

**Office of Population Research Princeton University**  
**WORKING PAPER SERIES**

The Design of a Multilevel Survey of  
Children, Families, and Communities:  
The Los Angeles Family and Neighborhood Survey

Working Paper No. 2003-06

Narayan Sastry  
RAND Santa Monica

Bonnie Ghosh-Dastidar  
RAND Santa Monica  
OPR, Princeton University

John Adams  
RAND, Santa Monica

Anne Pebley  
University of California, Los Angeles  
RAND Santa Monica

The Los Angeles Family and Neighborhood Survey was funded by NICHD grant R01HD35944. Additional funding was provided by Office of the Assistant Secretary for Planning and Evaluation (OASPE) of the U.S. Department of Health and Human Services, Los Angeles County, the National Institute of Health Office of Behavioral and Social Science Research (OBSSR), and the Russell Sage Foundation. We thank our colleagues at UCLA and RAND for comments and suggestions. Direct correspondence to Narayan Sastry, RAND Corporation, 1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138; [sastry@rand.org](mailto:sastry@rand.org).

Direct correspondence to Narayan Sastry, RAND Corporation, 1700 Main Street, P.O. Box 2138, Santa Monica, CA90407-2138; email: [sastry@rand.org](mailto:sastry@rand.org)

Papers published in the OPR Working Paper Series reflect the views of individual authors. They may be cited in other publications, but are intended to be work-in-progress. Comments are welcome. Additional copies are available by writing to the Office of Population Research Working Paper Series, Princeton University, Wallace Hall, Second Floor, Princeton, NJ 08544 Fax (609) 258-1039. Or on the web. <http://www.opr.princeton.edu/>

## Abstract

In the last fifteen years, there has been a growing interest in the role of neighborhoods in shaping a variety of outcomes for children, adults, and families. Although theoretical perspectives are well advanced and the basic statistical methods for modeling neighborhood effects are in place, a major shortcoming concerns the limitations of existing datasets. Past surveys concerned with understanding children's outcomes have not been designed with the explicit goal of supporting multilevel modeling. This makes it difficult to address the most important unresolved research issue in this area, which is to develop an understanding of the *causal* effects of neighborhood factors. In this paper, we describe the development and implementation of the sampling design for the Los Angeles Family and Neighborhood Study (L.A.FANS), a survey of children, adults, families, and neighborhoods in Los Angeles County. This survey was designed to support multilevel studies on a number of topics, including child development, residential mobility, and welfare reform. We describe the design of the baseline wave, highlighting the analytical and statistical issues that shaped the study. We also present the results of an in-depth statistical investigation of the survey's ability to support multilevel analyses that was carried out as part of the study design. The results of this study provide important guideposts for future studies of neighborhoods and their effects on adults and children.

## 1. Introduction

In the last fifteen years, there has been a growing interest in the role of neighborhoods in shaping a variety of outcomes for children, adults, and families.<sup>1</sup> The broad set of outcomes that have been studied includes violent crime (Sampson, Raudenbush, and Earls, 1997), educational attainment (Garner and Raudenbush, 1991), domestic violence (O'Campo et al., 1995), fertility (Billy and Moore, 1992), residential mobility (Lee, Oropesa, and Kanan 1994), children's development (Duncan, Brooks-Gunn, and Klebanov, 1994), teenage pregnancy (South and Baumer, 2000), and health status (Robert, 1999). Studies examining neighborhood effects have considered not only the U.S., but also countries overseas (e.g., Pebley, Goldman, and Rodriguez, 1996; Sampson and Groves, 1989; Sastry, 1996). Understanding the effects of neighborhoods is potentially quite important, from both a research and a policy perspective. For example, neighborhoods may provide a more cost-effective focus for interventions than individuals or families. Research in this area has, however, failed to produce persuasive and consistent results about the nature and strength of neighborhood effects, especially regarding children's outcomes (Jencks and Mayer, 1990; Duncan and Raudenbush, 1998; Furstenberg and Hughes, 1997; Gephart, 1997). Although theoretical perspectives are well advanced and statistical methods for modeling neighborhood effects are in place, a major shortcoming concerns the limitations of existing datasets. The Los Angeles Family and Neighborhood Survey (L.A.FANS) was designed to address many of these limitations.

