MEASURING SOCIAL CAPITAL AND ITS ECONOMIC IMPACT

June, 2000

John T. Durkin, Jr
NORC
Harris Graduate School of Public Policy Studies
University of Chicago
1155 E. 60th St.
Chicago, IL 60637
313-577-2693
durkin-sean@norcmail.uchicago.edu

ABSTRACT

This paper assesses the quality of various micro level measures of social capital. It begins by examining a simple model of social capital accumulation in which households acquire social capital by devoting resources to forming relationships with other households. Through these relationships households acquire access to social resources which raises utility for any level of consumption. Tests of the model reveal that there is no statistically significant relationship between widely used proxy variables for social capital, measures of trust and group membership, and access to social resources. However, variation in measures of the frequency of contact with family members and friends is positively related to differences in access to social resources. The paper then tests whether evidence that social capital has a significant economic impact depends on the use of these different measures of social capital. It does so by investigating whether the impact of the average level of human capital on individual wages is an increasing function of a person's social capital. The impact on wages of the share of the population within a city or county that has a bachelor's degree is indeed an increasing function of the frequency of contact measure, but it is unrelated to the measures of trust and group membership.

1. Introduction

Despite the increase in interest among economists in the notion of social capital, many remain skeptical that it will ever prove to be a useful analytic tool. Others have been more critical of research in this area. The criticism has taken three main forms. First, Durlauf (1999), Bowles (1999), Manski (2000) and others have argued that there still exists a great deal of ambiguity about what the term represents and whether it conveys ideas that are conceptually new. Within the sociology literature at least, Portes (1998, p. 6)) argues that there is a growing concensus that "social capital stands for the ability of actors to secure benefits by virtue of membership in social networks or other social structures." A second set of criticisms has to do with the measurability of social capital. Some, such as Solow (1995), have questioned whether social capital is measurable, while many have focused on the particular proxy variables used, survey responses to questions about trust and group membership. A final criticism, based in part at least on the previous one, is that the empirical research has not presented convincing evidence that social capital has any significant economic effects.

Human capital, like social capital, is an unobservable force, so it too is subject, in theory at least, to the same kind of criticisms. This raises an interesting question. Why do economists have so much more confidence in the notion of human capital than in social capital? The answer seems to stem from two sets of results. First, tests of simple theoretical models of human capital formation using variables which should be related to the stock of human capital, wages, are related in a consistent way with variables which should be related to previous investment, education and experience. Second, a significant amount of research has found, using education as a proxy, that human capital has effects on economic outcomes other than simply wages.

While the results from Mincerian regressions have played a large role in persuading economists that human capital is a useful analytic tool, the social capital research has provided no such convincing results. There has been some theoretical modeling of social capital as a choice variable, some efforts to explain variation in proxies for social capital and some tests of whether variation in proxies for social capital have economic effects. There has not, however, been any effort to test whether a variable related to investment in social capital is related to some

measure of the stock of social capital consistent with a theoretical model of social capital formation in the same way that wages, education and experience are related.

The failure to run these kind of tests has had two implications. First, social capital can be defined in any way one pleases if one is not going to actually test a model of social capital accumulation. Tests of such a model would presumably help reduce the ambiguity about what the term represents by constraining the definition to what is testable. Second, because there is no evidence that the various proxies which have been used are related to social resources, it is hard to convince others that the use of the proxy as a right hand side variable indicates that social capital has economic effects. Again, presenting such evidence that would go a long way toward convincing the skeptics.

This paper makes two contributions. It presents evidence that the widely used proxies for social capital have no explanatory power in a test of a simple model of social capital accumulation but shows that some alternative measures are consistent with the theory. Second, using one of these alternative measures, it present results suggesting the importance of peer effects is an increasing function of one's social capital. The particular peer effect that is considered is the impact of the average level of human capital on individual wages.

The paper is organized as follows. The first section writes down of simple model of social capital accumulation. The model helps clarify exactly what is meant by social capital and suggests how one might hope to test these relationships. Social capital provides households with access to social resources which raises utility for any level of consumption. People acquire social capital by forming relationships with others. Investment in social capital reduces consumption because time devoted to social interaction reduces time spent working. The optimal share of resources devoted to social capital investment equates the utility loss from foregone consumption to the utility gain from higher social capital in the future.

The next section focuses on the data used to test the theory. It begins by describing the data used by other economists to proxy for social capital and then characterizes the ideal data. It concludes by describing data contained in the 1986 General Social Survey which is used to test

the theory. Access to social resources is proxied for by survey responses to hypothetical questions about who respondents would turn to for help in a variety of different situations. Social capital is in turn proxied for by answers to questions about the number of family members and friends and the frequency of contact with family members and friends as well as the questions about trust and group membership used by others.

The next section presents the results of testing the theory. There is no evidence of any systematic relationship between the measures of trust and group membership and access to resources through their social relationships. This is true in both single variable regressions and in regressions with a set of other control variables. Using the number of friends variables and, to a lesser extent, the number of family members produces similar results. However, the variables constructed from responses to questions about the frequency of contact with friends and relatives do have explanatory power in both the single and multiple variable regressions. Stated differently, the results suggest a positive relationship between the frequency of contact with family members and friends and access to resources through social relationships.

The final section examines whether there is evidence that social capital has an economic impact using the various measures of social capital employed in the previous section. One way that social capital might have an economic impact is by affecting the importance of peer effects. This issue is examined in the context of the research into whether individual wages depend on the average level of human capital or that the social returns to education exceed the private returns. More specifically, the analysis examines whether the impact of the average level of human capital within a city or country on individual wages is an increasing function of a person's social capital. Using the share of people with at least a bachelor's degree as a measure of the average level of human capital, the results indicate that the impact of the bachelor's degree percentage on individual wages is positively related to a person's frequency of contact with friends and relatives. These results are robust to the inclusion of a variety of different control variables including differences in the cost of living and to alternative instruments for the bachelor's degree share, and they hold even for a sample consisting only of college graduates.

There is, however, no evidence that these relationships hold when the measures of trust or group membership are used to proxy for social capital.

2. The Theory

Following Becker (1974), suppose households "produce" a single, composite household commodity using a composite consumption good and social resources. Let instantaneous utility at time t be given by[1]:

$$U_i(t) = Z_i(t) = Z(c_i(t), S_i(t)) \tag{1}$$

where $c_i(t)$ and $S_i(t)$ denote the consumption good and social resources respectively.

Next, suppose household have access to these resources through their social capital which is embodied in their relationships with other households.[2] Define $n_i(t)$ as the number of households with whom household i has a relationship at time t and write:

$$S_i(t) = S(n_i(t)) \tag{2}$$

with $S' \geq 0$ and $S'' \leq 0$.

Assume households acquire relationships by devoting resources to social interaction, and define $\sigma(t)$ as the share of resources devoted to social interaction. Let the time path of n follow:

$$\dot{n}_i(t) = n_i(t)g(\sigma_i(t)) - \delta_s n_i(t) \tag{3}$$

[1] . One could just as easily refer to (1) as a utility function. However, the fact that S is interpreted as services derived from a capital stock suggests, perhaps, that a more natural interpretation is that it represents a household production function.

[2] . Coleman (1990, Chapter 12) emphasizes the fact that social capital is embodied in relationships rather than individuals.

where δ_s denotes the constant rate of depreciation.

Suppose labor is the only source of household income and that households have one unit of time. Define w_i as the wage, and ignoring for now any economic effects of social capital, assume that the wage is constant. This means that consumption is given by:

$$c_i(t) = w_i(1 - \sigma_i(t)) \tag{4}$$

where $(1 - \sigma)$ is the share of time devoted to working.

Equations (1)-(4) allow a reasonably precise definition of social capital. Social capital represents those features of relationships that provide individuals/households access to social resources which raises household utility/output for any level of consumption.[3] Implicit in this specification is the notion that current interaction increases the probability of future interaction. This is true because people would only be willing to provide resources to others with whom they have relationships if they could reasonably expect that they will be compensated in the future. As the probability of future interaction increases, people become more confident that there will be opportunities for reciprocation and, therefore, more willing to share resources with each other. Thus, a relationship exists when people can reasonably expect that there will be future interactions which would provide opportunities for reciprocation. While not modeling these effects explicitly, these assertions are consistent with the Folk theorem for repeated games which indicates that repeated interaction can yield cooperative behavior provided that the probability of future interaction is sufficiently high.

