

# On The Empirics of Social Capital

Steven N. Durlauf

Department of Economics, University of Wisconsin

This Draft: January 20, 2002

I thank William Brock, Adam Gamoran, Michael Sobel, and Chih Ming Tan for useful conversations on the issues in this paper. Artur Minkin, Eldar Nigmatullin and Chih Ming Tan have provided excellent research assistance. I thank Stephen Machin and three referees for comments on an earlier draft. The John D. and Catherine T. MacArthur Foundation, Romnes Trust and Vilas Trust have generously supported this research. All errors are mine.

## 1. Introduction

One of the most striking developments in social science over the last decade is the rise of interest in social capital as a mechanism for understanding socioeconomic phenomena. This enthusiasm has not been restricted to sociologists; within economics and other disciplines interest in social capital has grown dramatically as well. Social capital has been treated as a key feature of phenomena ranging from the mortality crisis in Russia (Kennedy, Kawachi, and Brainerd (1998)) to political participation (DiPasquale and Glaeser (1999)) to children's welfare (Putnam (2000)) to the development traps (Woolcock (1995)) to judicial efficiency (LaPorta et al (1997)) to the spread of secondary education (Goldin and Katz (2001)).

This paper is designed to explore a number of difficulties with the empirical analysis of social capital. In doing this, my purpose is not to challenge the importance of addressing the "undersocialized conception of man" that Granovetter (1985) and many others have described as a limitation of neoclassical economic theory. Rather, my purpose is to question whether the empirical evidence so far adduced to support a significant explanatory role for social capital achieves this objective. In order to do this, I will focus on three prominent papers in the social science literature. Each of these papers was written by a distinguished author or group of authors and has achieved a prominent place in various social capital debates. Each of these papers is a serious and valuable contribution to the empirical literature because each establishes "stylized facts" that structural analyses of social capital should address. Yet I will argue that these papers suffer from a set of common flaws that renders any inferences concerning the explanatory significance of social capital dubious at best. Hence, from the perspective of asking whether these studies establish the empirical importance of social capital in understanding various socioeconomic outcomes, my conclusion is no. The sources of these criticisms apply, in my judgment, very widely to the empirical social capital literature.

Social capital is an elusive concept, as reflected in the fact that its definition differs across studies; Portes (1998) has developed a strong critique of the social capital literature because of this definitional ambiguity. That being said, the following definition by Fukuyama (1999) seems to capture many of the intuitions that have driven this literature.

*Social capital* can be defined simply as an instantiated set of informal values or norms shared among members of a group that permits them to cooperate with one another. If members of the group come to expect that others will behave reliably and honestly, then they will come to *trust* one another. Trust acts like a lubricant that makes any group or organization run more efficiently. (pg. 16)

Similarly, Bowles and Gintis (2002) state

Social capital generally refers to trust, concern for one's associates, a willingness to live by the norms of one's community and to punish those who do not. (pg. 1)

and Putnam (2000), in what is clearly the most exhaustive defense of the importance of social capital yet to appear, states

...social capital refers to connections among individuals - social networks and the norms of reciprocity and trustworthiness that arise from them...A society of many virtuous but isolated individuals is not necessarily rich in social capital. (pg. 19)

One important feature of these general definitions, to which I will return, is how they mix a number of disparate ideas. One such combination is the mixing of functional and causal conceptions of social capital. When social capital is defined as a set of norms or values that facilitate cooperation and efficiency, this is a functional notion. In contrast, when one argues that the cooperative behavior of others leads to expectations under which cooperation is individually rational, this is a causal notion. It seems self evident that causal definitions of social capital are necessary for successful empirical analysis, so the extent to which studies are careful about this distinction is important. Further, as these quotations illustrate, standard social capital definitions also mix very different notions of individual motivation. Many social capital papers criticize standard notions of *homo economicus*, arguing that conventional notions of inner-directed preferences fail to capture the roles of altruism and fairness in motivating individual decisions. Yet other analyses accept the standard model of preferences, and (at least implicitly) employ ideas from repeated game theory in which behaviors which are not rational in one-shot games become rational via the potential for retaliatory strategies. Hence one question for

particular empirical studies is whether they are based on a consistent and coherent definition of social capital.

Section 2 of this paper outlines three benchmark studies. Section 3 discusses some requirements for structural inference, namely, partial exchangeability of observations and identifiability of structural parameters. Section 4 uses these ideas to assess the benchmark studies. Section 5 makes some suggestions for future research. Section 6 contains summary and conclusions.

## **2. Benchmark studies**

The social capital literature contains many dozens of empirical studies that purport to show the importance of social capital to a very wide-ranging set of socioeconomic phenomena. For the purposes of this paper, I work with three "exemplar" studies. These studies have been chosen for several reasons. First, each is, in my judgment, well regarded in the social science literature. Second, these studies span very different levels at which social capital has been analyzed.

### ***i.* Individual level analysis: US data**

Much of the early use of social capital in empirical work has been in the analysis of individual level outcomes. A typical example is Furstenberg and Hughes (1995). This paper explores the relationship between social capital as experienced by youths and subsequent outcomes. The study employs a data set taken from a 20-year longitudinal study that tracked 252 children of teenage mothers originally residing in the Baltimore area. This data set contains extensive attitudinal and family information and so allows one to explore the relationship between various outcomes for these children and a number of interesting measures. The outcome measures are binary variables including high school graduation, college enrollment, and avoidance of criminality.

In order to measure social capital, Furstenberg and Hughes construct two variables that are designed to measure within-family social capital and community-based social capital, respectively. Within-family social capital is measured by a number of proxy variables including 1) presence of father in home, 2) frequency of activities involving child and parents, 3) parents'

expectations of school performance, 4) mother's encouragement of child, 5) mother's attendance at school meetings, and 6) number of child's friends a mother knows. The community-based social capital variable is measured by proxy variables including 1) whether child ever changed schools due to a move. (The idea behind the variable is that an individual's social capital builds up over time within a community, so that more frequent moves mean that an individual experiences less social capital on average), 2) membership and frequency of attendance in church, 3) whether one sees a close friend weekly, 4) whether one believes there is someone to turn to for help under various scenarios, 5) friend's educational expectations, 6) school quality, and 7) parent's perception of neighborhood quality.

For each of the community- and family-based social capital measures, a separate regression is run. In their analysis, Furstenberg and Hughes estimate equations of the form

$$\Pr(\omega_{i,j}) \propto \exp(X_i\gamma_j + SC_{i,k}\beta_{j,k}) \quad (1)$$

where  $\Pr()$  denotes a probability measure,  $\omega_{i,j}$  denotes outcome  $j$  for individual  $i$ ,  $X_i$  denotes some family control variables (notably two family human capital measures), and  $SC_{i,k}$  denotes social capital measure  $k$  for family  $i$ . Statistical significance of  $\beta_{j,k}$  is interpreted to mean that social capital measure  $k$  helps explain outcome  $j$ .

