Does Social Capital reduce moral hazard? A network model for non–life insurance demand

Giovanni Millo*, Giacomo Pasini†

This version: October 15, 2006

Abstract

We start from the Arnott and Stiglitz model on the coexistence of market and non–market insurance contracts. We focus our attention on the effect of moral hazard involved in non market contracts on the demand for marketed contracts. We extend the model to allow for the presence of Social Capital and Trust as determinants of moral hazard in informal contracts, and provide a rigorous definition of those concepts by means of an equilibrium concept typical of the Network literature. Such a formal approach gives us a clear guidance for measuring Social Capital and validate the model on empirical data. The model is estimated on a new panel dataset, supporting our claim that Social Capital increases the demand for non–life insurance. We test for the presence of spatial correlation, and conclude that the the spatial structure of demand for non–life insurance contracts is completely determined by the spatial distribution of Social Capital.
JEL Classification: C21, D85, G22, Z13
Keywords: Insurance, Social Capital, Network stability, Spatial Panel data model

1 Introduction

Social Capital is a concept not limited to sociology: during the last 20 years it spread out and has been used across almost all social sciences. Despite such a great interest and huge amount of research on it, it's still a suggestive word that reminds of many different but related research fields, rather than a precise concept. Further on, the study of social capital has a lot to do with

*Research Dept., Assicurazioni Generali SpA. e–mail giovanni_millo@generali.com.
†Economics Dept., University of Venice and SSAV, School for Advanced Studies in Venice; corresponding author. Dipartimento di Scienze Economiche, Cannaregio 873 – 30121 Venice, Italy. Tel: +39 041 2349231; e–mail giacomo.pasini@unive.it; homepage: http://venus.unive.it/pasinigi

Italy: the seminal book by Putnam [16] about democracy and institutions' efficiency across Italy is a source of overwhelming empirical evidence on the relevance of social capital in Italian social life. Focusing on economics, recently Guiso et al. [13] found that social capital matters on asset allocation choices of Italian households: they started from the idea that any financial contract involves trust, which is strongly correlated to Social Capital, and found empirical results on this relation.

Our question is whether it matters also on individual choices about insurance expenditure. In particular, we are interested in demand for non–life insurance contracts. While life insurance can be assimilated to pension funds and other financial assets in terms of economic rationale - it's an investment which gives a return - for non–life insurance things are different. Households buy a non–life insurance contract to avoid the risk of suffering losses in some future state of the world: they pay a fixed price (the premium) to transfer money from a future uncertain state of the world to a certain one. Arnott and Stiglitz [3] set up a model where together with market insurance, individuals can enter in non–market mutual insurance contracts. In their model the role played by non market insurance is related to peer monitoring: if informational asymmetry between the insurer and the customer still holds in non–market contracts, they are dysfunctional and non–market insurance displaces market contracts reducing social welfare. Vice versa, if individuals can observe other individuals' effort, non–market contracts are welfare enhancing since they provide extra insurance coverage at the market price set by the insurance company. What they call peer monitoring is actually the severity of moral hazard in non–market agreements. We will show that the lower the level of moral hazard, the higher the aggregate demand for market insurance. We will also formally link moral hazard and social capital, concluding that social capital itself increases the aggregate demand for insurance. A careful definition of Social Capital and its role in the model allows us to test our conclusions empirically.

Previous studies on Italy leave space for such a model. Millo and Lenzi [14] found that the Italian insurance market exhibits spatial heterogeneity and spatial correlation at the province level even after controlling for a number of demographics. If heterogeneity in the diffusion of insurance contracts is due to differences in the degree of Social Capital, it is reasonable to think that its diffusion does not follow administrative province boundaries: therefore, our explanation is coherent with the presence of spatial correlation at the province level. The social capital interpretation is suggestive also for another reason: Durlauf and Fafchamps [8] point out that a possible role for social capital in economic models is to limit market inefficiencies when institutions fail to resolve them: In Italy family ties are frequently substitutes for inefficient institutions. Religious (mainly catholic) communities as well as some other professional and voluntary associations play a role in supplementing part of the social welfare not provided by the State: disabled and elder people assistance or scholarships are some examples.

The paper is structured as follows: the second section describes Arnott and Stiglitz's [3] model. The following one extends the model to provide a formal definition of Social Capital and to include it as a determinant of demand for market insurance. Such an extension will be done within a Network approach. Before going to empirical validation of our model we will describe the ISVAP (Italian authority on Insurance market) dataset. The fifth section is dedicated to the measurement of trust and social capital. The sixth part describes the estimation procedure and results. In the seventh section we carry on the analysis of the spatial structure of the model. Last section concludes.

2 The model

Arnott and Stiglitz [3] were interested in the general equilibrium and welfare effects of non market insurance and peer monitoring. Their model provides the background to study the effect of moral hazard and therefore - as we will see in the next section - of Social Capital on the demand for market insurance.

The starting point is the canonical moral–hazard model without non market insurance. There is a single and fixed damage accident. The probability of its occurrence, $p(e)$, is strictly convex and decreasing in the individual's effort at accident avoidance, e, which is not observable to the insurer. Individual wealth is w, the damage caused by the accident d. Individuals pay a premium β and receive a net payout α in case the accident occurs. Assuming a well behaved (increasing and strictly concave), separable and event–independent utility function, individuals maximize their expected utility

$$\begin{aligned} EU^M &= (1-p(e))U(w-\beta) + p(e)U(w-d+\alpha) - e \\ &= (1-p(e))u_0 + p(e)u_1 - e \end{aligned}$$

At the competitive constrained equilibrium, the insurer offers less than full insurance to induce the clients to augment their effort at accident avoidance, i.e. $d - \alpha > \beta$, meaning that the ordering of states of the world in terms of utility is not altered: the wealth reduction in the "good" state of the world, β, must be lower than the wealth reduction in the "bad" state, $d-\alpha$. This equilibrium is stable only if clients purchase no additional insurance. Such a condition must be enforceable by the insurer. This exclusivity condition is not far from what happens in the real world: insurance companies cannot force their clients to buy just one contract, but they ask them to reveal which other contracts they have covering the same risk, and in case of accident occurrence payout is divided proportionally among insurers.

Non-market insurance is introduced as follows: a couple of symmetric individuals, i and j, agree that if one of them has an accident and the other doesn't, the latter will transfer δ to the former. Each of them realizes that the extra insurance will pay out if they have an accident and their partner doesn't, therefore their expected utility changes:

$$\begin{aligned} EU_i &= (1-p(e_i))(1-p(e_j))U(w-\beta) + p(e_i)p(e_j)U(w-d+\alpha) \\ &\quad + (1-p(e_i))p(e_j)U(w-\beta-\delta) \\ &\quad + p(e_i)(1-p(e_j))U(w-d+\alpha+\delta) \\ &\quad - e_i \\ &= (1-p(e_i))(1-p(e_j))u_0 + p(e_i)p(e_j)u_1 \\ &\quad + (1-p(e_i))p(e_j)u_2 + p(e_i)(1-p(e_j))u_3 - e_i \end{aligned} \tag{1}$$

Individuals maximize their utility considering α and β and therefore the contract's price q as fixed: they perceive that if they enter a mutual contract they can buy extra insurance at the market price q. They consider as given also δ, which is the premium but also the payoff of the non–market agreement. Further on each of them consider her partner as rational and assumes she will choose the level of effort which maximizes her own utility.

If each individual does not observe the others' effort, the exclusivity provision cannot be enforced: each client pays an extra premium δ if the partner has an accident and he doesn't, while he receives an extra payoff δ in the opposite case. It is optimal for them to reduce the effort while the insurance company is still offering the same contract. This is a partial equilibrium result since it doesn't consider the reaction of insurance companies to agents' behavior. In a General Equilibrium context the company knows that the required level of effort for the offered contract cannot be enforced: non market insurance crowds out market insurance and individuals substitute insurance provided by a risk neutral insurer with a contract provided by a risk averse one. Individual's expected utility, EU^{NMU}, is lower than without non–market insurance, EU^{M}[1].

