Social Capital: A Standard Method of Measurement

(15-12-00)

(*) Lene Hjøllund, Research assistant, cand.oecon, Department of Economics, The Aarhus School of Business, lehj@asb.dk

Gert Tinggaard Svendsen, gts@asb.dk, http://www.hha.dk/eok/nat/staff/gts_form.htm
Associate Professor, Department of Economics
The Aarhus School of Business
Fuglesangs Allé 20
8210 Aarhus V
Denmark

Disposition:

1. Introduction
2. Concept, review
3. Measurement, review
4. Questionnaire
5. Factor analysis
    5.1 Introduction
    5.2: Common factor analysis
    5.3: Component analysis
    5.4: Summary
    5.5: Method of measurement, choice
6. Conclusion

Abstract:
The purpose of this paper is to suggest a standard method of measurement for social capital. Various authors have investigated the influence of social capital on economic growth but still social capital has not been measured in any satisfactory way. So far, each survey has used its own ad hoc methodology due to the heterogeneity of the very definition of the concept of social capital. A consensus concerning a standardized method of measurement, has not yet been reached. Based on the existing theoretical and empirical approaches, we suggest the use of a questionnaire where the principal component analysis should be applied as the most appropriate method of measurement.

JEL classification: A12, C71, D23, D70

Keywords: Social capital, capitalism, communism, factor analysis, principal component analysis, questionnaire.

*) This work is part of the book project Paldam, M. and Svendsen, G.T.: *Trust, Social Capital and Economic Growth: An International Comparison*. Edward Elgar, Cheltenham, UK. Forthcoming. We are grateful to the SSF for funding our part of the project concerning social capital investigations in Denmark and the United States. We also thank the Rockwool Foundation which funds Martin Paldam's part of the project, i.e. to undertake surveys on social capital in Russia, the three Baltic States and Poland.

1: Introduction

The purpose of this paper is to develop a method of measurement for social capital. Paldam (2000) describes social capital as "the glue that holds societies together." He divides the theoretical approaches into three families of social capital, namely trust, cooperation, and network. Using the key assumption that trust implies ease of voluntary cooperation, Paldam shows that a similar relationship exists between trust and network. Thus, Paldam concludes, the network definition fits rather well into the trust-cooperation definitions, and "everything might be shades of and approaches to the very same basic phenomenon" (Paldam, 2000: 11). Consequently, this social glue or social capital can be defined as mutual trust in a group. The presence of social capital determines how easily people can work together as transaction costs are lowered - informal self-enforcement of contracts is now possible without third party enforcement. In this way, the presence of social capital may lubricate collective action and diminish the incentive to free ride on any contract or collective good provided.

Social capital has not been measured in any satisfactory way yet, see Paldam and Svendsen (2000a). Thus, our contribution is a first attempt to propose the best methodological approach based on existing theoretical and empirical evidence. A standardized methodology for measuring social capital is strongly needed.

As described below, studies of social capital have been conducted mainly in developing countries at the micro level, or in the United States. The studies that have included other developed countries are mainly cross-national studies using a single measure, for instance that of generalized trust, and examine the correlation with growth rates, i.e. at the macro level. Therefore, we aim to close this gap between purely micro level and "ad hoc" studies of social capital on one side, and the (too) simple macro economic studies on the other. The underlying considerations of constructing the questionnaire for this purpose can be most accurately illustrated by the point that "social capital is an aggregate concept that has its basis in individual behavior, attitudes, and predispositions" (as noted by Brehm and Rahn, 1997). Our questionnaire is enclosed in Appendix I.

After establishing this method of measurement, our next step will be to apply this method to establish comparative studies between Western and Eastern Europe. The general idea in Paldam and Svendsen (forthcoming) is to compare the effect on the level of social capital from two different political systems, namely those of capitalism in Western Europe and communism in Eastern Europe. Though the former East Bloc countries have started implementing market-based reforms since 1989, the stock of social capital has presumably not changed yet as it probably takes centuries to build it (Paldam and Svendsen, 2000a). In a preliminary investigation by Paldam and Svendsen (2000b), they argue that the slowness of the transition of the old communist countries of East and Central Europe is caused by the lack of social capital. Thus, the concept should be particularly relevant to the comparison of the economic performance in Western and Eastern Europe. Our measures for social capital will be applied to a number of countries. In particular, we will focus on Denmark, the United States, Russia, Poland and the three Baltic States and the United States. Except for the US, social capital has not been thoroughly investigated in these countries yet and therefore we will collect data by using the mentioned questionnaire for household interviews.

This comparison may catch different aspects of political culture (shared norms and values) and corruption (quality of institutions) implying different levels of social capital. Consequently, one may predict differences in institutional development and economic performance. However, in order to do this, we first need to develop a method of measurement and an appropriate questionnaire and, as stated above, this is the purpose of this work.

First, Section 2 introduces the concept of social capital. Next, Section 3 surveys the already existing methodological approaches found in literature so far. Section 4 presents our questionnaire. Based on this questionnaire, Section 5 argues that principal component analysis is the most appropriate method of measurement for simplifying data.