## **ii. Individual level analysis: developing country data**

Narayan and Pritchett (1999) study the role of social capital in influencing household outcomes in rural Tanzania. They do this using an interesting data set that explicitly surveyed social capital and socioeconomic outcomes across a number of rural villages. Letting  $g(i)$  denote the village of residence of household  $i$ , these authors model household expenditures (which is treated interchangeably with household income)  $\omega_i$  as

$$\omega_i = X_i\gamma + Z_{g(i)}\pi + SC_{g(i)}\beta + \varepsilon_i \quad (2)$$

Here,  $X_i$  denotes a set of family-specific controls,  $Z_{g(i)}$  a vector of village-specific controls and  $SC_{g(i)}$  denotes social capital in village  $g(i)$ . The social capital variable is constructed by a weighted average of answers to questions that measure 1) memberships in various groups, 2) the characteristics of these groups, and 3) general trust-related attitudes. Variables from the first two classes of questions were used to construct an overall level of "associational" activity that is used as the social capital measure.

Narayan and Pritchett recognize that their social capital measure is endogenous and so used trust variables in their survey as an instrumental variable. In their words

We posit that certain of these "trust variables", particularly an individual's trust in strangers and trust in various government officials, are not affected directly by household income, nor do they affect income directly, but that greater levels of trust do lead to higher village social capital. (pg. 880)

The available instruments include survey measures of degree of trust in strangers, fellow tribesmen, village chairmen, district officials and the central government. Instrumental variables estimates confirm their findings that social capital is positively associated with higher expenditures. In their analysis, they find that instrumental variables estimates of social capital are substantially higher than OLS estimates and that, when social capital and instruments are measured at the village level, the instruments pass a standard exogeneity test, whereas the individual analogs do not. I will therefore focus on the group social capital specification in subsequent discussion.

### **iii. Cross-country analysis**

Social capital has also been asserted to influence aggregate economic growth. A standard citation in this literature is Knack and Keefer (1997). In this paper, a cross-country growth regression of the following type is estimated,

$$g_i = X_i\gamma + Z_i\pi + CIVIC_i\alpha + TRUST_i\beta + \varepsilon_i \quad (3)$$

where  $g_i$  is real per capita growth in economy  $i$  over a given time period,  $X_i$  is a vector of country-specific control variables suggested by the Solow growth model,  $Z_i$  is a vector of country-specific controls suggested by recent growth theories,  $CIVIC_i$  is an index for the level of civic cooperation in country  $i$ , (measured by averaging answers to questions such as "Is it ever justified to cheat on taxes if you have the chance?") and  $TRUST_i$  represents the percentage of survey respondents in country  $i$  (after omitting those responding "don't know") who, when queried about the trustworthiness of others, replied that "most people can be trusted," and  $\varepsilon_i$  is an unexplained residual.

In this paper, the importance of social capital is assessed by the positive sign and statistical significance of the  $\alpha$  and  $\beta$  coefficients. Positive and statistically significant coefficients are shown to hold for a range of alternative specifications as well as when the  $TRUST_i$  variable is instrumented, using a measure of ethnolinguistic diversity.

### 3. Issues in causal analysis

In order to evaluate disparate studies like these, it is useful to identify some baseline ideas concerning the identification of structural factors in socioeconomic data. In this section, I discuss two issues that are important in evaluating studies of social capital of the type described above.

#### *i. exchangeability*

One way to link many of the requirements for the structural interpretation of regressions is that a structural regression presupposes that the data under study possess conditionally exchangeable errors. For our purposes, suppose that one has information  $F_i$  on each of the  $i$  observations in the sample, and each of these observations is assumed to obey a linear structural model of the form

$$\omega_i = Z_i\gamma + \eta_i \tag{4}$$

The errors  $\eta_i$  are said to be  $F_i$ -conditionally exchangeable if

$$\Pr(\eta_1 = a_1, \dots, \eta_K = a_K | F_1 \dots F_I) = \Pr(\eta_{\rho(1)} = a_1, \dots, \eta_{\rho(K)} = a_K | F_1 \dots F_I) \quad (5)$$

where  $\rho(\cdot)$  is any operator which permutes the  $K$  indices. For this same linear model, the  $\omega$ 's are said to be partially exchangeable if, were  $Z_i = Z_j \forall i, j$  they would be  $F$ -conditionally exchangeable.

Intuitively, exchangeability says that a researcher has no prior way of distinguishing errors in a structural model; partial exchangeability says a researcher has no way of distinguishing two dependent variables were they to be associated with the same controls. Many of the standard sources of inconsistency in parameter estimation are interpretable as exchangeability violations. For example, an omitted variable may be interpreted as an element of  $F_i$  that causes the errors in a linear regression such as (2) or (3) to be distinguishable. Similarly, selection bias or parameter heterogeneity may be interpreted this way. A key advantage of using exchangeability as a criterion for evaluating regressions is that it is relatively easy to link exchangeability to substantive socioeconomic features of the processes under study.<sup>1</sup>

---

<sup>1</sup> Exchangeability judgments are an essential part of virtually every statistical exercise. Draper et al (1993) remark

Statistical methods are concerned with combining information from different observational units and with making inferences from the resulting summaries to prospective measurements on the same or other units. These operations will be useful only when the units to be combined are judged to be *similar* (comparable or homogeneous)...judgments of similarity involve concepts more primitive than probability, and these judgments are central to preliminary activities that all statisticians must perform, even though probability specifications are absent or contrived at such a preliminary stage. (pg. 1)

Brock and Durlauf (2001b) make an extended argument on the importance of exchangeability judgments in empirical economics.

Why is exchangeability a useful benchmark in assessing social capital analyses? One reason why this is so is that statistical analyses of social capital typically compare outcomes for individuals or aggregates who have social capital versus those who do not. These studies, in turn, typically do not incorporate a separate theory of the determinants of social capital formation, although they do often employ instrumental variables to account for the endogeneity of social capital. However, without a theory as to why one observes differences in social capital formation, one cannot have much confidence that unobserved heterogeneity is absent in the samples under study.

To be clear, exchangeability conditions are typically stronger than those necessary for causal inference. For example, heteroskedastic errors in a regression are an exchangeability violation, but have no bearing on the interpretation of the regression coefficients. Nor does exchangeability guarantee a regression is causal. What exchangeability does is provide a standard by which to evaluate whether the units of analysis under comparison are in fact comparable in the ways assumed by the exercise.

## **ii. identification**

One question for social capital models is whether, under plausible assumptions, they are econometrically identified. The potential for nonidentification in models of this type has been established in the context of the neighborhood effects literature, in which individual outcomes are influenced by the behaviors and characteristics of those with whom they live. Manski's seminal (1993) paper provided the first general analysis of identification in such environments; further results may be found in Brock and Durlauf (2001a). When social capital is a characteristic of groups, it may be analyzed using the tools found in these papers. In this discussion I use Manski (1993) and Brock and Durlauf (2001a) to make some basic claims concerning when social capital effects are or are not identified from the data.