Relationships and, therefore, social capital can depreciate in the following sense. If households do not devote sufficient resources to an existing relationship, then they are less

[3] . Note the similarity between this definition and Portes' definition cited above. Schiff (1992) has a similar definition in terms of utility functions. It should also be noted that the idea that social capital is a product of the individual is quite different from Putnam's (1993) argument that social capital is a feature of communities or larger political entities. There is some disagreement that whether it is more appropriate to think about social capital as a macro or micro feature. See Glaeser et. al. (2000) for a discussion.

certain that future interactions will occur and less certain that they can count on each other for social resources in the present. While it is impossible to quantify the rates at which social capital and human capital depreciate, intuition suggests that social capital probably depreciates at a faster rate. You may not be able to get much in the way of resources from someone that you have not had any contact with for a few years, while one's human capital is not likely to depreciate much over the same time frame.

The resources one can have access to because of social capital might be something as esoteric as friendship or companionship or as concrete as having someone watch your home when you are out of town. Equation (2) suggest that having more relationships with neighbors means a household is more likely to be able to count on one of them to look after its house. As in the home watching example and many others like it, social capital is often associated with resources for which the provision implies very low costs to the providers. Thus, while a market could and in some cases even may exist for purchasing the resources, the market price for these resources often greatly exceeds the cost to provider.

An example of resources provided by social capital that Coleman (1990, Chapter 12) emphasizes is information. Individuals can often rely on others within their social network for information that they themselves have not devoted time to acquiring. If more information raises household output for a given level of consumption, then the theory suggests that those households which devote more time to social interaction should have more relationships, more information and higher household output for any level of consumption.

It is worth noting that writing social capital as a function of a one dimensional factor in (2) mask several important features of a person's social relationships that have been emphasized in the social networks literature. For example, access to information depends not just on the total number of relationships but also on such factors as the breath of those relationships, the strength of those relationships, and the knowledge of those with whom one has relationships. On the other hand, thinking about social capital as function of the total number of relationships considerably enhances analytic and empirical tractability. Coleman (1990, Chapter 12) makes an argument

along these line by suggesting that the concept of social capital is most useful as an analytic tool when one is more interested in the functions of the social structures than the details of the structures themselves. He argues that the notion of social capital might allow economists to consider the impact of social structures on economic behavior without getting caught up in the details of the structures themselves or the details of how these structures are formed. Moreover, a similar type of exercise has proved very helpful in the human capital literature. There are many different components of a worker's skills which are ignored in thinking about productivity as a function of a one dimensional factor such as human capital.

To complete the model, suppose households begin with a given number of relationships, $n(0) = n_0$ and live and work T years. Let the intertemporal utility function be given by:

$$V(t) = \int_0^T F(Z(t))e^{-\rho t}dt \tag{5}$$

Given an initial n, households choose a time path of σ to maximize (5) subject to (4).

Using (1)-(5) and standard techniques of optimal control yields the following first order condition for the optimal σ:

$$F'(Z(t))Z_2(t)w = \lambda(t)n(t)g'(\sigma(t)) \tag{6}$$

The shadow price of human capital in utility terms is given by $\lambda(t)$, and Z_2 denotes the derivative of Z with respect to its second argument. The left hand side of (6) represents the marginal cost of investment in social capital. It is the utility loss from foregone consumption. The right hand side represents the marginal benefit which is the future utility gain from higher levels of social capital.[4]

[4] . Glaeser et. al. (2000) derive a related first order condition.

The solution for the optimal time path of σ and n is obvious once one recognizes that simply renaming the variables in (1)-(4) reproduces a very familiar model. If S, n and c are renamed income (consumption), human capital and leisure respectively, then the model represents a version of the standard life cycle model of human capital accumulation with leisure.[5] As such, the optimal time paths of σ and n are shown in Figure 1. The share of time devoted to social interaction falls, though not necessarily linearly, over time and reaches 0 at t = T, and n rises initially but eventually falls when σ reaches a sufficiently low level that net investment becomes negative.

The next task in the paper is to construct variables related to social resources S to see whether existing measures of social capital, trust and group membership variables, can explain variation in S. One obvious implication of the theory is that if these variables reflect differences across individuals in their previous investment or the stock of social capital then they should explain variation in S independent of anything else. However, if these variables reflect differences in current investment in social capital, then one might only find a statistically significant relationship between these variables and S conditional on other factors. To see this note that two households with the same S can have very different σ depending on where they are in the life cycle, differences in S_0 and differences in other factors affecting the costs and benefits of investment in social capital. This is important because it is not at all obvious whether these proxies for social capital should be thought of as capturing differences in stocks or differences in current rates of investment.

3. The Data

One of the most commonly used proxies for social capital has been the number of group or organization memberships. The General Social Survey (GSS) has been asking respondents on

[5] . The main difference between this model and the life cycle earnings models which feature endogenous labor supply is that there is only one choice variable. For a survey of this literature, see Weiss (1986).

an almost yearly basis since 1972 whether they belong to a variety of different groups.[6] Responses to these questions are used to construct a variable representing the total number of group memberships.[7] This variable was used initially by Putnam (1993) and has also been used by Knack and Keefer (1997), diPasquale and Glaeser (1999), Alesina and LaFerrara (1999) and Glaeser et. al. (2000).

The strength of this measure lies in the fact that one of the ways in which people interact is through organizations. The main weakness of this measure, which has been noted by many, is that it does not capture variation in the extent to which people are active in the various groups. For example, most economists belong to the AEA, but few of them get any social capital as a result of their membership per se. However, it may be the case that those who are not members are less likely, other things equal, to interact with others.

The other main measure of social capital used by economists has been responses to survey questions about trust. The GSS has been asking on an almost yearly basis whether respondents believe that, in general, people can be trusted.[8] This variable is also used by Knack and Keefer (1997), and by LaPorta, et. al (1997), Glaeser, et. al. (1999), Alesina and LaFerrara (2000) and Guiso et. al. (2000).

The strength of this variable is that, as described above, trust is a crucial element in access to social resources. People provide resources to those with whom they have relationships because they trust that they will be reciprocated. The obvious weakness of this measure is that the question is so ambiguously worded, it is not at all clear what is being measured. Indeed, Glaeser, et. al. (2000) find that those who respond by agreeing that people in general can be trusted are less likely to trust in experiments.[9]

---

[6] . The same questions have also been asked on versions of the International Values Survey (IVS). For information on the IVS see http://wvs.isr.umich.edu/index.html.
[7] . See the data appendix for detailed description of this variable.
[8] . This variable is also available in some versions of the IVS.
[9] . They do, however, find a positive relationship between trustworthiness and responses to this questions.

As with human capital, social capital is an unobservable force. If one were interested in testing the model described above, how might one go about doing this and what would be the ideal data? Ideally, one would like to have some measure of the resources people have access to through their relationships. If one could then find measures of either the number of relationships that people have or the amount of time they have devoted to interacting with their family members, friends, neighbors and co-workers, one could test whether these measures of social capital help explain variation in access to social resources.

While far from ideal, the 1986 GSS does contain two sets of variables relevant for these kind of exercises. The survey elicits detailed information about the family structure of respondents from which it is possible to construct a variable indicting the number of adult relatives for each respondent (RELNUM). Respondents were also asked how many close friends they have (FRINUM). There are also a number of variables which indicate how frequently respondents have contact with these relatives (RELFRE), either directly or by phone or letter, and with their closest friend (FRIFRE). These variables and their sums (SC = RELNUM + FRINUM and SCFRE = RELFRE + FRIFRE) are used as alternative measures of social capital. A detailed description of all the variables used in the paper can be found in the data appendix.

The second set of relevant variables are used to proxy for access to social resources or S. In the survey, respondents were asked a series of hypothetical questions about who they would turn to for help a variety of situations. For example, they were asked who they would turn to first for help with a household chore or garden job that they could not do alone. Possible responses include family members, friends, neighbors, co-workers, no one, some you pay or someone else. Respondents are also asked who they would turn to second for help. Responses to these questions are recoded into a dummy variable (CHORE) which takes the value of 1 if the respondent has at least two family members, friends, neighbors or co-workers that he or she could turn to for help and 0 otherwise. Respondents are also asked who they would turn to for help if they were sick and needed help around the house, if they needed to borrow money, if they needed to discuss a problem in their relationship with their spouse or partner, if they were

depressed and if they needed advice on a major life change. Responses to these questions are recoded in the same way into the dummy variables SICK, BORROW, UPSET, DOWN and CHANGE respectively.

A final feature of the 1986 GSS which is attractive is that it also contains the questions on group membership (MEMNUM) and the trust (TRUST) question. Table 1 contains descriptive statistics on all these variables. Table 2 shows the correlation coefficients among all of the measures of social capital. The relatively low correlation coefficients among these variable, except in the cases where one of the variables is used to construct the other, and the fact that many of the coefficients are negative suggests that these variables are measuring different things. On the other hand, the correlation coefficients between the variables used to proxy for S in Table 3 are in 14 of the 15 cases positive and significant at the 1% level.