2: Concept, review

The concept of social capital is new within the economics area but has old roots within the other social sciences, see Paldam and Svendsen (2000a). Having said that, a slight distinction between sociologists and economists can be claimed though, as sociologists tend to be more diverse in their interpretation of the concept, whereas economists are less so. But even when staying within the science of economics, there is still no complete agreement on the definition of the concept of social capital. Dasgupta (1999) describes in an essay both the sociological and the more economical background for the use and interpretation of social capital.

First of all, social capital can be analyzed at various levels in society, i.e. at the micro level (individual/household/neighborhood level), the meso level (institutions) and the macro level (national level). Or as Narayan and Pritchett (2000) put it, the literature on social capital can be separated into three "streams", the first of which is concerned with country level politics (including the growth aspect), whereas the second focuses the meso-level (efficacy of institutions in a broad sence of the word). Finally the third part of the literature views social capital as a resolution of market failures at the micro level. Concerning measurement, this distinction does not tend to cause any major problems as various authors seem to agree upon this distinction into three levels. At the micro level social capital consists of networks and norms that govern the interactions among individuals, households and communities. At the meso level and the macro level, social capital is made up by the functioning of institutions, rule of law, government etc.

However, when it comes to the question of aggregation, problems arise. How does the individual level of social capital affect the national level – and vice versa? Moreover, the facets of social interaction can be separated into two parts, namely the structural dimension (which facilitates social interaction) and the cognitive dimension (which predisposes people to act in a socially beneficial way). Trust belongs to the latter. These two parts work interactively, and are mutually reinforcing.

Rose (1999) separates the analysis of social capital into three alternative approaches.

1) Situational theories

Social capital is defined in situational and instrumental terms, i.e. it varies from person to person and from situation to situation. This, again, implies that social capital cannot be reduced to a single unit of account and then aggregated into a summary statistic characterizing the whole of society (Coleman, 1990).

2) Social psychological approach

Social capital is a set of cultural beliefs and norms. Supporters of this approach argue that voluntary organizations emerge as a consequence of trust, rather than the reverse, i.e. social capital is in essence equal to trust. Social capital varies from person to person but is situationally consistent (Inglehart, 1997).

3) Culture theory

Culture is the source of trust and cooperation. Social capital is homogeneous among individuals belonging to the same culture (society), as well as consistent from situation to situation. This implies that it should be very simple to identify social capital for a specific country (Fukuyama, 1995).

As can be seen from the above discussion, economists differ in their approach to the theoretical explanation (and therefore treatment) of social capital, reflecting the fact that it is a multifaceted concept. Again, this inevitably means that it will form a big challenge to not only operationalize but also standardize the concept. According to Rose's categorization, we primarily follow the social psychological approach.

3: Measurement, review

Defining social capital at basically three different levels (micro, meso and macro) leads to different measures of social capital. In this section, we will review eight surveys of social capital in terms of their measurement approaches.

The simplest measure was introduced by Putnam (1993) in an analysis of the differences in institutional efficiency (and its influence on economic development) between Northern and Southern Italy. Putnam found that to a large extent this could be explained by the (historically determined) differences in the number of memberships of voluntary organizations. This basic and easily accessible measure has been a point of departure for many of the social capital analyses, and is almost always included as one of the explanatory factors. See the discussion below, and for a more elaborated economic discussion, please consult Paldam (2000).

Krishna and Shrader (1999) have constructed a rather extensive questionnaire in an attempt to encompass all the possible aspects of social capital.[1] At the same time Krishna and Shrader provide an overview of the existing literature within the area of social capital (Krishna and Shrader (1999), annex A). They have divided the surveyed studies into four categories (individual/household level,

[1] It has been developed for the use of the researchers of the World Bank but is available to others who may be interested.

neighborhood/community level, regional level, and national level) according to the above mentioned primary separation of societal levels, and although not exhaustive, it can provide us with a good idea of how the research on social capital has been carried out so far. They undertake 13 studies, out of which the measures of three studies are not specified (only the method is described). Out of the ten studies "left", seven have included the concept of membership of a voluntary organization in some way. Furthermore, four out of the ten studies can be seen to have used the trust measure as an indicator of social capital.

Let us look in more detail at some empirical studies to identify the methods of measurements applied. The studies have been selected to represent a few of the many different approaches to measurement of social capital and to illustrate the variations and/or similarities between the studies of different levels of society, although the micro orientated studies do tend to dominate the literature. For most of the reviewed studies, data have been produced from household surveys (interviews of individuals) which is due to the predominantly micro economic orientation, as well as the fact that they have primarily been carried out in developing countries where less data are readily available. But for a couple of the studies, the data have been retrieved from existing databases of macro and/or micro data.[2]

Rose (1999) defines social capital as the stock of formal or informal social networks that individuals use to produce or allocate goods and services. Accordingly, his purpose is to identify the extent of formal and informal networks as well as their interaction. Using Rose's own categorization, he commits himself to the situational approach, while at the same time he considers membership of voluntary organizations to be an important indicator of a society's level of social capital. But this is not (necessarily) a contradiction. Finally, he emphasizes the lack of valid empirical indicators of social capital, even in the data-rich OECD countries.