Vice versa, the authors show that if individuals can observe perfectly each other's effort, it is optimal for the individuals to provide non market insurance up to full coverage to augment the risk sharing opportunity. The extra coverage still reduces the effort. Therefore, the insurance company will reduce its coverage but won't be displaced: peer monitoring enforces a certain level of effort. Non market insurance increases utility since it increases risk sharing capabilities for risk adverse individuals; on the other hand, it reduces utility since it substitutes part of the coverage provided by the insurance company, which is risk neutral, with coverage provided by other risk averse individuals: the two effects offset each other. Arnott and Stiglitz provide a

[1] We are in a symmetric world, therefore $e_i = e_{-i} \Rightarrow EU^i = EU^{-i} = EU^{NMU} < EU^M$

sufficient condition for the first to be greater than the second: the probability of accident occurrence has to be small enough, $p(e) < \frac{1}{2}$.

Such a result is crucial once heterogeneity among individuals is introduced. Whatever the peer monitoring involved in non–market agreement, the insurance company is never able to observe individual efforts, so it is forced to pool potential clients and offer a unique contract. More realistically, insurers offer different contracts based on observed characteristics of individuals such as age or marital status and on past statistics as loss ratios in a particular region[2]. What they are not able to do, is to offer different contracts based on individual effort, due to information asymmetry. What Arnott and Stiglitz result tells us is that is the probability of accident occurrence is small, the new marketed contract combined with the non–market agreement will increase individual expected utility:

$$E_j^{NMO}[U|\boldsymbol{X}_j] > E_j^M[U|\boldsymbol{X}_j]$$

Where $\boldsymbol{X}_j$ is a vector of observable individual characteristics, EU^{NMO} is expected utility with non market contracts and perfect peer monitoring, EU^M is expected utility with only marketed insurance.

Individual expected utility without any insurance contract is

$$E_j[U|\boldsymbol{X}_j] = (1 - p(e_j))U_j(w_j|\boldsymbol{X}_j) + p(e_j)U_j(w_j - d|\boldsymbol{X}_j)$$

The previous inequality tells us how demand for insurance changes introducing non market insurance. The strict inequality sign and heterogeneity among individuals guarantee that in a large enough population there will be someone that is willing to buy an offered market contract only if non market insurance is provided as well. To be specific, this is true for all those js such that

$$E_j^{NMO}[U|\boldsymbol{X}_j] > E_j[U|\boldsymbol{X}_j] > E_j^M[U|\boldsymbol{X}_j]$$

These individuals would buy an insurance contract regardless of non market agreement. All the others won't change their choices: those for which

$$E_j[U|\boldsymbol{X}_j] > E_j^{NMO}[U|\boldsymbol{X}_j] > E_j^M[U|\boldsymbol{X}_j]$$

won't buy any insurance contract anyway, while those for which

$$E_j^{NMO}[U|\boldsymbol{X}_j] > E_j^M[U|\boldsymbol{X}_j] > E_j[U|\boldsymbol{X}_j]$$

are always willing to pay for insurance coverage. If we can assume Arnott and Stiglitz condition $p(e) < \frac{1}{2}$ is always satisfied for non–life insured risks,

[2] the loss ratio for a type of accident is the ratio between payouts received by clients over premium payed to the insurer

which seems to be an innocuous assumption given the contracts offered on the market, we have an empirical implication about the presence of non–market agreement: given perfect peer monitoring, we should observe a marginal effect on demand for market insurance with the introduction of non–market agreements.

Nevertheless there is still something to do in order to achieve a testable implication: we would like to discriminate peers of individuals endowed with non–market agreements and to measure the level of peer monitoring within those communities.

A possible strategy to do it comes from a deeper understanding of peer monitoring. What Arnott and Stiglitz name peer monitoring and that discriminate the case in which market insurance is displaced with the one in which it fosters the aggregate demand for insurance is essentially the severity of moral hazard within the peer entering a non market contract. As the authors point out in the discussion it depends on reciprocal observability of the effort but also on the duration of the partnership, the level of trust between them, the severity of punishment when deviating from an agreement, the power of reputation and social pressure: in one word, the severity of moral hazard depends on the stock of social capital a community is endowed with.

3 A network–based definition of Social Capital

As already pointed out in the introduction, there isn't a clear–cut definition of Social Capital. It is an elusive concept that declines in particular meaning depending on the context where it is used. Social Capital is a suggestive idea, but to have a testable model we need to formalize concepts of social capital and trust. Durlauf and Fafchamps [8] point out as a common feature of many definitions of Social Capital the focus on interpersonal relationships and social networks. This is the reason why we use a network approach proposed by Vega–Redondo [17].

Suppose that couples of individuals that enter a non market insurance agreement with a given δ can choose in each period whether to put an effort e_{NMU}, which is the one with moral hazard in the Arnott Stiglitz framework, or e^{NMO}, effort without moral hazard. If expected utility is decreasing in the effort, such a game is a Repeated Prisoner's dilemma. From (1),

$$\begin{aligned}\frac{\partial EU^i}{\partial e_i} &= \big[-(1-p(e_{-i}))u_0 + p(e_{-i})u_1 \\ &\quad - p(e_{-i})u_2 + (1-p(e_{-i}))u_3\big]p'(e_i) - 1 \\ &= [(u_3-u_0)(1-p(e_{-i}) + (u_1-u_2)p(e_{-i})]\,p'(e_i) - 1\end{aligned} \tag{2}$$

which is decreasing in e_i if $\beta+\delta < d-\alpha-\delta$, i.e. the total cost of insurance, $\beta+\delta$ must be lower than the loss suffered when the accident occurs. If this condition holds (together with $p(e) < \frac{1}{2}$), the game rewritten in strategic form with expected utilities as payoffs is of the Prisoner's dilemma type (see figure 1). Since marginal utility is decreasing in the (own) effort, for individual i we can write

$$EU_{ij}^{H} = EU(e_i = e_{NMU}, e_j = e^{NMO}) > EU_{ij}^{NMO}$$

$$EU_{ij}^{L} = EU(e_i = e^{NMO}, e_j = e_{NMU}) < EU_{ij}^{NMU}$$

		Player j: e^{NMO}	Player j: e_{NMU}
Player i	e^{NMO}	$EU_{ij}^{NMO}, EU_{ji}^{NMO}$	EU_{ij}^{L}, EU_{ji}^{H}
	e_{NMU}	EU_{ij}^{H}, EU_{ji}^{L}	$EU_{ij}^{NMU}, EU_{ji}^{NMU}$

Figure 1: the non–market insurance game in strategic form

Once this game is put in a dynamic setting, the social network can be described as in Vega–Redondo [17]: we have a finite population of agents $N = \{1, 2, \ldots, n\}$ where each pair of interacting agents i, j is involved in an infinite repetition of the described game. Players' connecting decision is captured by a directed graph $\vec{g} \subset N \times N$, where each directed link $(i, j) \in \vec{g}$ is player i decision to connect with player j. Suppose now that every linking decision lead to play. We have a definition for social network:

Definition 1 (Social Network) *The social network induced by the linking decision $\vec{g}$ is the undirected graph $g \subset N \times N$ defined as*

$$\forall i, j \in N, \quad (i, j) \in g \Longleftrightarrow [(i, j) \in \vec{g} \vee (j, i) \in \vec{g}]$$

and for any player i the set of her neighbors is

$$N_i = \{j \in N : (i, j) \in g\}$$

In order to complete the repeated game model we need a rule for information diffusion within the network: in our model information spread around the network only gradually. To be specific, at each round before playing i, j share information about their behavior with their neighbors, i.e. whether they deviated from the cooperative strategy. To sustain a cooperative equilibrium it's also necessary that each agent adopts a strategy that punish defiance: i force herself to play a trigger strategy, i.e. she will switch to defection with j as soon as she knows j deviated with some of her neighbors. More formally, for any agent i the strategy $s^g = (s_1^g, \ldots, s_n^g)$ is of the following type:

1. first, player i chooses whether to start her interaction with j putting effort e^{NMO} (which is to cooperate) or to put effort e_{NMU};
2. in the following rounds, she reacts immediately to the news j did not start with e^{NMO} with some $k \in N_j$ switching irreversibly to e_{NMU} in her game with j.