4. Results

One can use, with some degree of confidence, years of education as a proxy for human capital in testing for the impact of human capital on something other than wages in large part because of the results of Mincerian regressions. If one wants to use a variable to proxy for social capital to test whether social capital has an economic impact, then one would presumably want to show that the variable is related to a model of social capital formation.

Table 4 contains the first set of results which attempt to do so. Each cell contains the results of a single variable logistic regressions in which one of the measures of S is the dependent variable and one of the measures of social capital is used as the regressor. For example, the cell in the first column and first row contains the results of the following regression:

$$\text{Prob}(\text{CHORE}_i = 1) = a + b\text{TRUST}_i + u_i$$

For the cell in the second row of the first column, CHORE is again the dependent variable and LMEMNUM (1 + the natural log of MEMNUM) is the regressor.

The first two rows of the Table report the results when TRUST and LMEMNUM are used as dependent variables. In only one of the 12 regressions, the regression of SICK on LMEMNUM, is the estimate positive and significant. The estimates are negative in three of the regressions when TRUST is the regressor and in four when LMEMNUM is used. In two of the cases in which the estimates of the LMEMNUM coefficient are negative, they are statistically significant. In the remaining four regressions, the estimates are positive but not significant. Moreover, the very low pseudo R^2's indicate that these variables have very little explanatory power independent of their sign.

The next three rows report the results when the number of family members, friends and the sum of the two variables are used as explanatory variables. The results with LFRINUM are very similar to the results with LMEMNUM, while the results with LRELNUM are more consistent with the model. In this case, only one of the coefficients is negative, and in three of the five regressions with positive coefficients, the estimates are significant. Not surprisingly, the results with LSC are not as good as with LRELNUM but better than with LFRINUM.

The final three rows report the results with the frequency of contact variables. In each of the 18 regressions, the estimates of the coefficients are positive. They are also significant except when CHANGE is the dependent variable. The pseudo R^2's, while still relatively low, are in general much higher than in the previous regressions.

The main findings from Table 4 can be summarized as follows. First, the two main variables which have been used to proxy for social capital have very little power to explain variation in access to social resources. Given the criticisms of these variables discussed above, the results are not particularly surprising. Second, the number of friends variable does not do much better. This may be due to the subjective nature of the question. How people classify those they know as "close friends" could differ significantly. However, using the number of family members to predict variation in social resources does produce some results which are consistent

with the model perhaps due to the less subjective nature of the question. This is also not inconsistent with the results from the social networks literature which emphasizes the importance of family members in people's social networks. Third, the frequency of contact with family members and friends variables generate signs consistent with the model, and have more explanatory power than the other variables. Access to social resources depends less on simply having many family members than how much contact you have with them. In addition, the definition of how frequently one has contact with one's closest friend is less likely to suffer from idiosyncratic definitions than the number of friends question. Fourth, the coefficients and the pseudo R^2's are highest when LSCFRE is used as the regressor.

It is also worth noting that there are significant differences in the results across dependent variables. The dependent variables which seem to work the best are those that reflect access to more tangible assistance like chores, help around the house or borrowing money than with more emotional/psychological assistance. This is not particularly surprising given the way social capital is being measured. Those that have frequent contact with a large number of different people might have greater access to more tangible resources, while those with a small number of close relationships may have greater access to emotional/psychological assistance. There are also differences across variables in the extent to which organized markets exist for the services. The markets for services such as CHORE and SICK are not particularly well organized, while the markets for borrowing money and getting help with depression are organized. The fact that the results are stronger for services in which markets are less organized is consistent with the earlier argument that social capital often provides access to resources for which the market price greatly exceeds the cost to the providers.

These differences in results across dependent variables are important. Economists tend to be very suspicious of responses to subjective questions let alone hypothetical questions.[10] The fact that there are plausible explanations for why the relationship between the social capital

[10] . See Manski (2000) for a argument against this bias.

variables and the social resource variables are stronger in the cases of more tangible resources and in cases in which markets are less organized lends some credence to the contention that variation in the dependent variables is indeed consistent with variation in access to social resources.

One possible explanation for why the trust and group membership variables are unrelated to the social resources variables is that they are better flow than stock measures. One could make a plausible argument to this effect particularly for the group membership variable. If they are actually measuring variation in current investment in social capital rather than the stock of social capital, then the theory indicates that they should only be related to the social resources variables conditional on where people are in the life cycle and other factors affecting (6).

Table 5 contains the results of rerunning some of the regressions in Table 4 while controlling for life cycle effects and other idosycratic differences in costs and benefits to investment in social capital. Column (1) of Table 5A, for example, contain the results of the following regression:

$$\text{Prob}(\text{CHORE}_i = 1) = a + b_1 X_i + b_2 \text{TRUST}_i + u_i$$

Elements of the X vector include linear and quadratic age terms, years of schooling, a black dummy, a male dummy, a married dummy and Catholic, other religion and no religion dummies.[11]

There are two important features of the results in Table 5. First, adding these additional variables has virtually no effect on the sign, size and signficance of the coefficients for TRUST, LMEMNUM AND LSCFRE. Once again, the trust and group membership variables produce weak and sometimes inconsistent results, while the frequency of contact with family members

[11]. These regressions were also run using wages and household income as controls with almost no change in the results. The regressions without these variables reported because there are many more missing values with these variables, particularly the wage variable, which complicates the comparison with the results in Table 4.

and friends is always positive and is significant except when CHANGE is used as the dependent variables. Hence, it does not matter whether the trust and group membership variables are thought of as stock or flow terms. They still have no explanatory power.

The fact that the estimates of the LSCFRE coefficient are not much different in Tables 4 and 5 is, perhaps, somewhat surprising. This is because measures of frequency of contact seem better proxies for flow rather than stock terms. One possible explanation for the similarity in the results in the two tables is, as suggested earlier, that social capital has a particularly high rate of depreciation. If so, then stocks and flows are closely related, and a good measure of the rate of investment might serve as a better proxy for the stock than a bad measure of the stock.

The second important feature of Table 5 is that there is little if any evidence of the life cycle kind of effects predicted by the theory. While there is not a consistent pattern to the estimates of the coefficients on the linear and quadratic age terms, the evidence suggests, if anything, a u-shaped relationship between age and access to social resources.

In one sense, the lack of any significant life cycle effects consistent with the theory is not unexpected. The simple model described above assumes that wages are constant and ignores the decision to invest in human capital. This may be particularly important to the extent that time devoted to investment in human capital is time that could otherwise be devoted to investment in social capital. Since both are subject to life cycle effects, then the life cycle effects of human capital may offset the life cycle effects of social capital. Furthermore, as suggested by many, social capital itself may affect the return to investment in human capital. These issues are dealt with in the next section.

It is also worth investigating how the results change if the responses to the social resource questions are aggregated into a single variable. The simplest way of doing this is to create a new variable (SOCCAP) by summing the 6 dummy variables used as dependent variables `in Tables 4 and 5. Running the following OLS regression:

$$SOCCAP = \alpha_0 + \alpha_1 LSCFRE_i + v_i \qquad (7)$$

generates an estimate of α_1 of 3.5 with a t-statistic of 5.[12] Similar regressions with LMEMNUM and TRUST as the explanatory variables generated negative and significant estimates of α_1.

Finally, it is worth noting that the estimates of the coefficients on SCFRE overstate the impact of additional contact with family members and friends if there is social ability bias. In other words, people who are more socially able may devote more time to social interaction. To check whether the conclusions are affected by social ability bias requires finding an instrument which is correlated with LSCFRE but not correlated with social ability. One possible instrument is the number of relatives (RELEX) such as parents, grandparents, aunts and uncles and siblings, but excluding children, grandchildren, nieces and nephews. These relatives are certainly exogenous, and the correlation coefficient between the natural log of RELEX (LRELEX) and LSCFRE is .15 which is significant at the 1% level. Running an IV version of (7) with LRELEX as the instrument produced an estimate of α_1 of 2.9 with a t-statistic of 3. Thus, while social ability bias does reduced the size and significance of the α_1, it does not affect the basic contention that SCFRE should be a good proxy for social capital.

5. Social Capital and the Social Returns to Education

The main result from the previous section is that a measure of the frequency of contact with family members and friends does a much better job of explaining variation in access to social resources than measures of trust and group membership. This section addresses the question of whether having a better measure of social capital matters. It might not matter for two reasons. Social capital may have no significant economic effects. Alternatively, bad proxies of social capital might work as well as good proxies in identifying these effects. This section address these issues by examining an extension of the model which allows for investment in

[12] . Running multinomial logistic regressions produced similar results.

human capital and complementarity between human and social capital. This model suggests a simple test of whether social capital has a significant economic effect.