Suppose that each agent  $i$  is a member of some group  $g(i)$ . Each individual chooses an outcome variable  $\omega_i$  that is linearly dependent on some control variables. Assume these variables are of four types: an  $r$ -dimension vector of variables that are measured at the individual level,  $X_i$ ; an  $s$ -dimension vector of variables (often called contextual effects) that are measured at the group level and are predetermined at the time that choices are made,  $Y_{g(i)}$ ; an

individual's expectation of the average choice of others,  $E(\omega_{g(i)} | F_{g(i)})$  (often called an endogenous effect), where this expectation is made conditional on some information set  $F_{g(i)}$ ,<sup>2</sup> and some measure of social capital,  $SC_{g(i)}$ .

$$\omega_i = k + cX_i + dY_{g(i)} + J_1 E(\omega_{g(i)} | F_{g(i)}) + J_2 SC_{g(i)} + \varepsilon_i. \quad (6)$$

The identification problem amounts to determining whether, for data on various observables in this equation, one can recover the behavioral parameters  $k$ ,  $c$ ,  $d$ ,  $J_1$  and  $J_2$ .

### **a. social capital predetermined**

I first analyze (6) under the assumption that the social capital variable is predetermined and orthogonal to  $\varepsilon_i$ . Analyzing identification of a model such as eq. (6) is nontrivial because of the presence of the contextual effects  $Y_{g(i)}$  and the endogenous effect  $E(\omega_{g(i)} | F_{g(i)})$  in the same equation. Intuitively, if expectations are rational, these variables will be closely related; if the structure of the model means that they are collinear, then identification fails. Thus, unless one is willing to rule out *a priori* that there are no endogenous effects, identification needs to be analyzed. In the exogenous social capital case, social capital acts as another contextual effect and so identification of its impact is also problematic.

Under what conditions is eq. (6) identified? This question has been analyzed in Manski (1993) and Brock and Durlauf (2001a). Following the analyses in those papers, observe that, by taking expected values of both side, eq. (6) implies that

$$E(\omega_{g(i)} | F_{g(i)}) = \frac{k}{1-J_1} + \frac{c}{1-J_1} E(X_i | F_{g(i)}) + \frac{d}{1-J_1} E(Y_{g(i)} | F_{g(i)}) + \frac{J_2}{1-J_1} E(SC_{g(i)} | F_{g(i)}). \quad (7)$$

---

<sup>2</sup> Information for any neighborhood  $g(i)$  is assumed to be common across all individuals.

Substituting this back into (6), one therefore has the reduced form expression for individual choices,

$$\omega_i = \frac{k}{1-J_1} + cX_i + dY_{g(i)} + J_2 SC_{g(i)} + \frac{J_1 c}{1-J_1} E(X_i | F_{g(i)}) + \frac{J_1 d}{1-J_1} E(Y_{g(i)} | F_{g(i)}) + \frac{J_1 J_2}{1-J_1} E(SC_{g(i)} | F_{g(i)}) + \varepsilon_i \quad (8)$$

which forms the basis on which empirical work may be conducted. Identification of the behavioral parameters in (6) amounts to asking whether these parameters can be recovered from the regression parameters in (8). There are  $2r + 2s + 3$  regressors in (8) (and hence  $2r + 2s + 3$  regression coefficients) and only  $r + s + 3$  parameters to recover, so if the regressors are not collinear in the reduced form regression, one would expect that identification can be achieved.

A benchmark case for identification analysis in linear models of this type (cf. Brock and Durlauf (2001a)) assumes that  $F_{g(i)}$  consists of the elements  $1, X_j \forall j \in g(i), Y_{g(i)}$ , and  $SC_{g(i)}$ . This means that  $E(X_i | F_{g(i)}) = X_{g(i)}$  (where  $X_{g(i)}$  denotes the average of value of  $X_i$  in group  $g(i)$ ),  $E(Y_{g(i)} | F_{g(i)}) = Y_{g(i)}$ , and  $E(SC_{g(i)} | F_{g(i)}) = SC_{g(i)}$ . In this case, the reduced form equation for individual behavior is

$$\omega_i = \frac{k}{1-J_1} + cX_i + \frac{J_1 c}{1-J_1} X_{g(i)} + \frac{d}{1-J_1} Y_{g(i)} + \frac{J_2}{1-J_1} SC_{g(i)} + \varepsilon_i \quad (9)$$

Hence, there are now  $2r + s + 2$  regressors with which to uncover the  $r + s + 3$  parameters. However, the identification problem is still not resolved, without assumptions on the relationship between the exogenous variables in this model. Brock and Durlauf (2001a) Theorem 6 immediately implies

**Proposition 1. Identification of predetermined social capital effect**

Identification of the parameters of eq. (6) from a reduced form regression requires

- i.* The dimension of the linear space spanned by elements of  $(1, X_i, Y_{g(i)}, SC_{g(i)})$  is  $r + s + 2$ .
- ii.* The dimension of the linear space spanned by the elements of  $(1, X_i, X_{g(i)}, Y_{g(i)}, SC_{g(i)})$  is at least  $r + s + 3$ .

The first part of the theorem requires that the predetermined variables in (6) are not linearly dependent. The second part of the theorem requires that the variables that comprise  $X_{g(i)}$  not be linearly dependent on these predetermined variables. When this is so, one can form an instrumental variable estimate of  $E(\omega_{g(i)} | F_{g(i)})$  that is not perfectly collinear with the predetermined variables in (6). What does this require operationally? A necessary condition is that there exists at least one individual level variable whose group level average does not causally affect individuals. When this fails to hold,  $X_{g(i)}$  will be a subset of  $Y_{g(i)}$  and  $SC_{g(i)}$ , so  $X_{g(i)}$  is perfectly collinear with some combination of the other regressors in (6). It is precisely this type of collinearity that generates nonidentification in this class of models. Notice that, as suggested above, when social capital is predetermined, it acts as another contextual effect.

Even when identification of the parameters in eq. (6) cannot be achieved, one can still test for the presence of some social capital effect, since one can test whether the social capital coefficient in the reduced form (9),  $\frac{J_2}{1 - J_1}$ , is zero. This result follows from a straightforward simple rewriting of Manski's (1993) Proposition 3.

**Proposition 2. Testability of predetermined social capital effect**

If the dimension of the linear space spanned by elements of  $(1, X_i, Y_{g(i)}, SC_{g(i)})$  exceeds that of the space spanned by elements of  $(1, X_i, Y_{g(i)})$ , then the presence of a social capital effect is testable using a reduced form regression.