In order to give a definition of an equilibrium, some additional notation is needed: $\pi_i(s^g)$ is the overall payoff from the link (i, j) given the strategy s^g; for every agent i s_C^g and s_D^g are the strategies that starts respectively with cooperation and defection with all the agents $k \in N_i$.

Definition 2 (Pairwise–stable Network (PSN)) *a PSN is a network where for every separate link, the two players have incentives to sustain the cooperative equilibrium, i.e.*

$$\forall (i,j) \in g \quad \pi_i(s_C^g) \geq \pi_i(s_D^g)$$

The connection of this definition with the Social Capital literature is clear once the PSN is characterized in terms of cohesiveness. Let define

Definition 3 (i–excluding distance) *$d^i(j,k)$, the i–excluding distance between j and k is the shortest path joining j and k which does not involve player i. In other words, it is the number of steps needed for any information held by j to reach k (and vice versa) without the concourse of i.*

Then

Proposition 1 *Let g be a Social Network where agents play the described game, and they all face a common discount factor $\delta \in (0,1)$. Define $\nu_{ik} = EU_{ik}^{NMO} - EU_{ik}^{L}$ Then, g is a PSN if and only if for all $(i,j) \in g$*

$$EU_{ij}^{NMO} + \sum_{k \in N_i/\{j\}} \delta^{d^i(j,k)} \left[\delta EU_{ik}^{NMO} + (1-\delta)\nu_{ik}\right] \geq (1-\delta) EU_{ij}^{H}$$

Proof of proposition 1 is in the appendix and follows the one in Vega–Redondo [17]. The implications of this proposition are:

- Stability is more likely in *large span* networks, i.e. in networks where each agent i has a large neighborhood N_i;
- Stability is more likely in *cohesive* networks, i.e. in networks with small excluding distances $d^i(j,k)$.

It is also clear that, since payoffs are uncertain, the level of volatility in the model is inversely related with stability. Given this formalization,

Definition 4 (Trust and Social Capital) *The stock of trust embedded in a community is the level of stability of its social network. The stock of Social Capital of the network g is the* density[3] *of g.*

Going back to the first part of the model, we showed that demand for market insurance increases as moral hazard involved in non–market insurance falls. In a pairwise stable network agents have no incentives to reduce the effort, i.e. moral hazard is inversely related to network stability. Therefore, from definition 4 the empirical implication of the model is that demand for market insurance is increasing in Trust and Social Capital. Further on, as Vega–Redondo pointed out definition 3 and the following concept of cohesiveness is the counterpart of Coleman's concept of closure of a Social Network. We have a second empirical implication: demand for market insurance is higher in closed networks.

4 Demographics and insurance data

In order to identify the effect of social capital on insurance purchases, we have to control for the determinants of insurance development. Theoretical models of non-life insurance demand, starting from the seminal paper of Mossin ([15]), predict that for a given level of risk exposure insurance demand is increasing with risk aversion, probability of loss and total wealth . Empirical studies identify some observable counterparts. Wealth, when not observable, is generally proxied by means of income; so it is risk exposure, which is in turn related to total wealth and the level of economic activity. Loss probability may too be related to income as a measure of economic activity; urbanization has also been suggested for this purpose (Browne et al. [5]). Loss ratios[4] have also been suggested as a proxy for the probability of loss. Aspects of risk aversion may be captured by education or the age structure of the population, even though the expected sign of the effect is unclear (see Browne and Kim [6], Grace and Skipper [12] and the discussion in Browne et al. [5]).

4.1 Controlling for supply side variables

We stated in section 2 that insurance company has a limited discriminating power, i.e. it can offer different contracts (which means different prices) based on observable characteristics of individuals and of particular subpopulation, but it can't offer individual contracts based on effort, which is always unobserved by the insurer. This means that in an empirical investigation on demand for insurance it is crucial to control for supply side changes, meaning for offered prices, in order to be sure that the marginal effects of interest which we exploit on the demand equation are not completely absorbed by

[3]The density is the average number of links per agent (degree) in the network.

[4]Loss ratios are defined as the ratio of claims incurred to premiums earned.

equilibrium prices. This is a non–trivial problem: as Schlesinger (in [7]) notes, "it is often difficult to determine what is meant by the price and the quantity of insurance. [...] the fundamental two building blocks of economic theory have no direct counterparts for insurance". In practice we can usually only observe insurance consumption, the product between equilibrium price and quantity, jointly determined by the interplay of supply and demand. The choice of a price variable, when available at all, is therefore far from being obvious. We cannot observe the amounts insured, therefore inclusion of medium premium rates, which would probably be best, is ruled out. We resort therefore to the loss ratio, as e.g. in [10], observing that the role of this index as a proxy for market riskiness could lead to some ambiguity. Due to unavailability of data on losses for the non-life market as a whole, we include the aggregate loss ratio for the property sector only (Fire, Motor non-TPL, Other material loss).

Lastly, given the importance of tied agents in the distribution of insurance products (this channel did account in 2000 for 88.3 of non-life premium volume)[5], the number of agencies per capita has been included as a supply-side driver, inversely related to the opportunity-cost of searching for insurance covers.

Our dataset consists mainly of an excerpt for the years 1998-2000 from the Geo-Starter database provided by Istituto Tagliacarne, an institution inside SiStaN (the Italian national statistical system). It provides both first-hand data and an organized collection of data from various institutional sources. Data on insurance premiums, in particular, are collected on a provincial basis by ISVAP, the Italian insurance Authority, divided into three categories: life, compulsory third party liability, the vast majority of which regarding motor vehicles, and other non-life. While motor third party liability is a homogeneous class, both life and other non-life comprise very different kinds of policies.

4.2 Measuring insurance consumption

As observed above, we are only able to observe the equilibrium value of insurance consumption, and neither the quantity nor the price of insurance. Furthermore, measuring insurance consumption across administrative regions of different economic and demographic "size" requires resorting to some kind of relativization. Two common normalized measures are used in the literature as well as among practitioners: insurance penetration, defined as the ratio of insurance premiums on GDP, measures the importance of the insurance sector with respect to the total economy; insurance density, defined as premiums per capita, measures average per capita expenditure. We focus henceforth on premiums per capita. In the same fashion, all variables sub-

[5]Including motor TPL.

ject to a size bias in the information set have been normalized with respect to the relevant benchmark.

4.3 Locational issues

Premium data are registered according to the location of sales point as communicated by the companies. Besides the inevitable aggregation bias due to the arbitrarinesses of administrative boundaries with respect to the geographic dimension of economic phenomena (see Anselin [1]), some important additional biases may arise if the location of sales point is different from the actual location of the insured.

First, mostly for big contracts negotiated by brokers but also for some distribution agreements, e.g., in bancassurance, some big units, usually located in an important industrial or financial center, are accountable for all business nationwide. This happens, for example, for marine insurance premiums collected by business units located in the main harbours for customers located and doing business elsewhere, or for some nationwide salesmen network whose business goes through a single agency, typically located at the company headquarters.

Second, collective policies purchased by the firms as a mandatory cover or as a fringe benefit for their employees, most typically in the accident, health and life classes, are bound to one sales point location even if they are actually insuring risks spread over a wider territory.

4.4 Administrative boundaries in Italy

In the following, we refer to the Italian administrative units called *province*, corresponding to level 3 in the NUTS (Nomenclature of Territorial Units for Statistics) classification by Eurostat, using the generic name of regions, and to the classification used by Istat, the Italian statistical office, when speaking of macro-regions. Macro-regions divide the 20 NUTS2 Italian regions (*regioni*) into 5 aggregates: North-West, North-East, Centre, South and Islands.

