Reconsidering Social Capital: A Latent Class Approach

Ann L. Owen

aowen@hamilton.edu

Julio Videras

jvideras@hamilton.edu

Hamilton College

June 2007

Abstract

Social capital has proven to be a useful concept, but has not been well-measured in the economics literature. We propose a different empirical method for measuring social capital, latent class analysis, based on the idea that social capital is an unobservable multidimensional construct. We explain and demonstrate the construction of latent classes that measure an individual's social capital using data from the General Social Survey. Our proposed method generates meaningfully different conclusions about the accumulation of social capital than those obtained by previous research. We present evidence consistent with the hypothesis that higher income influences social capital accumulation because of a higher opportunity cost of time. We also find evidence for complementarities in social capital accumulation within an individual's peer group. Finally, we show that community heterogeneity influences the likelihood that individuals hold beneficial social norms independently of their propensity to participate in voluntary organizations.

1 Introduction

In the last few years, the concept of social capital has proven to be useful to economists because it helps to explain how social norms and networks influence economic behavior and the outcomes of economic policies.[1] However, as many researchers have pointed out, the concept has been ill-defined and imperfectly measured. Motivated by the fact that social capital is an unobservable multidimensional construct, we propose in this paper a novel application of latent class models to the measurement of social capital.

There are varying definitions and uses of the term social capital in the economics literature. Sobel (2002) describes social capital as circumstances in which individuals can benefit from group membership. The World Bank, however, focuses on social capital as an aggregate, stating on its web site that it is the "norms and networks that enable collective action." Other researchers focus not on social connections and group membership but on trust among individuals. The seminal paper by Knack and Keefer (1997) showed an important role for trust at the country level in explaining economic growth, while, at the individual level, Karlan (2005) uses laboratory experiments to measure trust, arguing that these experiments measure social capital and are helpful in explaining individual behavior. However, Guinnane (2005) argues that trust is not a useful concept at all because it cannot be separated from the quality of institutions. Finally, Durlauf and Fafchamps (2004) discuss the many different definitions of social capital in empirical research and assert that research on social capital is plagued by "conceptual vagueness." Thus, the concept of social capital has been broadly defined and applied, running the risk of becoming a useless catch-all concept.

Rather than proposing another definition of social capital, we focus on two basic components of the term to motivate our methodology. First, social capital is a multidimensional concept: it embeds multiple manifestations of civic engagement as well as trust and fairness. A useful measurement of social capital must account for as many of these dimensions as possible. Second, unlike physical or human capital, social capital is not necessarily beneficial to society at large (Durlauf and Fafchamps, 2004) – terrorists networks is one example. More generally, individuals' values and socio-economic characteristics can attract them to different networks that have different effects on the economic system. An emphasis on the types of social capital is

[1] See for example, Easterly and Levine (1997), Golden and Katz (1999), Narayan and Pritchett (1999), or Guiso, Sapienza aand Zingales (2002) for a broad range of examples. Interested readers should also see Durlauf and Fafchamps (2004) who provide a survey and critical analysis of this literature.

important from the point of view of public policy since, as Durlauf and Fafchamps (2004) argue, "it is social structures, not their consequences, which can be influenced by policymakers." This emphasis on the typology of social capital also becomes important in empirical applications. Much of the discussion about Putnam's claim in *Bowling Alone* (2000) that civil engagement in the United States has declined centers on how this concept is measured. Durlauf and Fafchamps (2004) and Skocpol (2003) have argued that it is more appropriate to talk about changes in the nature rather than the quantity of civic engagement. Thus, we must apply a methodology that generates a typology of social capital that accounts for the different incentives that networks provide and the various and even divergent effects that those networks might have on economic outcomes.

We propose applying latent class models. This methodology approaches social capital as a multidimensional concept that embeds both group membership and social norms and, in doing so, allows for nuanced conclusions about the amount and type of social capital possessed by individuals, how it is accumulated, and its impact on economic behavior. While our paper focuses primarily on identifying social capital in individuals, we discuss in the conclusion how our approach could be applied to the measurement of social capital at the country level.

In two examples that compare our approach to recent findings, we show that measuring social capital as a multidimensional construct provides new meaningful insights about the determinants of social capital. First, we replicate the models in Glaeser, Laibson, and Sacerdote (2002) and find results that the authors' theoretical model predicts but they were unable to show empirically. Our second application elaborates on the research by Alesina and La Ferrara (2000) on the effects of heterogeneity on the likelihood that a person belongs to one voluntary organization. Our methodology incorporates social norms (fairness and trust) that Alesina and La Ferrara did not consider and we estimate the likelihood of choosing a set of organizations and social norms rather than assuming membership choices are independent of each other. Section 2 describes in more detail the differences in the empirical approach and underlying theoretical motivation between this paper and Glaeser, Laibson, and Sacerdote (2002) and Alesina and La Ferrara (2000).

Although many researchers agree social capital is a multidimensional concept, few have applied multivariate methods. Sabatini (2005) uses principal components analysis to reduce a set of indicators of social capital (including memberships) to a single variable. Paxton (1999)

applies confirmatory factor analysis to several indicators of trust. Our approach shares the same motivation and strengths of these analyses because we use multiple indicators of social capital. However, latent class models have some advantages over principal components analysis (PCA) and factor analysis (FA). First, the results of FA and PCA are generally not unique as researchers can rotate factor loadings to obtain a meaningful interpretation of the solution. Latent class analysis, on the other hand, is model-based and there exist formal criteria to decide on the dimensionality of the latent variable, that is, there are formal criteria to decide on the number of types of social capital that are present in the data. In addition, the interpretation of latent classes is generally less arbitrary than the interpretation of principal components. More importantly, PCA and FA assume that the observed indicators and the unobserved underlying factors are continuous and normally distributed. In some cases, however, it is more reasonable and in accordance with the theory to assume that social capital is a categorical variable and to characterize individuals according to the type of social-capital class they belong to.[2]

In what follows, we first discuss briefly the research by Glaeser, Laibson, and Sacerdote (2002) and Alesina and La Ferrara (2000). We explain our empirical approach to measuring social capital using latent class analysis and then discuss the results of our latent class models. In Section 5 we use the results of the latent class models in an estimation that mirrors that of Glaeser, Laibson, and Sacerdote (2002) to demonstrate that our proposed approach generates new insights about social capital formation. Section 5 also discusses the application of our methodology to research by Alesina and La Ferrara (2000) on the effects of community heterogeneity. Section 6 summarizes the findings and discusses further applications of latent class models to the economics of social capital.

2 Related Literature

This section discusses how our research expands two well-known investigations on the determinants of social capital at the individual level. In particular, we apply our methodology to the models discussed in Glaeser, Laibson, and Sacerdote (2002) and Alesina and La Ferrara (2000). Glaeser, Laibson, and Sacerdote (2002), henceforth GLS, develop a model in which

[2] A typology of social capital from FA or PCA can be obtained, however, the process is more arbitrary than that used by latent class analysis. The researcher would need to calculate factor scores of each individual for each factor loading. Then, the researcher would need to determine a cut-off point for the factor scores and cross-tabulate below and above the cut-off point in order to obtain clusters of respondents.

individuals invest time to accumulate social capital. Their theoretical approach is an adaptation of human capital investment models. Some of the predictions of the model are that social capital investment declines with the opportunity cost of time, age, and mobility, and increases with the occupational returns to social skills and the level of social capital in a community. Using the number of group memberships as a proxy for social capital, the authors find evidence for many of their predictions. However, there are two key predictions for which they are unable to find evidence. First, GLS find a positive association between income and group membership, contrary to the prediction that higher wages should be associated with less social capital because of the higher opportunity cost of accumulating it. Second, GLS do not find evidence that social capital correlates within peer groups.