Rose does not undertake any econometric analysis of the data, as such. He simply reports the marginal distribution of the answers for the various questions, and from there he draws his conclusions. However, he does describe the considerations on formulating and selecting the questions (or situations) to ask (about). Rose claims that the situations should be relevant to a majority of households regardless of economic status etc. and it should be situations in which formal organizations are expected to deliver the goods or services. Finally, in every question focus should be on a particular good or service, and it should be left open whether the respondent relies on a formal organization or an

[2] The General Social Survey provides an extensive coverage of the U.S. individual data. Similarly, the World Value Survey provides data on slightly different aspects for 43 countries worldwide.

informal organization to produce the good/service. With regards to the choice of method, this study has an interesting point (particularly relevant to our studies) of the existence of informal networks – working against (or instead of) the constituted formal networks in anti-modern societies, such as Russia. And therefore, the ideal would be to correct any measure of social capital for these societies with an indicator of this "negative" social capital.

In a study of a watershed conservation/development programme in India, Krishna and Uphoff (1999) have demonstrated that social capital is positively related with development performance. They start out with constructing an index of *development orientated collective action*, and test this against alternative hypotheses that might explain the collective actions, thus testing the validity of this measure. Following this, they construct a social capital index from 6 variables (interview questions), - three structural and three cognitive, - using factor analysis. And this index is positively and significantly related with the index of development orientated collective action.

The authors defend an "ad hoc" approach to measuring social capital, as they state that "indicators of social capital need to correspond with the pattern of life in this largely agrarian setting". Thus, they commit themselves to Rose's category of situational theories. The data for the study are based on household surveys of 2397 individuals, and Krishna and Uphoff included questions related to a large number of what they considered as locally relevant activities. Afterwards, the questions for which more than 80 percent of the respondents indicated that this was considered an individualistic activity, were excluded.

As can be seen from the above, Krishna and Uphoff (1999) approach their estimations and inferences with a critical attitude. They test their specifications against alternative explanations in order to be able to eliminate irrelevant (though initially included) variables. This ensures validity of the measures and invokes a high degree of trust in the results.

A more "classical" economic approach can be seen in Whiteley (2000) who incorporates a measure of social capital as an explanatory variable, into an *endogenous growth model*. The measure is constructed by a principal components analysis of three trust variables from the World Value Survey (1990-93) and can, in Rose's categorization, be placed under social psychological approaches. Along with several other explanatory variables (e.g. investments, education etc.) Whiteley regresses this measure on GDP per capita in a sample of 34 countries (for the period 1970-92).

A potential problem lies in the fact that the economic variables (including the explained variable) precede the social capital measure (one of the explanatory variables) in time. This could give rise to a discussion of the causality between social capital and economic performance. Whiteley is aware of this and therefore performs a test for it. By including older (less extensive) indicators for the measure of trust (from 1981) in a regression using economic time series from 1981-92, he obtains similar results.

Further, the correlation between the trust measure from 1981 and the one from 1990-93 is 0,90. This confirms Putnam's view that social capital changes very slowly; in fact, it may take centuries (Putnam, 1993). The only point at which this concern is truly relevant is in the case of sudden radical changes in the level of social capital (e.g. due to revolutions, outbreak of war, radical changes of the political system, etc.).

Brehm and Rahn (1997) specify a *structural model* of social capital, consisting of the interaction between three concepts, namely civic engagement, interpersonal trust, and confidence in the government, whereby they emphasize the existence of an endogenous (and dynamic) property of the concept.[3] Using data from the General Social Survey from 1972 to 1994, they estimate their model in a pooled cross-sectional analysis combining latent variables (civic engagement, interpersonal trust, and confidence in government) for the key concepts and exogenous variables, all measured at an individual level. More specifically they estimate the model using a *covariance structure analysis* which means using the correlation matrix as input. One of the implications (and advantages) of this approach is that missing data are deleted pair-wise rather than list-wise which reduces the possibility of biases.[4] First, they estimate measurement models for each of the three endogenous (or latent) variables, using factor analysis on exogenous explanatory variables, secondly they estimate the structural model using the three latent variables and some structural components.

Brehm and Rahn obtain results that show that civic engagement and interpersonal trust are in a tight reciprocal relationship, where the connection is stronger from participation to interpersonal trust, rather than the reverse. Thereby, they reject the approach termed the "social psychological approach" by Rose, an approach which actually appears to be followed by a large part of the studies done within the social capital area.

[3] Referring to the discussion in section 2, Brehm and Rahn have a good point that "social capital is an aggregate concept that has its basis in individual behavior, attitudes, and predispositions".
[4] A list-wise deletion of missing data would favor the respondents who answer all questions, and therefore are most likely to be the most involved citizens.