5 How to measure Trust and Social Capital?

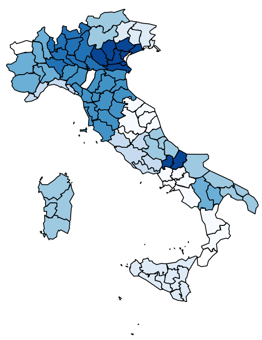
In the third section we tackled one of the major problems pointed out by Durlauf and Fafchamps [8], which is to give a funded economical meaning to Social Capital and Trust. Now we have to address a second controversial issue: a reasonable empirical measure of these sociological concepts. As far as Trust s concerned, like Guiso, Sapienza and Zingales [13] we used results from a survey called "World Value Survey" run in 1999 in Italy. The question asked was

> “Using the responses of this card, could you tell me how much you trust other Italians in general? (5) Trust them completely, (4) Trust them a little, (3) Neither trust them, nor distrust, (2) Do not trust them very much, (1) Do not trust them at all”

Answers provided by the "World Value Survey" are aggregated at regional level. This could generate a potential collinearity problem with the macro–areas dummies, nevertheless Trust index values don’t seem to follow exactly a north–south gradient:

trust	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
North West	3.172	3.313	3.313	3.316	3.371	3.371
North East	3.132	3.22	3.352	3.302	3.386	3.398
Centre	3.068	3.11	3.185	3.239	3.351	3.351
South	3.029	3.091	3.244	3.201	3.247	3.625
Islands	3.172	3.172	3.172	3.191	3.236	3.236

Figure 2: geographical distribution of Trust



Obtaining a measure for the stock of Social Capital is more controversial. Nevertheless, our definition suggests a somewhat natural way to measure it. As we stated in the previous section, what matter is closure of the Social Network, meaning the density and cohesivness of social networks of each province. In other words, we need to measure to what extent each province is characterized by closed and dense social networks. We are not the first to try to measure closure with this kind of data: Goldin and Katz [11] based their empirical measure of Social Capital intensity directly on Coleman’s definition of closure. They have a dataset on schooling and some economic variables on Iowa, USA in 1915. The detail is at county level, comparable to Italian provinces. Their measure was the proportion of county population living in small towns. Their claim was that

> Small town in America was a locus of associations (religious, fraternal/sororal, business, and political organizations) that could have played an important role in galvanizing support for the provision of local publicly provided goods [...]. These associations [...] provide another indicator of community cohesion.

As they did, we want to measure closure of social networks with the dimension and isolation of communities. Goldin and Katz measure can be replicated for our data (as we did), but it's not sufficient to identify isolated communities: we have to deal with a high population density compared with Iowa in 1915. Therefore, our claim is that the degree of closure of social networks characterizing an Italian province is identified by three variables: the percentage of population living in towns with less than 1000 citizens (pupop1000); the fraction of province's hill territory (percsup.c) and the fraction of mountainous territory (percsup.m); the fraction of territory devoted to agriculture (percsup.agr). As for trust, all those variables do not follow a north–south gradient:

pupop1000	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
North West	0.3158	0.5512	0.7080	0.6492	0.7805	0.8443
North East	0	0.1183	0.4085	1.6490	2.0880	13.780
Centre	0	0.4006	0.7385	1.6300	1.6120	14.430
South	0	0	1.936	2.901	2.612	20.520
Islands	0	0	0.2445	2.0190	1.9270	12.670

percsup.m	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
North West	0	9.078	44.960	43.180	64.310	100
North East	0	0	24.540	29.170	40.200	100
Centre	0	7.080	31.680	31.020	42.480	85.320
South	0	3.990	29.730	32.120	54.200	100
Islands	0	0	11.100	16.860	30.680	66.300

percsup.c	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
North West	0	6.503	18.700	25.240	38.250	97.290
North East	0	0	20.380	23.120	35.910	100
Centre	0	47.310	65.500	60.580	74.140	100
South	0	32.100	52.950	47.590	60.980	80.910
Islands	33.700	53.520	65.200	64.610	73.880	86.970

percsup.agr	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
North West	0.0684	0.1911	0.3766	0.4254	0.6884	0.9101
North East	0.1173	0.4370	0.6626	0.5735	0.7328	0.8843
Centre	0.1717	0.4133	0.5166	0.5035	0.6147	0.7603
South	0.2202	0.5632	0.6638	0.6372	0.7545	0.9197
Islands	0.3158	0.5512	0.7080	0.6492	0.7805	0.8443

Figure 3: geographical distribution of pupop1000 and agricultural land

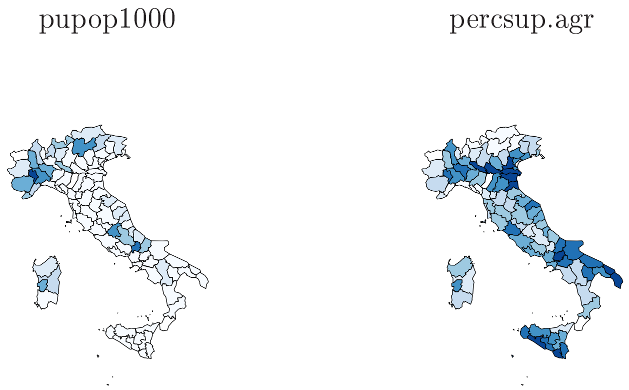
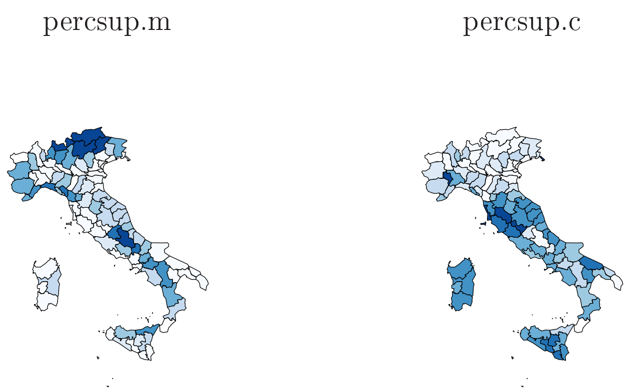


Figure 4: geographical distribution of mountainous and hill territory



6 Model estimation and results

The dataset we use is a balanced panel: we have 103 observations (one for each province) observed over three years. A pooled OLS is likely to be inefficient, since the IID hypthesis on the error terms is usually violated in panel data. Once the longitudinal dimension of the dataset is taken into account, such a hypothesis can be tested. If it is rejected, the choice is then between a fixed–effects and a random–effects model. In our case we are forced to choose a random–effects model: fixed–effects estimators are based on within–group heterogeneity, i.e. they require all the explanatory variables to vary within each group (in our case, within each province). Two of our key–explanatory variables are based on province's territory shape, which is clearly invariant. Even excluding these regressors, many other variables have a low variability across years and within each province, reducing efficiency of a FE estimator[6].

6.1 The panel model

Therefore, the econometric model to be estimated in its most general form is the following error components model:

$$y_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + \nu_i + \epsilon_{it} \quad i = 1, \ldots, 103;\ t = 0, \ldots, 2 \tag{3}$$

Where $\boldsymbol{X}$, u_i and ϵ_{it} are independent of each other. y_{it} is the log of non–life insurance premium paid per capita in province i in year $1998 + t$.

Defining $\xi_{it} = \nu_i + \epsilon_{it}$, the assumption that shocks are independent can be be rewritten as

$$\begin{array}{rcll} \mathrm{Var}(\xi_{it}) & = & \sigma_\nu^2 + \sigma_\epsilon^2 & \\ \mathrm{Cov}(\xi_{it}, \xi_{is}) & = & \sigma_\nu^2 & \forall t \neq s \\ \mathrm{Cov}(\xi_{it}, \xi_{js}) & = & 0 & \forall t \neq s, i \neq j \end{array}$$

A test for the Random Effects model against a pooled OLS is a test for

$$H_0 : \sigma_\nu^2 = 0$$
$$H_1 : \sigma_\nu^2 > 0$$

Assuming normality of the errors, a parsimonious testing strategy can be based on the Lagrange Multiplier principle: the OLS model is estimated and then maintained, while it is compared to the more general alternative in a maximum likelihood framework. Test statistics are based on the OLS residuals without need to estimate the panel model. Baltagi [4] reports the original LM test derived by Breusch and Pagan together with some refinements. We ran the King and Wu modification, which is distributed as a

[6] see the summary table in the appendix

standard normal[7]. The result of the test is 0.8895, with p–value equal to 0.1869, thus not providing any evidence in favor of the random effects model.