Alesina and La Ferrara (2000) also examine the determinants of individual social capital. The focus of their paper is the effects of community heterogeneity. Noting that it is difficult to measure social capital, Alesina and La Ferrara examine membership in voluntary organizations only. The authors develop a model in which individuals prefer to interact with people from their same racial, ethnic, and social-class group and utility from participating in a voluntary organization depends on the composition of the organization and traveling costs. The key prediction of this model is that increased community heterogeneity reduces participation in mixed organizations. Empirically, Alesina and La Ferrara first estimate probit models of membership in any organization. Second, they estimate the likelihood of participating in specific organizations. The findings provide consistent evidence for the predictions of the model: racial diversity, in particular, strongly reduces the likelihood of participation and the effects are most marked for organizations that require frequent interactions.

While GLS measure social capital as the number of voluntary organization memberships and Alesina and La Ferrara estimate a binary choice (membership or no membership in one organization), we use latent class analysis to classify individuals into distinct types of social capital using both memberships and indicators of trust and fairness. We argue that the number of organizations an individual belongs to and whether or not a person belongs to any organization are not as important as the nature of the organizations. In formulating policy implications, it is important to investigate whether different types of social capital influence economic decisions differently. We should expect that the socio-demographic profile of individuals joining a union and a fraternal organization, for example, is different than the profile of individuals joining

literary and hobby clubs. Likewise, the economic impact of social networks formed around labor unions and fraternal organizations is likely different than the influence of networks of literary and hobby clubs. Thus, researchers need a categorical variable that distinguishes among types of social networks. Such typology can be constructed by applying latent class modeling to the usual indicators of social capital.

Latent class models allow us to use trust and fairness in addition to voluntary group membership to generate a typology of social capital. As GLS and Alesina and LA Ferrara we use data from the General Social Survey. We find peer-group effects that Glaeser, Laibson and Sacerdote predict in their theoretical model but cannot identify empirically. Importantly, we find results consistent with the hypothesis that individuals with high opportunity cost of time choose classes of social capital with relatively low time commitment and perhaps high monetary cost. Our results also indicate that socio-economic determinants sort individuals into types of social capital that vary in the nature of the social capital class, but not necessarily in the number of memberships forming the classes. We also show that community heterogeneity influences the likelihood that individuals hold beneficial social norms independently of their propensity to participate in voluntary organizations. This finding suggests that focusing on a single component of social capital can generate results that might not apply to the concept of social capital as a whole.

The economic framework we adopt differs from that adopted by Glaeser, Laibson, and Sacerdote (2002) because we view the individual's choice of social capital to be a decision about the *type* of social capital rather than the *amount* of social capital. It also differs from the model by Alesina and La Ferrara (2000) in that we do not assume the decision to participate in a given organization is independent from membership in other organizations. Rather the underlying theoretical framework of our empirical analysis considers individuals who select a type of social capital in order to maximize utility, given time and budget constraints. An individual choosing social capital type s joins a combination of groups and organizations and adopts norms of behavior that others of type s are also likely to join and adopt. Different types of social capital provide different resources that individuals can use to produce new goods for consumption or as consumption goods themselves. The individual's maximization problem can then be analyzed

using the framework of discrete choice models.[3] In our empirical analysis, it is assumed that the individual chooses the type of social capital that has the highest indirect utility among the set of all available types. We assume the respondent's utility can be decomposed into utility derived from observable characteristics and the impact of all unobservable factors. Thus, the probability that the researcher observes the individual choosing social class *k* over social class *j* is equal to the probability that the difference in observable utilities is greater than the difference in the unobservable components. The first step is to define the alternative social capital structures individuals can choose from. We use latent class models to derive a categorical construct that characterizes individuals as holding different (and exclusive) types of social capital. Then, we assume the researcher observes the choice with error so that the choice of social capital class can be analyzed in the context of a random utility model and estimated via a multinomial logit model.

Although we believe our approach has many advantages, naturally, it also has limitations. In particular, data availability restricts us to considering the number of memberships rather than the strength of engagement in groups when classifying individuals into different social capital groups. However, this criticism can be rightly levied at many empirical studies of social capital. To some extent, our treatment of social capital - which incorporates attitudes towards trust and fairness and uses membership in groups only as indicator variables of the latent class - addresses to some degree this concern, though we admit it does so imperfectly.

3 Using Latent Class Analysis to Measure Social Capital

In this section we motivate and describe latent class analysis. Then, in Section4, we apply this methodology to eighteen questions from the General Social Survey (GSS) that are common proxies for social capital: sixteen questions about membership in voluntary organizations and two questions regarding whether people can generally be trusted and whether other people are fair or try to take advantage of others.

Intuitively, one might think of latent class analysis as an alternative to using dummy variables that correspond to unique response patterns. For example, in the GSS data, we observe the answers to eighteen different questions with binary outcomes. Consequently, there are 2^{18} or

[3] The properties of discrete choice models are well-known. See Train (2003), McFadden (2000), or McFadden and Train (1978), among many other works on discrete choice methods. Our exposition here follows closely McFadden (2000).

262,144 possible unique response patterns, of which 3,027 are represented in our sample of approximately 14,500 respondents. If we were to consider each observed response pattern a unique type of social capital, then we would need to include 3,026 dummy variables in regression models explaining economic behavior. This exercise would provide regression models that are cumbersome to estimate and results that are very difficult to interpret. Instead, latent class analysis examines response patterns and groups individuals by these patterns. In our application, it reduces the possible response patterns down to seven distinct classes or types of social capital.

Latent class models are part of the family of generalized finite-mixture models. The methodology is grounded in a probability framework that allows model testing and the calculation of goodness-of-fit measures. Although latent class analysis has been applied to several social issues (see, for example, Patterson et al. 2002; and Biemer and Wiesen, 2002), it is still a fairly novel methodology in the economics literature. (See Boxall and Adamowicz, 2002; Greene and Hensher, 2003; Clark, Etile, Postel-Vinay, Senik and Van der Straeten, 2005; and Morey, Thacher, and Breffle, 2005). On theoretical and practical grounds, however, latent class analysis is a promising approach to the study of social capital.

In particular, let $i = 1,\ldots, I$, denote the respondents. For each individual we observe the responses to a set of eighteen questions denoted $k = 1,\ldots, 18$. $Y_{ik} = 1$ if the individual responds "yes" to question k and $Y_{ik} = 0$ otherwise. The response pattern of an individual is represented by the vector Y_i. We assume a finite number of social capital classes denoted $s = 1,\ldots, S$. The discrete latent variable X represents the social capital class:

$$P(Y_i) = \sum_{s=1}^{S} P(X_i = s) \times \prod_{k=1}^{18} P(Y_{ik} \mid X_i = s). \qquad (1)$$

The conditional probability that an individual in latent class s responds "yes" to indicator k is modeled as a logit equation:

$$P(Y_{ik} = 1 \mid X_i = s) = \frac{\exp(\beta_{ks})}{1 + \exp(\beta_{ks})}, \qquad (2)$$

where β_{ks} is a free parameter. It is possible to include observed variables to predict class membership. In our application, we allow the probability that an individual is assigned to a

particular social capital class to depend on the year the individual responded to the survey.[4] Letting $\mathbf{z}_i$ be a vector of year dummies, the conditional probability is:

$$P(Y_{ik} = 1 \mid \mathbf{z}_i, X_i = s) = \frac{\exp(\beta_{ks} + \sum_{t=1}^{T} \beta_{kt} z_{it})}{1 + \exp(\beta_{ks} + \sum_{t=1}^{T} \beta_{kt} z_{it})}. \qquad (3)$$

We discuss the structure of the latent classes by comparing conditional response probabilities for each indicator given latent class membership. In addition, we calculate posterior probabilities using Bayes theorem:

$$P(X_i = s \mid Y_i) = \frac{P(X_i = s) \prod_k P(Y_{ik} \mid X_i = s)}{\sum_{s'} P(X_i = s') \prod_k P(Y_{ik} \mid X_i = s')}. \qquad (4)$$

After calculating posterior probabilities, we assign each individual to the latent class for which she has the highest posterior probability (empirical Bayes modal classification rule). The resulting characterization of each individual as belonging to one of *S* classes constitutes our typology of social capital.