In a more “straight forward” attempt, Grootaert (1999) considers social capital as a production factor of the household, similar to human or physical capital.[5] He analyses the link between social capital and household welfare and poverty in Indonesia by undertaking a multivariate analysis of the role of institutions, in affecting household welfare and poverty outcomes and in determining access to services. The data are generated by 1200 household interviews mapping the different levels of society (household, community, and district), and interviewing respondents from identified focus groups.[6] Grootaert investigates six dimensions of social capital (as represented by “local associations”), namely the density of associations, the internal heterogeneity, frequency of meeting attendance, members’ effective participation in decision making, payment of dues, and the community orientation of associations. From the values of these six variables, he constructs a social capital index, which turns out to be positively related to household welfare – measured by indicators of expenditure per capita, assets, access to credit, savings, school attendance, etc.[7]. Using instrument variables Grootaert has tested that the causality goes from social capital to income, and not the other way round. This result is robust to several sets of instruments. Finally, Grootaert compares the impact of household memberships in local associations with the impact of human capital, on household welfare, and finds that at low incomes, the returns to social capital are higher than returns to human capital. At higher incomes, the reverse is true.

Grootaert accredits Narayan and Pritchett (1997) for their “pioneering effort” in a study measuring social capital at the community level in rural Tanzania, to which Grootaert’s study is supposed to be an extension to the Indonesian case. Narayan and Pritchett conduct a household survey (of 1376 households in 87 clusters) to examine the link between social capital and village-level economic outcome. They ask questions about the households’ memberships in groups, the characteristics of these groups, and the individuals values and attitudes (particularly their definition and expressed level of trust), thus using both of the two most commonly used indicators (organizations and trust) in a combined quantitative measure of social capital.

So, broadly speaking the two dominant (operational) features describing social capital are “trust” and “memberships of voluntary organizations”.

[5] Interpreting the concept as “capital” basically means that it is a stock that yields a flow.
[6] This method of categorizing is similar to the one recommended by Krishna and Shrader (1999).

4: Questionnaire

In the proposed study of the differences in levels of social capital between Western and Eastern Europe - represented by Denmark, United States, Russia, Poland and the three Baltic States - we will follow an approach similar to that of Narayan and Pritchett (1997). More specifically, this means that the questionnaire contains a combination of "trust" indicators and "voluntary organizations" indicators. Moreover, the questionnaire includes a measure of relationship to the local area, networks and civic action.[8]

Our key explanatory variables have been selected from two existing questionnaires. The main source of inspiration was the previously mentioned questionnaire constructed by Krishna and Shrader (1999), from which **Questions 1-2 and 5 through 9** have been taken – these questions cover both structural and cognitive aspects of social capital. The remaining questions (no. 3 and 4) are taken from the World Values Survey.[9] The questionnaire is enclosed as Appendix I and it is structured according to the three main hypotheses suggested by Paldam and Svendsen (2000b).

Questions 1 and 2 concern Putnam's Instrument which tries to measure the level of trust by measuring the density of voluntary organizations. **Questions 3, 4 and 5** ask about general trust regarding other people and institutions. Trust in other people is important to our *first hypothesis* concerning the role of dictatorship. We suggest that communism destroyed social capital because the state made all decisions without leaving room for entrepreneurship and voluntary organizations. Therefore, we suggest that dictatorship destroys positive social capital measurable as a low density of voluntary organizations, ibid.

Trust in institutions concerns corruption. In Paldam and Svendsen (2000b), it is argued that corruption may lead to negative social capital, that is social network constellations harmful to economic growth. This is so because corruption, in the sense of low-quality institutions, hampers non-personal enforcement of rules. If it pays to break the rules, it is hard to build up social capital and the high level

---

[7] Grootaert claims, that using the additive index of the number of memberships and the index of active participation in decision making (with equal weights) explains just as much as using all the variables. So, this is what he does.
[8] The questionnaire is included in Appendix.
[9] In Inglehart et al. (1998).

of transaction costs in the market will *not* lead to the experience of significant aggregated benefits as agents would need more formal contracts. When citizens cannot trust institutions in society and when everyone is not equal to the law, this unpredictability blocks the building of social capital. This idea is relevant to our *second hypothesis:* an agents' trust in the quality of formal institutions (degree of corruption) will affect the level of positive social capital beneficial to economic growth because a high level of corruption means non-enforcement of contracts and as such prevents building of trust among trading parties.

Questions no. 6 through 9 concern local networks and also links to entrepreneurship and voluntary action such as civic actions. These local networks can both imply positive and negative social capital. In this specific setting of comparing Western and Eastern Europe, our *third hypothesis* states that the communist model develops negative social capital. This idea originates from Rose (1999: p156) who shows that citizens in Russia have formed many private networks for getting things done. In response *against* the State, individuals can invoke informal networks: begging or cajoling public officials, using connections to »bend« rules or paying bribes that break rules; social networks compensate for organizational failure. Furthermore, the persistence of such networks is a formidable barrier to Russia's transition from an antimodern to a modern society with a well-functioning market economy. These informal interest groups maintain status quo by exploiting their personal contacts, barter or bribes to get what they want within the existing system and they range widely from ordinary households to the Mafia (ibid., p.148-49).

