). Economists have also begun to develop a theoretical economic framework for thinking about the determinants of investment in social capital (Glaeser et al, 2002 and more recently Guiso et al, 2008), concluding that social capital fits well into standard microeconomic reasoning. Recently, Guiso et al. (2008, 2010) have convincingly redefined social capital in terms of civic capital and have shown that it can be measured, accumulated, transmitted and that it has non-negative returns. In doing so, Guiso et al. (2008, 2010) have fully addressed Solow's (1995) criticism of the concept of social capital. Yet, the relationship between social capital and health has come to the attention of economists only very recently, in some contrast to a steadily growing public health literature on the subject since the mid-1990ies (Scheffler and Brown, 2008).

The existing empirical evidence in both economics and public health documents the close association between social capital and health (see Cooper et al, 1999; Lochner et al, 1999; Machinco and Starfield, 2001 for early reviews in public health and, more recently, Engstrom et al, 2008 and Islam et al, 2006, for an excellent survey). It is much less clear though to what extent the observed association also implies a causal relationship. Both Durlauf (2002) and Durlauf and Fafchamps (2005) have demonstrated how the early literature on social capital (not much of which was applied to health) generally failed to identify the causal effect of social capital, especially because most studies did not separate the effect of social capital itself from that of individual preferences as well as other community characteristics. Only a few recent contributions have tried to address the causality problem in the relationship between social capital and health by means of instrumental variables (Folland, 2007; D'Hombres et al, 2010; Ronconi et al, 2011 among others). In addition, although the previous literature had acknowledged that the relation between social capital and health can be circular, a more complete empirical evaluation of this complex, potentially bi-directional relationship has hitherto been lacking.

Besides, there is a need to specify the relevant dimension of social capital that may be responsible for the link between social capital and health: is it the social capital at the individual level or at some more aggregate community level, or both, that matters? This ambiguity is partly related to the way social capital is defined. The existing literature is divided into those papers that consider social capital primarily as an individualistic property (Glaeser, 2002) and those that perceive it as a feature of the community (Putnam et al. 1993). It is crucial to assess the relative importance of both dimensions, as this would suggest the appropriate entry point for policy interventions: should policy target social capital at the individual level (e.g. by providing social support through individualised "befriending", a practice that has gained considerable interest in treatment of mental health problems[1]), or should policy rather seek to promote social capital at the aggregate level via community development strategies, e.g. urban regeneration programmes (Thomson et al, 2006)?

The purpose of the paper is twofold: first, we evaluate for the first time the simultaneous two-way causal link between social capital and health. Second, we analyze at which level the relationship operates. We use data from 26 European countries that were part of the first four waves of the European Social Survey (ESS) collected in the years 2002, 2004, 2006 and 2008.

Our "baseline" empirical modelling approach derives from an agnostic view about the direction of causality: we have specified a simultaneous equation model to allow for a general feedback relationship between social capital and health. In doing so, instead of pre-imposing a given direction of causality, we let the data judge whether it is social capital that influences health, whether causation runs in the opposite direction or whether the relationship

1 "Befriending" is defined in the public health literature as the forming of "a relationship between two or more individuals which is initiated, supported and monitored by an agency that has defined one or more parties as likely to benefit. Ideally the relationship is non-judgemental, mutual, and purposeful, and there is a commitment over time" (Mead et al 2010, p. 96) Many charitable and voluntary sector organizations offer befriending in the UK, for instance, and clinical guidelines suggest a role for befriending for people with chronic depression (Mead et al 2010).

is indeed bi-directional. More specifically, we follow the two-stage procedure proposed by Stern (1989), paying attention to the exclusions required to identify the model. In a further elaboration (and robustness check) of this simultaneous equation model, we analyze whether and how possible mis-reporting in social capital and health might alter our baseline estimates. To the best of our knowledge this is the first paper in this literature that explicitly deals with the issue of mis-reporting, generally considered as a major empirical challenge in the social science research on social capital (Guiso et al., 2010).

We find that the relationship between social capital and health is indeed circular and mutually beneficial, both in the baseline model and when we account for mis-reporting. This result is robust to a series of alternative specifications. Moreover, social capital measured at the individual level has both a stronger and a statistically more significant impact than community social capital, although the effects at either level remain positive.

The paper is organized as follows. Section 2 reviews the concept of social capital, the plausible channels through which social capital influences health and the relevant empirical literature. Section 3 introduces the model, section 4 describes the data, section 5 presents our results and finally section 6 concludes. Two technical appendices are attached.

2. Social capital and health

2.1 Definitions

A key limitation that has arguably stymied an even greater research interest relates to the considerable ambiguity as to what is actually meant by the notion of social capital (Durlauf and Fafchamps, 2005). One prominent, if indeed very encompassing definition by the World Bank attempts to capture virtually all possible dimensions of social capital, when proposing that it "*refers to the institutions, relationships, and norms that shape the quality and quantity of a society's social interactions. […] Social capital is not just the sum of institutions [that] underpin a society - it is the glue that holds them together*" (Rossing Feldman and Assaf 1999). Short of even trying to provide an exhaustive account of the numerous other definitions that have been proposed, one basic distinction that can be made is between those authors who perceive social capital as a property of some aggregate (or "community") level, and those who consider social capital primarily as a characteristic of the individual. Sociological and political science definitions have tended - though by no means exclusively - to emphasise the aggregate level perspective. For instance, the American political scientist Robert Putnam referred to 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).

In some contrast, economists have tended to express a preference for a more individualistic notion of social capital. Glaeser et al. (2002) emphasise the individual level perspective by defining social capital as the individual's social characteristics that enable private returns via interaction with others. Glaeser et al. (2002) also point out that 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 (which is what they provide in their paper). By contrast, a concept mainly defined as a communitarian phenomenon would hardly be manageable.

Yet even among economists there is no consensus about a common definition. Guiso et al. (2008) define social capital as "good" culture—in other words, a set of beliefs and values that facilitate cooperation among the members of a community. Essentially, they consider social capital as a capital of civism and convincingly show that this can be measured by both direct indicators (e.g. generalized trust) and indirect ones (e.g. blood donations and referenda turnout). Since there is no economic payoff and there is no legal obligation to donate or participate, the decision can be seen as a direct measure of how much people internalize the common good. This decision must be closely related to values and beliefs and it is therefore hard to explain by standard neoclassical economics. Such "good culture" can be accumulated and transmitted across generations and it can be distinguished from human capital because its returns are contingent on the norms and beliefs of the other members of the community.

As we cannot capture all the potential dimensions of social capital in our empirical analysis, we focus on one widely used proxy for social capital: the generalised level of trust. According to Knack and Keefer trust ''*reflects the percentage of people in a society who expect that most others will act cooperatively in a prisoner's dilemma context*'' (Knack and Keefer, 1997, p. 1258). Trust, as it is measured in many attitudinal surveys, including the World Value Surveys, has been shown to have predictive validity for a number of other objective social capital indicators, e.g. the share of wallets returned in wallet drop experiments from capitals around the world (Zak and Knack, 2001), and with indicators of corruption and violent crime (Uslaner, 2002). We also take comfort in using trust from studies that find a close correlation between generalized trust as measured in attitudinal (and hence hypothetical) surveys and actual behavior as observed in experiments (e.g., Fehr and Fischbacher, 2003; Glaeser et al, 2000). Our selection is also supported by Guiso et al. (2010) who, after having discussed the pros and cons of alternative measures of beliefs related to social capital, conclude that "*If we want to measure a country's or a community's civic capital, which is the right measure of trust? From a theoretical point of view, the right measure is generalized trust*." (Guiso et al 2010, p.23).

2.2 Mechanisms

Although all definitions refer directly or indirectly to social connections or social network, as elements of social capital, Putnam (1993) and Guiso et al. (2008, 2010) point to the role of social capital as a catalyst of coordination/cooperation, an essential device to achieve better social or economic outcomes. Cooperation allows to reduce transaction costs, to overcome difficulties due to incomplete or asymmetric information, and to establish efficient transactions in the presence of incomplete contracts (Alesina and La Ferrara, 2002). With this in mind, our selection of trust as an indicator of social capital appears to be appropriate, as trust favours cooperation, easing the need for long standing personalized relationships and processes of reputation building. Moreover, trust is a determinant of social connections, as a minimum amount of trust is required to initiate a social interaction (Ghosh and Ray, 1996; Kranton, 1996).

By favouring cooperation, social capital could indirectly also benefit individual health. Several mechanisms may account for the potential positive influence of social capital on individual health.

(1) The most intuitive one is the easier access to health relevant information, which is related to the intensity of social interaction (i.e. the frequent meeting of friends and relatives, participation in social events and meetings, membership in formal and informal organization, etc.) (Berkmann and Glass, 2000). In fact, the more an individual is involved in continuous social interaction, the easier and cheaper her access to information on diseases, remedies, past experiences with hospitals, health personnel, doctors or drugs tends to be. When asymmetric information between health suppliers and consumers is pervasive, obtaining better and more complete information is certainly important, as it allows to carefully choose the most suited health care. Beyond health care, the influence of the social network may also affect the extent to which (more or less) healthy behaviours have been adopted within a community.