The Theory

Assume first that wages depend on both own human capital and on the average level of human capital. Define H as the average level of human capital in the economy in which household i resides, and let:

$$w_i(t) = w(h_i(t), H(t)(S_i(t)) \quad (8)$$

Assume also that $(\partial w_i/\partial H) > 0$ and that $(\partial w_i/\partial H)(\partial H/\partial S_i) > 0$. The average level of human capital has a positive effect on wages for any level of own human capital, and this effect is an increasing function of one's social capital.

There are at least two stories told about why individual wages might depend on the average level of human capital in an economy. The first, most closely associated with Lucas (1988), is that people learn through their interactions with others. If you are interacting with people who on average have higher levels of human capital, then you should be more productive for any level of your own human capital. The other story, associated with Acemoglu (1996), is that firms are uncertain about the human capital levels of the workers with whom they will be matched, so they make physical capital investment decisions based on expected or average human capital levels. The assumption that the impact of H on w_i depends on S_i is motivated by the former story. In other words, if human capital spillovers do occur through interactions, then it ought to be true that those who interact more are better able to capture the spillover benefits.

The exact mechanisms through which this learning takes place have rarely been modeled.[13] Nonetheless, one might imagine many different ways in which a worker can improve

[13]. For an exception, see Jovanovic and Rob (1989).

his or her productivity and wages by learning from interactions with others. For example, workers can learn by interacting with others on the job. They might also learn about innovations occurring at other firms from their social interactions with others with whom they do not work. Finally, they might simply learn from others about different job opportunities which, if it leads to more efficient matching, could increase productivity.

Next, assume that the time path of human capital is given by:

$$\dot{h}_i(t) = h_i(t) f(1 - \sigma_i(t)) - \delta_h h_i(t) \tag{9}$$

where δ_h is the constant rate of depreciation of human capital. Since $(1 - \sigma)$ is the share of time devoted to working, (9) essentially implies that human capital grows as a result of learning by doing. In other words, there is no distinction between working and on the job training.

Assuming that all households begin with h_0 units of human capital in period zero and that they live and work T years,[14] households choose a time path of σ to maximize (4) subject to (3) and (9) and the initial levels of S and h. Again, using standard techniques of optimal control yields the following first order condition for the optimal σ:

$$F'(Z(t))Z_2(t)w + \lambda_h h f'(1 - \sigma(t)) = \lambda_s(t) n(t) g'(\sigma(t)) \tag{10}$$

where λ_h denotes the shadow price of human capital in utility terms. Comparing (10) with (6) indicates that the marginal cost of an increase in σ includes both the cost of foregone consumption and the cost of foregone wage growth. In other words, taking time out of the labor market to interact socially reduces both current consumption and wages at each point in the future.

[14] . More realistically, one would want to allow for differences in years lived and years worked, but the simpler approach is adopted as a first pass.

One immediate implication of (10) is that the life cycle path of the optimal n is ambiguous. To see this note that if $g' = 0$, then the model is nearly identical to the one described in Section 2, so a decrease in λ_h along the optimal path must be accompanied by an increase in σ to satisfy (10). However, if $g' > 0$ and both shadow prices are falling as t approaches T, σ can either rise or fall along the transitional path.[15] If the model is taken literally, one explanation for observed paths of hours worked and wages over the life cycle is that λ_h falls more rapidly than λ_s. If so, then the model could predict the kind of u-shaped relationship between S and t suggested by the evidence in Table 5.

Another implication is that the model does not generate an unambiguous sign prediction for the relationship between social capital and wages when controlling for education and experience. To see this note that for any level of h, those with more social capital should have higher wages because of the indirect of social capital on wages through the average h. However, for any education and experience levels and for the same n_0, the direct effect of social capital on wages is negative because those with more social capital must have lower h. This is true because the only way to have a higher number of relationships is to have devoted less time to working in the past. However, if one controls for the interaction between social capital and the average level of human capital, the model predicts that the direct effect of social capital on wages should be negative.

The Data and Econometric Specification

Previous research which has examined the social returns to education has focused on estimating versions of the following human capital earnings function:

$$\ln w_{ijt} = \alpha + \alpha_j + \alpha_t + \beta_1 X_{ijt} + \beta_2 E_{ijt} + \beta_3 \overline{H}_{jt} + u_{jt} + \varepsilon_i$$

[15] . Complementarity between individual and average h in (8) could also make the optimal time path of σ ambiguous.

The dependent variable is the log of the wage of individual i living in geographic area j at time t, and α_j and α_t denote city and time specific effects. Years of schooling is given by E_{ijt}, while X_{ijt} represents a vector of individual characteristics such as race, sex and linear and quadratic potential experience terms. The average level of human capital in geographic area j in year t is given by $\overline{H}_{jt}$, while u_{jt} and ε_i denote the area/time and individual error components respectively. Positive and significant estimates of β_3 signify social returns to education in excess of the private returns to education because increases in own human capital raise $\overline{H}_j$ which has a positive effect on wages of all workers in area j.

Previous research in this area differs along four dimensions. The first is the relevant geographic area. Rauch (1993), Moretti (1999) and others tested for social effects using data on US cities, Topel (1999) uses cross country data and Angrist and Acemoglu (1999) using state level data. The second dimension is the measure of $\overline{H}_j$. Angrist and Acemoglu use average years of schooling, Rauch uses average schooling and experience, Topel uses data on education and labor productivity and Moretti uses the share of population with a bachelor's degree. The papers also differ in the extent to which they try to separately identify β_2 and β_3. Rauch treats both E_{ij} and $\overline{H}_j$ as exogenous, Moretti instruments for $\overline{H}_j$ while Angrist and Acemoglu instrument for both. They also differ in terms of their results. Rauch, Moretti and Topel find evidence of social returns, while Angrist and Acemoglu do not.

Manski (2000) is highly critical of the peer effects literature in general and for these kind of studies in particular. He is especially critical of the practice of simply inferring patterns of interaction without information on the characteristics of the people with whom one is interacting or how frequently one is interacting with them.[16] The data described above do not allow one to address the characteristics issue, but it does allow one to address the issue of how much they are

[16] . One paper not subject to this criticism is Goolsbee and Klenow (1999). They ask whether the probability of computer adoption depends not just on the average level of computer adoption within one's city but also on the number of family members and friends who already own computers. They find that the self reported probability of adoption by non computer owners is strongly related to the number of family members and friends that already own one.

interacting. In the context of the model discussed above, this means that one can test whether the impact of $\overline{H}_j$ on w_{ij} depends on how much a person interacts with others within geographic area j.

This suggests estimating the following type of model.

$$\ln w_{ijt} = \alpha + \alpha_j + \alpha_t + \beta_1 X_{ijt} + \beta_2 E_{ijt} + \beta_3 \overline{H}_{jt} + \beta_4 S_{ijt} + \beta_5 \left(\overline{H}_{jt} * S_{ijt}\right) + u_{jt} + \varepsilon_i \qquad (11)$$

The model described in (1)-(9) predicts that β_4 should be negative and β_5 should be positive.

There are two ways that one could use the data described above to estimate a version of (11). The first is to interpret the individual social resource variables or, possibly, the sum of them as a direct measure of the stock of social capital and use them in the estimation. Alternatively, since the results of Tables 4 and 5 indicate a strong relationship between the frequency of contact with family and friends and S, one could use SCFRE as a proxy for S in the same way that years of education is used as a right hand side variable to proxy for human capital in a wide variety of contexts. One advantage of the later approach is that asking someone about the frequency of their contact with family members and friends is much easier than asking them about their access to social resources. Thus, if one can find positive and significant economic effects from social capital using SCFRE this may suggest that collecting data on individuals' social capital is less problematic.

The fact that the data for SCFRE comes from the 1986 GSS influences the particular econometric specification used in the analysis. First, because there is only a single year of data, one cannot control for time or area specific effects. Second, respondents in the 1986 GSS live in one of 84 Primary Sampling Units (PSUs). These are either counties, MSA'S or PMSA's, and it

is possible to identify the PSU for each respondent. Thus, the relevant geographic area for this study is the PSU.[17]

The choice of what variable to use to represent $\overline{H}_j$ is motivated in part by the availability of a suitable instrument. The need to instrument arises because there may be unobserved factors which are correlated with both wages and average human capital levels which bias OLS estimates.[18] Moreover, if average years of education is used to represent $\overline{H}_j$ then the simple fact that $\overline{H}_j$ is the average of E_{ij} implies that OLS estimates are biased upward. Finally, since the 1986 GSS has wage data from 1985 which is not a census year, any variable used to represent $\overline{H}_j$ is likely to suffer from significant measurement error.