Very few existing datasets have been designed for the purpose of studying neighborhood

---

<sup>1</sup> For reviews see Jencks and Mayer (1990), Gephart (1997), Aneshensel and Sucoff (1996), Robert (1999), Sampson, Morenoff, and Gannon-Rowley (2002), Leventhal and Brooks-Gunn (2000), Ginther, Haveman, and Wolfe (2000), and Pebley and Sastry (2003).

effects on child and family outcomes.<sup>2</sup> Rather, the existing neighborhood effects literature has largely used standard household surveys—such as the 1979 National Longitudinal Study of Youth and the Panel Study of Income Dynamics—linked with contextual data obtained primarily from the decennial census. The near-universal practice of employing multistage, clustered sampling schemes in the design of standard household surveys has supported the widespread use of multilevel models in this research. These sampling schemes are designed to achieve a balance between efficiency of fieldwork operations, which is facilitated by concentrating the sample in a relatively small number of compact areas, and the magnitude of design effects, which increase with number of respondents per cluster to reduce the effective sample size.

There are a number of limitations of the existing datasets for studying neighborhood effects on child and family outcomes. Contextual data are generally restricted to published sources. Extremely few surveys have collected neighborhood measures as part of the fieldwork, either from the respondents themselves (for example, on their perceptions of neighborhood characteristics and conditions) or by interviewers (for example, based on their observation of neighborhood conditions during visits to conduct interviews). Furthermore, sample sizes per cluster are generally too small to estimate community-level explanatory variables with sufficient reliability and precision. The small cluster sample sizes in many of these datasets also make it difficult to distinguish family effects from neighborhood effects.<sup>3</sup> Existing datasets lack detailed measures on several topics that are important for studying neighborhood effects. These topics

---

<sup>2</sup> The most notable exception is the Project on Human Development in Chicago Neighborhoods.

<sup>3</sup> This is especially true in longer-running panel surveys, such as the Panel Study of Income Dynamics, in which residential mobility of respondents has resulted in the originally clustered sample being much more geographically dispersed.

include past residential mobility, neighborhood choice, neighborhood definitions, neighborhood characteristics and conditions, and the salience of neighborhoods for daily life. Finally, although longitudinal information on individuals is available from several existing datasets, prospective information on sampled neighborhoods has not been collected. In particular, people moving into sampled neighborhoods are not interviewed. Therefore, the surveys do not capture a representative cross-section of residents at each wave, which is essential for understanding patterns of residential mobility and tracking neighborhood change.

These limitations in the designs of existing datasets make it difficult to address the most pressing unresolved research issue in understanding the effects of neighborhoods on children's outcomes, which is to uncover the extent to which these effects are *causal*. The prime process leading to endogenous neighborhood characteristics is that families choose the neighborhoods in which they live—and this choice may be related to the outcome of interest, such as children's development or well-being. It is necessary to understand this process in order to understand the effects of neighborhood characteristics on outcomes. Another implication is that few analyses to date have examined the effects of a richer set of contextual characteristics, particularly those relating to neighborhood social processes. Finally, little is known about how residential mobility and neighborhood choice work to influence how neighborhoods affect children and families.

We return to these issues in the conclusions, where we discuss methods for studying neighborhood effects while addressing the issue that neighborhood characteristics may be endogenous. In the body of the paper we describe the sampling design of the Los Angeles Family and Neighborhood Study (L.A.FANS), a survey of children, families, and neighborhoods in Los Angeles County. This survey was designed to support multilevel studies on a number of topics relating to child and family well-being. The baseline wave was completed in 2000-2001.

We highlight the main design and analytical considerations that shaped the baseline wave and describe the results of an in-depth statistical investigation of the survey's ability to support multilevel analyses that was carried out as part of the study design. We begin, in the next section, with a brief review of the literature on sample design for multilevel models.

## **2. Sample Design for Multilevel Models**

As described above, a central issue in sample design for standard social surveys is to maximize the efficiency of fieldwork operations by concentrating the sample in the smallest number of primary sampling units while minimizing design effects by limiting the concentration of the sample within primary sampling units. Given sufficient funding, the optimal design would be for the sample to be chosen completely at random from the population with no spatial clustering. In surveys designed to support multilevel analyses, these issues become more complex.