***b.* endogenous social capital**

A second way to formulate social capital is to assume that it is an endogenous outcome of decisions that are contemporary to the behavioral choice  $\omega_i$ . So, for example, one can think of the level of trustworthy behavior one exhibits as codetermined with the level of effort or output of an individual; an analysis of social capital as a choice variable may be found in Glaeser, Laibson, and Sacerdote (2001). From this perspective, one can think of each individual as choosing some  $SC_i$  with the individuals then influenced by the expected community social capital  $E(SC_{g(i)}|F_{g(i)})$ . (Nothing changes if one assumes that realized social capital rather than expected social capital affects individuals.) In this case, the community social capital variable must be treated in a way similar to the endogenous effect  $E(\omega_{g(i)}|F_{g(i)})$  rather than in a way similar to the contextual effects  $Y_{g(i)}$ . A joint linear-in-means model for the codetermination of  $\omega_i$  and  $SC_i$  is

$$\omega_i = k + cX_i + dY_{g(i)} + J_1E(\omega_{g(i)}|F_{g(i)}) + J_2E(SC_{g(i)}|F_{g(i)}) + \varepsilon_i \quad (10)$$

and

$$SC_i = \bar{k} + \bar{c}X_i + \bar{d}Y_{g(i)} + \bar{J}_1E(\omega_{g(i)}|F_{g(i)}) + \bar{J}_2E(SC_{g(i)}|F_{g(i)}) + \eta_i \quad (11)$$

Identification of eq. (10) is most easily analyzed by recognizing that one needs to form estimates of  $E(\omega_{g(i)}|F_{g(i)})$  and  $E(SC_{g(i)}|F_{g(i)})$  as these are unobservable. The reduced forms of these expressions are cumbersome and omitted. For our purposes, it is easiest to simply recognize that under my assumptions  $E(\omega_{g(i)}|F_{g(i)})$  and  $E(SC_{g(i)}|F_{g(i)})$  are linear functions of  $E(X_{g(i)}|F_{g(i)})$  and  $E(Y_{g(i)}|F_{g(i)})$ . Writing these as

$$E(\omega_{g(i)}|F_{g(i)}) = \pi_0 + \pi_1E(X_{g(i)}|F_{g(i)}) + \pi_2E(Y_{g(i)}|F_{g(i)}) \quad (12)$$

and

$$E(SC_{g(i)}|F_{g(i)}) = \bar{\pi}_0 + \bar{\pi}_1E(X_{g(i)}|F_{g(i)}) + \bar{\pi}_2E(Y_{g(i)}|F_{g(i)}) \quad (13)$$

it is straightforward to show that the identification problem is equivalent to asking whether these variables are or are not collinear with each other and the other variables in eq. (10). Put differently, since one can construct empirical analogs of the unobservables  $E(\omega_{g(i)}|F_{g(i)})$  and  $E(SC_{g(i)}|F_{g(i)})$  via projections of  $\omega_{g(i)}$  and  $SC_{g(i)}$  on  $F_{g(i)}$  respectively, one may reinterpret (10) as a second stage regression once this is done. When these regressors are used in (10), identification will depend in the standard way on whether the regressors are linearly independent.

In order to state a usable identification condition, assume now that  $F_{g(i)}$  consists of the elements  $1, X_{g(i)}, Y_{g(i)}$ . It is now easy to verify

**Proposition 3. Identification of endogenous social capital effect**

Identification of the behavioral parameters in eq. (10), when expectations follow (12) and (13), requires

- i.* The dimension of the linear space spanned by elements of  $(1, X_i, Y_{g(i)})$  is  $r + s + 1$ .
- ii.* The dimension of the linear space spanned by the elements of  $(1, X_i, X_{g(i)}, Y_{g(i)})$  is at least  $r + s + 3$ .

Like the exogenous case, identification can be achieved via the  $X_{g(i)}$  variables. Intuitively, one needs two elements of  $X_{g(i)}$  not to be elements of  $Y_{g(i)}$  so as to provide instrumental variable estimates of  $E(\omega_{g(i)}|F_{g(i)})$  and  $E(SC_{g(i)}|F_{g(i)})$ . Hence one needs two individual characteristics that affect individual behavior yet whose group level analogues are excluded from the behavioral equation (10).

Notice that if there are prior exclusion restrictions for  $Y_{g(i)}$  relative to (10) and (11), they can facilitate identification. For example, if there is a contextual effect that affects social capital but not the behavioral outcome, that will help with identification of (10) whereas a contextual

effect that only appears in (10) will not. The way in which identification is facilitated by exclusion restrictions is analogous to what one finds in bivariate simultaneous equations models with exogenous variables.

### c. aggregation

The discussion so far has focused on the question of what may be identified using individual level data. In many contexts, social capital is employed to study larger units of analysis. For example, Putnam (2000) argues that differences in the stocks of social capital help explain variations in phenomena ranging from health to murder rates to the amount of TV watched.

The difficulties in identifying social capital effects from aggregate data are, perhaps unsurprisingly, even greater than from individual level data. To see this, I focus on the endogenous social capital case and consider the outcome and social capital equations when one only has data that has been aggregated across individuals within each  $g$  :

$$\omega_g = k + dY_g + J_1 E(\omega_g | F_g) + J_2 E(SC_g | F_g) + \varepsilon_g \quad (14)$$

$$SC_g = \bar{k} + \bar{d}Y_g + \bar{J}_1 E(\omega_g | F_g) + \bar{J}_2 E(SC_g | F_g) + \eta_g \quad (15)$$

Identification for this system is standard since it equates to the textbook case of determining whether there are instruments available for endogenous variables  $E(\omega_g | F_g)$  and  $E(SC_g | F_g)$ . Assuming that each individual in a given  $g$  possesses a common information set  $F_g$  consisting of 1 and  $Y_g$ , identification will depend on whether one has prior information on which elements of  $Y_g$  have zero coefficients in (14) and (15). To see this, let  $H_{\omega,g}$  and  $H_{SC,g}$  denote the linear spaces spanned by those regressors  $Y_g$  with nonzero coefficients in equations (14) and (15) respectively. Let  $H_{SC,g}^c$  denote that part of  $H_{SC,g}$  that is orthogonal to  $H_{\omega,g}$  (i.e. the linear space formed by the orthogonal complements of any basis of  $H_{SC,g}$  after being projected on  $H_{\omega,g}$ ).

Under the various assumptions that have been made, one can replace  $E(\omega_g | F_g)$  and  $E(SC_g | F_g)$  in (14) with the projections defined by the reduced forms.

$$\omega_g = \pi_0 + \pi_1 Y_g + \varepsilon_g \quad (16)$$

and

$$SC_g = \bar{\pi}_0 + \bar{\pi}_1 Y_g + \eta_g \quad (17)$$

In order for a second-stage regression using these observable analogs to be identified, it is necessary that neither  $\pi_0 + \pi_1 Y_g$  nor  $\bar{\pi}_0 + \bar{\pi}_1 Y_g$  is collinear with those variables that appear in (14). This will require that these terms contain elements of  $Y_g$  whose associated coefficients in (14) are zero and are thereby excluded from the second stage regression. The following proposition is immediate.