Moretti suggests that changes in educational levels across cities depend in part on the fact that there are long run trends resulting from younger more educated cohorts entering the labor force. These changes also depend on the difference between national trends and local educational attainment and on changes caused by migration. He argues that the first of these effects is exogenous, and constructs an instrument which he contends is correlated with this first effect and uncorrelated with the others.

The instrument used here is a version of the one used by Moretti. This is calculated by taking the share of people in 1980 within a PSU in a particular age category, and then multiplying this share by the change in the mean educational level between 1980 and 1985 for that age group in the country as a whole. These products are then summed up over all age categories. In other words, for some exogenous reasons cities had different age structures in the past. As such, differences in average human capital levels across cities and counties in the present should depend on these differences because of variation in the increase in educational

[17] . One could also identify the state of residence for each respondent, so state could also have been the relevant geographic area. In addition to the conceptual advantages of using a smaller geographic, the GSS is not a representative sample at the level of the state.

[18] . Moretti argues that the direction of the bias is ambiguous. If there are unobserved demand factors such that an increase in the demand for skilled workers increases wages and average human capital levels, then OLS estimates are bias upward. If there are unobserved supply factors such that an increase in the supply of skilled workers increases amenities and makes a city a more desirable place to live, then OLS estimates are biased downward.

levels across age groups. Detailed information on how this instrument, AGEDIS, is calculated for all of the PSU's is contained in the data appendix.

The variable used to represent $\overline{H}_j$ is also the same one used by Moretti, the share of the population in the PSU with at least a bachelor's degree (BAPCT). The correlation coefficient between AGEDIS and BAPCT is .46 which is significant at the 1% level. One advantage of using this variable is that it will allow better comparison of the results because a version of Moretti's instrument is being used. Furthermore, this proxy for $\overline{H}_j$ is not the average of the E_{ij}. On a conceptual level, one may object that the use of BAPCT implies that spillovers only occur through interaction with those with college degrees, but at the same time, one may argue that the kind of spillovers occuring through social interaction are more important in industries employing more skilled workers. Nonetheless, Moretti reports that using average years of education rather than the share with BA's had no effect on his results.

As mentioned above, another difficulty raised by the use of the 1986 GSS is that there is no census in 1985. As such, BAPCT is calculated by linearly extrapolating data from the both the 1980 and 1990 census. This introduces an additional element of measurement error. To account for this, another instrument used is the share that comes from the 1985 GSS (BAPCTGSS).[19] The GSS is a random sample of each PSU, but the number of cases within most PSUs are quite small particularly for some of the rural counties. This suggests that BAPCTGSS is itself measured with a considerable degree of error. However, the measurement error in the GSS measure ought not be correlated with the measurement error in the extrapolated census measure, so it should be a useful instrument. The correlation coefficient between BAPCTGSS and BAPCT is .42 which is significant at the 1% level.

OLS and IV Estimation

Table 6 reports the results from estimating the following base equation:

[19] . The PSU's are the same for both the 1985 and 1986 GSS.

$$\ln \mathrm{WAGE}_{ij} = \alpha + \beta_1 \mathrm{X}_{ij} + \beta_2 \mathrm{EDUC}_{ij} + \beta_3 \mathrm{BAPCT}_{ij} + \beta_4 \mathrm{SCFRE} + \beta_5 (\mathrm{BAPCT} * \mathrm{SCFRE}) + \eta_{ij} \quad (12)$$

The dependent variable is the log of annual wages,[20] and EDUC denotes years of schooling. The elements in X include linear and quadratic potential experience terms (EXPER), a female dummy (FEMALE) and a black dummy (BLACK). Preliminary analysis indicated a negative and significant correlation between the residuals and PSU size (POP), so the all the regressions are weighted by the square root POP.

Column (1) contains the results of an OLS estimation of (12) without SCFRE or its interaction. The private returns to education are about 10%. The estimate of β_3 is positive (1.53) and significant (se = .612). This is similar to the estimates reported by Moretti. In column (2), the main SCFRE and the interaction variables are added. The estimate of β_3 increases to 2.39 and is still significant. The estimates of β_4 and β_5 are positive and negative respectively but neither is statistically significant.

Column (3) contains the results of the IV estimation in which AGEDIS, BAPCTGSS and interactions between these variables and SCFRE are used as instruments. The estimate of β_3 increases further to 2.9 though so does its standard error. As a result it is only significant at the 10% level. The estimate of coefficient on the interaction term is now positive with a t statistic of 1.5. Though neither the main effect on BAPCT nor the interaction effects are individually significant at the 5% level, the estimates of β_3 and β_5 are jointly significant.[21]

These results, combined with those in Tables 3 and 4, suggest that the human capital spillover effects are indeed larger for those with more social capital. The magnitude of the estimates indicates that for those that do not interact at all with family members and friends a one percentage point increase in the share of the population with a BA raises wages by 2.9%.

---

[20] . The GSS asks respondents how much they earned last year, but not how many hours or how many weeks they worked last year.

[21] . Testing the joint significance of β_3 and β_5 yields an F = 10.6, and the 99% critical value is 4.6.

However, at the mean SCFRE a similar increase raises wages by over 4%.[22] Thus, comparing the results in columns (1)-(3) indicates that controlling for patterns of social interaction and instrumenting more than doubles the estimate of the social returns to education.

In column (3) the main SCFRE coefficient now has a negative estimate as predicted with a t statistic of 1.5. Again, though neither the main effects or the interaction effects are individually significant at the 5% level, the estimates of β_3 and β_5 are jointly significant. While the main effect is negative, the interaction effect is positive. This raises the question of whether an increase in SCFRE raises wages. The magnitude of the coefficients indicates that it does so as long as one lives in a PSU in which at least 21% of the residents have a bachelors degree. At the mean BA share, 19%, an increase in SCFRE has a negligible net effect on wages.

The next step is to test whether there is any evidence that social capital has significant economic effects when TRUST and MEMNUM are used to proxy for social capital. Columns (4) and (5) of Table 6 replace SCFRE and the interactions with these other proxies and the respective interactions. In both cases, the estimate of β_3 is positive and significant and larger in magnitude to that reported in column (3). The main effects for TRUST and MEMNUM have negative and positive estimates respectively but very high standard errors. The estimates of the interactive terms have the opposite signs and high standard errors. These results indicate that the choice of a particular proxy for social capital does matter. There is no evidence that the importance of the average level of human capital on own wages depends on one's social capital when TRUST and MEMNUM are used to proxy for social capital.

Sensitivity Analysis

The next task is to examine the extent to which the results in Table 6 are robust to alternative specifications. One possible explanation for the results is that they are due to spurious correlation related to differences in the cost of living. In other words, wages may be higher in

[22] . This estimate is higher than the estimate that Moretti reports, 2.2%, using a similar instrument.

places with a large share of college graduates in order to compensate them for higher costs of living. One way to examining this hypothesis is to estimate the following.

$$\ln \mathrm{WAGE}_{ij} = \alpha + \beta_1 \mathrm{X}_{ij} + \beta_2 \mathrm{EDUC}_{ij} + \beta_3 \mathrm{BAPCT}_{ij} + \beta_4 \mathrm{SCFRE} + \beta_5 (\mathrm{BAPCT}^{*}\mathrm{SCFRE}) + \beta_6 \mathrm{Z}_j + \eta_{ij} \quad (13)$$

where Z_j denotes the cost of living in PSU j. If the results are due to spurious correlation, the estimate of β_6 should be positive, and the inclusion of Z_j should reduce the size and significance of the estimate of β_3.

Column (1) of Table 7 contains the results of estimating (13) using differences in housing costs as a control variable. Most of the variation in cost of living across cities is due to differences in the cost of housing, and the particular variable used is the "fair market rent" (FMR) for a two bedroom apartment calculated by HUD. There are two surprising features of these results. First, the estimate of β_6 is negative and marginally significant (t=1.5). Second, the inclusion of FMR raises the size and significance of both β_3 and β_5. One possible explanation for these results is that housing costs are correlated with other amenities that make a place more desirable to live. If places with more college graduates have more amenities and places with more amenities have higher housing costs, then people may be willing to accept lower wages to live in places with higher housing costs.

In addition to differences in housing costs, there are many other observable features of cities or counties which are correlated with both wages and the BA share which may be driving the results in Table 6. To test this, column (2) of Table 7 contains a regression in which Z_j in (13) denotes a vector of control variables including the percent black (PCTBLACK), percent other race (PCTOTHER), percent unemployed (PCTUNEMP), percent native born (PCTNAT) and percent urban (PCTURBAN). The estimates on the control variables, while not reported, all have the expected signs, and the estimates of β_3 and β_5 are larger and more significant. Thus, the results in Table 6 are robust to the inclusion of additional control variables.