Snijders and Bosker (1999) provide an introduction to the literature on sample design in multilevel studies, which builds on standard sample survey methods (e.g., Cochran, 1977; Kish, 1965) to consider several additional concerns relevant for estimating multilevel models. Principal among these is the ability to estimate, with sufficient statistical power, the effects of neighborhood level characteristics as well as the effects of unmeasured neighborhood factors (i.e., variance components in multilevel models). The goal is generally to do this at the lowest feasible cost (Snijders and Bosker, 1993; Cohen, 1998). Costs are determined by the number of groups or macro units to be chosen (which we denote as  $N$ ) and the number of observations to be selected per group (which are generally assumed to be constant across macro units and which we denote as  $n$ ). There are fixed expenses associated with each community in the sample, reflecting costs for setting up fieldwork operations and collecting contextual data. There are also variable

costs based on the number of respondents in each group. Note that research to date has considered only the *average* costs for each macro or micro unit. This is presumably due to the difficulty of obtaining accurate estimates of *marginal* costs, which are more relevant for decisions on whether to increase or decrease the number of macro units.

An alternative constraint is to select the optimal design while holding constant the overall sample size ( $N \times n = M$ ). It is easy to show that increasing the number of macro units will reduce the standard errors of the estimated parameters. Thus, the most precise estimates will be obtained when there is only one individual observation per macro unit. However, if the distribution of the random effects is also of interest, at least two observations per cluster must be sampled. Another alternative is to determine the *smallest* number of macro units needed to evaluate a specific hypothesis with sufficient power (in order to minimize costs) while keeping sample size fixed.<sup>4</sup>

Snijders and Bosker (1993) and Cohen (1998) present analytical investigations of cost-constrained optimal sample sizes. Snijders and Bosker (1993) developed approximate formulas for standard errors of fixed regression parameter estimates in two-level models as a function of the sample design. They showed that with a fixed budget, as the number of macro units increases the total sample size decreases due to the high cost of adding macro units to the sample. Thus there is often a clear optimal range of macro and micro units, above and below which the standard errors increase. The difficulty in using their approach is the considerable

---

<sup>4</sup> One issue not considered yet is the need for sufficient observations within each cluster to *estimate* certain community-level explanatory variables. For many potentially important neighborhood variables, averaging individual responses is the only way to construct neighborhood measures; larger samples per cluster result in more precise estimates of these measures.

information that is required by the program, including the means, variances, and covariances of all explanatory variables as well as the variances and covariances of the random effects. In addition, estimates are needed regarding the unit costs for each macro and micro unit in the sample. Cohen's (1998) study was similar, but focused on sample size considerations for estimating variance components. From our perspective, the focus on the variance components is of limited interest because the magnitude of these random effects depends on the completeness of variables included in the model. Thus, the only interesting case is the null model, in which no explanatory variables are included. Unfortunately, as Snijders and Bosker (1999) show, the optimum group size for estimating the variance components depends strongly on the magnitude of the intraclass correlation—which is usually unknown in the design phase. Finally, in two recent studies Moerbeek and Wong (2002) considered optimal designs for hierarchical linear models when there are multiple objectives and Normand and Zou (2002) estimated sample sizes for the analysis of binary outcomes using hierarchical models when interest centers on comparative inferences about the clusters (rather than individuals within the clusters).

Multilevel designs have also been examined using simulation (see Mok, 1995; Afshartous, 1995; Maas and Hox, 2002). These studies are perhaps most useful for establishing simple rules of thumb for multilevel survey designs. Maas and Hox (2002) suggested that dividing the total sample of individuals among 30 groups is reasonable, except for obtaining precise estimates of random effects for which 50 or more groups is necessary. Mok's (1995) and Afshartous' (1995) simulation studies drew on large existing data sets. The simulation study by Mok (1995) found that gains in efficiency by increasing the total sample size beyond 2,500 were relatively small. She also found that, for a given sample size, designs with more groups and fewer observations per group tended to yield the least biased and most efficient parameter

estimates. This is clearly due to the larger effective sample sizes from these types of designs. However, when comparing designs with the same effective sample size, those with large number of observations per group appear to be more efficient, at least in estimating the random components. Afshartous (1995) investigated the minimum number of groups needed to obtain unbiased, stable, and efficient parameter estimates, holding group size fixed. He found that as few as 40 groups were sufficient for estimating regression coefficients but that as many as 320 groups were needed to estimate variance components. Work to date has not given sufficient attention to the estimation of models with random slopes. Snijders and Bosker (1999) suggested that “if precise estimation of such parameters is required, it seems advisable to use large sample sizes (30 or higher) at either level.” The same conclusion was reached by two unpublished studies (Bassiri, 1988; van der Leeden and Busing, 1994) cited by Kreft and de Leeuw (1998). Hox (1998) suggested that a good rule of thumb for estimating models with random slopes is to have about 50 groups with 20 individuals per group.