Although equation (1) implies responses to the indicators are independent given latent class membership, it is possible to relax this assumption of local independence by including direct effects, that is, by combining any pair of dichotomous variables into one item and modeling the four potential response patterns (0,0), (1,0), (0,1), and (1,1), as one multinomial response (Skrondal and Rabe-Hesketh, 2004). For example, if we want to model a direct dependency between item 2 and item 3, we modify equation (2) as follows:

$$P(Y_{i2} = 1, Y_{i3} = 1 \mid X_i = s) = \frac{\exp(\beta_{2s} + \beta_{3s} + \beta_{23s})}{1 + \exp(\beta_{2s} + \beta_{3s} + \beta_{23s})}. \qquad (5)$$

To identify the need for direct effects we use bivariate residuals: Pearson chi-squared statistics divided by the degrees of freedom.[5]

The parameters of the models are estimated using maximum likelihood with the likelihood function that is derived from the unconditional probability in equation (2).[6] To evaluate the goodness-of-fit of the models, we use the Pearson statistic that compares the

[4] The samples were taken in 1975, 1978, 1980, 1983, 184, 1986, 1987 through 1991, 1993, and 1994.
[5] For each pair-wise combination of indicators, this statistic compares expected counts to actual counts in a two-way table. Large residuals indicate that the model cannot explained well the association between those two indicators.
[6] We use Latent GOLD to estimate the latent class models.

observed frequencies of the response patterns to the expected frequencies of the model.[7] Because of sparse tables, we apply bootstrapping to calculate the Pearson statistic.[8] We estimate each model for 500 replication samples using maximum likelihood estimates as starting values. We reject the model if the bootstrap p-value of the Pearson statistic is smaller than .05 (Eid, Langeheine, and Diener, 2003). To select among models that cannot be rejected using the Pearson statistic, we use the minimum Bayesian information criterion (BIC) rule to select a model. The criterion is based on the log-likelihood of the model (LL) and accounts for the number of observations N and parameters to be estimated: BIC = -2LL + (logN)Npar.[9] A common problem estimating latent class models is the presence of local maxima. To address this problem, we estimated each model 10 times with 10 different starting values. For the model with the lowest BIC, the seven-class model, we found two maxima and the solution with the largest log-likelihood appeared 7 times.[10] We use this solution to obtain the latent classes.

4 Results of Latent Class Models

We apply latent class analysis to the responses to eighteen indicators by 14,527 individuals from the General Social Survey. Table 1 summarizes these variables. We choose our sample to be identical to that used by Glaeser, Laibson, and Sacerdote (2002).[11] Table 2 presents fit statistics from the latent class models (from one to nine groupings). We present each model's log-likelihood, the associated BIC, number of parameters, and bootstrapped p-value of the Pearson statistic. Using the bootstrap p-value, we can reject models with fewer than 6 classes. The minimum BIC rule indicates that the best model uses 7 classes. Once we selected a model with 7 classes, we relaxed the assumption of local independence. We modeled the responses to the indicators with the largest bivariate residual as a joint response and estimated

---

[7] The Pearson statistic is usually defined as $\sum_j (n_j - e_j)^2 / e_j$ where n is the observed frequency of pattern j and e is the expected frequency. The statistic is asymptotically distributed as a χ^2 distribution with degrees of freedom equal to the number of response patterns minus the number of estimated parameters minus the number of categories.

[8] Sparse tables occur when small and zero observed frequencies are common. With sparse tables, fit statistics such as the Pearson statistic or the likelihood-ratio test cannot be guaranteed to follow the assumed χ^2 distribution.

[9] We also computed the Akaike information criterion (AIC = -2LL+2*Npar) and the consistent Akaike information criterion (CAIC = -2LL+[log(N)+1]*Npar). The AIC favors models with larger number of classes, a tendency that increases with the sample size. The CAIC favors models with fewer classes but this under-fitting declines with sample size. In our application, both the BIC and the CAIC indicate a model with seven latent classes is the best among the models that fit the data.

[10] The second maximum for the seven-class model produces very similar conditional probabilities and class sizes.

[11] We thank Bruce Sacerdote for graciously supplying the data.

the model. We repeated this process until the last direct effect was not statistically significant at the 5 percent level.[12]

Table 3A presents class sizes (as a proportion of the total sample). To construct the class sizes, we calculate the probability of membership in each class for each individual and assign them to the class for which they have the highest probability. Table 3 also presents the conditional response probabilities for each indicator given latent class membership. We interpret the classes by observing the differences in the patterns of these response probabilities. We discuss below the characteristics of each class and provide a summary of these characteristics in Table 3B.

Class 1, the largest class (about 41 percent of the sample), corresponds to individuals with very low probabilities of membership in any type of voluntary organization as well as by low probabilities of FAIR and TRUST. Arguably, these individuals have low levels of social capital. Compared to individuals in latent Class 1, respondents in Class 2 are much more likely to state they trust other people and believe other people are fair than individuals in Class 1. Descriptive statistics by class show that individuals in Class 1 and 2 belong on average to .60 and .64 organizations, respectively. However, in Class 1 only 38 percent of individuals believe people are fair and none state people can be trusted while in Class 2, 80 percent of the individuals believe other people are fair and 99 percent claim others can be trusted.

Individuals in Class 3 have higher probabilities of FAIR and TRUST than individuals in any other class but Class 2. In addition, the probability that an individual assigned to Class 3 belongs to a professional organization is relatively high, 56 percent. Individuals in Class 4 also have high probabilities of FAIR and TRUST but are not as likely to belong to any type of organization except a church group (almost 76 percent probability). Among all individuals in the sample, individuals in Class 5 have the largest probabilities of belonging to unions, veteran, and fraternal groups (39, 34, and 39 percent respectively). Individuals in Class 6 have very low probabilities of FAIR and TRUST but, unlike Class 1, a relatively high probability of belonging to a church organization (around 52 percent) and between 20 percent and 30 percent probability

[12] We include a total of eight direct effects. Membership in church organizations and membership in labor unions as well as membership in church organizations and membership in sport clubs show a negative correlation. The other six local dependencies show a positive correlation between sport and labor unions, youth clubs and sport clubs, school groups and youth clubs, hobby and sport clubs, school fraternities and fraternal organizations, and literary and hobby clubs. These direct effects are statistically significant at the 1 percent level.

of belonging to a youth organization, school organization, sport group, and professional organization.

Class 7, the smallest class with 4 percent of the individuals in the sample, corresponds to individuals with large amounts of social capital, with high probabilities of group membership as well as high probability of trusting others. Individuals in Class 7 are the most likely to belong to all types of organizations except veterans groups, unions, fraternal organizations, and nationality groups. In addition, individuals in this class have fairly high probabilities of stating other people can be trusted and people are fair. The average number of memberships in this latent class is 7.04 (with a standard deviation of 1.75 groups). By most measures, Class 7 would be characterized by high levels of social capital.

Three conclusions are particularly worth noticing from the results above. First, FAIR and TRUST help to distinguish individuals across classes in a nontrivial manner. Although in the literature it is often implied that trust is a consequence of social capital, we find that there are individuals who are very likely to express trust but are unlikely to belong to any voluntary organization (individuals in Class 2). On the other hand, we find that some individuals who are likely to belong to church and other groups are not necessarily stating other people are fair and can be trusted (individuals in Class 6).

Second, individuals with the same number of memberships are sorted into different types of social capital. For example, of all individuals with three memberships who trust others and believe people are fair, 20 percent are classified in latent Class 2, 32 percent in Class 3, and 34 percent in Class 4. Similarly, individuals in Class 3 have the highest probability of belonging to a professional organization, individuals in Class 4 have relatively high probabilities of belonging to church and sports groups, while individuals in Class 5 have relatively high probabilities for belonging to veterans groups and unions. Thus, in spite of the fact that all the individuals in these classes belong to about three groups, their social networks are different. A key dimension along which membership in these groups differ is that they may require different levels of involvement of time and/or money. The importance of this last observation will become apparent when we discuss the determinants of social capital formation.