**Proposition 4. Identification of social capital effect with aggregated data**

- i.* Identification of eq. (14) from a reduced form regression requires that the dimension of the linear space  $H_{SC,g}^c$  is at least 2.
- ii.* If  $J_1$  is known to equal 0, then identification of (14) from a reduced form regression requires that the dimension of the linear space  $H_{SC,g}^c$  is at least 1.

When compared to their individual level analogs, one can see that the identification problem for aggregated data is quite different, as one no longer has any individual level variables  $X_i$  to provide instruments for the group level variables  $E(\omega_g | F_g)$  and  $E(SC_g | F_g)$ . One must rely on exclusion restrictions, i.e.  $Y_g$  variables that appear in the social capital equation but not the behavioral outcome equation, for identification. Hence, without a theory as to why an aggregate variable only matters for average social capital and not for the average behavioral choice under study, there is no hope for identification.

The basic messages from this discussion are threefold. First, there is nothing distinct about testing for social capital, as opposed to any other sort of group interaction effect. When social capital is exogenous, from the perspective of identification it is nothing more than a contextual effect, and so the identification results in Manski (1993) and Brock and Durlauf (2001a) immediately apply. When social capital is endogenous, these techniques easily generalize. Second, identification of the magnitude of social capital effects from individual data requires a particular sort of prior information, namely knowledge of individual determinants of behavior whose group-level analogues do not causally influence individual outcomes. Third, identification of social capital effects from aggregate data is especially problematic as it must rely on exclusion restrictions concerning variables that affect social capital formation but not other behaviors.

#### **4. Assessment of empirical studies**

In this section, I apply the basic ideas of exchangeability and identification to the three studies I have described. In each case, I will argue that there are good reasons to question whether the studies establish a causal role for social capital. These arguments will amount to questioning whether the regressions in these papers can be interpreted in the usual ways regressions are used for causal claims. This questioning, does not, however, constitute a proof that their claims are incorrect. Rather, it demonstrates what additional information would be needed to justify the structural claims made in these papers. Further, one should not evaluate empirical studies such as these as either "right" or "wrong"; instead such studies provide evidence of varying strength for propositions such as the causal influence of social capital on socioeconomic outcomes. One should also keep in mind that my discussion only covers a subset of the results in each paper, although the criticisms I make are not invalidated by those statistical exercises I ignore.

##### ***i.* Furstenberg and Hughes**

Furstenberg and Hughes' evidence for the role of social capital in individual outcomes relies on the use of large numbers of alternative social capital variables. In their various

exercises, they see how each of these social capital variables (considered in isolation) helps predict individual outcomes. Based on the discussion in Section 3, there are good reasons to question their analysis.

First, the various social capital proxies employed would seem very likely to pick up violations of exchangeability in the model errors for Furstenberg and Hughes' different binary regressions. Put differently, Furstenberg and Hughes fail to provide any justification for thinking that their small set of individual controls is adequate for rendering their model errors exchangeable. And it is easy to construct reasons why these exchangeability violations will produce spurious evidence of social capital effects.

Consider their use of the variable, "changed schools during childhood due to moving" as a proxy for social capital formation. In order for their regression results to be causal, this would require that differences in omitted family influences are not systematically different across those families that did and did not move and cause such a change. Since moves are choice variables, it is difficult to see how they would be unrelated to the unobserved heterogeneity across families as embedded in  $\varepsilon_i$ . To see how fragile any claims about social capital are based on this variable, suppose that one had an alternative hypothesis that unobserved heterogeneity exists with respect to the level of nurturing that each family gives to a child. Suppose one ran the identical regression as Furstenberg and Hughes, and then made the argument that having conditioned on socioeconomic status, as they do, a move by a family resulting in a change of schools is a proxy for a relatively small weight placed by parents on their children's welfare. In other words, moves induce costs to children, so the fact that a child changed schools due to a move would suggest a family assigns these costs relatively less weight. One then claims that evidence that such moves predict higher incidences of high school dropping out or unstable economic status among children later in life (both of which Furstenberg and Hughes find, at 10% significance levels) shows the importance of parental caring and investment in children. Is this interpretation any less well supported than that given by Furstenberg and Hughes? No reason is given to prefer one interpretation over another.

In terms of identification, it seems fair to argue that Furstenberg and Hughes (1995) results cannot really be interpreted as providing evidence of social capital versus the claim that group variables help predict individual outcomes. The reason for this is their failure to distinguish between social capital and other individual and contextual effects. A number of the

variables used by the authors can be argued to have causal influences on individual behavior that are independent of what is meant by social capital. It is easy to imagine that the presence of a father in a household matters for childrearing in any number of ways that affect child outcomes. Similarly, school quality can affect individuals without any social influences at work. Hence, one cannot argue that the instruments employed for social capital do not lie in the space spanned by individual and contextual effects, and so the identification questions raised by Propositions 1 and 2, have not been satisfactorily resolved. Interestingly, as shown in Brock and Durlauf (2001a), many of the identification problems associated with the linear-in-means model do not apply to parametric binary choice models; binary choice models contain nonlinear relationships between expected average behavior and various neighborhood averages which eliminate the collinearity problems which drive the nonidentification results in the linear-in-means model. Hence, while the Furstenberg and Hughes analysis may be faulted for failing to distinguish contextual, social capital and endogenous effects, it is possible to do so in the binary outcome context they employed.

This failure of Furstenberg and Hughes reflects the general problem that they do not provide or employ a consistent definition of social capital. As suggested earlier, social capital has been used as a rubric for very different phenomena; this definitional ambiguity is quite apparent in Furstenberg and Hughes' social capital measures. Does regular church attendance denote a rich social network to rely on in times of crisis? Or does it mean acceptance of strong ethical constraints on behavior, which Putnam (2000) evidently denies is a form of social capital (cf. the quotation in the Introduction)? Do neighbors' educational expectations constitute an imitative influence or a community norm? Definitional ambiguity makes identification impossible.

## **ii. Narayan and Pritchett**

Narayan and Pritchett's paper is concerned about issues of exchangeability and identification. Unlike Furstenberg and Hughes, they are careful about choosing control variables that render relatively plausible the assumption that their model errors are conditionally exchangeable. Further, they are sensitive to the issue that social capital may be endogenous and attempt to construct instrumental variables to account for this. Of course, they do not address the issue of the role of endogenous versus social capital and other contextual effects in their

expenditure model, so by Proposition 1, their analysis does not identify the magnitude of the social capital effect. However, by Proposition 2, they can claim, if one accepts the assumption that their social capital instruments are exogenous and uncorrelated with the regression errors in the causal expenditure equation, that some social capital effect exists.