5: Factor analysis

5.1: Introduction

We will now move on to discuss the methodology of factor analysis which can be used to analyze the data generated from the questionnaire described in the previous section. Our review of the literature so far showed that factor analysis is a commonly used tool for constructing a measurement index of social capital. The notion of factor analysis is often used as a common description of several different methods, which can be separated into two areas – statistical models for inference, and simplifying procedures for the approximate description of data. We focus on the latter here.

The theory rests on the explanation of correlation between two (or more) variables to be that of "measuring the same thing" – as opposed to explaining each other, for example. From the regression theory we have the concept of partial correlation (explaining correlation) and multiple correlation (explaining variance). These two concepts form the basis of *common factor analysis* and *principal component theory*, respectively. These are the two most commonly used methods of factor analysis and let us look at them in turn.

5.2: Common factor analysis

The theory behind *common factor analysis* emerges from the recognition of empirical correlation matrices that look as if their correlations could be explained by the correlations of the given variables with an unknown variable. Or as the fundamental theorem of factor analysis states "there exists a number of unobserved variables (common factors) that explain our observed correlations, in the sense that when these (the common factors, that is) are partialled out, the partial correlations of our observed variables all become zero".[10]

In this way, the statistical model can be expressed as:

(1) $y_j = f_{j1}x_1 + f_{j2}x_2 + \ldots + e_j$

where y_j is the dependent variable, the x_is are the common factors and f_{ji} indicate the regression weight of the j^{th} variable on the i^{th} factor – also called the factor loadings.

In direct extension, the correlation between two observed variables (y_j and y_k) can be expressed mathematically as:

(2) $r_{jk} = f_{j1}f_{k1} + f_{j2}f_{k2} + \ldots + f_{jm}f_{km}$,

where m is the number of common factors.[11] This is actually the testable numerical formulation of the problem.

[10] McDonald (1985), p. 50.
[11] Note that due to the assumption of uncorrelated factors, the factor loadings also express the correlation between the dependent variable and the individual factors.

It is not important whether the common factors explain a lot (or a little) of the variances of the variables. What is important is that they explain the *correlation* of the (dependent) variables completely. Intuitively, what we are after is a set of factors that make the partial correlation matrix as close to zero (or strictly speaking, identity) as possible. In practical terms this can be done by following two different paths, i.e. by specifying the testable hypotheses in two distinct ways. The two methods are called *exploratory* factor analysis and *confirmatory* factor analysis.

In short, *exploratory* factor analysis means hypothesizing a certain *number* of common factors to explain the correlations without worrying about which set of factor loadings is obtained, and then transforming (rotating) this set until it approaches simple structure (this requires uncorrelated factors).[12] This is necessary, as specifying only the number of common factors is not sufficient to identify the factor loadings. Moreover, as the fit will (inevitably) become perfect for a sufficiently large number of factors, what is desirable is to affirm the most restrictive hypothesis, i.e. the hypothesis that specifies the smallest number of factors that will account for the correlations.

Using the *confirmatory* factor analysis approach, a more detailed hypothesis is specified, including both the number of common factors *and* an a priori assumption/guess that specified variables are unrelated with specified factors.[13] This means that a specific matrix of variable-factor correlations is tested, which demands more from the analyst in terms of presupposed knowledge about the data set – on the other hand this method may retain more of the intuitive feeling of what is happening to the data.

5.3: Principal component analysis

Related to common factor analysis is the *principal component analysis*. As mentioned above, it is based on the multiple correlation principle and seeks to explain the *variance* of the dependent variables. This is done by the identification of a single (or a few) index - named scores - constructed from weighted combinations of the dependent variables.

[12] A factor pattern has simple structure when each variable has nonzero loadings (regression weights) on as few of the factors as possible.
[13] Either from analysis of the nature of the measures or from past experience, or a similar qualified "guess".

A set of *principal component scores* is obtained by adding together the weighted dependent variables, i.e.

(3) $s = w_1 y_1 + w_2 y_2 + \ldots + w_n y_n$

The set of weights hereby obtained should maximize the total sum of squared correlations of s with y_1, …, y_n.