Finally, examining the average number of group memberships by class also reveals that simply adding number of memberships does not necessarily provide a meaningful measure of social capital. For example, it is difficult to claim that individuals in Class 6 have more social

capital than individuals in Class 3. The results of the latent class model suggest that it is more appropriate to say that individuals in Class 6 have a different type of social capital than individuals in Class 3.

Table 4 presents the estimates of the β parameters in equation 3. We use dummy coding for identification where the reference class is Class 1, the low social capital class. A positive estimate for class s means that an individual in class s is more likely to answers "yes" to the indicator than an individual in Class 1. In order to compare the magnitude of the effects we can calculate the exponential value of the estimate and interpret the result as the odds of answering "yes" in class s relative to the reference class. For example, the probability of membership in a church organization is approximately 3.4 (or $e^{1.22}$) times higher in Class 3 than in Class 1; while the probability of membership in a church organization is more than 15 times higher in Class 4 than in Class 1. Table 4 also presents Wald statistics and p-values for the null hypothesis that each estimate for a given indicator equals zero. In this model, we can strongly reject the null hypothesis for the eighteen indicators. Thus, all indicators discriminate between classes in a statistically significant manner.

Table 5 presents the effects of year dummies. The omitted year is 1994. The results indicate that the year dummies influence significantly latent class membership only for 1975, 1978, 1980, and 1984. Compared to respondents in 1994, individuals who took the survey in 1975 and 1978 are twice as likely to be in Class 4 as in Class 1. Compared to respondents in 1994, individuals who took the survey in 1980 and 1984 are twice as likely to belong to Class 2 as to Class 1. These results suggest that, relative to the most recent year of data in the sample, respondents in 1975, 1978, 1980, and 1984 are more likely to be in classes with high conditional probabilities for FAIR and TRUST than in the no social capital class. All other effects on the odds are not particularly large.[13]

5 Using Latent Classes to Investigate the Determinants of Social Capital

In the previous section, we estimated latent classes that measure an individual's social capital. In this section, we compare the empirical results of Glaeser, Laibson, and Sacerdote (2002) to results we obtain using the latent classes. We also apply our typology to the research

[13] See Paxton (1999) for a thorough analysis of Putnam's claim. Our results are consistent with Paxton's finding that trust in individuals, but not degree of association, has declined in the United States.

by Alesina and La Ferrara (2000) on the effects of community heterogeneity on individual social capital. We show that the use of latent classes allows meaningful new insights about the determinants of social capital.

Social Capital Formation

As we discussed in Section 2, GLS do not find evidence for two important results of their theoretical model. First, GLS find a positive association between income and group membership, contrary to the prediction that higher wages should be associated with less social capital. Second, GLS do not find evidence that social capital correlates within peer groups. We use the same independent variables used by GLS (descriptive statistics are presented in Table 6) and present the coefficients from multinomial logit estimation in Table 7. The specification in Table 7 closely resembles a base specification of GLS in which they investigate how demographic characteristics are associated with the number of group memberships.[14] Before discussing the results of the multinomial logit model, we note that Hausman tests strongly reject the null hypothesis that any pair of categories (that is, the latent classes) can be collapsed and that all categories are indistinguishable with respect to the independent variables in the model. These tests support the idea that the types of social capital we identified using latent class analysis are distinct and economically meaningful.[15]

Examined directly, the coefficients in Table 7 have limited interpretive value. In Table 8, we report the change in the odds ratio that results from a one standard deviation increase in the independent variable. For example, the first row of column 3 indicates that the odds ratio of being in Class 1 vs. 2, $\frac{P(class1)}{P(class2)}$, is .58 times larger when education increases by one standard deviation. In other words, higher education levels reduce the probability of being in Class 1 relative to Class 2. Similarly, the first row of column 5 indicates that $\frac{P(class1)}{P(class2)}$ is 1.41 times larger when the variable Black is increased by one standard deviation, suggesting that being black is associated with a higher probability of being in Class 1 relative to Class 2. As can be seen from these examples, cells in which the change in the odds ratio is less than one indicate a

[14] We report results of only one of the several GLS specifications. We receive qualitatively similar results for all other specifications reported in GLS.
[15] A Hausman test also confirms that the assumption of the independence of irrelevant alternatives is valid.

reduced probability of being in the class in column 1 relative to the class in column 2, while a change in the odds ratio that is greater than one indicates that the probability of being in the class in column 1 has increased relative to the probability of being in the class in column 2.[16] In Table 8, we report only changes that are significant at the 10 percent significance level. Insignificant changes are associated with blank cells in the table.

The results in Table 8 indicate how individual-specific characteristics sort people into different types of social capital. Some particularly interesting comparisons are between Class 1 and Class 2 or between Class 1 and Class 6. For example, the odds comparing Class 1 (low membership, low trust group) and Class 2 (low membership, high trust group) indicate that individuals with high levels of education and income are less likely to be in Class 1 than in Class 2, while being black or female or over the age of 49 are associated with higher probabilities of being in Class 1 relative to Class 2. Looking at the comparison between Class 1 and Class 6, we see that higher income and education are associated with higher probabilities of being in Class 6 (low trust, relatively high membership) relative to Class 1 (low trust, low membership), but being married and younger increase the probabilities of being in Class 1.

Comparing among classes 3, 4, and 5, we see that, unsurprisingly, women are much more likely to be in Class 3 than Class 5 (recall that Class 5 had much higher probability of belonging to a veterans group or a union). Education has also a strong substantive effect: individuals with higher levels of education are much more likely to be in Class 3 (the class with high conditional probability of membership in professional groups) than in classes 4, 5, or 6. This finding is particularly interesting because these classes are similar in terms of number of memberships. Thus, an overall point to make about the results in Table 8 is that there are socio-economic determinants that sort individuals into different types of social capital classes that have very similar average group memberships.

Our results can also speak to the commonalities between the classes. For example, individuals in both classes 1 and 6 have low values of TRUST and FAIR. The results in Table 8 suggest that one demographic characteristic that sorts individuals into these two "low-trust" classes is being black. Interestingly, however, this same demographic characteristic also sorts

[16] These changes are calculated directly from the coefficients in Table 7. For example, the Class 2 – Class 3 comparison in Table 8 for the variable education (column 3, row 7) is $e^{change\,in\,b*std\,(education)}$ where the change in b is the difference between the coefficients for education for classes 2 and 3 in Table 7. In other words, with a standard deviation for education of 3.15, the change in the odds ratio reported in Table 8 is $e^{(.1710-.6083)*3.15}=.25$.

people into class 7, the high social-capital class, vs. classes 2, 3, or 4. The fact that being black sorts into a "low-trust" class is broadly consistent with the argument developed by Akerlof and Kranton (2000) that individuals in isolated social structures adopt an "oppositional identity" as a way to reject the dominant culture when economic opportunities are lacking. In their model, a mixed equilibrium is possible in which some members of the minority group conform their economic behavior to that of the dominant group, depending on the degree of social exclusion (i.e., some African Americans choose the high social capital class in a mixed equilibrium).

One of the hypotheses GLS obtain from the theoretical model is a negative association between social capital and income: if individuals with high incomes have a high opportunity cost of time, then they should belong to few groups. However, GLS find that there is a positive association between social capital and income. Our typology of social capital sheds some light on this issue. We note that some voluntary groups have high time commitments but low monetary costs (church or sport groups) while other groups (professional associations) have relatively low time commitments and high monetary costs in the form of membership fees. If individuals do take into account the opportunity cost of their time when accumulating social capital, then we should see that higher-income individuals should be more likely to join social networks with lower time commitments but higher monetary costs. Our results support this conjecture. When we compare classes 3, 4, and 5, we see that the odds that individuals are in Class 3, the class with the highest probability of professional memberships, rather than in Class 4 or Class 5 increase with income. Furthermore, the probability of being in Class 4, the class described by high probabilities of being in a sports or church group, declines with income. These results are consistent with the fact that individuals consider the opportunity cost of their time in decisions regarding social capital accumulation, a result that is impossible to uncover simply using number of group memberships as a measure of social capital.[17] These results are also broadly consistent with the argument in Skocpol (2003) that elite Americans have shifted the burden of civic engagement toward managed professionally-staffed organizations.