A second group of studies considers the optimal design of cluster interventions and multisite trials. These are less relevant to the design of observational studies such as L.A.FANS. Nevertheless, a number of issues are similar—such as approaches to deciding on the number of clusters or sites and the sample size per site.<sup>5</sup> In addition, these papers highlight the importance of conducting simulation studies that closely parallel the planned analyses and which vary important model assumptions.

In summary, there is a growing literature on the design of surveys to support multilevel

---

<sup>5</sup> Recent studies include Donner and Klar (2000), Raudenbush (1997), Raudenbush and Liu (2000; 2001), Zou and Normand (2001), Moerbeek, van Breukelen, and Berger (2000), and Hedeker, Gibbons, and Waternaux (1999).

analyses. However, it generally provides limited guidance for real-world applications. In particular, to apply the available analytical techniques one requires a considerable amount of information regarding the model parameters and their covariances, as well as a number of significant assumptions regarding models, data structure, and costs. Although the rules-of-thumb that have emerged are helpful, perhaps the most important practical advice this literature offers is the value of conducting a realistic simulation.

### **3. The Los Angeles Family and Neighborhood Survey**

The Los Angeles Family and Neighborhood Survey is a study of families in Los Angeles County and of the neighborhoods in which they live. The survey was designed to support research in three main areas: the effects of neighborhoods and families on children's development and well-being; the effects of welfare reform at the neighborhood level; and the process of residential mobility and neighborhood change. There are, however, a large number of other topics that can be examined with these data, not all of which have a neighborhood focus. As discussed in the introduction, there are shortcomings of studying these three main research issues with other existing datasets, primarily because of incomplete or missing information on certain important child, family, and neighborhood characteristics.

Fieldwork for the baseline wave of L.A.FANS began in April 2000 and, with the exception of a small number of cases completed in January 2002, ended in late 2001. Wave 1 of L.A.FANS included the collection and assembly of three interrelated data sets: (1) a household survey, (2) a neighborhood survey, and (3) a file of neighborhood services and characteristics (NSC) based on census data and administrative and other records. Our focus in this paper is to describe the multilevel part of the sampling plan used for the household survey.

#### 4. L.A.FANS Sample Design

L.A.FANS was designed as a multilevel survey, first sampling neighborhoods, then sampling blocks within these neighborhoods, next selecting families within these blocks, and finally sampling children and adults within these families. As discussed below in greater detail, the multilevel sampling scheme has several strengths, as well as certain limitations. Two strengths are worth mentioning here. First, it provides an efficient and cost-effective method for collecting detailed information about respondents and neighborhoods because the sample is concentrated in a relatively small number of locations. Second, family and neighborhood clustering provides researchers with the opportunity to construct neighborhood-level measures from the data and to control for unmeasured or unmeasurable factors at the family and neighborhood levels using, for example, fixed effects or random effects models.

In this section, we discuss the definition of neighborhoods that was used for sampling purposes; the total sample size; the number of neighborhoods in the sample; stratification; the sampling of tracts, blocks, and households; the selection of household respondents; and the fieldwork results. Note that although L.A.FANS sampled blocks within tracts, the simulation used block groups instead. This was because the census data necessary for the simulation were not available for blocks. Block groups are comprised of a small number of blocks and are relatively homogeneous; hence, this decision is unlikely to affect our results.

The setting for this study is Los Angeles County, California. The county is a large and highly diverse area that includes 88 separate cities and many unincorporated areas, spread over 4,083 square miles.<sup>6</sup> No neighborhoods in Los Angeles County were excluded from the L.A.FANS sample. The total 2000 county population of about 9.5 million was 45 percent

---

<sup>6</sup> The City of Los Angeles contains only about 40 percent of the County's population.