(2) A second mechanism relates to the provision of informal health care and/or psychological support in case of illness. Even in developed countries, where formal health care is ubiquitous, there still exists a substantial demand for informal assistance, housing services and babysitting in case of temporary illness. Even financial support may be occasionally required to cover the out-of-pocket costs of health care. The market or the public health system are usually unable to provide such services, either because of the short duration of the illness periods, which makes organization difficult, or because the costs of provision might exceed people's budget. Therefore, 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. Reciprocal support and assistance are possible only in a context of reciprocal trust, as there is no enforceable contract guaranteeing obligations. It has also been supposed that individuals from a disadvantaged socioeconomic background stand to reap greater health benefits from social capital compared to wealthier individuals, because the former tend to be

at a comparative loss when it comes to acquiring and understanding relevant health information and obtaining social support (Scheffler and Brown, 2008).

(3) At an aggregate level, social capital may coordinate people's lobbying efforts and power to obtain health-enhancing goods and services from public authorities, including health infrastructure, traffic regulations, and sport facilities etc. As these investments are essentially non-excludable public goods that will indistinctly benefit both citizens who do and do not participate in lobbying, such coordination is easier to realize at village or borough level and less likely at the regional or the city or even more aggregate level, where coordination costs may become prohibitive.

Not all the effects of social capital on health have to be unambiguously positive, however. Social relationships may even increase susceptibility to infectious diseases, or to the adoption of unhealthy behaviours, driven by existing norms among peers (Kawachi and Berkman, 2001). Nonetheless, Brown et al. (2006) find a negative association between community social capital and smoking.

While the bulk of the above potential mechanisms suggests a health improving effect of social capital, the actual net impact remains to be empirically assessed.

2.3 Previous empirical literature

Most existing studies adopt multilevel estimators and focus only on the *association* between health and alternative measures of social capital, either at the individual or at the community level (or both). Typically, these studies estimate reduced form models. Many proxies of social capital (e.g. social engagement, socialization, social support and social-psychological constructs) have been correlated with self-reported health, but not all of them turned out to be significant after controlling for individual characteristics, mainly income and educational level (see e.g. Veenstra, 2000, using a cross-sectional sample of around 500 people in Sweden).

However, several other studies do find a positive and significant correlation between self reported health (or EQ-5D) and self reported trust. In particular, Petrou and Kupek (2008), using the Health Survey of England and Fujisawa et al. (2009), using data from Japan, focus on the differential impact of individual social capital vs. community level social capital. They found that while the former has a positive effect on health, the influence of the latter turns out to be statistically insignificant. The finding is confirmed by Hurtado et al. (2011), using a representative sample of the Colombian population[2]. By contrast, according to Snelgrove et al. (2009) and Iversen (2008) both the individual and the communitarian levels matter for health. Snelgrove et al (2009) tested

2 Individual social capital is measured both by cognitive and by structural social capital

the hypothesis in the UK using the British Household Panel Survey (BHPS) data, while Iversen (2008) used Norwegian data. Finally, the analysis by Engrstrom et al. (2008) in Sweden shows that none of the levels of social capital was strongly related to health. A similar result is obtained by Ferlander and Makinen (2009), who find that civic participation and social contacts explain health for Russian males but not for women.

Some authors examine the relationship using multi-country datasets, as do Mansyur et al, (2008) who analyse the World Value Survey, and Poortinga (2006), who focuses on European countries using the European Social Survey. The multilevel models used in both studies suggest that only individual social capital is important in explaining health.

The main limitation of the studies presented so far lies in their failure to explicitly address the problem of causality. A small number of recent studies have tried to fill this gap. As for the causal link from social capital to health, Folland (2007), Ronconi et al. (2010) and D'Hombres et al. (2010) cannot reject the hypothesis of a causal impact of social capital on health after a careful analysis, using data from different sets of countries. They all share the feature of employing an instrumental variable strategy: Folland (2007), by exploiting cross-state variability in the US, instruments the social capital variable with employment rate, latitude and the state government contribution to colleges per capita; Ronconi et al (2010) use information about the access to public transportation and D'Hombres et al. (2010) use indicators of community heterogeneity in terms of schooling, health and religion.

Finally, Sirven and Debrand (2011) identify the circular relationship between social capital and health among elderly Europeans covered in the Survey of Health, Aging and Retirement in Europe (SHARE) by exploiting the information about social capital and health in early individual life. Other papers also recognize the potential existence of a circular relationship between the two variables but do not attempt a formal evaluations (see e.g. von dem Knesebeck et al.2005; Islam et al, 2006).

3. Empirical modeling strategy

We follow two complementary modeling strategies: our baseline model consists of two simultaneous equations allowing for a general feedback relationship between social capital and health. In order to examine whether the results from this model are robust to the potential bias resulting from mis-reporting in social capital and health, our second estimation strategy adds specifications of the reporting bias to baseline model.

3.1 Baseline model

A straightforward, natural way to describe a circular relationship between social capital and health is defining a simultaneous equation model where each side of the relationship is described by one equation. We set up the following two equations model, broadly following Stern (1989):

$$h^{**} = \alpha_0 + \alpha_1 SC^{**} + \alpha_2 CSC^{**} + \alpha_3 Z_h + X\alpha_4 + \varepsilon_h \quad (1)$$

$$SC^{**} = \beta_0 + \beta_1 h^{**} + \beta_2 CSC^{**} + \beta_3 Z_{SC} + X\beta_4 + \varepsilon_S \quad (2)$$

In equation (1) individual health h^{**} is made dependent on individual social capital (SC^{**}) and community social capital (CSC^{**}) and on two separate sets of controls, one specific to equation (1), Z_h and one, X common to both equations. In equation (2) individual social capital SC^{**} is explained by individual health h^{**}, community social capital CSC^{**}, a specific (Z_{SC}) and a common (*X)* set of controls. Finally, ε_h and ε_S are i.i.d. error terms.

The model contains three endogenous variables: h^{**}, SC^{**} and CSC^{**}. The double asterisks indicate that they are continuous latent variables, not directly observed, but truly reported and correctly measured. Rather than the latent variables, we observe the categorical variables h, S and $\overline{S}$. The relationship between latent and observed variables is described by the following equations:

$$h = i \text{ if } \Xi_{i-1} < h^{**} \le \Xi_i,\ i = 1,2..,5 \quad (3)$$

$$S = i \text{ if } \Sigma_{i-1} < SC^{**} \le \Sigma_i,\ i = 0,1..,10 \quad (4)$$

$$\overline{S} = i \text{ if } \overline{\Sigma}_{i-1} < CSC^{**} \le \overline{\Sigma}_i,\ i = 0,1..,10 \quad (5)$$

where Ξ_i, Σ_i and $\overline{\Sigma}_i$ are threshold parameters with Ξ_0, Σ_{-1} and $\overline{\Sigma}_{-1}$ equal to $-\infty$ and Ξ_5, Σ_{10} and $\overline{\Sigma}_{10}$ equal to $+\infty$.

For each individual, community SC is defined as the average individual SC of her reference group. The individual's reference group is composed of the pool of people more likely to be related to her. While it is easy to imagine that people are more likely to relate with peers of similar age who live nearby, the precise definition of a reference group is ultimately arbitrary. We therefore examine the robustness of our results against a range of different definitions of the "reference group", as discussed further below[3]. In each definition we follow the reference groups are partially overlapped as in De Giorgi et al. (2010), implying that the reference group of a member of individual *i*'s reference group does not coincide with individual *i*'s reference group. Not only does this render the notion of reference group more credible, as it is tailored to the individuals' specific

[3] We anticipate that our results are very stable across alternative definitions of the reference group.

characteristics, but in addition it alleviates some of the identification problems raised in Durlauf (2002) and in Acemoglu and Angrist (2001), whose concerns and results, developed for fully overlapping reference groups do not apply in our context.

Following Stern (1989) and Maddala (1983), we estimate the theoretical model using a reduced form two-stage procedure.

We first derive the reduced form of the system (1)-(2) as:

$$h^{**} = W\Pi_h + \mu_h \tag{6}$$

$$SC^{**} = W\Pi_{SC} + \mu_{SC} \tag{7}$$

$$CSC^{**} = W\Pi_{CSC} + \mu_{CSC} \tag{8}$$

where $W = [Z_h, Z_{SC}, Z_{CSC}\ X]$ and Π_j , $j = h, SC, CSC$ are the reduced form parameters. Variable Z_{CSC} is an external-to-the-model exogenous variable required to identify the effect of CSC^{**}. The use of external information is necessary since *CSC*** is included in both equations.

In equations (6) to (8) the three endogenous variables depend on all the exogenous ones. They express the equilibrium levels of health, individual and community SC. Parameters Π_j indicate the contribution of each exogenous variable to such equilibrium levels. Taking into account equations (3) to (5), equations (6) to (8) are separately estimated by means of ordered probit. The predicted values $\hat{h}^{**}$, $\hat{SC}^{**}$ and $\hat{CSC}^{**}$ of the latent variables obtained in this first stage are replaced in equations (1)-(2) which are next estimated in the second stage by ordered probit as well. Standard errors are estimated by bootstrapping the entire procedure, given that the usual 2SLS standard error correction does not apply to non-linear models.