"Setting aside" this best combination (and label it s_1) and searching for the second-best, the third-best etc. up to the n^{th}-best combination, gives us a set of n component scores. Each of these scores maximize the sum of squared correlations (with the y_is) subject to it being *uncorrelated* to all of the previous scores. Formally

(4)
$$
\begin{aligned}
s_1 &= w_{11}y_1 + w_{12}y_2 + \ldots + w_{1n}y_n \\
s_2 &= w_{21}y_1 + w_{22}y_2 + \ldots + w_{2n}y_n \\
&\vdots \\
s_n &= w_{n1}y_1 + w_{n2}y_2 + \ldots + w_{nn}y_n
\end{aligned}
$$

Introducing p_{il} indicating the correlation of the i^{th} variable with the l^{th} score (s_l), the above relationships (equations) can be inverted to obtain expressions for the dependent variables

(5)
$$
\begin{aligned}
y_1 &= p_{11}s_1 + p_{12}s_2 + \ldots + p_{1n}s_n \\
y_2 &= p_{21}s_1 + p_{22}s_2 + \ldots + p_{2n}s_n \\
&\vdots \\
y_n &= p_{n1}s_1 + p_{n2}s_2 + \ldots + p_{nn}s_n
\end{aligned}
$$

From which it can be seen that - provided that they are in standard measures – the variances of the dependent variables are given by[14]

(6) $\sigma_i^2 = 1 = p_{i1}^2 + p_{i2}^2 + \ldots + p_{in}^2$

This means that the variance of y_i is explained by n additive parts (components), where each component explains the maximum possible proportion of the variance of all n variables. We need all n components to fully explain the variance but if we want only one combined measurement, s_1 should be used, as per definition it explains the most. And of course any number in between can be chosen. Again, the loading of the variables on the components determines which variables should be used in a composite index for the variable in question.

5.4: Summary

In summary, principal component analysis is considered explorative whereas common factor analysis is confirmative. The principal components are known weighted sums of the given variables, chosen to explain *variance* in terms of multiple correlation principles. The common factors are unknown variables, chosen to explain *correlations* in terms of partial correlation principles. Furthermore, all components are required to fully explain the correlations in the principal components analysis, in contrast to the common factor analysis where only m factors (much smaller than n) explain the correlations (but not the variances).

As previously mentioned, for the application of both methods it is a question of reducing the number of explanatory variables. In principle, one follows the same procedure for both factor analysis and the principal components approach, although the statistical mathematics differ substantially. The difference between the two (as it was explained above) lies in the explanation of correlation versus variance.

5.5: Method of measurement, choice

As we intend to investigate data in order to determine whether any underlying explanations/relations exist, we will follow the explorative approach. The principal component analysis is the relevant choice for this. In practical terms, the analysis is performed using computer software, and the procedure can be summarized into the following steps:

[14] Again, this hinges on the restriction that the scores are uncorrelated.

1. The relevant number of principal components is identified, by how much (of the variance they explain (cumulatively). In a computer analysis output, this can be read from a column of "cumulative explanation".

2. For the relevant components, the loadings of all of the variables are calculated. All the variables that have high loadings onto a specific component can (with reasonable probability) be explained by that component.

3. Construct an index from these variables, for instance by using a weighted sum.[15] Krishna and Uphoff (1999) suggest normalizing the range of each variable to be between 0 and 1 by dividing each variable by its range and then summing across the variables.

In order to illustrate the discussion and focus the presentation, we will present an example from an investigation of the level of social capital in Russia (data collected by Martin Paldam). One section of the questionnaire (see Appendix I, Question 4, below) contains four sub-questions regarding "trust in institutions". Factorizing the observations on these variables yields the following printout from the STATA statistics program.

Table 1: Printout from STATA (Trust in institutions)

(obs=2498)

(principal components; 4 components retained)

Component	Eigenvalue	Difference	Proportion	Cumulative
1	2.13790	1.33653	0.5345	0.5345
2	0.80137	0.24736	0.2003	0.7348
3	0.55401	0.04729	0.1385	0.8733
4	0.50672	.	0.1267	1.0000

Eigenvectors

Variable	1	2	3	4
tr_legal	0.49718	-0.50935	0.53677	-0.45305
tr_polce	0.52199	-0.39392	-0.25658	0.71170
tr_adm	0.52755	0.25525	-0.65222	-0.48078
tr_gov	0.44948	0.72128	0.46973	0.23890

It can be seen that the first component explains 53,5% of the variance within the observations, the first two components explain 73,5%, and so forth. It only makes sense, however, to keep the components that have an eigenvalue higher than 1 – if this is not the case, the component will actually have a variance that is larger than the variance it is meant to explain. The eigenvectors indicate the correlation

[15] Another possibility would be to choose just one variable as a "representative" variable.

of the four variables with the four components, i.e. they are the p_{il}s in (5). In this case, only one component should be retained. It explains just above half of the total variance and is somewhat correlated with all the investigated variables.

Let us then move on to investigate the practical implications for the use of principal components analysis in this project. As previously mentioned, the central hypothesis of the project is that the type of political system in a country can explain the level of social capital. This means that the questionnaire is constructed to measure the dependent variable in a relation that could be expressed as

(7) $y_i = b_{i1} + b_{i2}x_i + e_i$

where - in the usual terminology - for country i, y_i is the level of social capital, x_i is the type of political system with coefficient b_{i2}, b_{i1} is a constant and e_i the error term. In order to avoid confusion, a few remarks on the relation to the above presentation of factor analysis might be necessary.