We turn now to another prediction of the theory that GLS were unable to support - that the social capital of individuals should be positively correlated with the social capital of their

[17] Of course, like GLS, we can only comment on probabilities and cannot make a statement about causation. In fact, in our theoretical framework and in that of GLS, social capital also influences income so we cannot emphasize the interpretation of this result. We simply point out that our findings are at least consistent with a theory of social capital accumulation that takes into account the opportunity cost of time.

peer group because of interpersonal complementarities in social capital accumulation. In a simple discrete choice framework, this phenomenon would occur if peer group membership in a social capital class is an attribute of that class that yields utility or if selecting a social capital class that many peers also select reduces the cost of joining.

Following GLS, we define an individual's peer group by geographic location (primary sampling unit) and religion. For example, a peer group might be Baptists in Memphis or Catholics in Cleveland. We calculate the average probability of membership in each class for that individual's peer group, excluding the specific individual and dropping any observations for which there are less than five people in a peer group. We then use the average probabilities of membership for the individual's peer group as an explanatory variable. Now, the probabilities of peer group membership in each class vary by both individual and class so that for each individual there will be a separate probability of peer group membership associated with each class, in contrast to the other independent variables (e.g., age, education, income, etc.) that do not vary by class for each individual. Therefore, rather than estimating a multinomial logistic regression, we estimate a conditional logistic regression.

An additional concern is that omitted variables might create a spurious correlation between peer group values and individual values. We follow GLS and instrument for the probabilities of peer group membership with the average education, age, marital status and income of the peer group. (In this estimation the standard errors are bootstrapped.) Unlike GLS, however, even after employing instrumental variables estimation we still find significant effects of the peer group effects, indicating that there are interpersonal complementarities in social capital accumulation. Table 9 displays the results of the conditional logit model. In the conditional logit model we estimate six coefficients for the individual-specific variables (education, income, and so on) but only a single coefficient for the class-specific variable (average class-membership probability). The positive and statistically significant effect of the class-membership probability for peer group provides evidence that interpersonal complementarities in social capital accumulation do exist. We believe that our stronger empirical results are a direct consequence of our improved definition of social capital that considers not just the number of groups, but the types of social capital individuals hold.[18]

[18] The specification reported in Table 9 does not include state dummies because including them does not allow the maximum likelihood estimation to converge. GLS do include state dummies in the IV estimations they report.

Community Heterogeneity and Individual Social Capital

Alesina and La Ferrara (2000) also use GSS data to examine the determinants of individual social capital. The focus of their paper is the effects of community heterogeneity. Alesina and La Ferrara first estimate probit models of membership in any organization. Then, to explore whether heterogeneity affects differently the propensity to participate in organizations that have different requirements in terms of interaction among their members, they estimate the likelihood of participating in specific organizations. The authors find that racial diversity, in particular, strongly reduces the likelihood of participation and that the strongest effects are for organizations that require frequent interactions such as church groups and sports clubs.

Compared to Alesina and La Ferrara (2000), we do not assume the decision to participate in a given organization is independent from membership in other organizations. Because we implicitly assume that individuals belong to a set of memberships and norms, our empirical models investigate whether differences in the degree of racial and ethnic heterogeneity in the community where individuals live are related to the type of social capital individuals choose. We continue using the sample whose descriptive statistics are presented in Table 6 but we now add community-level variables following Alesina and La Ferrara: the logarithm of the size of the MSA/PMSA where the individual lives (per thousand people), the logarithm of median household income of the MSA/PMSA and the logarithm of median household income squared, and two indices of racial and ethnic heterogeneity of the MSA/PMSA where the individual lives. We also follow Alesina and La Ferrara to construct these indices. We use 1990 Census data of five ethnic groups: (1) white, (2) black, (3) American Indian, Eskimo, Aleutian, (4) Asian, Pacific Islander, and (5) other (that mostly corresponds to Hispanics). We group ancestries into ten groups (see Alesina and La Ferrara, 2000). The value of the indices can be interpreted as the probability that two individuals selected at random belong to different racial (or ethnic) groups.

We estimate two multinomial logit models: one specification includes the racial index (as well as the group shares that form the index), the second specification includes the ethnic index (and ancestry groups shares as well). Table 10 presents the change in odds ratio for one standard deviation increase in the racial and ethnic indices.[19] We only present changes that are significant

---

However the inclusion of state dummies is not driving the results: when we exclude state dummies from the IV specification reported in GLS, we do not find materially different results for the coefficient on peer memberships. Thus, the difference in results is not attributable to the exclusion of state dummies.

[19] Results of the multinomial logit models are available upon request.

at the 10 percent level or better (after clustering standard errors by MSA/PMSA). The results indicate that the more heterogeneous the community is in terms of race, the probability that the individual is assigned to the low social capital class is 1.13 higher than the probability that the individuals is assigned to a class characterized by high levels of trust and fairness (Class 2) and a class characterized by high levels of trust and fairness and memberships in professional and church organizations (Class 3). On the other hand, as racial heterogeneity increases by one standard deviation, the probability that the individual is assigned to Class 2 is .91 lower than the probability is assigned to Class 4, a class that is defined by medium/high levels of TRUST and FAIR and high probability of church membership. As ethnic heterogeneity increases by one standard deviation, the probability that the individual is assigned to Class 2 is .86 lower than the probability is assigned to Class 4 and .81 lower than the probability of being assigned to Class 5, characterized by medium/high levels of TRUST and FAIR and the largest probabilities of memberships in unions, veterans, and fraternal groups.

These results add new interesting insights to those found by Alesina and La Ferrara. First, community heterogeneity influences the likelihood of holding beneficial social norms independently of memberships in organizations. Individuals in Class 1 and Class 2 have very similar average memberships but more heterogeneity makes individuals less likely to trust others and believe other people try to be fair. This effect also occurs between Class 1 and Class 3, where individuals in Class 3 are likely to belong to organizations with both high and low interaction (church and professional groups). Second, our results show that more racial and ethnic heterogeneity is associated with higher probabilities of being assigned to Class 4 than Class 2 and to Class 5 than to Class 2. Thus, for equal high levels of trust and fairness, more diversity is associated with higher probabilities of membership in church groups and unions, veterans, and fraternal groups. Overall, these results suggest that community heterogeneity influences social capital accumulation not only because of aversion to interact with people of different backgrounds.

6 Conclusions

In this paper we motivate and demonstrate the use of latent class models to measure social capital. The empirical treatment of social capital that we propose links more closely the measurement of social capital with the concepts of social networks that underlie it. Latent class

analysis treats different kinds of social capital as being qualitatively different and is consistent with an interpretation of social capital as an unobservable multidimensional construct. We show that this new approach generates empirical results consistent with a theory of social capital accumulation that the standard treatment of social capital could not reveal. We replicate the models in Glaeser, Laibson, and Sacerdote (2002). While Glaeser, Laibson, and Sacerdote measure social capital as the number of voluntary organization memberships individuals hold, we use latent class analysis to classify individuals into distinct types of social capital using both memberships and indicators of trust and fairness. Using the same data, we find peer-group effects that could not be identified previously. We also find results consistent with the hypothesis that individuals with high opportunity cost of time choose networks with relatively low time commitment and perhaps high monetary cost. Finally, we find that socio-economic determinants sort individuals into types of social capital that vary in the nature of the social network but not necessarily in the number of memberships forming the network. We also apply this new measurement of social capital to the research by Alesina and La Ferrara (2000) on the effects of community heterogeneity on social capital. Our results suggest that community heterogeneity influences the likelihood of holding beneficial social norms independently of memberships in organizations and that heterogeneity might influence social capital accumulation not only because of aversion to interact with people of different backgrounds.