An identification not depending only on the functional form, but more substantial of the structural parameters of equations (1)-(2) depends on the exclusion restrictions and the fact that we have supplemented Z_{CSC} to the reduced forms. The crucial non-testable assumptions are that excluded variables Z_h have no autonomous effect on *SC*** , Z_{SC} have no autonomous effect on *h*** and Z_{CSC} have no autonomous effect on both *SC*** and *h***. Conversely, Z_h must be relevant to equation (1) and, symmetrically, Z_{SC} must be relevant in equation (2). These conditions are equivalent to the IV conditions of excludability and relevance of the instruments in the single-equation IV models (see Wooldridge, 2002, chapter 9). If the conditions hold, this two-stage procedure will produce consistent estimates.

3.2 Endogenous variables, inclusions and exclusions

Individual social capital is measured by an indicator extensively used in the literature since Putnam (1993), i.e.

the individual degree of generalized trust. The wording of the question posed in the ESS, very similar to the one 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 likert-type scale ranging from 0 to 10. Compared to other surveys, which often include generalized trust measures on a five-point scale, the ESS measure is considerably finer, allowing us to exploit a greater degree of variability. Community social capital associated with individual *i* is measured as the average generalized trust of *i*'s reference group (see section 5 for details).

Regarding health, people are asked to rate their current health on a five-step scale ranging from very bad (1) to very good (5). Although self-reported health is of course a noisy measure of true health, it has proven a remarkably reliable health proxy. In particular, it has been shown that self-reported health is highly correlated with subsequent mortality at the individual level (Ferraro, 1999).

We condition the estimates of both equations (1)-(2) on a set of controls X (see Tables 2A for the complete list):

1) *Parental background indicators*: respondent father's and mother's educational attainment, father's and mother's employment status when respondent was 14, whether either father or mother died before respondent was 14 year-olds, whether either father or mother were born in the country of the respondent's residence.

2) *Individual characteristics*: gender, age, age squared, years of education, type of occupation, religion, respondent's household income level, whether the respondent was born in the country of current residence, respondent's marital status, place of residence (urban/rural)

3) *Reference group controls*: average years of education, average household income, proportion of males

4) *Regional controls*: working age (15-64) activity rate and employment rate, young adults (25-34) activity rate and employment rate, youth (15-24) unemployment rate, proportion of college graduates, GDP per capita, GDP growth rate, age structure of the population, population density, kilometers of motorways, proportion of households without internet access, proportion of immigrant residents, proportion of citizens out of total residents and, finally, a measure of the level of crime[4].

We include these controls to capture possible confounders of the effect of individual and community SC. Parental background has long been recognized as being closely correlated with individual unobserved ability and preferences, which in turn influence the individual's perception and behaviors. Furthermore, education and type of occupation have an obvious influence on both SC and health[5]. In addition, both reference group and regional characteristics are included to account for possible confounders of CSC. For instance, CSC could be lower in more sparsely populated regions, due to higher mobility costs of meeting friends and fellows. Simultaneously,

[4] More precisely we use the proportion of residents who report having been victim of a burglary or an assault in the past five years. This choice of crime measure is important for our identification strategy, as we shall explain further below..

[5] Education and occupation are choices and typically correlated with unobservable characteristics (e.g. ability). The inclusion of extensive family background controls increases our confidence about the unbiasedness of the estimated parameters.

health could be better in sparsely populated areas because they are typically rural, less industrialized and less polluted.

Variable Z_h included in equation (1) is the number of doctors per 1000 inhabitants, a measure of health-care supply at the regional level (NUTS 2). The relevance of this variable rests on the fact that health care services are inputs into individual health. Symmetrically, conditional on the common set of controls X, and specifically on the population age structure which captures residents' average health, the supply of doctors is unlikely to have a direct impact on individual social capital. Further, we neutralize the possibility that the quality of the health care system is indirectly correlated with individual social capital via its correlation with more general good public governance (which could facilitate social cohesion) by conditioning on individual, peer and regional characteristics (e.g. individual and reference group education, regional employment rate, development and economic cycle). These considerations motivate the exclusion of Z_h from equation (2).

Variable Z_{SC} included in equation (2) is a measure of crime victimization. Respondents are asked whether they have been victim of a burglary or an assault in the past five years, i.e. petty crimes, which hardly have any lasting medium or longer term direct physical or mental health consequences. Conditional on regional characteristics, regional crime level, individual wealth, education, gender and age, having being victim of a burglary can be considered a random event outside individual control which has no direct effect on individual health. In addition, variable Z_{SC} is not an individual perception, possibly correlated with individual characteristics, but an actual experience which leads individuals to be more skeptical and less trusting vis-à-vis the rest of the society, at least with respect to those outside the inner circle of close relatives and friends.

Finally, variable Z_{CSC} is the proportion of members of the reference group who have been victim of a burglary or an assault, i.e. it is the average of Z_{SC} at the reference group level. The choice of using a transformation of Z_{SC} to identify the effect of *CSC*** is motivated by the relation between *SC*** and *CSC***. Given that we are controlling for the overall crime rate in the region of residence (at the level NUTS 2), Z_{CSC} can be excluded in both equations (1) and (2).

3.3 Extended Model

The model above is based on the strong assumption that we observe true, objective, measures of both social capital and health. Unfortunately, our measures are self-reported and thus possibly distorted. To account for mis-reporting we add two further equations to the baseline model. For the readers' convenience we report the complete extended model below. The model is composed of equations (9)-(10) – i.e. the baseline model – and equations (11) and (12), which propose a simple representation of mis-reporting in SC and health, respectively.

$$h^{**} = \alpha_0 + \alpha_1 SC^{**} + \alpha_2 CSC^{**} + \alpha_3 Z_h + X\alpha_4 + \varepsilon_h \quad (9)$$

$$SC^{**} = \beta_0 + \beta_1 h^{**} + \beta_2 CSC^{**} + \beta_3 Z_S + X\beta_4 + \varepsilon_S \quad (10)$$

$$SC^{*} = SC^{**} + (1-\lambda)(CSC^{*} - SC^{**}) + \gamma h^{**} + \theta_S \quad (11)$$

$$h^{*} = h^{**} + \theta_h \quad (12)$$

Reported individual social capital SC^{*} is a latent variable equal to truthfully reported individual social capital SC^{**} plus a bias which depends on 1) the average level of social capital *reported* by the reference group, 2) the true individual health conditions and 3) a random error possibly correlated across members of the same reference group. In other words, we allow reports of individuals belonging to a given reference group to be strategic complements. If peers report a level of social capital higher or lower than respondent's true SC, the respondent will tend to conform to their reporting. This specification of the mis-reporting accounts for the possibility that individuals "anchor" their situation to the average value reported by the members of their reference group (see Winkelmann and Winkelmann, 1998 and Senik, 2004)

Our specification can also be viewed from a different perspective, by rewriting (11) as

$$SC^{*} = \lambda SC^{**} + (1-\lambda) CSC^{*} + \gamma h^{**} + \theta_S \quad (13)$$

Here SC^{*} is (partly) defined as a weighted average of SC^{**} and CSC^{*} with weight equal to λ. This might be motivated by the desire of the respondent to appear more society-oriented than she actually is, or to conform to the answers of her reference group in front of the interviewer. Conformism could be the outcome of the so-called *"social desirability bias",* i.e. the tendency of individuals to deny socially undesirable traits and behaviors and to profess socially desirable ones (Randall and Fernandes, 1992), where social desirability is reflected in social behavior. In addition, individual health conditions are also likely to affect the way social capital is reported, though the sign of this influence is difficult to predict. Sick people could report a lower level of social capital because their transitory illness temporarily alters their perception of social capital. Alternatively, they could report a higher level of social capital because they may be experiencing the support of relatives, friends and neighbors[6].

Finally, equation (12) simply assumes that reported health is equal to true health plus a random independently distributed error term. Note however that individual variables, such as age, gender, education, etc. (i.e. those included in the set X) could easily be included in (12) without affecting the estimate and the identification of the

[6] It is difficult to develop an unambiguous hypothesis on the sign and magnitude of the effects of the variables assumed to influence SC reporting. However, the specification is written in a "neutral way", in the sense that if reporting bias does not exist (i.e. λ and γ both equal to zero), reported SC equals true SC. No artificial and systematic bias is introduced.

structural parameters, provided that they are orthogonal to θ_h.

By averaging (11) over the members of individual's reference group we (approximately) obtain

$$CSC^{*} = CSC^{**} + (1-\lambda)(CSC^{*} - CSC^{**}) + \gamma \overline{h}^{**} + \overline{\theta}_S \tag{14}$$

(see Appendix A for details) where $\overline{h}^{**}$ is the average health status of the reference group and $\overline{\theta}_S$ is the average error term which can interpreted as an unobservable group effect. By solving (14) for CSC^{**} we obtain

$$CSC^{**} = CSC^{*} - \frac{\gamma}{\lambda}\overline{h}^{**} - \frac{\overline{\theta}_S}{\lambda} \tag{15}$$

From (11) SC^{**} can be written as

$$SC^{**} = \frac{1}{\lambda}SC^{*} - \frac{1-\lambda}{\lambda}CSC^{*} - \frac{\gamma}{\lambda}h^{**} - \frac{\theta_S}{\lambda} \tag{16}$$

Finally, by substituting first (15) and (16) and next (12) in both (9) and (10) we obtain a specification of the extended model expressed exclusively in terms of the observed variables.