Relaxing the assumption of "well behaved" residuals (see 5) and (6 below), another test for the random effects hypothesis feasible in short panels is given in Wooldridge ([18]). This is based on estimation of ${\sigma_\nu}^2$ from the upper triangle of the N empirical Ω blocks given by the outer product of the residuals vectors $\tilde{v}_i = (\tilde{v}_{i1}, \ldots, \tilde{v}_{iT})$. The result of the test is 5.4713, with p–value smaller than 10^{-7}, this time favoring the random effects model. As random effects estimators remain consistent under the OLS specification, we proceed estimating a random effects model.

6.2 The random effects model

Under the random effects specification and under homoskedasticity in both ν_i and ϵ_{it} and no serial correlation in ϵ_{it}, the variance-covariance matrix of the errors becomes

$$V = {\sigma_\nu}^2 (I_N \otimes \boldsymbol{i}_T \boldsymbol{i}'_T) + {\sigma_\epsilon}^2 (I_N \otimes I_T) \tag{4}$$

where I_N is the $N \times N$ identity matrix and $\boldsymbol{i}_N$ is a $N \times 1$ vector of 1. Therefore, V is block-diagonal with

$$V = I_N \otimes \Omega \tag{5}$$

where

$$\Omega = \begin{bmatrix} \sigma^2_\epsilon + \sigma^2_\nu & \sigma^2_\nu & \ldots & \sigma^2_\nu \\ \sigma^2_\nu & \sigma^2_\epsilon + \sigma^2_\nu & \ldots & \vdots \\ \ldots & & \ddots & \sigma^2_\nu \\ \sigma^2_\nu & & & \sigma^2_\epsilon + \sigma^2_\nu \end{bmatrix} \tag{6}$$

Observations regarding the same province share the same ν_i effect, thus the relative errors are autocorrelated, with $Corr(v_{is} v_{it}) = \frac{\sigma^2_\nu}{(\sigma^2_\epsilon + \sigma^2_\nu)}$. Ordinary least squares estimates for β in model (3) are therefore inefficient, though consistent. Generalized least squares (GLS) are the efficient solution if Ω is known. Various feasible GLS procedures exist drawing on consistent estimators of Ω.

The standard approach to random effects panels is to assume both (5) and (6). In "large N" panels a less restrictive approach is possible, termed *general GLS* estimator [18], which allows for arbitrary *intra-group* heteroskedasticity and serial correlation of errors , i.e. inside the Ω covariance blocks, provided

[7] This is a locally mean most powerful refinement of the usual Breusch–Pagan χ^2 test. Breusch and Pagan test $H_0 : \sigma^2_\nu = 0$ against $H_1 : \sigma^2_\nu \neq 0$, thus rejecting for $\sigma^2_\nu < 0$, which should be excluded by the model restrictions. The original Breusch and Pagan test strongly rejects the null.

that these remain the same for every individual. Formally, it assumes (5), while the only requisite for Ω is, of course, symmetry:

$$[\Omega]_{ij} = [\Omega]_{ji} = \sigma_{ij} \tag{7}$$

Wooldridge describes this as an attractive alternative, provided that $N \gg T$, because, by Gauss' formula, the number of variance parameters to be estimated with NT data points is $T(T+1)/2$. This fits our case very well. He suggests estimating the pooled OLS residuals, then averaging over all the N empirical Ω blocks given by the outer product of the residuals vectors $\tilde{v}_i = (\tilde{v}_{i1}, \ldots, \tilde{v}_{iT})$:

$$\hat{\Omega} = \frac{1}{N} \sum_{i=1}^{N} \tilde{v}_i \tilde{v}_i' \tag{8}$$

For the sake of robustness, we try out both estimators. Results are much alike; GGLS are reported in Table 1.

Marginal effects for Trust and the Social Capital variables are[8]:

	marg.eff	std.dev	t–stat	p–value
trust	0.510440	0.125794	4.057747	0.000064
pupop1000	0.009716	0.003731	2.604263	0.009708
percsup.m	0.001926	0.000976	1.973811	0.049404
percsup.c	0.000347	0.000723	0.479762	0.631778
percagricol	0.026837	0.118823	0.225861	0.821478

Trust and Social Capital effect are not completely absorbed by equilibrium prices: supply side proxies (in particular log(ag/pop) do have a positive effect but Trust and Social Capital marginal effects are significant (at least three out of five of them), while all the signs are positive as expected. Further on, significance of the interaction parameters suggest for a non–linear dependence on our Social Capital proxies. Those results, though supporting our model, are not conclusive. In order to have a deeper understanding of the role of trust and social capital we further investigate the spatial structure of the empirical model.

7 Spatial structure

As we said while describing insurance data, there are good reasons to think that non–life insurance may not be bounded by provincial administrative boundaries. For example, these may overlap with operational areas of the sales force, or there may be any other kind of cross-border purchase. As

[8]marginal effects for Social Capital variables are computed over the mean of the relevant variables

Table 1: GGLS estimates of the RE model

	coef	se	t	pt
(Intercept)	−7.232032	1.720103	−4.204418	0.000035
log(Ydproc)	1.156512	0.165342	6.994670	0.000000
I(pop25.54/popover60)	0.268546	0.119293	2.251143	0.025149
inef	−0.051496	0.011580	−4.447068	0.000013
NO	0.045594	0.055319	0.824209	0.410520
NE	0.084450	0.049043	1.721975	0.086174
SU	−0.255475	0.054661	−4.673777	0.000005
IS	−0.288996	0.064639	−4.470903	0.000011
dum98	−0.094306	0.012889	−7.316894	0.000000
dum99	−0.036897	0.009717	−3.797031	0.000179
I(den/1000)	0.102903	0.051126	2.012731	0.045097
I(va.indutot/va)	0.405786	0.442339	0.917366	0.359738
I(va.serv/va)	0.368092	0.434076	0.847990	0.397165
u	−0.000067	0.001900	−0.035110	0.972017
qexport	0.022075	0.092149	0.239558	0.810848
numcompfam	0.016147	0.109742	0.147132	0.883133
trust	0.510440	0.125794	4.057747	0.000064
pupop1000	0.134755	0.033928	3.971815	0.000091
percsup.m	0.002976	0.001112	2.676534	0.007876
percsup.c	0.000725	0.000772	0.938466	0.348811
log(dep/pop)	0.167496	0.051477	3.253771	0.001278
I(va/1000)	0.002882	0.001207	2.386959	0.017650
lrpro	0.014176	0.023118	0.613210	0.540234
percsup.agr	0.117385	0.134327	0.873875	0.382933
log(ag/pop)	0.167453	0.054103	3.095078	0.002166
pupop1000:percsup.m	−0.001326	0.000303	−4.377530	0.000017
pupop1000:percsup.c	−0.000477	0.000135	−3.522188	0.000499
pupop1000:percsup.agr	−0.114299	0.036413	−3.138982	0.001876

in many other studies about the spatial distribution of an economic phenomenon, this problem cannot be neglected. In particular, Lenzi and Millo [14] found evidence of spatial correlation for several specifications of regressions of insurance on a set of demographics, based on the very same dataset.

In econometric applications, proximity between data points in space is usually characterized by means of a *proximity matrix*, say, W, containing a measure of proximity for every pair of data points and, by convention, setting the diagonal to zero. Hence a *spatial lag operator* is defined such that Wy, the *spatial lag* of y, stands for "the values of y at *neighboring* locations"[9]. Anselin [1] warns about the relevant consequences on estimation (and, to a lesser extent, on testing) of the choice of W. Here we resorted to the simplest form of proximity matrix: a binary matrix with values set to 1 if the corresponding provinces share a common border, 0 otherwise. This has been row-standardized, so that the spatial lag of y, Wy, is simply the average of values of y at neighboring locations. Many agree on the perception that this kind of matrix be somehow the most "neutral", imposing the minimum amount of *a priori* structure to the data, and in fact it is perhaps the most widely used in empirical applications.

The two standard specifications for spatial effects in regression models are the *spatial lag* (SAR) model:

$$y = \rho W y + X\beta + \epsilon \tag{9}$$

and the *spatial error* (SEM) model:

$$\begin{aligned} y &= X\beta + e \\ e &= \lambda W e + \epsilon \end{aligned} \tag{10}$$

Spatial effects may be related to the exclusion from the model of some unobservable latent variable that be spatially correlated. Omission of the latter may reflect in a spatial structure in the residuals, or it can be successfully modeled by inclusion of the lagged dependent variable (which is formally equivalent to including the sum of a power series of the lagged explanatory variables: see [2]).