What we want is to identify the relevant variable(s) that should be used to construct the y_i in (7), i.e. the measure of social capital. The measure of x_i, we will worry about later.

Starting out with the full data set for this project, the principal components analysis can be employed in order to reduce the questions containing more than one observation per respondent (questions 4, 6, 8 and 9) into one measure (variable) each, - according to the previously indicated three steps. For each case (question) n (in eq. (4)) would equal the number of sub-questions in each question, so for instance 4 scores would have to be constructed for question 4, and so forth. Accordingly, each y can be fully explained by "its" number of scores, as in (5), and the preferred "amount" of explanation decides the number of scores we choose to use.

The main purpose of factor analysis (both common factor and principal component analysis) is to reduce a number of explanatory variables into one or a few variables and it is an obvious way of obtaining a simpler measure of social capital (as the above review of studies also indicates). Constructing a questionnaire that encompasses a broader range of aspects of social capital and then narrowing it down to a few measures will also guarantee that no possible (and theoretically acceptable) key explanatory variable is missed out. This is so because the important variables are identified by their "loading" on the factors. Whether the method is actually suitable for our purpose still remains to be seen.[16] There is a possibility that too many factors turn out to be important – or not important enough.

6: Conclusion

The purpose of this paper was to develop a method of measurement for social capital. As argued, this was most relevant for e.g. international comparison of Western and Eastern Europe. Various authors had contributed to the investigation of the influence of social capital on economic growth. However, when focusing on the methodological question of how to establish an operational measure of this concept of social capital, it appeared that consensus had not yet been reached. Generally speaking, each survey used its own ad hoc method of measurement which again partly could be explained by the heterogeneity of the very definition of the concept of social capital. Thus, a gap in literature exists.

After introducing the concept of social capital and surveying the already existing methodological approaches found in literature, we presented a questionnaire. Based on this questionnaire, we argued that factor analysis and in particular the principal component analysis was the most appropriate method of measurement. This is due to the explorative nature of this approach (as mentioned, we do not yet know for sure whether the data can be reduced), and the qualitative type of the variables that we intend to measure.

[16] At the time of writing, we have not started working with the data.

Literature

Brehm, J. and Rahn, W. (1997). Individual-Level Evidence for the Causes and Consequences of Social Capital. *American Journal of Political Science*, vol. 41(3), p. 999-1023.

Coleman, J. S. (1990). Foundations of Social Theory. Cambridge, MA: Harvard University Press.

Dasgupta, P. (1999). *Economic progress and the idea of social capital*. I: Social Capital, A Multifaceted Perspective, eds. Dasgupta, P. og Serageldin, I., Washington D.C.: The World Bank, 1999.

Fukuyama, F. (1995). Trust: The Social Virtues and the Creation of Prosperity. New York: Free Press.

Grootaert, C. (1999). Social Capital, Household Welfare, and Poverty in Indonesia. *Policy Research Working Paper No. 2148*. The World Bank Social Development Department.

Inglehart, R. (1997). Modernization and Postmodernization: Cultural, Economic and Political Change in 41 Societies. Princeton: Princeton University Press.

Inglehart, R., Basañez, M., and Moreno, A. (1998). Human Values and Beliefs: A Cross-Cultural Sourcebook. Political, religious, sexual, and economic norms in 43 societies. Findings from the 1990-93 World Values Survey. University of Michigan Press: Ann Arbor, MI.

Krishna, A. and Shrader, E. (1999). Social Capital Assessment Tool. *Conference Paper for Conference on Social Capital and Poverty Reduction*, The World Bank, June 1999.

Krishna, A. and Uphoff, N. (1999). Mapping and Measuring Social Capital: A Conceptual and Empirical Study of Collective Action for Conserving and Developing Watersheds in Rajasthan, India. *Social Capital Initiative Working Paper No. 13*. The World Bank, Washington, D.C.

McDonald, R. P. (1985). Factor Analysis and Related Methods. New Jersey: Lawrence Erlbaum Ass.

Narayan, D. and Pritchett, L. (1997). Cents and Sociability. *World Bank Policy Research Working Paper No. 1796*. World Bank, Washington, D.C.

Paldam, M. (2000): Social Capital: One or many? - Definition and measurement. *Journal of Economic Surveys* – special issue on political economy.

Paldam, M. and Svendsen, G. T. (2000a). An essay on social capital: looking for the fire behind the smoke. *European Journal of Political Economy*, vol. 16, p. 339-366.

Paldam, M. and Svendsen, G. T. (2000b). Missing Social Capital and the Transition in Eastern Europe. *Journal for Institutional Innovation, Development and Transition*, forthcoming.

Paldam, M. and Svendsen, G.T. (2002): *Trust, Social Capital and Economic Growth: An International Comparison.* Edward Elgar, Cheltenham, UK. Forthcoming.

Putnam, R. D., Leonardi, R. og Nanetti, R. Y. *Making Democracy Work, civic traditions in modern Italy*. Princeton: Princeton University Press, 1993.