Precisely, the health equation now becomes:

$$h^{*} = \frac{\alpha_0}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)} + \frac{\frac{\alpha_1}{\lambda}}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)}SC^{*} + \frac{\left(\alpha_2 - \alpha_1 \frac{1-\lambda}{\lambda}\right)}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)}CSC^{*} + \frac{-\alpha_2 \frac{\gamma}{\lambda}}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)}\overline{h}^{*} +$$

$$+ Z_h \frac{\alpha_3}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)} + X\frac{\alpha_4}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)} + \left[\theta_h + \frac{\varepsilon_h - \alpha_1 \frac{\theta_S}{\lambda} - \alpha_2 \frac{\overline{\theta}_S}{\lambda}}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)}\right] \tag{17}$$

and the social capital equation can be written as:

$$SC^{*} = \lambda\beta_0 + \lambda\left(\beta_1 + \frac{\gamma}{\lambda}\right)h^{*} + \left(1-\lambda+\lambda\beta_2\right)CSC^{*} - \gamma\beta_2\overline{h}^{*} +$$

$$+ Z_S\lambda\beta_3 + X\lambda\beta_4 + \left[\theta_S - \theta_h + \beta_2\overline{\theta}_S + \lambda\varepsilon_S\right] \tag{18}$$

where we used the fact that $\overline{\theta}_h = 0$ since θ_h is iid.

Note that, compared to equations (1)-(2), equations (17) and (18) need to include the average health of the reference group. As $\theta_S \perp h^{**}$, and since each individual contribution to the reference group averages is negligible, $\overline{h}^{*}$ can be considered exogenous.

By considering the counterparts of equations (3)-(5), defined over the one-asterisk-latent variables, we shall estimate the system (17)-(18) by means of the procedure described for the baseline model.

Importantly, all the structural parameters associated with the endogenous variables of the system (11)-(14) can be recovered by estimating the system (17)-(18), as shown in Appendix B.

4. Data and descriptive statistics

We use the first four waves of the European Social Survey (ESS), a repeated cross-country survey which covers many European countries[7]. The ESS provides information on individual social behavior and perception, such as political opinions, political participation, exposure to media and news, social relationships, trust in other people and institutions. In addition, the 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 general health status and whether they are hampered in daily activities by illness or disability. More than 130,000 people answered the questionnaire, equally shared between the four rounds collected in 2002/03, 2004/05, 2006/7 and 2008/9. Each round of the survey defines a representative sample for each of the European countries covered. On average about two thousand residents are interviewed in each country per round (see Table 1). The full list of variables used and simple summary statistics are reported in tables 2A and 2B.

[TABLE 1 ABOUT HERE]

[TABLE 2A ABOUT HERE]

[TABLE 2B ABOUT HERE]

Information on the region of residence, at NUTS2-level (Nomenclature of Territorial Units for Statistics)[8] in most cases, is also available. This feature has allowed us to supplement additional data about regional characteristics from the Eurostat REGIO dataset. In particular, we have added regional indicators of development (GDP per capita, GDP growth and employment), 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 graduates.

The variables of primary interest are health on the one hand and individual and community social capital on the other hand. Regarding health, people are asked to rate their current health on a five-step scale ranging from very bad (1) to very good (5). Individual social capital is captured by the individual degree of generalized trust. The

---

7 Our analysis is based on data from the following countries: Austria, Belgium, Bulgaria, Switzerland, Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, Great Britain, Greece, Croatia, Hungary, Ireland, Italy, Latvia, Netherland, Norway, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia

8 The European Union (EU) has developed a geocode standard for referencing the subdivisions of countries for statistical purposes. For each EU member country, a hierarchy of three NUTS levels is established; the subdivisions in some levels do not necessarily correspond to administrative divisions in a country. NUTS1 defines major socioeconomic regions; NUTS2 captures basic regions for the application of regional policies; NUTS3 relates to small regions for specific diagnoses.

question posed in the ESS, very similar to the original one included in the World Values Survey, 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 likert-type scale ranging from 0 to 10. We have recoded this variable to allow it to take values between -5 and 5. The community social capital associated with individual *i* is measured as the mean trust of his/her reference group. Average measures of individual social capital are frequently used in the literature (see Islam et al., 2006). However, the definition of what is the "right" community or the relevant reference group is debatable. In our analysis, the reference group is defined within the region of residence (coded as NUTS 2), which is the most disaggregated regional level at which data is available in the ESS for most countries. In order to examine the robustness of our estimates to different definitions of "reference group", we used three different definitions:

Definition 0 (baseline definition): the reference group of individual *i* includes all people living in the same region of *i* (defined according to the European classification NUTS 2) and who are at most 10 years older or 10 year younger than *i*.

Definition 1: members of the reference group of individual *i* are the residents of the same region as individual *i*, whose age belongs to the interval $[age_i – (2+0.2*age_i),\ age_i + (2+0.2*age_i)]$.

Definition 1 adopts the plausible notion that younger people tend to interact more with people of similar age, compared to older people. For instance, according to this definition the reference group of a 20 year-old individual is composed of people aged between 14 and 26, while the reference group of a 60 year-old person includes people aged between 46 and 74. Both Definition 0 and Definition 1 imply a sharp discontinuity at the age boundary: only people within the age interval matter, while those just outside the interval do not. In reality, we cannot exclude that all residents are part of an individual's reference group, although age differences certainly matter. A less restrictive definition, taken up in Definition 2, assigns a weight to all residents, related to the age distance with individual *i*:

Definition 2: members of the reference group of individual *i* are all the residents of the same region as individual *i* . We assign a specific weight to each member of the reference group: a value of 1 is assigned to people whose age differs from *i*'s age by less than three years and a smaller weight, decreasing at increasing rates with age difference, is assigned to those aged 3 older or younger than *i*.

Note that we have always defined reference groups based only on age, an exogenous individual characteristic. This is admittedly a rather crude approach, but it has been motivated by the need to avoid the problem of self-sorting of individuals into reference groups (e.g. trustworthy peers self-sort into the reference groups of a trustworthy individual and vice-versa).

5. Results

Baseline model estimates based on reference group Definition 0 are displayed in Table 3. Only the coefficients of endogenous variables and exclusions Z_{SC}, Z_{CSC} and Z_h have been reported[9]. First-stage estimates, i.e. those corresponding to the reduced form equations (6)-(8) are reported in columns 1 to 3, in this order. Second stage estimates of the SC equation (2) and the health equation (1) are reported in columns 4 and 5, respectively.

[TABLE 3 ABOUT HERE]

The reduced form equations (6)-(8) are the "solutions" of the simultaneous system (1)-(2), hence they represent the equilibrium levels of individual health, SC and CSC. Thus their parameters capture the total effect of the exogenous on the endogenous variables in equilibrium.

Instead, if taken alone, structural equation (1) describes how SC, CSC and the exogenous variables influence health, by treating SC and CSC as if they were not jointly determined with health. Symmetrically, if taken in isolation, structural equation (2) indicates how health, CSC and the other explanatory variables influence SC, by treating health and CSC as if they were not jointly determined with SC. Therefore, their parameters capture partial effects, which is what we are ultimately looking for, because they allow us to understand the complex interaction between SC and health, and how other exogenous variables enter this process.

The estimates of the structural equations (columns 4 and 5) indicate that there exists a circular relationship between individual SC and health: higher individual SC improves health conditions and better health is in turn beneficial to individual SC. Moreover, CSC has a positive influence on both SC and health, though only marginally significant in the second case. The size of the coefficients of individual and community SC can be meaningfully compared, as they are measured on a common scale: the beneficial impact of individual SC on health turns out to be at least an order of magnitude larger.

Exclusion restrictions are always highly statistically significant and have the expected sign. Increasing the number of doctors is beneficial to individual health (column 5). Importantly, its impact on health is positive and statistically significant also at equilibrium (column 1), and, given the reinforcing two-way relationship between SC and health, it is positively correlated with the equilibrium value of individual SC (column 2). By averaging equation (6) and (7) at the reference group level and by subtracting the latter from the former, we obtain an equilibrium equation for CSC which allows us to predict the sign of the effect of the exogenous variables. After solving for CSC, the resulting sign of the coefficient on Z_h is negative as it clearly results from the estimate of

[9] Complete estimate tables are available upon request.

the number of doctors reported in column (3). Having been victim of a crime (variable Z_{SC}) significantly reduces individual SC (column 4) and the same effect prevails at equilibrium for both individual SC and health (columns 1 and 2). Finally, the proportion of crime victims in the reference group significantly reduces CSC (column 3) as expected.

As a first robustness check we have tested whether results qualitatively change by adopting alternative definitions of the reference group. The stability of the point estimates across alternative definitions of the reference groups is remarkable, as it emerges from Table and Table 5, where we have adopted Definition 1 and 2 respectively. This indicates that the evidence of a circular relationship between SC and health is robust to the way reference groups are defined, thus reinforcing our confidence in this result.