The consequences on estimation of omitting the lagged dependent variable are inconsistency and biasness of parameter estimates. Neglecting a spatial error structure has less serious consequences: estimates, while still consistent, are inefficient, thus reducing test power.

Stacking the data as one cross section for every point in time and assuming $\epsilon \sim IID$, in the case of panel data, the SAR models becomes

$$y = \rho(I_T \otimes W)y + X\beta + (i_T \otimes \mu) + \epsilon$$

and for the SEM case,

$$y = X\beta + (i_T \otimes \mu) + e$$

$$e = \lambda(I_T \otimes W)e + \epsilon$$

7.1 Testing for spatial effects in the panel model

Testing for spatial effects may be based on one of the three classical likelihood-based test procedures. Anselin [1] suggests the LM approach, which has the

[9] See [1], Ch.3, for a classic treatment

advantage of requiring only estimation of the non-spatial model. The check for spatial dependence may thus be based on the OLS residuals.

Those tests should be reformulated for panel models. Observing that panel estimation is a special case of GLS and thus is equivalent to OLS estimation on suitably transformed data (see e.g. [4]), one may run the standard tests on the OLS residuals of the transformed models.

Testing on a specification without social capital variables, the LM test *à la Anselin* (see [1]) rejects the null hypothesis at the 10 percent level, therefore finding some evidence in favor of the SAR alternative (test value 2.8466, p–value 0.09157). The augmented specification, on the contrary, passes the test at any level (test value 0.1226, p-value 0.7262). The non-spatial specification is instead not rejected against an alternative SEM specification (test value 0.5014, p-value 0.4789) and this result is much unchanged augmenting the model with the social capital variables. These results suggest to further explore the possibility of a SAR specification.

7.2 Spatial panel estimation

As estimators for spatial panel models are now available, we also proceeded from a general to specific perspective, estimating (on the full specification, with Social Capital variables) a RE model with a SAR structure and one with a SEM structure following Elhorst [9]. Given the estimation results (see the Appendix), tests for the SAR and, respectively, the SMA form are simple t–tests on ρ and λ:

	est.	std.dev	t–stat	p–value
ρ	-0.005248	0.019870	-0.264119	0.791688
λ	-0.1481	0.091801	-1.613794	0.106572

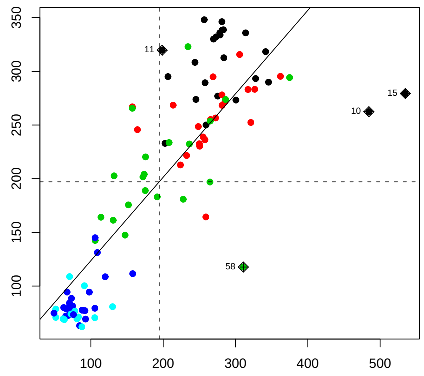
Both parameters are deemed insignificant. Such a result could rise from the fact that actually there is no spatial correlation in non–life insurance, which is at odds with Lenzi and Millo [14], or, as we think, this means that the spatial correlation of non–life insurance is due to the spatial distribution of Social Capital. A first evidence in this sense comes from the moran plots of non–life insurance and the social capital variables we chose (see figure 5[10]).

Moran I statistic is a spatial correlation measure. If (as in our case) the proximity matrix is a row–standardized dichotomic matrix it is equivalent to a regression coefficient of the variable of interest over its spatial lag (see [1]). The Moran plot is its scatter plot, which means on the x–axis there is the variable of interest and on the y–axis its spatial lag. The straight line is the OLS estimated one. Therefore graphs show that both the variable

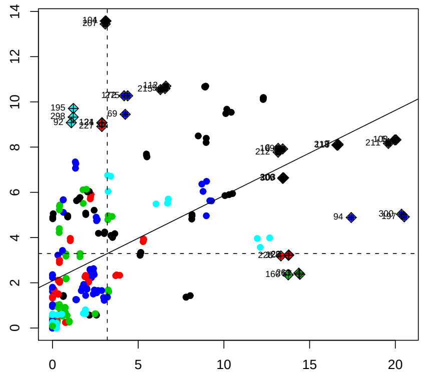
[10]different colors represents different macro–regions: black is North-West, red is North-East, green is Center, blue is South and light blue Islands

Figure 5: Moran plots

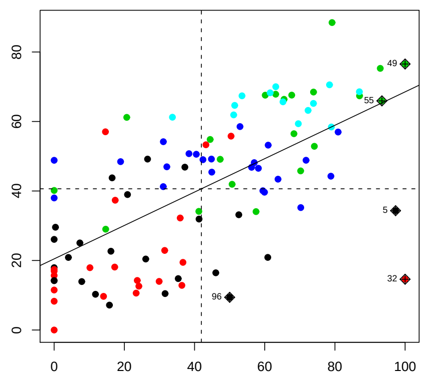
ppcd



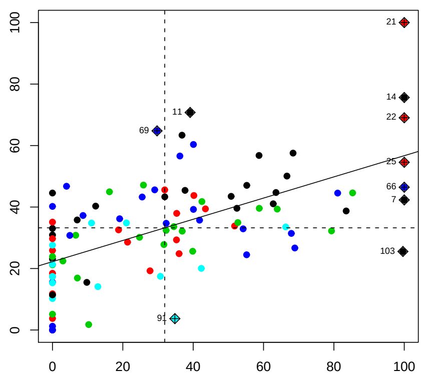
pupop1000



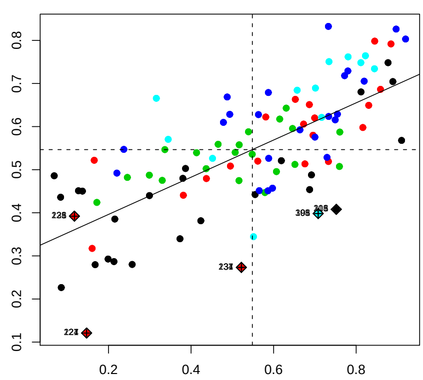
percsup.c



percsup.m



percsup.agr



of interest (ppcd, which are log premium per capita) and the social capital variables exhibit spatial correlation, which is not completely explained by the geographical dummies.

Another piece of evidence comes from estimation of the model once social capital variables are omitted:

	est.	std.dev	t–stat	p–value
ρ^*	0.065138	0.021854	2.980633	0.002877
λ^*	-0.133542	0.092655	-1.441271	0.149508

Results of these tests are in line with Lenzi and Millo [14]: a panel model without social capital effects exhibits a significant Spatial autocorrelation structure ($\rho \neq 0$). Augmenting the model with social capital variables eliminates the need for a spatial specification. These results do confirm that the spatial autocorrelation of non–life insurance is due to Social Capital spatial distribution.

7.3 Issues in characterizing proximity

As we said in the previous section, Anselin [1] points out the possible bias introduced by a wrong choice of the proximity matrix W. We performed a robustness check employing two different matrices W based on the inverse of distance measures: one on distance in space between provinces' administrative centres and the other on road travelling distance. A cutoff has been introduced setting the distance to infinity if it was over a certain threshold, thus constraining the relative influence to zero. It should be noted that this is only affecting the speed of distance decay in the intensity of spillovers, as the very structure of both SAR and SEM models allow for the influence to spread across *all* of the regions anyway, as it is proved by Anselin [2].

The results of the two alternative distance-based specifications are much alike each other, and somewhat stronger than those of the binary contiguity case. We report those for a coordinates' distance matrix with cutoff point at 100 km. The LM tests *à la Anselin* fail to find significant spatial effects in the complete model (p-values lag: 0.5668, error: 0.1258), while the one without social capital variables is rejected against the spatial lag (p-value: 0.03677) but not against the spatial error alternative (p-value: 0.3743), again favoring the lag specification. We then turn to the Wald tests.

	est.	std.dev	t–stat	p–value
ρ	0.007373	0.019961	0.369347	0.711869
λ	-0.192933	0.103535	-1.863448	0.062399
ρ^*	0.085277	0.021729	3.924634	0.000087
λ^*	-0.182558	0.104732	-1.743096	0.081317

Parameters with asterisks refer to the model without social capital variables. The estimates of the spatial models confirm the finding that significant spatial lag effects arise from the omission of social capital variables. In contrast to the binary contiguity specification, the spatial process in the errors is found significant at the 10 percent level.