The next question is whether the results are robust to an alternative instrument. It is well know that educational levels tend to rise as the number of colleges or universities rises. However, the number of colleges may respond to the demand for higher education, so the number of colleges or universities would not necessarily be suitably exogenous to serve as a valid instrument. Moretti argues that the presence of a land grant college is, however, a suitable instrument since these colleges were established by government edicts in the 19th century. Of the 84 PSU's, 13 had a land grant college. The correlation coefficient between the land grant college dummy (LGCDUM) and BAPCT is .44 which is significant at the 1% level.

In column (3) of Table 7, LGCDUM, BAPCTGSS and the respective interaction terms are used as instrument. In this case, the estimate of β_3 is 2.8 which is much smaller than in column (2) and more in line with that reported in column (3) of Table 6. The estimates of the main SCFRE term and the interaction terms are also smaller and less significant than in column 2. Nonetheless the main and interaction effects for BAPCT are jointly significant.[23]

An alternative explanation for the positive relationship between the BA share and wages is that it is due to diminishing returns to individual factors of production and complementarily between factors of production. In other words, an increase in the share of college educated workers may decrease the marginal product of high skilled workers and raise the marginal product of low skilled workers. If the latter effect is sufficiently strong, then on average an increase in the share of high skilled workers would raise wages independent of any spillover effects associated with human capital.

If this were true, then estimating (13) with only college educated workers should yield a negative estimate of β_3 and an insignificant estimate for β_5. In other words, an increase in the share of college educated workers should reduce wages of college educated workers, and social capital should have no impact on this effect. To test for this, column (4) of Table 7 reports results from the same regression reported in column (2) but one that includes only those with

[23] . Testing the joint significance of β_3 and β_5 yields an F = 14.4, and the 99% critical value is 4.6.

bachelor's degrees or higher in the sample. The estimate of β_3 is still positive though much smaller than in column (2), and it has a t-statistic of 1.5. The estimate of β_5 is also still positive, but it less precisely estimated. One thing to bear in mind is that there are less than 200 observations in the regression in column (4) compared to nearly 900 in the other regressions.

Before concluding, it is worth noting that in all of the regressions reported in Table 6 years of schooling is treated as an exogenous variable. The main reason for doing so is the lack of suitably valid instruments in part due to the use of the GSS.[24] Angrist and Acemoglu argue that failing to instrument for years of schooling can in theory bias the estimate of both β_2 and β_3. However, they find that instrumenting for years of schooling has almost no effect on their estimates of the social returns to education. Moreover, Moretti finds that estimates of the social returns to education in a cross sectional regression are very similar to those found using panel data to control for individual specific fixed effects such as differences in ability. Thus, even if suitably valid instruments were available, the evidence suggests that the results in Table 3 would likely be robust to the use of these instruments.

6. Concluding Comments

Convincing the skeptics that social capital can be a useful analytic tool for economists will require, at the minimum, three developments. First, researchers must begin to be more specific and more uniform about exactly what the term represents. Second, they must develop better ways of measuring variation across individuals in their social capital. Third, they must show using these measures that social capital has significant economic effects.

This paper makes contributions along all three lines. First, the most appropriate definition of social capital should ultimately depend on what is measurable and what is testable. By analyzing and then testing a simple model of social capital formation, the paper suggests a

[24] . For example, using differences in institutional features of the school systems is very problematic because GSS does not have enough respondents in each city or state that were affected by the same institutional features.

reasonably precise definition. Social capital represents those features of relationships that provide individuals/households access to the social resources which raise utility/output for any level of consumption. Second, the paper shows that measures of trust and group memberships do not explain variation in access to social resources. It also shows that measures of the frequency of contact with family members and friends do explain variation in access to social resources. Third, it presents evidence that the importance of peer effects may be an increasing function of a person's social capital by showing that the impact of the average level of human capital within a city or county on own wages is an increasing function of the frequency of contact with family members and friends.

In addition to the three developments mentioned above, research into the economic effects of social capital is constrained by the availability of data on patterns of social interaction. Asking people about their patterns of social interaction will never be as simple as asking about their earnings, age and schooling. However, the results in the paper provide some insights into how survey researchers might ask questions about patterns of social interaction which will allow one to easily construct measures of social capital.

REFERENCES

Acemoglu, Darren. "A Microfoundation for Social Increasing Returns in Human Capital Accumulation." Quarterly Journal of Economics 111 (1996): 779-804.

Acemoglu, Darren and Joshua. "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws." NBER Working Paper #7621 (March, 2000).

Alesina, Alberto and Eliana La Ferrara. "Participation in Heterogeneous Communities." NBER Working Paper #7444 (December, 1999).

_____. "The Determinants of Trust." NBER Working Paper #7621 (March, 2000).

Angrist, Joshua D. and Alan B. Krueger. "Does Compulsory School Attendance Affect Schooling and Earning?" Quarterly Journal of Economics 106 (1991): 979-1014.

Becker, Gary S. “A Theory of Social Interactions.” Journal of Political Economy 82 (Dec, 1974): 1063-1093.

Bowles, Samuel. “’Social Captial’ and Community Governance”. Focus 20 (Fall, 1999): 6-10.

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

DiPasquale, Denise and Edward L. Glaeser. “Incentives and Social Capital: Are Homeowners Better Citizens?” Journal of Urban Economics. 45 (May, 1999): 354-384.

Durlauf, Steven. “The Case ‘Against’ Social Capital”. Focus 20 (Fall, 1999): 1-5.

Glaeser, Edward L. “The Formation of Social Capital.” Manuscript, Harvard University (2000).

Glaeser, Edward L., David Laibson, Jose A. Scheinkman and Christine L. Soutter. “What is Social Capital: The Determinants of Trust and Trustworthiness.” NBER Working Paper #7216 (July, 1999).

Glaeser, Edward L., David Laibson, and Bruce Sacerdote. “The Economic Approach to Social Capital.” NBER Working Paper #7728 (June, 2000).

Goolsbee, Austan and Peter J. Klenow. “Evidence on Learning and Network Externalities in the Diffusion of Home Computers.” NBER Working Paper #7329 (September, 1999).

Guiso, Luigi, Paola Spaienza and Luigi Zingales. “The Role of Social Capital in Financial Development.” NBER Working Paper #7563 (February, 2000).

Jovanovich, Boyan and Rafael Rob. “The Growth and Diffusion of Knowledge.” Review of Economic Studies. 56 (October, 1989): 569-582.

Knack, Stephen and Philip Keefer. “Does Social Capital Have an Economic Payoff.? A Cross Country Investigation.” Quarterly Journal of Economics. 107 (November, 1997): 1251-88.

La Porta, Rafael, Florencio Lopez-De-Silanes, Andrei Shleifer and Robert Vishny. “The Quality of Governments.” Journal of Law and Economics. (January, 1999): 222-278.

Lucas, R. E. "On the Mechanics of Economic Development." Journal of Monetary Economics 22 (1988): 3-42.

Manski, Charles F. “Economic Analysis of Social Interactions.” NBER Working Paper #7580 (March, 2000).

Moretti, Enrico. “Estimating the Social Return to Education: Evidence from Repeated Cross-Sectional and Longitudinal Data.” Center for Labor Economics, University of California Working Paper #22 (November, 1999).

Portes, Alejandro. "Social Capital: Its Origins and Applications in Modern Sociology." American Review of Sociology. 24 (1998): 1-24.

Putnam, Robert D. Making Democracy Work: Civic Traditions in Modern Italy. Princeton: Princeton University Press (1993).

Rauch, James E.. "Productivity Gains from Geographic Concentration of Human Capital: Evidence From the Cities." Journal of Urban Economics 34 (1993): 380-400.

Schiff, Maurice. "Social Capital, Labor Mobility and Welfare." Rationality and Society 4 (1992): 157-75.

Solow, Robert. "Trust: The Social Virtues and the Creation of Prosperity." The New Republic 213 (1995): 36-40.

Topel, Robert. "Labor Markets and Economic Growth." In Orley Ashenfelter and David Card, eds. Handbook of Labor Economics: Volume III. Amsterdam: North Holland (1999).

Weiss, Yoram. "The Determination of Life Cycle Earnings: A Survey." In Orley Ashenfelter and Richard Layard, eds. Handbook of Labor Economics: Volume I. Amsterdam: North Holland (1986).

DATA APPENDIX

Unless otherwise noted all variables come from the 1986 GSS. For more detailed information see Davis, Smith and Marsden (1998).

TRUST: A recode of responses to GSS variable TRUST with question wording "Generally speaking, would you say that most people can be trusted or that you can't be too careful in life?" Response *most people can be trusted* = 1, *can't be too careful* = 0 and *depends* is coded as missing.