However, if one fully considers the implications of the endogeneity of social capital, then one can question whether the instrumental variables chosen by Narayan and Pritchett fulfill the requirements for identification. Their instrument choice is predicated on the assumption that income is not directly causally influenced by levels of trust in strangers and/or trust in government officials. Consider first the variable that measures trust in strangers. From the perspective of Proposition 3, it is unclear why the trust variable is a valid instrument. Suppose that this trust instrument is determined by trustworthiness, i.e. that the trust variable reflects the actual behavior of others in a society. A natural way to make this operational is to assume that individual expenditures/income depend on the expected level of trust in society. This is hardly an uncontroversial assumption.<sup>3</sup> The growth and development literatures are filled with discussions of how investment and effort strategies depend on the honesty of others. Taken as a whole, these considerations suggest that an appropriate framework for studying social capital of this type is via equations (10) and (11). If so, then the Narayan and Pritchett analysis can only be completed by finding some  $X_{g(i)}$  variables that can be plausibly excluded a priori from the expenditure equation.

One can also question the use of the trust in government officials measure as an instrument. Unlike the trust in strangers instrument, one need not consider joint decisions on the part of individual households as a direct issue in assessing the instrument. Instead, the question of exchangeability is pertinent. One would generally think that differences in trust of officials have something to do with differences in their performances. But if so, would one assume that families who live in villages with high levels of trust in officials are partially exchangeable with families in villages that have low trust? And, given that there is presumably some correlation

---

<sup>3</sup>While controversial, the idea is also of ancient lineage. As pointed out by Landes (2000), Weber's classic argument on the role of Protestantism in the rise of capitalism was based on the idea that "Protestantism produced a new kind of businessman, one who aimed to live and work a certain way." (pg. 11).

across family incomes within villages due to common unobservables, it is also unclear that one can argue that these correlated elements are not related to whatever factors explain variation across villages in average trust in officials. In other words, it does not seem plausible to argue that differences in average trust across villages does not imply systematic differences in the unobserved components of individual expenditures.

Notice that the use of individual rather than group measures would not help even if the instrument validity test were passed. The individual variation within a village in assessing trust would, in order for the instrument to be valid, have to be uncorrelated with the residuals in the income equations. Now suppose that individual trust equals trustworthiness plus some idiosyncratic noise. Instrument consistency would require that this idiosyncratic noise is orthogonal to the residual in expenditure; it is difficult to see how one could claim this is a good basis for identification.

### ***iii. Knack and Keefer***

The Knack and Keefer study suffers from many of the problems that have plagued cross-country growth regressions in general (see Durlauf and Quah (1999) and Brock and Durlauf (2001b) for critiques of this literature). Drawing on the literature criticizing these studies, one may first argue that the Knack and Keefer analysis is based on an untenable set of exchangeability assumptions. This is true in two respects. First, Knack and Keefer assume that growth rates are partially exchangeable given their parsimonious specification of the determinants of growth (I focus here on their Tables 1 and 2). In order for this to be credible, one would have to believe both that their particular cross-country growth regressions capture all relevant causal determinants of growth and that the growth process that relates these determinants to a given country are described by a constant coefficient linear relationship.

Neither of these assumptions is tenable. With respect to the first requirement, if there is any message from the vast cross-country growth literature, it is that no consensus exists on which causal growth theories are empirically salient. As pointed out by Durlauf and Quah (1999) nearly 100 different variables have been used in cross-country growth regressions in order to capture different growth theories. While some progress has been made on variable selection in these contexts (cf. Levine and Renelt (1992), Doppelhofer, Miller and Sala-i-Martin (2000) and Fernandez, Ley, and Steel (2001)), the appropriate specification of cross-country regressions is

very much an open question. While Knack and Keefer do explore some aspects of model robustness (notably with respect to outliers), they do not establish that their choice of regressors is rich enough to avoid the problem that their findings of social capital effects may be resulting from omitted variables that both causally affect growth and are correlated with the social capital measures. As for the second, a number of studies have found evidence of multiple regimes in cross-country growth data (Durlauf and Johnson (1995), Canova (1999), Desdoigts (1999)). Hence, one cannot interpret Knack and Keefer's findings as saying something about common socioeconomic structure across countries.

From the perspective of identification, the ethnolinguistic heterogeneity variable also appears to be unsatisfactory. Notice that in order for the variable to facilitate identification, it would be necessary that it does not appear in the structural model for growth, the analog of eq. (10). This is the import of Proposition 4. But for this to be so, one would have to establish that it is not a variable in the reduced form for a complete structural growth model when some variables are not observed. Hence if such a model should include the endogenous variable political regime, and political regime is causally determined by ethnolinguistic heterogeneity, then that variable should appear as a variable in the reduced form for growth.

Further, this argument indicates how it is questionable to employ instrumental variables in this context. Suppose that one accepts that the ethnolinguistic variable is predetermined with respect to the growth process. Does this imply that it is a valid instrument? The answer is no, of course, unless one is prepared to assume that it is uncorrelated with the structural innovation in the growth process,  $\varepsilon_i$ . But what would this assumption mean? It would require that  $\varepsilon_i$  is uncorrelated with all the growth determinants that are omitted from the regression. Such a requirement is not tenable. As argued in Brock and Durlauf (2001b), growth theories are fundamentally open-ended, in the sense that the validity of one growth theory generally has no logical implications for the validity of another (although one can imagine covariation between them). Hence, the claim that social capital affects growth has no bearing on the statement that political regimes affect growth. Yet how does one argue that there is no relationship between ethnolinguistic heterogeneity and political regime, when the logic which justifies the use of the instrument (heterogeneity leads to lack of social cohesion) has well known implications for the stability of democracy, appeal of nationalism, and the like?

It is useful to compare their use of the ethnolinguistic heterogeneity variable to its use in Easterly and Levine (1997). Easterly and Levine (1997) argue that public good formation is difficult in the presence of ethnic conflict and therefore helps explain why growth in sub-Saharan Africa lags behind the rest of the world. Their instrument for ethnic conflict is the same variable used in Knack and Keefer. Granting that public good formation is related, at least vaguely, to social capital, these two papers nevertheless use the same variable as an instrument for different causal growth determinants.

## **5. What is to be done?**

The problems from which these empirical studies suffer are of course, at some level, endemic to all empirical work. As recognized by Heckman (2000), the establishment of causal relationships is fundamentally difficult:

Some of the disagreement that arises in interpreting a given body of data is intrinsic to the field of economics because of the conditional nature of causal knowledge. The information in any body of data is usually too weak to eliminate competing causal explanations of the same phenomenon. There is no mechanical algorithm for producing a set of "assumption free" facts or causal estimates based on those facts. (pg. 91)

That being said, the empirical social capital literature seems to be particularly plagued by vague definition of concepts, poorly measured data, absence of appropriate exchangeability conditions, and lack of information necessary to make identification claims plausible. These problems are especially important for social capital contexts as social capital arguments depend on underlying psychological and sociological relations that are difficult to quantify, let alone measure.