Rose, R., 1999, Getting Things Done in an Antimodern Society: Social Capital Networks in Russia. In Dasgupta & Serageldin (1999: 147-71)

Whiteley, P. F. (2000). Economic Growth and Social Capital. *Political Studies* vol. 48(3), p. 443-466.

APPENDIX I:
SOCIAL CAPITAL QUESTIONNAIRE AND MARGINALS

Putnam's Instrument

1 – 2 How many voluntary organizations are you a member of :

	Name	Code from*	Frequency of Contacts		
			week	month	year
1					
2					
3					
4					
5					
6					
7					
8 up					

Frequency should be filled in only for lowest possibility
A contact is when the respondent takes part in an activity of the group

* Type codes

	% of those, participated in a organization				
1	Farmers'/fisherman's group		11	Parent group	
2	Traders' association/business group		12	School committee	
3	Cooperative		13	Health committee	
4	Women's group		14	Water/waste	
5	Credit/finance group (formal)		15	Sports group	
6	Political group		16	NGO	
7	Youth group		17	Civic group (ie, Rotarian)	
8	Religious group		18	Professional Association	
9	Cultural Association		19	Trade Union	
10	Neighborhood/village association		98	Other	

Trust

3 Standard generalized trust

Generally speaking, do you believe that most people can be trusted or can't you be too careful in dealing with people ?

Most people can be trusted	
You can't be too careful	

4 Trust in institutions

How much confidence do you have in the following institution?

	Institution	(1) A great deal	(2) Quite a lot	(3) Not very much	(4) None at all	(5) Hard to answer
1	The legal system					
2	The police					
3	The administration					
4	The government					

5 Loan question

Do you think that in this neighbourhood/village people generally trust each other in matters of lending and borrowing ?

(1) Do trust	(2) Do not trust	(3) Don't know/ not sure	(4) No answer

6 Networks

Suppose your neighbor suffered an economic loss, say (RURAL: “crop failure”; URBAN “job loss”). In that situation, who do you think would assist him/her financially? [Record first three mentioned.]

1	No one would help	
2	Family	
3	Neighbors	
4	Friends	
5	Religious leader or group	
6	Community leader	
7	Business leader	
8	Police	
9	Family court judge	
10	Patron/employer/benefactor	
11	Political leader	
12	Mutual support group to which s/he belongs	
13	Assistance group to which s/he belongs	
14	Other	
15	Don’t know/not sure	
16	No answer	

7. People here look out mainly for the welfare of their own families and they are not much concerned with village/neighborhood welfare. Do you agree or disagree with this statement?

1	Strongly agree	
2	Agree	
3	Disagree	
4	Strongly disagree	
5	Don’t know/not sure	
6	No answer	

8

Please tell me whether *in general* you agree or disagree with the following statements:

		Strongly Agree	Agree	Disagree	Strongly Disagree
a	Most people in this village/neighborhood are basically honest and can be trusted				
b	People are always interested only in their own welfare				
c	Members in this village/neighborhood are always more trustworthy than others				
d	In this village/neighborhood one has to be alert or someone is likely to take advantage of you				
e	If I have a problem there is always someone to help you				
f	I do not pay attention to the opinions of others in the village/neighborhood				
g	Most people in this village/neighborhood are willing to help if you need it				
h	This village/neighborhood has prospered in the last five years				
i	I feel accepted as a member of this village/neighborhood				
j	If you drop your purse or wallet in the neighborhood, someone will see it and return it to you				

9 Civic actions:

In the last three years have you personally done any of the following things

		(1) Yes	(2) No	(3) Don't Know
a	Voted in the election			
b	Actively participated in an association			
c	Made a personal contact with an influential person			
d	Made newspapers, radio and TV interested in a problem			
e	Actively participated in an information campaign			
f	Actively participated in an election campaign			
g	Taken part in a protest march or demonstration			
h	Contacted your elected representative			
i	Taken part in a sit-in or disruption of government meetings/offices			
j	Talked with other people in your area about a problem			
k	Notified the court or police about a problem			
l	Made a donation of money or in-kind			
m	Volunteered for a charitable organization			

DEMOGRAPHICS

Gender

Male	
Female	

Age

18-24	
25-34	
35-44	
45-59	
60-	

Education

Incomplete middle	
Middle, specialized middle	
Incomplete higher, higher	

Social position

Businessman	
Manager	
Professional (with higher education)	
Blue collar worker	
Supporting personnel (no higher education)	
Military	
Student	
Unemployed	
Pensioner	
Housewife	

Monthly per capita income

Below 600 roubles	
600-1000	
1000-1500	
1500-2000	
Above 2000	

Type of population center

Cities 1mln and above	
300,000-1,000,000	
100,000-300,000	
Towns below 100,000	
Rural	

Regions (Federal Districts)

St.Petersburg	
Moscow	
Urals	
Privolzhski (Volga)	
Yuzhnyi (Southern)	
Severo-Zapadnyi (North-Western)	
Tsentralnyi (Central)	