8 Conclusions

We started from Arnott and Stiglitz model on the co–existence of marketed and non–marketed insurance contracts. We extended it to allow for Social Capital and Trust as potential explanatory variables. We chose a network approach: non–market agreement are described as strategic decisions of agents playing a prisoners' dilemma type of game with their neighbors. Each of them adopt a trigger strategy to punish neighbors deviating from the co-operative equilibrium in any game they are involved. Such a behavior lead to a Pairwise Stable Equilibrium which is more likely the higher the level of trust and social capital embedded in the social network. Here comes the first contribution of our paper: the network approach we chose provide us with a formal definition of social capital and trust, which is crucial to obtain a clear testable model. The empirical part is carried out on a province–level Italian dataset provided by istituto Tagliacarne. We estimated a RE panel model, and our testable implication, which was of a positive marginal effect both of trust and social capital on demand for marketed non–life insurance, is confirmed. Further on, we are able to explain the spatial correlation found by Lenzi and Millo on the very same dataset by means of the spatial structure of our new explanatory variables.

References

[1] L. Anselin. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publisher, University of California, Santa Barbara, 1988.

[2] L. Anselin. Spatial externalities, spatial multipliers, spatial econometrics. Technical report, University of Illinois, Urbana USA, August 2002.

[3] Richard Arnott and Joseph E. Stiglitz. Moral hazard and nonmarket institutions: Dysfunctional crowding out or peer monitoring? *The American Economic Review*, 81(1):179–190, March 1991.

[4] Badi H. Baltagi. *Econometric Analysis of Panel Data*. John Wiley & Sons, UK, 1995.

[5] M. Browne, J.W. Chung, and E. Frees. International property-liability insurance consumption. *The Journal of Risk and Insurance*, 67, N°1, 2000.

[6] M.J. Browne and K. Kim. An international analysis of life insurance demand. *The Journal of Risk and Insurance*, 60:617–634, 1993.

[7] G. Dionne. *Handbook of Insurance*. Kluwer Academic Publishers, Dordrecht Netherlands, 2000.

[8] Steven N. Durlauf and Marcel Fafchamps. Social capital. *NBER Working Paper series*, July 2004.

[9] J. Paul Elhorst. Panel data models extended to spatial error autocorrelation or a spatially lagged dependent variable. Technical report, 2001.

[10] N. Esho, R. Zurbruegg, A. Kirievsky, and D. Ward. Law and the determinants of property-casualty insurance. Technical report, Australian Prudential Regulation Authority, June 2001.

[11] Claudia Goldin and Lawrence F. Katz. Human capital and social capital: the rise of secondary schooling in america, 1910 to 1940. *NBER working paper series*, (6439), March 1998.

[12] M. Grace and H. Skipper. An analysis of the demand and supply determinants for non–life insurance internationally. Technical report, CRMIR, Georgia State University, 1991.

[13] Luigi Guiso, Paola Sapienza, and Luigi Zingales. The role of social capital in financial development. *American Economic Review*, 94(3):526–556, June 2004.

[14] Giovanni Millo and Annapaola Lenzi. Regional heterogeneity and spatial sillovers in the italian insurance market. *Assicurazioni Generali Research Dept. Working Paper series*, 1, 2005.

[15] J. Mossin. Aspects of rational insurance purchasing. *Journal of Political Economy*, 79:553–568, 1968.

[16] Robert D. Putnam, R. Leonardi, and R. Nanetti. *Making Democracy Work: Civic Traditions in modern Italy.* Princeton University Press, Princeton, USA, 1993.

[17] Fernando Vega–Redondo. Building up social capital in a changing world. *Journal of Economic Dynamics & Control*, 30:2305–2338, 2006.

[18] J.M. Wooldridge. *Econometric analysis of cross-section and panel data.* MIT Press, 2002.

A Proof of proposition 1

The normalized payoff functions in case i cooperates with j is

$$\begin{aligned}\pi_i(s_C^g) &= \sum_{k\in N_i}\left\{(1-\delta)\sum_{\tau=0}^{\infty}\delta^{\tau}EU_{ij}^{NMO}\right\}\\ &= \sum_{k\in N_i}EU_{ij}^{NMO}\end{aligned}$$

while if i deviates her anticipated payoff is

$$\begin{aligned}\pi_i(s_D^g) &= (1-\delta)EU_{ij}^H + \sum_{k\in N_i/\{j\}}\left\{\left[\sum_{s=0}^{d^i(j,k)-1}(1-\delta)\delta^s EU_{ik}^{NMO}\right] + (1-\delta)\delta^{d^i(j,k)}EU_{ik}^L\right\}\\ &= (1-\delta)EU_{ij}^H + \sum_{k\in N_i/\{j\}}\left\{\left[\sum_{s=0}^{d^i(j,k)-1}(1-\delta)\delta^s EU_{ik}^{NMO}\right] + (1-\delta)\delta^{d^i(j,k)}EU_{ik}^{NMO}\right.\\ &\quad\left. -(1-\delta)\delta^{d^i(j,k)}\nu_{ik}\right\}\\ &= (1-\delta)EU_{ij}^H + \sum_{k\in N_i/\{j\}}\left\{\left(1-\delta^{d^i(j,k)+1}\right)EU_{ik}^{NMO} - (1-\delta)\delta^{d^i(j,k)}\nu_{ik}\right\}\end{aligned}$$

Therefore, the stability condition

$$\pi_i(s_C^g) \geq \pi_i(s_D^g)$$

Can be rewritten as

$$\sum_{k\in N_i}EU_{ij}^{NMO} \geq (1-\delta)EU_{ij}^H + \sum_{k\in N_i/\{j\}}\left\{\left(1-\delta^{d^i(j,k)+1}\right)EU_{ik}^{NMO} - (1-\delta)\delta^{d^i(j,k)}\nu_{ik}\right\}$$

$$EU_{ij}^{NMO} + \sum_{k\in N_i/\{j\}}\left(1-1+\delta^{d^i(j,k)+1}\right)EU_{ik}^{NMO} \geq (1-\delta)\left[EU_{ij}^H - \sum_{k\in N_i/\{j\}}\delta^{d^i(j,k)}\nu_{ik}\right]$$

$$EU_{ij}^{NMO} + \sum_{k\in N_i/\{j\}}\left\{\delta^{d^i(j,k)+1}EU_{ik}^{NMO} + (1-\delta)\delta^{d^i(j,k)}\nu_{ik}\right\} \geq (1-\delta)EU_{ij}^H$$

$$EU_{ij}^{NMO} + \sum_{k\in N_i/\{j\}}\delta^{d^i(j,k)}\left[\delta EU_{ik}^{NMO} + (1-\delta)\nu_{ik}\right] \geq (1-\delta)EU_{ij}^H \qquad (11)$$

Which is in the form of proposition 1.