[TABLE 4 ABOUT HERE]

[TABLE 5 ABOUT HERE]

As a further robustness check we have tested for the presence of unobserved regional factors by estimating one side of the relationship (i.e. the effect of SC on health) by including regional fixed effects. (We can estimate only this side of the relationship because Z_h varies at regional level only very slowly over time.) Table reports the results of this exercise adopting our baseline reference group definition (i.e. Definition 0). The first two columns are the first stage estimates for SC and CSC and the third column displays the second stage estimates. Estimates are very similar to those in Table 3, indicating that our controls already capture most of the relevant heterogeneity across regions.

[TABLE 6 ABOUT HERE]

Table 7 reports estimates of the extended model that explicitly accounts for mis-reporting and serves as a robustness check of the findings from the baseline model. The average health status of the reference group is included in both individual SC and health equations and in all the first stage equations. This is the only variation from the baseline model that has been dictated by the specification of equation (11), which includes health as an explanatory variable. Unsurprisingly, since the model estimated in Table 7 is very close to that of Table 3, point estimates are similar.

[TABLE 7 ABOUT HERE]

However, now the estimated coefficients have to be understood as combinations of the structural parameters and cannot be directly compared to the results of the baseline model. Structural parameters have been derived by using the formulas reported in Appendix B and have been included in Table 8.

Noteworthy results emerge when we compare the structural parameters of the extended model to those of the baseline model (reported in Table 8 as well for immediate reference): first, there is evidence of mis-reporting, as the weight of reported CSC on reported individual SC ranges between 0.194 and 0.352. Second, the main result of the baseline model, i.e. the circular relation between SC and health, is confirmed, as well as the finding that individual SC contributes far more to individual health than does CSC. However, the size of the individual social capital effect is considerably lower (i.e. by about one third) in the extended model, while the effect of CSC is broadly unchanged and it is now very precisely estimated. Third, looking at the other direction of the relationship, health contributes considerably more to individual SC according to the extended model than according to the baseline one, thus reinforcing the indication that SC and health are part of a feedback relationship that works strongly in both directions.

[TABLE 8 ABOUT HERE]

Somewhat surprisingly, the results also suggest that people in worse health conditions tend to report a higher SC, and, unlike in the baseline model, the link between true CSC and true individual SC is insignificant in two out of three specifications. These findings should not be over-interpreted, as they ultimately depend on how equation (11) is specified. Alternative specifications could have led to different conclusions. We prefer to emphasize the role of robustness check of this exercise, which broadly confirms the main findings from the baseline model.

6. Conclusions

In this paper we have examined the relationship between social capital, measured by self reported generalized trust, and individual health in a large sample of European countries, using four cross-sectional rounds of the European Social Survey. To the best of our knowledge this is the first paper that explicitly explores and models the potential simultaneous relationship between social capital and health. In addition, this should be the first paper that explicitly takes into account the potential bias resulting from mis-reporting in social capital.

We find strong evidence for a circular, mutually reinforcing relationship between social capital and health. The results from our extended model show that reporting bias in social capital does exist. Yet, when mis-reporting is explicitly taken into account, the basic findings from the baseline model still hold. A second implication is that individual social capital is far more important than community social capital as a determinant of health. This is consistent with the Guiso et al. (2008) definition of social capital that emphasizes the importance of the individual level dimension of the concept, focusing on values, norms and beliefs. Although their degree of acceptance in society is relevant, only individuals who share and participate in such “good culture” could benefit from it.

On the basis of these results, it would thus be worthwhile exploring the concrete ways in which social capital can be promoted, as one, yet largely untapped means by which health can be improved, on top of the various other benefits social capital may entail, as mentioned in the introduction.

Our analysis suffers from three limitations, partly due to the constraints imposed by a non-linear multiple equation model and partly due to data availability.

First, the influence of social capital on health (and vice-versa) may in principle differ between countries. Given large sample size needed to achieve the required precision of the estimates, it was not possible to run the regressions by country.

Second, a more adequate specification of the model would include also an interaction between individual and community SC, to explicitly test the hypothesis of complementarity between the two. Unfortunately, in light of the already complex model structure, and especially given that social capital at both levels is endogenous, it would not be feasible to control for such interaction effect and achieve identification.

Third, our measure of community social capital is the average of the individual social capital in the relevant community. While this is a very common way of measuring aggregate level social capital in the literature, it would have been useful to check the robustness of our results by using other measures of community level social capital such as blood and organ donations. Unfortunately, this information, not available in the ESS, could not be imputed from other sources for each possible reference group.

Admittedly, even if our results suggest that social capital is good for health (and vice versa), by themselves they provide only limited direct policy implications, beyond the recommendation that policymakers should at least consider social capital, alongside several other factors, when developing health policy. There remains significant scope for improving our understanding of how social capital can be promoted. Szreter (2004) has studied what factors were important in building social capital and, as a consequence, improving health in Britain during the industrial revolution. He identifies three important developments. The first is a rejection of policies that concentrate wealth in the hands of a few, instead promoting what is termed ‘co-production’, in which the state, small and medium enterprises, and bodies representing residents and workers collaborate in pursuit of the common goal of sustainable economic growth. The second is a mechanism by which civic society can articulate its concerns and the political structures can respond to them. The third, which underpins the first two, is political participation. Clearly, all of these developments involve actions far beyond the health sector and should therefore complement more traditional health care investment.

There is also some evidence that much can be achieved more locally through programs designed to increase social interaction. Thus, a program in three neighborhoods in Portland, Oregon, involving measures to promote community participation in urban renewal, achieved improvements in mental health (Semenza et al., 2007). Another before and after study of a traffic calming scheme, which led to greater pedestrian activity, was associated with better physical health (Morrison et al., 2004). A recent systematic review found some evidence of gains in mental health linked to interventions giving employees' greater control in organizations (Egan et al., 2007).

If our conclusion is true that promoting individual level social capital has potentially more significant health effects than can be reaped from targeting community level social capital, then interventions may want to target the former rather than the latter. Examples of individual level social capital interventions include the provision of support for families and parenting (Halpern 2005), the be-friending approach described in the introduction, as well as similar mentoring approaches - with a view towards giving in particular young and disadvantaged people access to bridging social capital. These approaches have been applied to some extent and with some success in the context of education (see e.g. Johnson, 1999; Kahne and Bailey, 1999). Alternative approaches may include the promotion of volunteering, in light of the evidence that those who begin volunteering in youth tend to persist with it later in life, entailing positive longer term effects on social and political participation and pro-social attitudes (Halpern et al. 2002). Some evidence from the US with an experience year for young people (i.e. the AmeriCorps Scheme) has shown that members became more active in community groups and showed significant value shifts over the observation period, and that the program helped solve unmet human and community needs (Simon and Wang 2002).

While the above mentioned programs may well have resulted in positive health outcomes as well, this has typically not been evaluated. Moreover, even if some examples do exist as to how to promote social capital, a full economic evaluation of social capital interventions would still require the assessment of the associated costs and exact health (and possibly non-health) benefits.

References

Acemoglu D. and Angrist J (2001) "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws" in Ben S. Bernanke and Kenneth Rogoff, editors "NBER Macroeconomics Annual 2000", Volume 15, MIT Press

Alesina, A. and La Ferrara, E. (2002). "Who trusts others?," Journal of Public Economics, Elsevier, vol. 85(2), pages 207-234, August

Akcomak I.S. and ter Weel, B (2009) "Social capital, innovation and growth: Evidence from Europe". European Economic Review. Volume 53, Issue 5, July 2009, Pages 544-567

Berkman, L.F. and Glass, T. (2000). "Social integration, social networks, social support, and health". In L.F. Berkman and I. Kawachi (Eds.), "Social epidemiology". (pp 137-173). New York: Oxford University Press.

Brown T, Sheffler R, Seo S and Reed, M. (2006) "The empirical relationship between community social capital and the demand of cigarettes". Health Economics 15, 1159-1172

Cooper, H., Arber, S., Fee, L. and Ginn, J. (1999). "The Influence of Social Support and Social Capital on Health: a review and analysis of British data". London: Health Education Authority.

De Giorgi, G.; Pellizzari, M.; Redaelli, S. (2010) "Identification of Social Interactions through Partially Overlapping Peer Groups" American Economic Journal: Applied Economics, Volume 2(2), pp. 241-275(35)

D'Hombres B., Rocco L, Suhrcke M, and McKee M, (2010). "Does social capital determine health? Evidence from eight transition countries," Health Economics, vol. 19(1), pages 56-74.

Durlauf, .S.N. (2002). "On The Empirics Of Social Capital". The Economic Journal, 112, F459-F479.

Durlauf, S N. and Fafchamps, M, (2005). "Social Capital," Handbook of Economic Growth, in: Philippe Aghion and Steven Durlauf (ed.), Handbook of Economic Growth, edition 1, volume 1, chapter 26, pages 1639-1699 Elsevier.

Engström K, Mattsson F, Järleborg A, and Hallqvist J. (2008) "Contextual social capital as a risk factor for poor self-rated health: a multilevel analysis". Social Science and Medicine, 66:2268-80.