MEMNUM: GSS Variable MEMNUM. The total number of the following groups a respondent belongs to: fraternal, service, veterans, political, union, sports, youth, school service, hobby or garden, fraternity, nationality, farm, literary or arts, professional or academic, church affiliated or other.

LMEMNUM: Natural log of (1 + MEMNUM).

FRINUM: GSS variable FRINUM with question wording "How many close friends would you say you have?"

LFRINUM: Natural log of (1 + FRINUM).

RELNUM: The total number of all relative calculated from GSS variables MALIVE, PALIVE, SISNUM, BRONUM, DAUNUM, SONNUM, GRPARNUM, GRKIDNUM, UNANNUM, INLAWNUM, RELNUM.

LRELNUM: Natural log of (1 + RELNUM).

SC: Sum of RELNUM and FRINUM.

LSC: Natural log of (1 + SC).

FRIFRE: Sum of recodes to GSS variables FRIVISIT and FRICALL with question wordings "How often do you see or visit with the friend you feel closest to?" and "How often do you have contact with this friend besides visiting, either by telephone or letter?" For both variables responses *lives in same household* and *daily* = 365, *at least several times a week* = 208, *at least once a week* = 52, *at least once a mont*h = 12, *several times a yea*r = 6, *less often* = 3. If respondent has zero close friends both variables are coded as 0.

LFRIFRE: Natural log of (1 + FRIINV).

RELFRE: Sum of recodes to GSS variables MAVISIT, PACALL, PAVISIT, PACALL, SISVISIT, SISCALL, BROVISIT, BROCALL, DAUVISIT, DAUCALL, SONVISIT, SONCALL, RELVISIT AND RECALL. The question wordings are "How often do you see or visit with your mother (father, closest sister, closest brother, closest daughter, closest son, and closet other adult relative)?" and "How often do you have contact with your mother (father, closest sister, closest brother, closest daughter, closest son, and closet other adult relative) besides visiting, either by telephone or letter?" For both variables responses *lives in same household* and *daily* = 365, *at least several times a week* = 208, *at least once a week* = 52, *at least once a mont*h = 12, *several times a yea*r = 6, *less often* = 3. If respondent has no mother, father, sister, brother, daughter, son, or other adult relative all variables are coded as 0.

LRELFRE: Natural log of (1 + RELINV).

SCINV: Sum of RELNUM and FRINUM.

LSCFRE: Natural log of (1 + SC).

CHORE: Recode of GSS variables CHORES1 and CHORES2 with question wording "There are some household and garden jobs you cannot really do alone – for example, you may need someone to hold a ladder or move furniture. Who would you turn to first (second) for help?" If the respondent's first and second choice are either family members, friends, neighbors or people they work with the variable is coded 1. If the response to either question is either *no one*, *someone you pay*, other responses the variable is coded as 0.

SICK: Recode of GSS variables SICK1 and SICK2 with question wording "Suppose you had the flu and you had to stay in bed for a few days and needed help around the house with shopping and such. Who would you turn to first (second) for help?" If the respondent's first and second

choice are either family members, friends, neighbors or people they work with the variable is coded 1. If the response to either question is either *no one*, *someone you pay*, other responses the variable is coded as 0.

BORROW: Recode of GSS variables BORROW1 and BORROW2 with question wording "Suppose you needed to borrow a large sum of money. Who would you turn to first (second) for help?" If the respondent's first and second choice are either family members, friends, neighbors or people they work with the variable is coded 1. If the response to either question is either *no one*, *someone you pay*, other responses the variable is coded as 0.

UPSET: Recode of GSS variables UPSET1 and UPSET2 with question wording "Suppose you were very upset about a problem with your husband, wife or partner and hadn't been able to work it out with them. Who would you turn to first (second) for help?" If the respondent's first and second choice are either family members, friends, neighbors or people they work with the variable is coded 1. If the response to either question is either *no one*, *someone you pay*, other responses the variable is coded as 0.

DOWN: Recode of GSS variables DOWN1 and DOWN2 with question wording "Suppose you were down or depressed and you wanted to talk about it. Who would you turn to first (second) for help?" If the respondent's first and second choice are either family members, friends, neighbors or people they work with the variable is coded 1. If the response to either question is either *no one, someone you pay*, other responses the variable is coded as 0.

CHANGE: Recode of GSS variables UPSET1 and UPSET2 with question wording "Suppose you needed advise about an important change in your life—for example, about a job or moving to another part of the country. Who would you turn to first (second) for help?" If the respondent's first and second choice are either family members, friends, neighbors or people they work with the variable is coded 1. If the response to either question is either *no one*, *someone you pay,* other responses the variable is coded as 0.

BLACK: Recode of GSS variable RACE. Response *black* = 1, *white* = 0 and *other* = 0.

MARRIED: Recode of GSS variable MARITIAL. Response *married* = 1, responses *divorced*, *separated* and *never married* = 0.

HHINC: GSS variable REALINC in which responses to the question "In which of the following groups did your total family income, from all sources, fall last year before taxes?" are coded at the midpoint of the relevant range and topcoding is done assuming a Pareto distribution.

FEMALE: Recode of GSS variable SEX. Response *female* = 1, response *male* = 0.

CATHOLIC, OTHER, NONE: Dummy variables for whether the respondent is Catholic, other religion or no religion from GSS variable RELIG.

LWAGE: Natural log of the annual wage computed from GSS variable. REALRINC is itself a recode of the GSS variable INCOME86 with question wording "In which group did your

earnings from your occupation for the last year fall? That is before taxes or other deductions." in which responses are coded at the midpoint of the relevant range and topcoding is done assuming a Pareto distribution.

EDUC: Years of eduction from GSS variable EDUC.

EXPER: Potential experience calculated as AGE – EDUC – 6.

BAPCTGSS: Percentage of the population over 25 in the respondent's GSS Primary Sampling Unit (PSU) that has at least a bachelor's degree recoded from the GSS variable DEGREE.

AGEDIS: The percentage of the population in the respondent's PSU in 1990 belonging to each of the following age categories is calculated: < 21, 22-24, 25-29, 29-34, 53-44, 45-54, 54-59, 61-62, 63-64, 65-69, 70-74, 75-84, > 84. These percentages come from 1980 census data from the 1980 "CensusCD+Maps" by GeoLytics, Inc. Each of these percentages is multiplied by the change in the average number of years of education for that group from 1980-1985 using data from the 1980 and 1985 GSS. These products are then summed over all the different age categories.

FMR: The 45^{th} percentile of rents for a two bedroom apartment in each PSU calculated by the Department of Housing and Urban Development. The data is available at http://www.huduser.org/datasets/fmr.

The following variables are are linearly extrapolated using 1980 and 1990 census data from the 1990 and 1980 "CensusCD+Maps" by GeoLytics, Inc.. For the MSA's and PMSA's the variable is aggregated from the county level data. Only those entire counties in the MSA/PMSA in both 1980 and 1990 are included in the caclulation.

POP: Population in the respondents PSU.

BAPCT: Percentage of the population over 25 in the respondent's GSS Primary Sampling Unit (PSU) that has at least a bachelor's degree.

PCTBLACK: The percentage of the population in the PSU that is black.

PCTOTHER: The percentage of the population in the PSU that is neither black or white.

PCTNAT: The percentage of the population in the PSU that is native born.

PCTUNEMP: The percentage of the population in the PSU that is unemployed.

PCTURBAN: The percentage of the population in the PSU that lives in an urban area.