This paper has focused on the accumulation of social capital at the individual level. We note that the typology of social capital we obtain from latent class models can also be used as an independent variable by entering a series of dummy variables indicating class membership. In addition, these methods could be extended to create aggregate data by examining the proportions of individuals that make up each class to explain the importance of social capital at the country level. In doing so, it would be important to consider the distribution of social capital as well—a highly fractionalized society with many different classes suggests that any one individual cannot have a great deal of social capital. Because fewer classes suggest fewer different social networks, the potential to accumulate social capital in these societies may be greater. A second approach to examine social capital at the country level would be to perform multilevel latent class models that derive simultaneously both country and individual segments based on individual-level responses (Vermunt, 2003). Research generating macro-level measures of social capital is currently in progress.

Table 1: Indicators (N = 14,527)

Indicator	Description	Percent
Fair	= 1 if “people are fair” (= 0 if “people try to take advantage”)	59.72
Trust	= 1 if people can be trusted	39.87
Church	= 1 if membership in church organization	34.58
Service	= 1 if membership in service group	9.72
Veteran	= 1 if membership in veteran group	7.04
Union	= 1 if membership in labor union	13.20
Political	= 1 if membership in political club	3.97
Youth	= 1 if membership in youth group	9.48
School	= 1 if membership in school service	13.08
Farm	= 1 if membership in farm organization	3.71
Fraternal	= 1 if membership in fraternal group	9.40
Sport	= 1 if membership in sports club	19.50
Hobby	= 1 if membership in hobby club	9.31
Greek	= 1 if membership in school fraternity	4.80
Nationality	= 1 if membership in nationality group	3.30
Literary	= 1 if membership in literary or art group	8.78
Professional	= 1 if membership in professional society	14.70
Other	= 1 if membership in any other group	10.43

Table 2: Model Selection: Fit Statistics

	LL	BIC(LL)	Npar	p-value of Pearson statistic
1 Class	-94204.2209	188580.9495	18	<.001
2 Classes	-89483.0706	179435.7456	49	<.001
3 Classes	-88436.1251	177638.9513	80	<.001
4 Classes	-88044.7586	177153.3151	111	<.001
5 Classes	-87753.8485	176868.5916	142	.004 (.0028)
6 Classes	-87529.3428	176716.6768	173	.030 (.0076)
7 Classes	-87328.8866	176612.8611	204	.162 (.0165)
8 Classes	-87213.3294	176678.8434	235	.188 (.0175)
9 Classes	-87108.0368	176765.3549	266	.272 (.0199)
7 Classes with 8 local dependencies	-87038.8559	176087.4698	212	.200 (.018)

Table 3A: Conditional Response Probabilities

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Class Size (based on modal assignment)	.4136	.1945	.1128	.1062	.0620	.0704	.0405
Individuals	6,008	2,825	1,639	1,543	901	1,023	588
Fair	0.3235	0.8113	0.9020	0.8681	0.6906	0.0087	0.7974
Trust	0.0095	0.7696	0.7571	0.5209	0.5217	0.0142	0.6038
Church	0.1853	0.1564	0.4082	0.7583	0.3564	0.5248	0.8018
Service	0.0052	0.0001	0.1939	0.0977	0.2085	0.1741	0.6364
Veteran	0.0251	0.0288	0.0369	0.0394	0.3869	0.1082	0.1692
Union	0.1185	0.1021	0.0972	0.0943	0.3389	0.1596	0.1685
Political	0.0021	0.0035	0.0709	0.0289	0.0791	0.0762	0.3002
Youth	0.0135	0.0060	0.1066	0.2275	0.0746	0.2636	0.4751
School	0.0356	0.0342	0.2300	0.2671	0.0139	0.3008	0.5607
Farm	0.0112	0.0167	0.0184	0.0875	0.0680	0.0601	0.1485
Fraternal	0.0191	0.0324	0.1345	0.0596	0.3885	0.1186	0.3635
Sport	0.0807	0.0989	0.3309	0.2826	0.2086	0.3905	0.5398
Hobby	0.0209	0.0332	0.1613	0.1485	0.1214	0.2057	0.3334
Greek	0.0007	0.0121	0.1605	0.0065	0.0198	0.1039	0.3203
Nationality	0.0078	0.0057	0.0625	0.0260	0.0423	0.0786	0.1922
Literary	0.0058	0.0111	0.2319	0.1119	0.0159	0.1828	0.5515
Professional	0.0205	0.0584	0.5647	0.0260	0.0559	0.2467	0.6487
Other	0.0504	0.0869	0.1498	0.1566	0.1223	0.1476	0.2118
Average number of group memberships (standard deviation)	0.60 (0.70)	0.64 (0.68)	3.14 (1.26)	2.84 (1.04)	3.00 (1.11)	3.56 (1.42)	7.04 (1.75)

Table 3B: Qualitative Characteristics of the Classes

Class	Characteristics
Class 1	Low probabilities of FAIR and TRUST and low probabilities of group membership. By most measures of social capital, people in class 1 would be considered to have low social capital.
Class 2	High probabilities of FAIR and TRUST and low probabilities of group membership.
Class 3	High probabilities of FAIR and TRUST. Relatively high probabilities of memberships in professional organizations and church groups.
Class 4	High probabilities of FAIR and TRUST. Relatively high probabilities of memberships in church groups. Relatively low probabilities of membership in youth, school and sports groups.
Class 5	High probabilities of FAIR and TRUST. Largest probabilities of memberships in unions, veterans, and fraternal groups.
Class 6	Low probabilities of FAIR and TRUST. High probability of membership in a church group and relatively low probability of membership in school and sports groups.
Class 7	High probabilities of FAIR and TRUST. Largest membership probabilities for all groups except veterans, unions, and fraternal organizations. By most measures of social capital, people in class 7 would be considered to have high social capital.

Table 4: Parameter Estimates for Indicators*

	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Wald	p-value
Fair	2.1965	2.9576	2.6219	1.5410	-3.994	2.1081	254.2726	5.0e-52
Trust	5.8546	5.7858	4.7324	4.7355	0.4064	5.0702	82.3532	1.2e-15
Church	-0.202	1.2156	2.7252	0.9957	1.7332	3.1161	602.8167	5.8e-127
Service	-3.903	3.8205	3.0219	3.9115	3.6887	5.8053	322.1396	1.5e-66
Veteran	0.1424	0.3984	0.4670	3.2010	1.5526	2.0700	357.3839	4.0e-74
Union	-0.1845	-0.3104	-0.2360	1.3186	0.2588	0.3178	150.3091	6.7e-30
Political	0.5034	3.5919	2.6494	3.7109	3.6701	5.3187	263.3275	5.8e-54
Youth	-0.8502	1.5451	2.5347	1.6429	2.5788	3.1477	178.8749	5.9e-36
School	-0.0265	1.9633	2.0031	-1.0812	2.1313	3.0408	425.4804	9.3e-89
Farm	0.4059	0.5061	2.1384	1.8653	1.7338	2.7366	170.8837	2.9e-34
Fraternal	0.5184	1.8300	1.1668	3.4617	1.7662	2.9887	403.8265	4.2e-84
Sport	0.2293	1.6816	1.4558	0.9346	1.7724	2.2572	289.3706	1.5e-59
Hobby	0.4624	1.8972	1.9179	1.8012	2.2138	2.5536	197.3525	6.9e-40
Greek	2.9034	5.5120	2.2255	2.8814	5.0223	6.1608	186.8986	1.2e-37
Nationality	-0.3137	2.1391	1.2244	1.7289	2.3860	3.4114	274.4583	2.4e-56
Literary	0.6414	3.8382	2.9641	0.9235	3.4987	5.1465	406.1928	1.3e-84
Professional	1.0891	4.1293	0.2443	1.0427	2.7530	4.4826	708.7377	7.9e-150
Other	0.5837	1.2003	1.2529	0.9658	1.1826	1.6223	159.0193	9.6e-32

*Reference group is Class 1 (no social capital class). A positive (negative) estimate means that individuals in that class are more (less) likely to respond affirmatively to the corresponding indicator than individuals in Class 1.