Egan M, Bambra C, Thomas S, Petticrew M, Whitehead M, and Thomson H. (2007). “The psychosocial and health effects of workplace reorganisation. A systematic review of organisational-level interventions that aim to increase employee control”. Journal of Epidemiology and Community Health 6111: 945–954

Fehr, E. and Fischbacher, U. (2003) “The nature of human altruism”, Nature 425, 785 - 791

Ferlander S and Makinen I (2009) “Social capital, gender and self rated health. Evidence from the Moscow health survey 2004”. Social Science and Medicine, 69 1323-1332

Folland, S (2007), "Does 'Community Social Capital' Contribute to Population Health?", Social Science and Medicine, Vol. 64, pp. 2342-2354.

Fujisawa Y, Hamano T, Takegawa S. (2009) “Social capital and perceived health in Japan: an ecological and multilevel analysis”. Social Science and Medicine. 69(4):500-505

Ghosh, P. and Ray, D. (1996). "Cooperation in Community Interaction without Information Flows," Review of Economic Studies, vol. 63(3), pages 491-519.

Glaeser E, Laibson, D. I., Scheinkman J.A. and Soutter, C. L. (2000). "Measuring Trust," The Quarterly Journal of Economics, vol. 115(3), pages 811-846

Glaeser E,. Laibson D and Sacerdote B, (2002). "An Economic Approach to Social Capital," Economic Journal, vol. 112(483), pages 437-458

Goldin, C., and Katz L (2001). “Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910–1940.” In R. Rotberg (edited by) “Patterns of Social Capital”,. Cambridge University Press.

Guiso L; Sapienza P, Zingales L (2004) “The Role of Social Capital in Financial Development” The American Economic Review, 94(3):526-556

Guiso L; Sapienza P, Zingales L (2008). "Alfred Marshall Lecture Social Capital as Good Culture," Journal of the European Economic Association, vol. 6(2-3), pages 295-320

Guiso L; Sapienza P, Zingales L (2010) “Civic Capital As The Missing Link”, NBER Working Paper 15845

Halpern DS (2005) “Social capital”. Cambridge: Polity.

Halpern D.S., John P., Morris Z. (2002). “Before the citizenship order: a survey of citizenship practice in England”. Journal of Education Policy, 17(2): 217-28.

Hurtado D, Kawaki I and Sudarsky J (2011) “Social capital and self-rated health in Colombia: The good, the bad and the ugly” Social Science and Medicine 72:584-590

Islam MK, Merlo J, Kawachi I, Lindström M, and Gerdtham U (2006) “Social capital and health: Does egalitarianism matter? A literature review”. International Journal for Equity in Health.; 5: 3.

Iversen, T (2008). "An exploratory study of associations between social capital and self-assessed health in Norway," Health Economics, Policy and Law, vol. 3(04), pages 349-364

Johnson A.W. (1999). “Sponsor-a-scholar: long-term impacts of a youth mentoring program on student performance”. London: Commonwealth Fund.

Kahne J, and Bailey K (1999). “The role of social capital in youth development: the case of ‘I have a dream’ programs”. Educational Evaluation and Policy Analysis, 21(3):321-43.

Kawachi, I. and Berkman, L. (2001). “Social ties and mental health. Journal of Urban Health” Bulletin of the New York Academy of Medicine. 78 (3), 458-467.

Knack, S and Keefer, P. (1997). "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," The Quarterly Journal of Economics, vol. 112(4), pages 1251-88.

Zak, P.J.and Knack, S. (2001), “Trust and Growth”, The Economic Journal, 111(479):295-321

Kranton, R.E, (1996). "The Formation of Cooperative Relationships," Journal of Law, Economics and Organization, vol. 12(1), pages 214-33.

Lochner K, Kawachi I, and Kennedy, B.P. (1999) “Social capital: a guide to its measurement”. Health Place 5(4):259-70.

Macinko, J. and Starfield, B. (2001). “The utility of social capital in research on health determinants”. The Milbank Quarterly 79: 387-427.

Maddala, G.S. (1983) “Limited-Dependent and Qualitative Variables in Econometrics”. Cambridge: Cambridge University Press.

Mansyur, C., Amick, B.C., Harrist, R.B., and Franzini, L. (2008). “Social capital, income inequality, and self-rated health in 45 countries”, Social Science and Medicine 66: 43–56

Mead, N., Lester, H., Chew-Graham, C., Gask, L. and Bower, P. (2010) “Effects of befriending on depressive symptoms and distress: systematic review and meta-analysis”. British Journal of Psychiatry 196, 96–101

Morrison D, H. Thomson, and M. Petticrew (2004) “Evaluation of the health effects of a neighbourhood traffic calming scheme” Journal of Epidemiology and Community Health;58:837–840.

Petrou S and Kupek E. (2008) “Social capital and its relationship with measures of health status: evidence from the health survey for England 2003”. Health Economics 17, 127-143

Poortinga, W. (2006). “Social capital: An individual or collective resource for health?”, Social Science and Medicine 62, 292-302.

Putnam, R.D. (1993). “Making democracy work: civic traditions in modern Italy”. Princeton NJ: Princeton University Press.

Putnam, R.D. (2000). “Bowling along. The collapse and revival of American community”. London: Simon & Schuster.

Randall, D. M. and Fernandes, M.F.: (1992), “Social Desirability Bias in Ethics Research”, Business Ethics Quarterly 2(2), 183–205.

Ronconi, L., Brown, T., Scheffler, R. (2011). “Social capital and self-rated health in Argentina”. Health Economics, (forthcoming)

Scheffler, R.M., and Brown T.T.(2008) “Social capital, Economics, and health: New evidence”, Health Economics, Policy and Law ; 3(4); 321-331.

Senik, C., (2004). “When information dominates comparison – Learning from Russian subjective panel data”. Journal of Public Economics 88, 2099-2123

Semenza, J.C., and Krishnasamy, P.V. (2007) “Design of a health-promoting neighborhood intervention”. Health Promotion Practice, 8:243–256

Simon, C.A., and Wang, C.H. (2002). “The impact of AmeriCorps service on volunteer participants: results from a 2-year study in four western states”. Administration and Society, 34(5): 522-40.

Sirven, N. and Debrand, T. (2011). "Social Capital and Health of Older Europeans," Working Papers DT40, IRDES institut for research and information in health economics

Snelgrove, J.W., Pikhart, H. and Stafford, M. (2009) “A multilevel analysis of social capital and self-rated health: Evidence from the British Household Panel Survey”, Social Science and Medicine 68(11): 1993-2001

Solow, R (1995), “Trust: The Social Virtues and the Creation of Prosperity (Book Review)”, The New Republic, 213:36-40.

Stern, S. (1989). ”Measuring the effects of disability on labor force participation”, Journal of Human Resources, 24, 361-95.

Szreter, S. (2004). “Industrialization and health”. British Medical Bulletin 69: 75–86.

Thomson, H., Atkinson, R., Petticrew, M. and Kearns, A. (2006) “Do urban regeneration programmes improve public health and reduce health inequalities? A synthesis of the evidence from UK policy and practice (1980–2004)”. Journal of Epidemiology and Community Health, 60(2):pp. 108-115.

Uslaner,E.M. (2002) “The Moral Foundations of Trust”, Cambridge University Press.

Veenstra, Gerry. (2000). “Social capital, SES and health: An individual-level analysis”. Social Science and Medicine 50, 5, 619-629.

Von dem Knesebeck O, Dragano, N., and J. Siegrist (2005) “Social capital and self-rated health in 21 European countries” Psychosocial Medicine.; 2: Doc02.

Winkelmann, L and Winkelmann, R. (1998) “Why are the Unemployed So Unhappy? Evidence from Panel Data” Economica, vol. 65, No. 257, pp. 1-15

Wooldridge, J.M. (2002) “Econometric Analysis of Cross Sections and Panel Data”, The MIT Press

Rossing Feldman, T. and Assaf, S. (1999) “Social Capital: Conceptual Frameworks And Empirical Evidence. An Annotated Bibliography”, Social Capital Initiative Working Paper No. 5, The World Bank

Zak, P.J., and Knack, S. (2001). "Trust and Growth," Economic Journal, vol. 111(470), pages 295-321

Appendix A

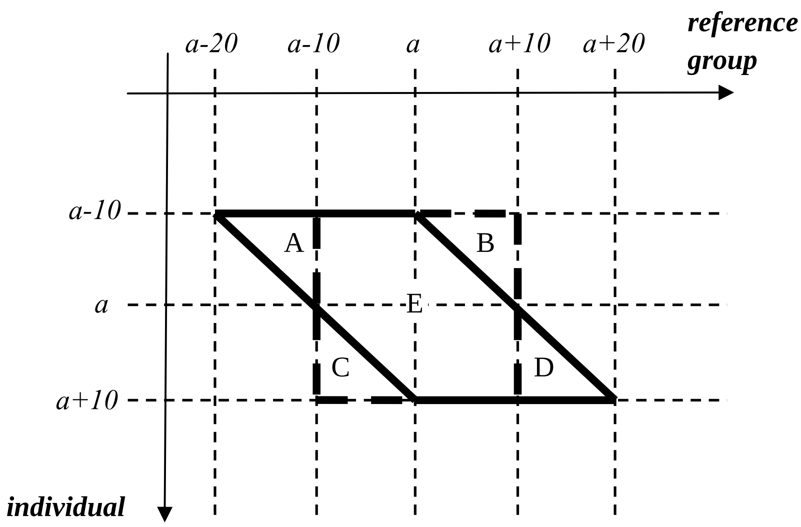
Rewrite equation (11) as

$$SC_i^* = SC_i^{**} + (1-\lambda)(CSC_i^* - SC_i^{**}) + \gamma h_i^{**} + \theta_{Si} \tag{A1}$$

to make explicit that all terms vary at individual level (*i*). For all members *j* of the reference group of individual *i*, denoted *R(i)* , a similar equation is defined. Suppose of adopting the baseline definition of reference group, i.e. Definition 0. Note that $CSC_i^* \neq CSC_j^*$ for each *j* because the reference group of *j* is not fully overlapped to the reference group of *i*, given the definition of reference group we have adopted.