TABLE 1

Descriptive Statistics

Variable	*Mean*	*Std. Deviation*	*N*
TRUST	.39	.49	1427
MEMNUM	1.87	1.86	1467
FRINUM	7.44	10.98	1467
RELNUM	39.71	34.06	1419
SC	47.3051	36.7199	1419
FRIFRE	259.0000	240.9547	1457
RELFRE	830.5042	698.7267	1416
SCFRE	1093.5583	769.8427	1406
BORROW	.4146	.4928	1382
CHANGE	.9015	.2981	1310
CHORE	.9121	.2833	1399
DOWN	.8076	.3944	1403
SICK	.9377	.2418	1396
UPSET	.6016	.4897	1378
AGE	45.43	17.80	1463
FEMALE	.58	.49	1470
BLACK	.1252	.3310	1470
CATH	.2584	.4379	1467
OTHER	4.703E-02	.2118	1467
NONE	6.680E-02	.2498	1467
MARRIED	.5619	.4963	1470
LWAGE	1.9917	.8244	919
EXPER	19.2802	12.4896	1405
BAPCT	.1917	6.187E-02	1470
POP	1900388.3319	2489182.2212	1470
BAPCTGSS	.1934	.1359	1470
AGEDIS	.2685	4.850E-03	1470
LGCDUM	.16	.37	1470
FMR	422	94	1470
PCTBLACK	.12	.10	1470
PCTOTHER	.07	.09	1470
PCTNAT	.89	.23	1470
PCTURBAN	.76	.23	1470
PCTUNEMP	.06	1.9	1470
EDUC	13.28	3.02	1416

TABLE 2

Correlation Coefficients Between Social Capital Variables

	TRUST	*MEMNUM*	*FRINUM*	*RELNUM*	*SC*	*FRIFRE*	*FAMFRE*	*SCFRE*
TRUST	1.000	.201**	.025	-.012	-.002	-.094**	-.103**	-.120

MEMNUM	1.000	.084**	.036	.054*	.032	-.044	-.028
FRINUM		1.000	.055*	.356**	.109**	.010	.041
RELNUM			1.000	.949**	.017	.093**	.086**
SC				1.000	.049	.088**	.092
FRIFRE					1.000	.134	.435**
FAMFRE						1.000	.950**
SCFRE							1.000

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

TABLE 3

Correlation Coefficients Between Social Resource Variables

	BORROW	*CHORE*	*CHANGE*	*DOWN*	*SICK*	*UPSET*
BORROW	1.000	.118**	.123**	.147**	.147**	.273**
CHORE		1.000	.044	.158**	.419**	.120**
CHANGE			1.000	.368**	.099**	.282**
DOWN				1.000	.197**	.354**
SICK					1.000	.094**
UPSET						1.000

** Correlation is significant at the 0.01 level (2-tailed).

TABLE 4

Logistic Regressions with Social Capital Measures

	Dependent					
Independent	CHORE	SICK	BORROW	UPSET	DOWN	CHANGE
TRUST	.181 (.200) *.001*	.097 (.230) *.000*	-.127 (.114) *.001*	-.187 (.114) *.003*	.044 (.142) *.000*	-.286 (.192) *.004*
LMEMNUM	.157 (.148) *.002*	.370 (.176) *.009*	-.121 (.085) *.002*	-.390 (.087) *.020*	-.063 (.106) *.000*	-.358 (.146) *.010*
LFRINUM	.103 (.118) *.001*	.079 (.141) *.001*	-.187 (.065) *.008*	-.128 (.065) *.004*	-.046 (.081) *.000*	-.110 (.107) *.002*
LRELNUM	.435 (.127) *.019*	.742 (.148) *.050*	-.028 (.074) *.000*	.064 (.074) *.001*	.298 (.092) *.012*	.038 (.129) *.000*

LSC	.476 (.144) *.018*	.817 (.170) *.046*	-.110 (.083) *.002*	.017 (.083) *.000*	.312 (.104) *.011*	-.109 (.146) *.001*
LFRIFRE	.174 (.050) *.019*	.282 (.056) *.046*	.135 (.033) *.017*	.123 (.032) *.015*	.153 (.037) *.019*	.044 (.055) *.001*
LRELFRE	.400 (.066) *.058*	.472 (.075) *.075*	.239 (.047) *.028*	.178 (.044) *.017*	.148 (.052) *.009*	.127 (.073) *.005*
LSCFRE	.474 (.080) *.056*	.628 (.095) *.091*	.340 (.061) *.034*	.261 (.056) *.023*	.211 (.063) *.013*	.096 (.095) *.002*

Notes: Each cell contains the results of a single variable logistic regression. The first term in each cell is the estimate of the regression coefficient. The standard errors of the estimates are in the parentheses. The Naglekerke r squared is in italics. The n's vary because of missing values and range from a high of 1401 to a low of 1254.

TABLE 5

A. Logistic Regressions with Trust Variable and other Control Variables

	Dependent					
Independent	CHORE	SICK	BORROW	UPSET	DOWN	CHANGE
TRUST	.171 (.222)	.182 (.249)	.060 (.127)	-.062 (.127)	.031 (.158)	-.370 (.217)
AGE	-.014 (.035)	.045 (.037)	-.044 (.021)	-.125 (.022)	-.131 (.027)	-.166 (.040)
AGE^2	-.0003 (.0003)	-.00074 (.00035)	-.0001 (.0002)	.0009 (.0002)	.0091 (.0026)	.0013 (.0004)
N	1352	1349	1336	1334	1356	1268
R^2	.150	.080	.128	.128	.106	.082

B. Logistic Regressions with Group Memberships Variable and other Control Variables

	Dependent					
Independent	CHORE	SICK	BORROW	UPSET	DOWN	CHANGE
LMEMNUM	.226 (.171)	.610 (.199)	-.042 (.096)	-.335 (.098)	-.024 (.120)	-.329 (.166)
AGE^2	-.020 (.035)	.045 (.037)	-.052 (.021)	-.125 (.022)	-.131 (.027)	-.164 (-.039)
AGE^2	-.00025 (.00032)	-.00076 (.00035)	.00023 (.00021)	.0010 (.0002)	.0092 (.0003)	.0013 (.0004)

N	1389	1386	1372	1370	1393	1302
R^2	.153	.096	.123	.135	.108	.078

C. Logistic Regressions with Frequency of Contact Variable and other Control Variables

	Dependent					
Independent	CHORE	SICK	BORROW	UPSET	DOWN	CHANGE
LSCFRE	.517	.663	.277	.244	.205	.122
	(.095)	(.104)	(.065)	(.062)	(.071)	(.105)
AGE	.019	.051	-.053	-.126	-.132	-.162
	(.037)	(.040)	(.021)	(.022)	(.027)	(.040)
AGE^2	-.0002	-.0007	.00026	.0095	.00094	.0013
	(.0003)	(.0004)	(.00021)	(.0022)	(.00026)	(.0004)
N	1333	1330	1317	1316	1337	1247
R^2	.189	.158	.137	.140	.120	.081

Notes: The standard errors of the estimates are in the parentheses. Additional explanatory variables in all regressions include EDUC, FEMALE, BLACK, MARRIED, CATH, OTHER and NONE.

TABLE 6

The Impact of Social Capital on the Social Returns to Education

	Dependent				
	(1)	(2)	(3)	(4)	(5)
Independent	LWAGE	LWAGE	LWAGE	LWAGE	LWAGE
EDUC	.098	.095	.089	.093	.091
	(.010)	(.012)	(.012)	(.021)	(.013)
BAPCT	1.53	2.39	2.94	4.88	5.81
	(.612)	(1.05)	(1.83)	(1.75)	(1.95)
SCFRE		.00014	-.00044		
		(.00017)	(.00029)		
BAPCT*SCFRE		-.00090	.0021		
		(.00088)	(.0014)		
TRUST				-.117	
				(.638)	
TRUST*SCFRE				.096	
				(2.95)	
MEMNUM					.063
					(.167)
MEMNUM*SCFRE					-.328
					(.791)
Instruments	None	None	AGEDIS	AGEDIS	AGEDIS

			BAPCTGSS	BAPCTGSS	BAPCTGSS
N	870	870	870	848	870
R^2	.282	.283	.281	.300	.283

Notes: The standard errors of the estimates are in the parentheses. Additional explanatory variables in all regressions include FEMALE, BLACK, EXPER and EXPER^2. Observations are weighted by the square root of POP in all regressions. In columns (3)-(5), additional instruments include the interactions between both AGEDIS and BAPCTGSS and the respective social capital proxies.

TABLE 7

Sensitivity Tests

	Dependent			
	(1)	(2)	(3)	(4)
Independent	LWAGE	LWAGE	LWAGE	LWAGE
EDUC	.087	.092	.094	.125
	(.013)	(.013)	(.012)	(.040)
BAPCT	4.59	7.81	2.83	5.31
	(2.37)	(2.78)	(2.06)	(3.71)
SCFRE	-.00049	-.0009	-.00047	-.00006
	(.00031)	(.0003)	(.00035)	(.00059)
BAPCT*SCFRE	.0024	.0044	.0021	.0007
	(.0014)	(.00017)	(.0017)	(.0027)
FMR	-.0012			
	(.0007)			
Controls	No	Yes	Yes	Yes
Instruments	AGEDIS	AGEDIS	LGCDUM	AGEDIS
	BAPCTGSS	BAPCTGSS	BAPCTGSS	BAPCTGSS
N	870	870	870	191
R^2	.276	.265	.284	.293

Notes: The standard errors of the estimates are in the parentheses. Additional explanatory variables in all regressions include FEMALE, BLACK, EXPER and EXPER^2. The control variables are PCTUMP, PCTBLACK, PCTNAT, PCTOTHER and PCTURBAN. Observations are weighted by the square root of POP in all regressions. In columns (1)-(4), additional instruments include the interactions between instruments listed and SCFRE. All of the observations in column (4) have at least a BA degree.

FIGURE 1