Latino, 31 percent white, 13 percent Asian-Pacific Islander, and 10 percent African American. Los Angeles is a major destination for immigrants: in 2000, about 30 percent of the population was foreign born.

### **Defining Neighborhoods**

As in many other urban areas, neighborhood boundaries are not clearly defined in Los Angeles County. This has important implications for both the sample design and the subsequent analyses. For sampling purposes, neighborhood units must be well-defined local geographic areas for which up-to-date population and poverty estimates are available. In Los Angeles County, these units include cities, zip codes, elementary school attendance areas (ESAAs), and census tracts, block groups, and blocks.<sup>7</sup> After examining maps, visiting several areas of the county, and consulting Los Angeles experts, we concluded that census tracts and ESAAs most closely approximate social definitions of neighborhoods, because they are of moderate size (an average of 8,000 inhabitants per ESAA and 5,600 per census tract), defined based on social ecological criteria, and are generally compact and not crossed by major geographic boundaries (e.g., freeways, major boulevards, and parks). We decided to use census tracts rather than ESAAs as the sampling unit, because tracts generally include children attending two or more elementary schools. Thus, the use of census tracts as sampling units will provide researchers greater scope for examining both neighborhood and school effects on children's development.

With information from the L.A.FANS household survey, researchers can examine variation in definitions of neighborhoods and identify the places where people live, work, shop,

---

<sup>7</sup> The boundaries of traditional "neighborhoods" such as Rancho Park or Pico Union in the City of Los Angeles are generally not well defined and not coded in administrative data needed to produce population and poverty estimates.

and attend school (Sastry, Pebley, and Zonta, 2001). With tract-level data for Los Angeles County, researchers can choose whether to use census tracts or consider larger “neighborhoods” in their analyses by combining data for adjacent census tracts.<sup>8</sup>

### **Total Sample Size**

The two primary design criteria for L.A.FANS were to support detailed multilevel statistical analyses and to generate reliable estimates for Los Angeles County. Drawing on information from the design of earlier multipurpose household surveys and imposing a constraint set by our budget, we initially established a total sample size for L.A.FANS to be approximately 3,250 households. We performed a second sample size calculation to verify this estimate. The second calculation was based on a generic test of proportions between two comparison groups, the equivalent of a logistic regression with a single explanatory variable. This represents a standard and routinely used approach to calculate the sample size for a survey. Note that this approach is not based on a particular hypothesis tied to a specific variable, since there are a large number of topics that the L.A.FANS was designed to address. Rather, it represents a generic approach that is both hypothetical in terms of the specific variable that is being tested and fairly conservative.

The hypothesis test for the power calculations was conceptualized as a between-strata comparison of proportions. We assumed a baseline proportion of 0.25 in the reference group and a minimum detectable difference of 0.1, or 40 percent, between the baseline group and a comparison group representing a small effect size of 0.16. Based on standard power calculations, we determined that a sample size of 325 per group was required to detect such a

---

<sup>8</sup> Given the relatively small geographic and population size of census tracts, most alternative neighborhood definitions are likely to be larger than census tracts.

difference with a type-1 error rate of 0.05 and power of 0.8.

If we were to draw a simple random sample (SRS) from the population of Los Angeles County, we would require 325 people per stratum. With a multilevel, clustered design, there are design effects that modify the actual sample size, to yield the effective sample size, in the following way:

$$\text{Effective sample size} = \text{Actual sample size} / \text{Design effect}.$$

Clustering produces a design effect greater than 1, and the design effect for estimates of summary statistics is an increasing function of the intraclass correlation (ICC) and the cluster size:

$$\text{Design effect} = 1 + \text{Intraclass correlation} \times (\text{Cluster size} - 1).$$

The inclusion of control variables in regression models will generally account for at least part of the correlation among observations in the same group, and hence lead to smaller design effects and more powerful tests than in simple comparisons without explanatory variables.

Based on previous studies, we expected the intra-community correlation to range from 0.01 to 0.05, while the intra-household correlation were likely to be substantially larger (0.2 to 0.5). Using these ICC estimates—and given a moderate cluster size of, say, 50 households per community—we calculate the design effects for our sample to range from 1.45 to 3.45. (Note that we conducted a detailed simulation study, described below, to evaluate the sensitivity of the sample design to cluster size.) The effective sample size calculations and design effects together suggest that the actual sample size for L.A.FANS should be between  $325 \times 1.45 = 471$  and  $325 \times 3.45 = 1,125$  households per stratum. Given that the L.A.FANS design calls for three strata (see below), the target sample size of 3,250 is quite large and is very conservative in its assessment of likely design effects.