Nevertheless the average of CSC_j^* over all $j \in R(i)$, denoted $\overline{CSC}^*_{R(i)}$, is approximately equal to CSC_i^*. Figure A1 provides the intuition. To simplify the matter, suppose that there is only one individual for each age a , so that we can label each individual with his/her age, i.e. $i = a$. Along the vertical axis we report individuals, each one represented by his/her age and along the horizontal axis we represent individual's reference group. For instance, the reference group of individual *a* is the set going from *a-10* to *a+10* on the horizontal axis. Similarly the reference group of individual *a-10* - who belongs to *R(a)* - is the set between *a-20* and *a*. The average of $CSC_{a'}^*$ over all $a' \in R(a)$, i.e. $\overline{CSC}^*_{R(a)}$ is the average *SC** over the support represented by the area A+E+D in Figure A1. Instead CSC_a^* is the average *SC** over the support represented by the area E+B+C. These two averages share three fourths of their support. Therefore they are likely to be very close each other.

Figure A1



Formally $\overline{CSC}^*_{R(a)}$ is

$$\overline{CSC}^*_{R(a)} = \frac{1}{21}\sum_{a'=a-10}^{a+10} CSC^*_{a'} = \frac{1}{21}\sum_{a'=a-10}^{a+10}\frac{1}{21}\sum_{a''=a'-10}^{a'+10} SC^*_{a''} = \frac{1}{441}\sum_{a'=a-10}^{a+10}\sum_{a''=a'-10}^{a'+10} SC^*_{a''} \quad \text{(A2)}$$

whose RHS which can be decomposed as

$$\frac{1}{441}\left[\sum_{a'=a-10}^{a+10}\sum_{a''=a-10}^{a+10} SC^*_{a''} + \left(\sum_{a'=a-10}^{a}\sum_{a''=a'-10}^{a-11} SC^*_{a''} - \sum_{a'=a-10}^{a}\sum_{a''=a'+11}^{a+10} SC^*_{a''}\right) + \left(\sum_{a'=a}^{a+10}\sum_{a''=a+11}^{a'+10} SC^*_{a''} - \sum_{a'=a}^{a+10}\sum_{a''=a-10}^{a'-11} SC^*_{a''}\right)\right] \quad \text{(A3)}$$

and compactly rewritten as

$$CSC^*_a + \frac{50}{441}\left(\overline{SC}^*_A - \overline{SC}^*_B\right) + \frac{50}{441}\left(\overline{SC}^*_C - \overline{SC}^*_D\right) \quad \text{(A4)}$$

where $\overline{SC}^*_k$, $k = A, B, C, D$ are the average individual reported SC over the areas k of the support represented in Figure A1.

Computations based on our sample indicates that for 90% of observations the difference between $\overline{CSC}^*_{R(i)}$ and CSC_i^* is between [-0.19,+0.19] to be compared with the fact that 90% of the values taken by $\overline{CSC}^*_{R(i)}$ ranges between [-1.47, +1.66].

Appendix B

Rewrite the system (17)-(18) more compactly as

$$h^* = a_0 + d_1 SC^* + d_2 CSC^* + d_3 \overline{h}^* + Z_h a_1 + X a_2 + \eta_h \quad \text{(B1)}$$

$$SC^* = a_3 + d_4 h^* + d_5 CSC^* + d_6 \overline{h}^* + Z_S a_4 + X a_5 + \eta_S \quad \text{(B2)}$$

where we have denoted by d_j, $j = 1,2,\ldots,6$ all parameters relevant for identification of the structural coefficients of interest. The system (B3) to (B8)

$$\frac{\frac{\alpha_1}{\lambda}}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)} = d_1 \quad \text{(B3)}$$

$$\frac{\left(\alpha_2 - \alpha_1 \frac{1-\lambda}{\lambda}\right)}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)} = d_2 \quad \text{(B4)}$$

$$\frac{-\alpha_2 \frac{\gamma}{\lambda}}{\left(1+\alpha_1 \frac{\gamma}{\lambda}\right)} = d_3 \qquad \text{(B5)}$$

$$\lambda\left(\beta_1 + \frac{\gamma}{\lambda}\right) = d_4 \qquad \text{(B6)}$$

$$\left(1-\lambda+\lambda\beta_2\right) = d_5 \qquad \text{(B7)}$$

$$-\gamma\beta_2 = d_6 \qquad \text{(B8)}$$

has a unique solution which is given by

$$\alpha_1 = \frac{d_1 AB}{(B+d_1 A)C} \qquad \text{(B9)}$$

$$\alpha_2 = \frac{d_3 B^2}{(B+d_1 A)C} \qquad \text{(B10)}$$

$$\beta_1 = \frac{(A+d_4 B)C}{BA} \qquad \text{(B11)}$$

$$\beta_2 = \frac{d_6 B}{A} \qquad \text{(B12)}$$

$$\gamma = \frac{A}{B} \qquad \text{(B13)}$$

$$\lambda = \frac{A}{C} \qquad \text{(B14)}$$

with $A = (d_3 - d_3 d_5 + d_1 d_6 + d_2 d_6)$, $B = (d_5 d_1 + d_2)$, and $C = (d_3 + d_1 d_6)$.

Table 1 Number of Observations

Country	Number of observations	Percentage of the sample
Austria	6,227	4.72
Belgium	6,764	5.12
Bulgaria	2,043	1.55
Czech Republic	5,425	4.11
Germany	10,321	7.82
Denmark	2,146	1.63
Estonia	4,553	3.45
Spain	5,581	4.23
Finland	4,789	3.63
France	4,685	3.55
Greece	6,367	4.82
Croatia	1,136	0.86
Hungary	4,999	3.79
Ireland	4,169	3.16
Italy	2,063	1.56
Latvia	1,358	1.03
Netherland	7,224	5.47
Norway	6,968	5.28
Poland	6,643	5.03
Portugal	7,423	5.62
Romania	1,456	1.10
Sweden	5,398	4.09
Slovenia	4,819	3.65
Slovakia	4,341	3.29
United Kingdom	7,880	5.97
Total	132,031	

Note: Data, ESS (first 4 waves). Sample size: 132,031

Table 2A Control variables included in the models. Descriptive Statistics. Source: ESS waves 1-4

Variable	Mean	Std. Dev.
Self reported health	3.785376	0.912855
Characteristics of peers		
Percentage of males (peers)	0.465388	0.081471
Years of education (peers)	12.03435	2.177994
Percentage of people in the first income quintile (peers)	0.606905	0.286844
Percentage of people in the second income quintile (peers)	0.077054	0.068425
Percentage of people in the third income quintile (peers)	0.251318	0.319599
Percentage of people in the fourth income quintile (peers)	0.019188	0.027234
Percentage of people in the fifth income quintile (peers)	0.02627	0.041884
Percentage of people in the sixth income quintile (peers)	0.005073	0.011542
percentage of citizens (peers)	0.961	0.056876
No internet (peers)	0.328085	0.162533
Percentage of Immigrants (peers)	0.097181	0.088344
Percentage of people victims of a crime (peers)	0.191145	0.102308

Parental Characteristics		
Father born in the country	0.893851	0.30803
Mother born in the country	0.895532	0.305868
Father in the 2nd quintile of education	0.242519	0.428608
Father in the 3rd quintile of education	0.21757	0.412595
Father in the 4th quintile of education	0.221244	0.415086
Father in the 5th quintile of education	0.034946	0.183645
Father in the 6th quintile of education	0.091244	0.287957
Mother in the 2nd quintile of education	0.272701	0.44535
Mother in the 3rd quintile of education	0.246268	0.430838
Mother in the 4th quintile of education	0.201604	0.4012
Mother in the 5th quintile of education	0.026782	0.161445
Mother in the 6th quintile of education	0.063894	0.244565
Employed father	0.663034	0.472675
In the 2nd income quantile	0.076785	0.266251
In the 3rd income quantile	0.255054	0.435893
In the 4th income quantile	0.01898	0.136456
In the 5th income quantile	0.026138	0.159546
In the 6th income quantile	0.005105	0.071266
In the 7th income quantile	0.014216	0.118382
Number of household members	2.777832	1.427084
Employed mother	0.424855	0.494323
Self-employed father	0.223978	0.416909
Self-employed mother	0.096546	0.295339
Father died	0.058411	0.234519
Mother died	0.020245	0.140839
Missing father's education	0.073559	0.261052
Missing mother's education	0.05885	0.235344
Individual Characteristics		
Male	0.465111	0.498783
Age	47.22327	18.24809
Age squared	2563.028	1816.175
Born in the country	0.924457	0.264266
Urban	0.621506	0.485014
Married	0.533117	0.498904
Years of education	11.97491	4.024108
Religious	4.787504	2.993591
Roman catholic	0.346744	0.475935
Protestant	0.151086	0.358134
Eastern orthodox	0.072809	0.259823