Table 5: Parameter Estimates for Covariates, Year Dummies*

Covariates	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Wald	p-value
1975	0.3231	-0.4509	0.7082	0.0267	-0.4127	-0.1367	16.2919	0.012
1978	0.2381	-0.1135	0.7316	-0.1286	-0.4559	-0.4848	15.9180	0.014
1980	0.7914	-0.1712	0.1037	0.2666	-0.2070	-0.2902	23.0640	0.00078
1983	0.0070	-0.0537	0.5058	0.0237	0.1996	-0.2287	3.6904	0.72
1984	0.7297	0.1233	0.3283	0.1854	-0.2687	0.1486	16.9302	0.0095
1986	0.1016	-0.0090	0.8367	-0.3011	-0.0524	-0.1786	10.3173	0.11
1987	0.2463	-0.4007	-0.1234	-0.2985	-0.1816	-0.5049	11.2087	0.082
1988	0.0679	-0.0236	0.3935	0.0657	-0.2061	-0.4531	4.8706	0.56
1989	0.2774	0.2301	0.2347	-0.1056	-0.0535	-0.3999	5.1187	0.53
1990	0.0779	-0.0797	0.4282	0.1878	-0.0708	0.2843	2.3091	0.89
1991	0.2027	0.0480	0.3947	-0.5406	-0.1177	-0.5987	8.8637	0.18
1993	-0.0537	0.1173	0.4109	-0.4323	0.2286	-0.0496	4.4589	0.61

*Reference group is Class 1 (no social capital class); reference year is 1994.

Table 6: Descriptive Statistics For Independent Variables

Variable	Definition	Mean	Std. Dev.
Education	Years Education	12.39	3.15
Log Income	Log Annual Income	2.09	0.65
Income Missing	=1 if income not provided	0.04	0.20
Black	=1 if black	0.14	0.34
Female	=1 if female	0.56	0.50
Birth Year	Year of birth	1940.18	18.28
Married	=1 if married	0.57	0.50
Babies	=1 if have young children	0.26	0.59
Preteen	=1 if have preteens in household	0.32	0.69
Teens	=1 if have teenagers	0.24	0.59
South	=1 if in south	0.34	0.47
East	=1 if in east	0.20	0.40
West	=1 if in west	0.19	0.39
Log of Population	= log of population in PSU	3.43	2.17
Age 18-29	=1 if between 18 and 29	0.24	0.43
Age 30-39	=1 if between 30 and 39	0.23	0.42
Age 40-49	=1 if 40 to 49	0.16	0.37
Baptist	=1 if Baptist	0.21	0.41
Methodist	=1 if Methodist	0.10	0.30
Lutheran	=1 if Lutheran	0.07	0.25
Presbyterian	=1 if Presbyterian	0.05	0.21
Episcopalian	=1 if Episcopalian	0.02	0.15
Other Protestant	=1 if belong to other Protestant religion	0.14	0.35
Non Denominational Protestant	=1 if non-denominational Protestant	0.04	0.19
Jewish	=1 if Jewish	0.02	0.13
Catholic	=1 if Catholic	0.25	0.43
Other Religion	=1 if other religious affiliation	0.02	0.14

Total observations used in estimations: 13,926.

Table 7: Multinomial logistic regression for class membership

	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Education	.1710**	.6083**	.2306**	.1530**	.3123**	.6263**
	(.0104)	(.0149)	(.0135)	(.0155)	(.0160)	(.0205)
Log	.1947**	.6225**	.4069**	.7566**	.2609**	.6308**
Income	(.0586)	(.1037)	(.0853)	(.1256)	(.0871)	(.1591)
Income	.4505**	.5007	.7654**	1.361**	.6876**	1.381**
Missing	(.1675)	(.3470)	(.2388)	(.3489)	(.2544)	(.4447)
Black	-1.004**	-1.322**	-.8539**	-.3625**	.1549	-.0985
	(.0975)	(8.73)	(.1208)	(.1559)	(.1073)	(.1761)
Female	-.1048**	.0878	.3595**	-1.478**	-.1078	.1387
	(.0504)	(.0678)	(.0658)	(.0880)	(.0734)	(.0995)
Birth Year	-.0215**	.0028	-.0197**	-.0263**	.0182**	-.0244**
	(.0032)	(.0047)	(.0041)	(.0047)	(.0052)	(.0063)
Married	.0708	-.0767	.2527**	.0583	-.1969**	-.1114
	(.0560)	(.0764)	(.0726)	(.0932)	(.0824)	(.1104)
Age 18-29	.3189**	-.4764**	.0002	-1.209**	-.4526**	-.0232
	(.1438)	(.1971)	(.1820)	(.2411)	(.2190)	(.2766)
Age 30-39	.2891**	-.3350**	-.0220	-.5446**	-.4334**	-.2221
	(.1228)	(.1653)	(.1515)	(.1899)	(.1838)	(.2301)
Age 40-49	.3448**	.0349	.1386	-.1374	-.1703	-.1349
	(.1021)	(.1364)	(.1255)	(.1511)	(.1559)	(.1925)

Base category is Class 1. Estimated with 13,926 observations. Standard errors are in parentheses. **significant at the 5% level, *significant at the 10% level. Estimations also include, state dummies, dummies for religious affiliation, regional dummies, log of population, and dummy variables for babies, preteens, and teens in the household.

Table 8: Change in Odds Ratio for a one standard deviation increase in independent variable

Odds Comparing Column 1 – Column 2		(3) Education	(4) Log Income	(5) Black	(6) Female	(7) Married	(8) Age 18-29	(9) Age 30-39	(10) Age 40-49
(1) Column 1	(2) Column 2								
Class1	Class2	.58	.88	1.41	1.05		.87	.88	.88
Class1	Class3	.15	.67	1.58			1.22	1.15	
Class1	Class4	.48	.77	1.34	.84	.88			
Class1	Class5	.62	.61	1.13	2.08		1.68	1.26	
Class1	Class6	.37	.85			1.10	1.21	1.20	
Class1	Class7	.14	.66						
Class2	Class3	.25	.76	1.11	.90	1.08	1.40	1.30	1.12
Class2	Class4	.83	.87		.79	.91		1.14	
Class2	Class5		.70	.80	1.98		1.92	1.42	1.20
Class2	Class6	.64		.67		1.14	1.39	1.36	1.21
Class2	Class7	.24	.75	.73	.88			1.24	1.19
Class3	Class4	3.29	1.15	.85	.87	.85	.82		
Class3	Class5	4.20		.71	2.17		1.37		
Class3	Class6	2.54	1.26	.60	1.10				
Class3	Class7			.66					
Class4	Class5	1.28	.80	.84	2.49	1.10	1.68	1.24	
Class4	Class6	0.77		.71	1.26	1.25	1.21	1.19	1.12
Class4	Class7	.29		.77	1.16	1.20			
Class5	Class6	.61	1.38	.84	.51	1.13	.72		
Class5	Class7	.22			.45		.60		
Class6	Class7	.37	.79		.88				

Each cell gives the change in the odds ratio that results from a one standard deviation increase in the selected independent variable. The odds ratio is defined as the probability of belonging to the class in the Group 1 column divided by the probability of belonging to the class in the Group 2 column. Only changes that are significant at the 10 percent level are reported. Blank cells indicate the estimated change is not statistically significant.