These problems suggest, in my view, two distinct points. First, in using observational studies, it seems clear that researchers need to provide explicit models of the codetermination of individual outcomes and social capital, so that the identification problems that have been analyzed may be rigorously assessed. Second, I would argue that worries about various exchangeability assumptions suggest the need for some new directions in data collection for social capital research.

## ***i. experiments***

Social capital seems intrinsically tied up with psychological notions such as trust, aspirations, and group identification. Interestingly, many of the ideas that are under debate in the social capital literature have received very extensive attention in the social psychology literature. My reading of this literature suggests that social psychological experiments have been successful in establishing evidence of specific social interaction effects. These effects can be used to sharpen the vague ideas surrounding the nature of social capital. Further, experiments hold the promise of constructing situations where questions of exchangeability and identification can be avoided via the appropriate use of randomization. In some respects, the Rubin (1974) formulation of causal inference in random experiments may be interpreted as a way of avoiding exchangeability and identification problems of the type often associated with observational data (see Rosenbaum (1995) for an outstanding overview of how randomization facilitates causal inference).

To see how the social psychology literature has successfully produced evidence of social interactions, consider the famous Robbers' Cave experiment, which has been called "the most successful field experiment ever conducted on intergroup conflict." In this experiment, psychologist Muzafer Sherif and colleagues (Sherif et al (1961)) studied the behavior of a group of teenage boys who were brought to an isolated retreat located in Robbers Cave State Park in Oklahoma. Prior screening on the basis of family background allowed the researchers to put together a homogeneous group of boys. The boys were broken into two groups who initially were not aware of the existence of one another. After a week, the groups were asked to assume names and chose Rattlers and Eagles respectively. A set of competitive activities was initiated between the two groups. Sherif et al (1961) documents in great detail how the two groups developed strong senses of group identity as well as differing internal behavior norms. Further, each group exhibited great animosity toward the other, animosity that carried over beyond the competitive activities. It became commonplace for the boys to attribute negative stereotypes to the other group; overt hostility bordering on violence even emerged.

This experiment illustrates how easy it is for group identification to influence behavior and trust. Further, the experiment suggests a much more precise causal mechanism than one finds in the social capital literature. To give a simple example, the mapping from group

identification to stereotyping (both positive and negative) that Sherif and coauthors established is far more useful for drawing specific claims about the effects of groups on individuals than the finding that church membership correlates with high school completion. And the design of the experiment renders inoperative worries about omitted control variables, differences between groups, etc.

Notice as well that an experiment of this type provides insights that have largely been ignored in the social capital literature. Relative to the social capital literature, the findings of the Robbers' Cave experiment suggest that the general presumption that strong social bonds are good may be mistaken. The intergroup hostility engendered between the Rattlers and Eagles suggests that the insouciant claims of Putnam (2000) on the virtues of social capital, are very likely exaggerated; see Portes (1998) and Durlauf (2000) for discussions of the negative effects of social capital on individuals and communities.

For these reasons, the use of experiments in Glaeser, Laibson, Scheinkman, and Soutter (2000) seems as compelling a study of the determinants of trust, and hence of an understanding of the causal mechanisms by which social capital-like phenomena emerge, as any which have appeared in the economics literature. What these authors do is conduct a range of experimental trust games for subjects where substantial background information is solicited. Among the interesting results the authors obtain is that trustworthy behavior is much more likely in games in which the players possess common ethnicity. This paper also shows how the behavior observed in experiments has implications for the interpretation of the sorts of conventional survey data used in the social capital literature to measure notions such as trust. For example, measures of trust from surveys correlate with trustworthy behavior in games. The finding of how trustworthy behavior is related to attitudes is the sort of evidence from which a model of individual decisionmaking can be built that explains how social capital emerges in the aggregate, cf. Glaeser, Laibson, and Sacerdote (2001).

## ***ii. survey data***<sup>4</sup>

---

<sup>4</sup> This discussion borrows from Durlauf (2000).

Survey data collection is a second area that can facilitate the study of social capital mechanisms and is a natural area for exploration given the links that exist between social capital ideas and social structure. In this regard, an extremely important project is currently underway, the Project on Human Development in Chicago Neighborhoods. The Project on Human Development in Chicago Neighborhoods is designed to produce a rich data set on attitudes among Chicago residents on a wide range of issues. In 1995, over 8000 individuals were surveyed across over 300 neighborhood clusters. What is critical in the study is the rich set of information that is produced which allows for the integration of information about individual characteristics with information on individual attitudes in order to study how these relate to communities-i.e. the social environment.

This data set has provided insights into a very wide range of phenomena. For our purposes, one aspect may be highlighted. Sampson, Morenoff, and Earls (1999), for example, find that even if one restricts attention to poorer neighborhoods, there is wide variation in the residents' expectations of the behavior of their neighbors and that this variation helps predict differences in neighborhood social problems. For example, for poor neighborhoods where individuals feel unable to rely on neighbors to report truancy or call the police in response to observing illegal activity, various social pathologies will be more serious. This sort of finding in turn is very suggestive of the role of community institutions in ameliorating social problems and indeed fulfills the authors' objective of moving beyond the typical vague formulations of social capital.

Relative to standard empirical analyses of social capital, this work has several advantages. First, the data set gathered in this project provides much richer controls for individual heterogeneity than are typically available. Second, the detailed attitudinal measurements in the study extend social capital analyses in directions that are far more conducive to the description of the causal mechanisms by which social capital is created. The expectation of neighbors' behavior which Sampson, Morenoff, and Earls describe gives a far more compelling vision of the role of community networks in influencing group outcomes than a cross-country regression of growth rates on vague measures of trust. Third, the detailed nature of the study may provide ways to characterize the endogenous formation of social capital, something that is critical for establishing identification of social capital effects, as illustrated in Section 3.

One should not, of course, regard survey data as any sort of panacea. As Manski (1990) has argued, there are many subtleties involved in the elicitation of subjective expectations from conventional survey data, particularly in mapping the sorts of qualitative measures often solicited into the sorts of quantitative measures needed for statistical analysis. However, as shown for example in Dominitz and Manski (1997), one can design surveys that permit one to elicit beliefs of the type that is needed. What Dominitz and Manski show is that by designing questions that require the respondent to give probabilistic assessments concerning his beliefs, one can overcome many of the problems that have traditionally disinclined economists to use survey data. While their approach has yet to be applied to the evaluation of beliefs about group behaviors and group norms that are key features of what is usually meant by social capital, it can in principle be done.

## **6. Conclusions**

Recent research on social capital has had a very salutary effect on economics through its introduction of richer sociology into standard reasoning. However, the concept itself has proven to be too vague to permit analysis whose clarity and precision matches the standards of the field. In this paper, I have attempted to illustrate some basic econometric difficulties with the social capital literature. In light of the vagueness of the concept, I believe that the use of observational data to identify substantive forms of social capital is unlikely to be successful. The relatively more compelling evidence from the social psychology literature, in contrast, suggests that economic experiments may be a more promising way to obtain empirical insights.