Other Christian denomination	0.01636	0.126855
Jewish	0.00097	0.031121
Islam	0.013429	0.115102
Eastern religions	0.003045	0.055095
Other non-Christian religions	0.002272	0.047614
Legislators, senior officials and managers and armed force	0.495384	0.499981
Professionals	0.05089	0.219773
Technicians and associate professionals	0.081564	0.273701
Clerks	0.089229	0.285075
Service workers and shop and market sales workers	0.052367	0.222766
Skilled agricultural and fishery workers	0.068688	0.252925
Craft and related trades workers	0.014701	0.120354
Plant and machine operators and assemblers	0.064629	0.245871
Elementary occupations	0.037756	0.190607
Victim of a crime	0.190319	0.392554
Regional Characteristics (NUTS 2)		
Activity rate 15-64 (Nuts 2)	71.91039	6.013201
Activity rate 25-34 (Nuts 2)	85.80376	4.005928
Employment rate 15-64 (Nuts 2)	66.67178	7.531689
Employment rate 25-34 (Nuts 2)	79.25461	6.117371
Unemployment rate 15-24 (Nuts2)	16.09077	8.696213
Unemployment rate 15-64 (Nuts 2)	7.392936	4.317816
Density (Nuts 2)	336.8073	710.632
Number of graduates (Nuts 2)	0.240185	0.08395
Length of motorways (Nuts 2)	30.08397	39.44004
GDP (Nuts 2)	23858.13	13338.5
GDP growth (Nuts 2)	4.683081	4.817901
Percentage of young (Nuts 2)	0.164728	0.023226
Percentage of adults (Nuts 2)	0.676178	0.022064
Percentage of old (Nuts 2)	0.150524	0.040908
Percentage of people victims of a crime (Nuts 2)	0.189839	0.080267
Number of doctors (Nuts 2)	3.272222	1.000858
Other dummy variables		
Round 2	0.246722	0.431105
Round 3	0.233415	0.423006
Round 4	0.270277	0.444105
North	0.237444	0.425518
Central-west	0.286213	0.451992
Central-east	0.148329	0.355427
East	0.130189	0.336513

Table2B . Characteristics of the groups of peers

Definition	Average number of peers in the group	Standard deviation
Definition 0	116.7231	108.7692
Definition 1	135.289	130.5744
Definition 2	116.8561	115.004

Note: Data, ESS (waves 1-4). Sample size: 132,031.

Table 3 Baseline model - Reference group definition 0

	First stage			Second stage	
	health	SC	CSC	SC	health
# doctors	0.051***	0.005	-0.025***		0.047***
	(0.005)	(0.005)	(0.006)		(0.008)
Victim	-0.128***	-0.095***	0.005	-0.074***	
	(0.008)	(0.007)	(0.008)	(0.012)	
% victims	-0.826***	-0.460***	-2.231***		
	(0.065)	(0.055)	(0.082)		
health				0.164**	
				(0.073)	
CSC				0.146***	0.090*
				(0.035)	(0.051)
SC					1.356***
					(0.127)
N	132031				

Note: Data, ESS (waves 1-4). Sample size: 132,031. *# doctors* is the number of doctors in the region of residence (NUTS 2); *Victim* is a dummy taking 1 in the respondent have been victim of a burglary or an assault in the past 5 years; *% victims* is the proportion of victims of a burglary or assault in the respondent's reference group; *health* is self-reported health status; *CSC* and *SC* are community and individual social capital respectively. The first 3 columns report first stage estimated parameters of the excluded exogenous variables. The last two columns report second stage estimates of exclusions and endogenous variables (Baseline model with reference group definition 0). Asterisks indicate level of statistical significance: ***p < 0.01, **p < 0.05, *p < 0.1. Boostrappend standard errors in parentheses.

Table 4 Baseline model - Reference group definition 1

	First stage			Second stage	
	health	SC	CSC	SC	health
# doctors	0.051***	0.005	-0.021***		0.047***
	(0.005)	(0.005)	(0.006)		(0.008)
Victim	-0.129***	-0.096***	0.006	-0.077***	
	(0.008)	(0.008)	(0.009)	(0.013)	
% victims	-0.864***	-0.399***	-1.946***		
	(0.067)	(0.056)	(0.089)		
health				0.154**	
				(0.078)	
CSC				0.137***	0.166***
				(0.041)	(0.062)
SC					1.355***
					(0.135)
N	132031				

Note: See Table 3 (Baseline model with reference group definition 1)

Table 5 Baseline model - Reference group definition 2

	First stage			Second stage	
	health	SC	CSC	SC	health
# doctors	0.050***	0.006	-0.022***		0.045***
	(0.005)	(0.005)	(0.006)		(0.009)
Victim	-0.123***	-0.092***	0.010	-0.068***	
	(0.008)	(0.008)	(0.009)	(0.013)	
% victims	-1.530***	-0.828***	-2.616***		
	(0.103)	(0.098)	(0.133)		
health				0.212***	
				(0.075)	
CSC				0.193***	0.156*
				(0.053)	(0.084)
SC					1.354***
					(0.138)
N	132031				

Note: See Table 3 (Baseline model with reference group definition 2).

Table 6 Baseline model - Reference group definition 0 with regional dummies

	First stage		Second stage
	SC	CSC	health
Victim	-0.096***	0.002	
	(0.007)	(0.009)	
% victims	-0.447***	-2.746***	
	(0.059)	(0.093)	
CSC			0.082*
			(0.045)
SC			1.350***
			(0.135)
N	132031		

Note: Data, ESS (waves 1-4). Sample size: 132,031. *Victim* is a dummy taking 1 in the respondent have been victim of a burglary or an assault in the past 5 years; *% victims* is the proportion of victims of a burglary or assault in the respondent's reference group; *health* is self-reported health status; *CSC* and *SC* are community and individual social capital respectively. The first two columns report first stage estimates the excluded exogenous variables of the health equation. The last column, reports second stage estimates of the endogenous variables (Baseline model with reference group definition 0 with regional dummies). Asterisks indicate level of statistical significance: ***p < 0.01, **p < 0.05, *p < 0.1. Boostrapped standard errors in parentheses)

Table 7 Extended model - Reference group definition 0

	First stage			Second stage	
	health	SC	CSC	SC	health
Mhealthtot	0.505***	0.046***	0.207***	-0.090*	0.452***
	(0.009)	(0.010)	(0.012)	(0.048)	(0.018)
# doctors	0.041***	0.004	-0.029***		0.034***
	(0.005)	(0.005)	(0.006)		(0.008)
Victim	-0.129***	-0.095***	0.005	-0.069***	
	(0.008)	(0.007)	(0.008)	(0.015)	
% victims	-0.497***	-0.428***	-2.086***		
	(0.061)	(0.053)	(0.080)		
Health				0.204**	
				(0.099)	
CSC				0.156***	-0.041
				(0.031)	(0.054)
SC					1.361***
					(0.133)
N	132031				

Note: See Table 3. (Extended model with reference group definition 0).

Table 8 Structural Parameters

	Reference gr. definition 0	Reference gr. definition 1	Reference gr. definition 2
Extended Model			
effect of true SC on health α_1	0.353***	0.426***	0.499***
	(0.037)	(0.042)	(0.041)
effect of true CSC on health α_2	0.077***	0.113***	0.272***
	(0.020)	(0.027)	(0.062)
effect of health on true SC β_1	2.170***	1.650***	1.255***
	(0.318)	(0.236)	(0.181)
effect of true CSC on true SC β_2	-0.059	-0.058	-0.211**
	(0.033)	(0.039)	(0.093)
effect of health on reported SC γ	-1.524***	-1.151***	-0.567***
	(0.243)	(0.186)	(0.111)
effect of reported CSC on reported SC $1\text{-}\lambda$	0.203***	0.194***	0.352***
	(0.028)	(0.037)	(0.055)
Baseline Model			
effect of true SC on health α_1	1.356***	1.355***	1.354***
	(0.127)	(0.135)	(0.138)
effect of true CSC on health α_2	0.090*	0.166***	0.156*
	(0.051)	(0.062)	(0.084)
effect of health on true SC β_1	0.164**	0.154**	0.212***
	(0.073)	(0.078)	(0.075)
effect of true CSC on true SC β_2	0.146***	0.137***	0.193***
	(0.035)	(0.041)	(0.053)

Note: Data, ESS (waves 1-4). Sample size: 132,031. Upper panel: structural parameters of the extended model, according to definitions 0,1,2, Lower panel: structural parameters of the baseline model (see Table 3). Asterisks indicate level of statistical significance: ***p < 0.01, **p < 0.05, *p < 0.1. Boostrapped standard errors in parentheses.