B Variables' description and descriptive statistics

Ydproc disposable income per capita

pop25.54/popover60 ratio of people aged 25-54 to people aged over 60

inef indicator of juridical system inefficiency: average duration of civil trials

den/1000 population density, inh. per sq. Km (scaled by a factor of 1000)

va.indutot/va share of industry on value added

va.serv/va share of services on value added

u unemployment rate

qexport share of export on total value added

numcompfam average number of family members

lrpro loss ratio of the property sector

trust trust indicator as defined by the World Values Survey (see above)

pupop500 share of population living in towns with less than 500 inhabitants

percsup.m share of mountainous territory

percsup.c share of hill territory

percsup.agr share of the land devoted to agriculture

dep/pop bank deposits per capita

va/1000 total value added (scaled by a factor of 1000)

ag/pop ratio of number of agencies over province's population

A table with some descriptive statistics follows.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
log(Ydproc)	9.00	9.27	9.54	9.47	9.63	9.84
I(pop25.54/popover60)	0.74	0.90	1.02	1.04	1.16	1.58
inef	1.44	2.74	3.47	3.79	4.59	8.32
I(den/1000)	0.04	0.10	0.17	0.24	0.26	2.66
I(va.indutot/va)	0.11	0.21	0.28	0.28	0.33	0.46
I(va.serv/va)	0.52	0.63	0.68	0.68	0.74	0.85
u	1.71	5.01	7.55	10.90	16.14	33.16
qexport	0.01	0.09	0.20	0.20	0.30	0.63
numcompfam	2.05	2.46	2.61	2.62	2.78	3.15
trust	3.03	3.17	3.25	3.26	3.35	3.63
pupop1000	0.00	0.30	1.38	3.20	3.28	20.52
percsup.m	0.00	0.00	30.68	31.92	52.43	100.00
percsup.c	0.00	17.25	42.40	41.95	63.14	100.00
log(dep/pop)	1.35	1.78	2.20	2.11	2.38	3.09
I(va/1000)	1.27	4.21	6.22	10.04	10.18	112.10
lrpro	0.25	0.43	0.49	0.52	0.59	1.82
percsup.agr	0.07	0.38	0.59	0.55	0.73	0.92
log(ag/pop)	−8.98	−8.01	−7.73	−7.83	−7.62	−7.32
pupop1000:percsup.m	0.00	0.00	28.90	158.40	121.10	1666.00
pupop1000:percsup.c	0.00	0.00	35.85	107.50	104.00	1949.00
pupop1000:percsup.agr	0.00	0.18	0.56	1.41	1.53	15.07

C Full results of spatial model estimation

In the following we report the full results from estimation of the spatial models.

C.1 Spatial lag model (SAR) with Social Capital variables

	coef	se	z	pz
log(Ydproc)	1.218524	0.172764	7.053097	0.000000
I(pop25.54/popover60)	0.217724	0.122657	1.775066	0.075887
inef	−0.052722	0.012918	−4.081142	0.000045
NO	0.066210	0.060111	1.101459	0.270697
NE	0.099271	0.053975	1.839213	0.065884
SU	−0.268589	0.060674	−4.426759	0.000010
IS	−0.289358	0.071175	−4.065432	0.000048
dum98	−0.080522	0.011590	−6.947653	0.000000
dum99	−0.027044	0.009116	−2.966616	0.003011
I(den/1000)	0.104357	0.056840	1.835985	0.066360
I(va.indutot/va)	0.439971	0.446897	0.984501	0.324869
I(va.serv/va)	0.352516	0.445951	0.790483	0.429246
u	−0.000852	0.001780	−0.478701	0.632151
qexport	0.049073	0.081996	0.598483	0.549518
numcompfam	0.040674	0.108317	0.375505	0.707285
trust	0.523101	0.139838	3.740763	0.000183
pupop1000	0.122108	0.037169	3.285181	0.001019
percsup.m	0.002703	0.001227	2.202421	0.027636
percsup.c	0.000758	0.000863	0.878720	0.379553
log(dep/pop)	0.086609	0.048288	1.793576	0.072881
I(va/1000)	0.003094	0.001306	2.368547	0.017858
lrpro	0.015925	0.021266	0.748877	0.453931
percsup.agr	0.056858	0.147064	0.386618	0.699039
log(ag/pop)	0.131509	0.050119	2.623931	0.008692
pupop1000:percsup.m	−0.001227	0.000332	−3.699087	0.000216
pupop1000:percsup.c	−0.000501	0.000151	−3.330730	0.000866
pupop1000:percsup.agr	−0.096373	0.039867	−2.417356	0.015634
rho	−0.005248	0.019870	−0.264119	0.791688

C.2 Spatial lag model (SAR) without Social Capital variables

	coef	se	z	pz
log(Ydproc)	1.360972	0.175751	7.743740	0.000000
I(pop25.54/popover60)	0.212975	0.122578	1.737473	0.082304
inef	−0.045435	0.013518	−3.361027	0.000777
NO	0.155460	0.050704	3.066038	0.002169
NE	0.091050	0.049143	1.852779	0.063914
SU	−0.213453	0.063736	−3.349030	0.000811
IS	−0.161358	0.073394	−2.198509	0.027913
dum98	−0.078543	0.011942	−6.577066	0.000000
dum99	−0.023713	0.009402	−2.522057	0.011667
I(den/1000)	0.061304	0.055386	1.106833	0.268366
I(va.indutot/va)	0.564502	0.450134	1.254077	0.209814
I(va.serv/va)	0.365946	0.439559	0.832529	0.405110
u	−0.000553	0.001818	−0.304221	0.760960
qexport	0.097812	0.083950	1.165124	0.243969
numcompfam	0.083204	0.107819	0.771706	0.440289
log(dep/pop)	0.088447	0.049374	1.791370	0.073234
I(va/1000)	0.002641	0.001326	1.991098	0.046470
log(ag/pop)	0.159350	0.050617	3.148162	0.001643
lrpro	0.013229	0.021557	0.613689	0.539421
rho	0.065138	0.021854	2.980633	0.002877

C.3 Spatial error model (SEM) with Social Capital variables

log(Ydproc)	1.207	0.1716	7.032	0
I(pop25.54/popover60)	0.2091	0.1209	1.729	0.08381
inef	−0.05227	0.01293	−4.042	5.3e−05
NO	0.06679	0.05983	1.116	0.2643
NE	0.1032	0.05375	1.92	0.05484
SU	−0.2698	0.06058	−4.454	8e−06
IS	−0.2896	0.07119	−4.068	4.7e−05
dum98	−0.07961	0.01113	−7.151	0
dum99	−0.02671	0.008636	−3.093	0.001982
I(den/1000)	0.1055	0.05689	1.854	0.06373
I(va.indutot/va)	0.3985	0.4416	0.9024	0.3668
I(va.serv/va)	0.2958	0.4367	0.6773	0.4982
u	−0.00123	0.001744	−0.7051	0.4807
qexport	0.02745	0.08157	0.3365	0.7365
numcompfam	0.04728	0.1068	0.4429	0.6578
trust	0.5242	0.1399	3.746	0.000179
pupop1000	0.1214	0.0372	3.263	0.001101
percsup.m	0.002681	0.001228	2.183	0.02905
percsup.c	0.000799	0.000864	0.925	0.355
log(dep/pop)	0.08077	0.0469	1.722	0.08507
I(va/1000)	0.003283	0.001302	2.521	0.0117
lrpro	0.0129	0.02085	0.6186	0.5362
percsup.agr	0.04581	0.1468	0.3121	0.755
log(ag/pop)	0.1294	0.0484	2.673	0.00752
pupop1000:percsup.m	−0.00122	0.000332	−3.677	0.000236
pupop1000:percsup.c	−0.000509	0.000151	−3.377	0.000732
pupop1000:percsup.agr	−0.09475	0.03987	−2.376	0.01749
lambda	−0.1481	0.0918	−1.614	0.1066

C.4 Spatial error model (SEM) without Social Capital variables

log(Ydproc)	1.386	0.1758	7.888	0
I(pop25.54/popover60)	0.2149	0.122	1.761	0.07821
inef	−0.04457	0.01369	−3.255	0.001135
NO	0.1781	0.05096	3.495	0.000474
NE	0.1037	0.04952	2.095	0.03617
SU	−0.2554	0.06436	−3.968	7.2e−05
IS	−0.2052	0.07418	−2.766	0.005672
dum98	−0.08324	0.01155	−7.207	0
dum99	−0.02587	0.008984	−2.879	0.003987
I(den/1000)	0.06174	0.05607	1.101	0.2708
I(va.indutot/va)	0.5461	0.4471	1.222	0.2219
I(va.serv/va)	0.3044	0.433	0.703	0.482
u	−0.001131	0.001784	−0.6337	0.5263
qexport	0.08111	0.0837	0.969	0.3325
numcompfam	0.08174	0.1069	0.7644	0.4446
log(dep/pop)	0.08701	0.04821	1.805	0.07109
I(va/1000)	0.00254	0.001337	1.899	0.05759
log(ag/pop)	0.1643	0.04917	3.342	0.000832
lrpro	0.01032	0.02114	0.488	0.6255
lambda	−0.1335	0.09266	−1.441	0.1495