Table 9: Conditional logistic regression for class membership

	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Membership probability of peer group	1.069** (.3869)	1.069** (.3869)	1.069** (.3869)	1.069** (.3869)	1.069** (.3869)	1.069** (.3869)
Education	.1722** (10.06)	.6160** (.0248)	.2399** (.0192)	.1525** (.0181)	.3068** (.0231)	.6149** (.0275)
Log Income	.1414 (.0924)	.5495** (.1574)	.3651** (.1147)	.5873** (.1539)	.2542** (.1124)	.5353** (.2463)
Income Missing	.4228 (.3211)	.3667 (.4308)	.6961** (.3151)	1.002** (.4343)	.7071* (.3946)	1.204** (.5950)
Black	-.9678** (.1274)	-1.294** (.1989)	-.8444** (.1878)	-.4801** (.2371)	.1680 (.1460)	-.1743 (.2293)
Female	-.1178* (.0703)	.1020 (1.03)	.3502** (.0853)	-1.411** (.1058)	-.1113 (.0858)	.0666 (.1301)
Birth Year	-.0199** (.0044)	.0033 (.0071)	-.0208** (.0058)	-.0258** (.0063)	.0156** (.0072)	-.0278** (.0078)
Married	.1061 (.0893)	-.0735 (.1436)	.2706** (.1141)	.1409 (.1343)	-.2246** (.1133)	-.0951 (.1233)
Age 18-29	.2924 (.2032)	-.4434* (.2696)	.0271 (.2772)	-1.144** (.2847)	-.3473 (.2827)	.1073 (.3160)
Age 30-39	.2710 (.1815)	-.3570 (.2466)	.0446 (.2009)	-.4862** (.2312)	-.3127 (.2542)	-.0962 (.2844)
Age 40-49	.3497** (.1513)	.0522 (.1818)	.1661 (.1942)	-.0004 (.1720)	-.1283 (0.58)	-.0437 (.2319)

Estimated via instrumental variables estimation. Bootstrapped standard errors are in parentheses. **significant at the 5% level, *significant at the 10% level. Estimations also include, dummies for religious affiliation, regional dummies, log of population, and dummy variables for babies, preteens, and teens in the household. Peer groups are defined as religion by primary sampling unit cell (e.g., Baptists in Memphis). Membership probability of peer group is the average probability of membership in a particular class for the individual's peer group, excluding the individual. Instruments for membership probability of peer group were average age, education, income and marital status for the peer group.

Table 10: Change in Odds Ratio for a one standard deviation increase in independent variable

Odds Comparing Column 1 – Column 2		(3) Race Index	(4) Ethnic Index
(1) Column 1	(2) Column 2		
Class1	Class2	1.13	
Class1	Class3	1.13	
Class2	Class4	.91	.86
Class2	Class5		.81

Each cell gives the change in the odds ratio that results from a one standard deviation increase in the selected independent variable. The odds ratio is defined as the probability of belonging to the class in the Group 1 column divided by the probability of belonging to the class in the Group 2 column. Only changes that are significant at the 10 percent level are reported. Blank cells indicate the estimated change is not statistically significant. In addition to the variables in Table 7, these specifications include size of the place the respondent live, median household income and household income squared (all in logs) and race group shares (column 3) and ancestry group shares (column 4).

References

Akerlof, George A., and Rachel E. Kranton, 2000, "Economics and Identity," *The Quarterly Journal of Economics* 115(3): 715-753.

Alesina, Alberto and Eliana La Ferrara, 2000, "Participation in Heterogeneous Communities," *The Quarterly Journal of Economics* 3: 847-904.

Biemer, Paul P., and Christopher Wiesen, 2002, "Measurement Error Evaluation of Self-reported Drug Use: A Latent Class Analysis of the US National Household Survey on Drug Abuse," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165(1): 97-119.

Boxall, Peter C., and Wiktor .L. Adamowicz, 2002, "Understanding Heterogeneous Preferences in Random Utility Models: The Use of Latent Class Analysis," *Environmental and Resource Economics*, 23(4): 421-446.

Clark, Andrew, Feabrice Etile, Fabien Postel-Vinay, Caludia Senik and karine Van der Straeten, 2005, "Heterogeneity in reported well-being: evidence from twelve European countries," *The Economic Journal* 115: c118-c132.

Dasgupta, Partha and Ismail, Serageldin, eds., 1999, *Social Capital: A Multifaceted Perspective*. Washington, D.C.: World Bank.

Durlauf, Steven N. and Marcel Fafchamps, 2004, "Social Capital," mimeo.

Easterly, William and Ross Levine, 1997, "Africa's Growth Tragedy: Politics and Ethnic Divisions," *Quarterly Journal of Economics*, 112: 1203-50.

Eid, Michael, Langeheine, Rolf, and Ed Diener, 2003: "Comparing Typological Structures Across Cultures by Multigroup Latent Class Analysis," *Journal of Cross-Cultural Psychology*, 34(2): 195-210.

Glaeser, Edward L., David Laibson, and Bruce Sacerdote, 2002, "An Economic Approach to Social Capital," *The Economic Journal* 112: F437-F458.

Golden, Claudia and Lawrence Katz, 1999, "Human Capital and Social Capital: the Rise of Secondary Schooling in America: 1910 to 1940," *Journal of Interdisciplinary History* 29: 683-723.

Greene, William H., and David A. Hensher, 2003, "A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit," *Transportation Research Part B* 37:681-698.

Guinnane, Timothy, 2005, "Trust: A Concept Too Many," Yale University Economic Growth Center Discussion Paper No. 907.

Guiso, Luigi, Paulo Sapienza and Luigi Zingales, 2004, "The Role of Social Capital in Financial Development," *American Economic Review* 94(3): 526-56.

Karlan, Dean S., 2005, "Using experimental economics to measure social capital and predict financial decisions," *American Economic Review* 95(5): 1688-1699.

Knack, Stephen, and Philip Keefer, 1997, "Does Social Capital Have an Economic Payoff?" *Quarterly Journal of Economics*, 112(4): 1251-88.

McFadden, Daniel, and Kenneth E. Train, 1978, "The Goods/Leisure Tradeoff and Disaggregate Work Trip Choice Models," *Transportation Research*, 12: 349-353.

McFadden, Daniel, 2000, "Disaggregate Behavioral Travel Demand RUM's Side. A 30-Year Retrospective," Department of Economics, University of California, Berkeley.

Morey, Edward, Thacher, Jennifer, and Bill Breffle, 2005, "Using Angler Characteristics and Attitudinal Data to Identify Environmental Preference Classes: A Latent-Class Model," *Environmental and Resource Economics* (in press).

Narayan, D. and L. Pritchett, 1999, "Cents and Socialbility: Household Income and Social Capital in Rural Tanzania," *Economic Development and Cultural Change* 47(4): 871-897.

Patterson, Blossom H., Dayton, Mitchell C., Barry I. Graubard, 2002, "Latent Class Analysis of Complex Sample Survey Data: Application to Dietary Data," *Journal of the American Statistical Association*, 97(459): 721-29

Paxton, Pamela, 1999, "Is Social Capital Declining in the United States? A Multiple Indicator Assessment," *American Journal of Sociology*, 105: 88-127.

Putnam, Robert, 2000, *Bowling Alone: The Collapse and Revival of American Community* (New York: Simon and Schuster).

Sabatini, Fabio, 2005, "Social Capital as Social Networks. A New Framework for Measurement," mimeo.

Skocpol, Theda, 2003, *Diminished Democracy: From Membership to Management in American Civic Life* (PLACE: University of Oklahoma Press)

Skrondal and Rabe-Hesketh, 2004, *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Chapman & Hall/CRC, Boca Raton, Florida.

Sobel, Joel, 2002, "Can we trust social capital?" *Journal of Economic Literature* 40(1): 139-154.

Train, Kenneth E., 2003, *Discrete Choice Methods with Simulation*, Cambridge University Press, New York, New York.

Vermunt, Jeroen K., "Multilevel latent class models," *Sociological Methodology*, 33: 213-239.