## **References**

Bowles, S. and H. Gintis. (2002). "Social capital and community governance." *Economic Journal*, this issue.

Brock, W. and S. Durlauf. (2001a). "Interactions-Based Models," in *Handbook of Econometrics*, Vol. 5 (ed. J. Heckman and E. Leamer), pp. 3297-3380. Amsterdam: North Holland.

Brock, W., and Durlauf, S. (2001b), "Growth empirics and reality." *World Bank Economic Review*, vol. 15, no. 3, pp. 229-272.

Brown, R. (1986). *Social Psychology: The Second Edition*. New York: Free Press.

Canova, F. (1999). "Testing for convergence clubs in income per-capita: a predictive density approach." Mimeo, University of Pompeu Fabra, Spain.

Desdoigts, A. (1999). "Patterns of economic development and the formation of clubs." *Journal of Economic Growth*, vol. 4, no. 3, pp. 305-330.

Durlauf, S. and P. Johnson. (1995). "Multiple regimes and cross-country growth behavior." *Journal of Applied Econometrics*, vol. 10, pp. 365-384.

DiPasquale, D. and E. Glaeser. (1999). "Incentives and social capital: do homeowners make better citizens?" *Journal of Urban Economics*, vol. 45, pp. 354-384.

Dominitz, J. and C. Manski. (1997). "Using expectations data to study subjective income expectations." *Journal of the American Statistical Association*, vol, 92, pp. 855-867.

Doppelhofer, G., Miller R. and X. Sala-i-Martin. (2000). "Determinants of long-term growth: a bayesian averaging of classical estimates (BACE) approach." National Bureau of Economic Research Working Paper no. 7750.

Draper, D., Hodges J., Mallows C. and Pregibon D. (1993). "Exchangeability and data analysis." *Journal of the Royal Statistical Society, series A*, vol. 156, pp. 9-28.

Durlauf, S. (1999). "The case "against" social capital." *Focus*, vol. 20, no. 3, pp. 1-4.

Durlauf, S. (2000). "The memberships theory of poverty: the role of group affiliations in determining socioeconomic outcomes." In *Understanding Poverty in America* (eds. S. Danziger and R. Haveman). Cambridge: Harvard University Press, forthcoming.

Durlauf, S. (2002). "Bowling alone: a review essay." *Journal of Economic Behavior and Organization*, forthcoming.

Durlauf, S. and P. Johnson. (1995). "Multiple regimes and cross-country growth behavior." *Journal of Applied Econometrics*, vol. 10, pp. 365-384.

Durlauf, S. and Quah, D. (1999). "The new empirics of economic growth." In *Handbook of Macroeconomics* (eds. J. Taylor and M. Woodford), pp. 235-310. Amsterdam: North Holland.

Easterly, W. and Levine, R. (1997). "Africa's growth tragedy: policies and ethnic divisions." *Quarterly Journal of Economics*, vol. CXII, pp. 1203-1250.

Fernandez, C., Ley E., and Steel, M. (2001). "Model uncertainty in cross-country growth regressions." *Journal of Applied Econometrics*.

Freedman, D. (1991). "Statistical models and shoe leather." In *Sociological Methodology 1991* (ed. P. Marsden). Cambridge: Basil Blackwell.

Furstenberg, F. and Hughes, M. (1995). "Social capital and successful development among at-risk youth." *Journal of Marriage and the Family*, vol. 57, pp. 580-592.

Fukuyama, F. (1999). *The Great Disruption*. New York: Simon and Schuster

Glaeser, E., Laibson, D., Scheinkman, J. and Soutter, C. (2000). "Measuring trust." *Quarterly Journal of Economics*, pp. 811-846.

Glaeser, E., Laibson, D. and Sacerdote, B. (2001). "The economic approach to social capital." *Economic Journal*, this issue.

Goldin, C. and Katz, L. (2001). "Human capital and social capital: the rise of secondary schooling in America, 1910-1940. In *Patterns of Social Capital* (ed. R. Rotberg), Cambridge: Cambridge University Press.

Granovetter, M. (1985). "Economic action and social structure: the problem of embeddedness." *American Journal of Sociology*, vol. 91, no. 3, pp. 481-510.

Heckman, J. (2000). "Causal parameters and policy analysis in economics: a twentieth century retrospective." *Quarterly Journal of Economics*, vol. CXV, pp. 45-97.

Helliwell, J. and Putnam, R. (1995). "Economic growth and social capital in Italy." *Eastern Economic Journal*, vol. 21, no. 3, pp. 295-307.

Kawachi, I., Kennedy, B., Lochner, K. and Prothrow-Smith, D. (1997). "Social capital, income inequality, and mortality." *American Journal of Public Health*, vol. 87, pp. 1491-1498.

Kennedy, B., Kawachi, I. and Brainerd, E. (1998). "The role of social capital in the Russian mortality crisis." *World Development*, vol. 26, no. 11, pp. 2029-2043.

Knack, S. and Keefer, P. (1997). "Does social capital have an economic impact? A cross-country investigation." *Quarterly Journal of Economics*, vol. CXII, no. 4, pp. 1252-88.

Landes, D. (2000). "Culture makes almost all the difference." In *Culture Matters* (eds. L. Harrison and S. Huntington), New York: Basic Books.

LaPorta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. (1997). "Trust in large organizations." *American Economic Review*, vol. 87, pp. 333-338.

Levine, R. and Renelt, D. (1992). "A sensitivity analysis of cross-country growth regressions." *American Economic Review*, vol. 82, pp. 942-963.

Manski, C. (1990). "The use of intentions data to predict behavior: a best case analysis." *Journal of the American Statistical Association*, vol. 85, pp. 934-940.

Manski, C., (1993), "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, vol. 60, pp. 531-542.

Narayan, D. and Pritchett, L. (1999). "Cents and sociability: household income and social capital in rural Tanzania." *Economic Development and Social Change*, vol. 47, no.4, pp. 871-897.

Portes, A. (1998). "Social capital: its origins and application in modern sociology." *Annual Review of Sociology*, 1-14.

Putnam, R.. (2000). *Bowling Alone*. New York: Simon and Schuster.

Rosenbaum, P. (1995). *Observational Studies*. New York: Springer-Verlag.

Rubin, D. (1974). "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of Educational Psychology*, vol. 66, pp. 688-701.

Sampson, R., Morenoff, J. and Earls, F. (1999). "Beyond social capital: collective efficacy for children." *American Sociological Review*, vol. 64, pp. 633-660.

Sherif, M., Harvey, O., White, B., Hood, W. and Sherif, C. (1961). *Intergroup Conflict and Cooperation: The Robbers Cave Experiment*. Norman: Institute of Group Relations, University of Oklahoma, reprinted by Wesleyan University Press, 1988.

Woolcock, M. (1995). "Social capital and economic development: towards a theoretical synthesis and policy framework." *Theory and Society*, vol. 27, pp. 151-208.