## Number of Neighborhood Clusters

Given a sample size of 3,250 households, the next decision was the number of clusters or neighborhoods to be selected for the sample. The key concern was the ability to reliably estimate the effects of neighborhood-level characteristics with sufficient statistical precision using a multilevel modeling approach. We wanted to meet this goal while sampling the smallest possible number of clusters in order to minimize fieldwork costs.

The trade-off between the number of clusters and increased variance—represented here using the effective sample size—is illustrated in Figure 1. This figure shows power curves for four different values of the intraclass correlation keeping the sample size fixed at 3,250. As the number of cluster increases along the  $x$ -axis, the effective sample size, shown on the  $y$ -axis, increases. However, the slopes—representing the gains in sample size—are quite different for the four curves; in particular, the gains in effective sample size as the number of clusters increases are greatest when intraclass correlation is weaker.

As discussed above, our *a priori* estimates of intra-community correlation range from 0.01 to 0.05. Thus our initial choice of dividing the sample of 3,250 households across 65 communities (equivalent to 50 households per cluster) should yield an effective sample size of between 1,000 and 2,200 households. This is likely to be sufficient for the generic types of analyses focusing on individual-level covariates that the L.A.FANS is designed to address, such as those illustrated in our power calculations.

We were concerned, however, that our simple power calculations may not have been adequate. In particular, we wanted to determine whether the design would support the estimation of multilevel models of the effects of neighborhood characteristics on individual outcomes. We therefore undertook a simulation study to examine in greater detail, and with

more confidence, the trade-offs associated with changing cluster size while holding sample size constant.

The multilevel simulations were conducted using data from the 1990 Census STF3A tables for Los Angeles County. We chose unemployment status as the dependent variable and only considered adults in the civilian labor force; we omitted those in the armed forces (0.3 percent) or out of the labor force (32.8 percent). We then ran a logistic regression model that corrected for tract-level clustering using the Huber-White approach, with unemployment status as the outcome and the following variables as explanatory factors: race (white/Asian versus black versus other), sex, percent Hispanic in the tract, percent receiving welfare assistance in the tract, and the block-group unemployment rate. Our goal was to assess the changes in the estimated standard errors of the regression coefficients across several competing designs that held the sample size fixed at 3,250 households but differed according to the number of sampled census tracts (and, consequently, the number of households per tracts).

For each design we simulated 200 samples and ran a logistic regression model for each sample; finally, we computed the variability of the regression coefficients and standard errors across the 200 analyses. We compared the results for the four designs with 51, 66, 75, and 81 total tracts.<sup>9</sup> Table 1 presents the main results, which are the standard errors for a representative group of regression coefficients, for each of the four different sample designs.

The results in the table show that there was a substantial decline in the standard errors of the estimated parameters when the number of sampled clusters increased from 51 to 66. Note that smaller standard errors indicate that the regression parameters were estimated with greater

---

<sup>9</sup> These numbers of census tracts were chosen so that the sample of tracts could be divided evenly into three groups.

precision and hence are preferable. In contrast, the declines were much smaller when the number of tracts was increased to 75 or 81. An analysis of variance comparing the standard errors across the four schemes for each coefficient shows that the differences are statistically significant for all four covariates (for sex and black race at the .01 significance level and for the unemployment rate and percent Hispanic at the .05 level). We also performed pair-wise tests between adjacent schemes. We found that the standard errors for all four covariates were significantly different when comparing the 51-tract design with the 66-tract design. The 66-tract design and 75-tract design differed significantly only for the black race variable, but not the other three variables. Finally, there were no statistically significant differences in estimated standard errors for any of the four variables when comparing the 75-tract design with the 81-tract design. These results suggest that the major gain in efficiency for the parameter estimates comes from increasing the number of tracts in the sample from 51 to 66, while there was very little gain (but higher costs) from having more than 66 tracts.

Figure 2 presents boxplots showing the distributions of the estimated standard errors under the four different designs. These plots are useful for examining the range of outcomes and, in particular, for seeing the likelihood of having an extreme (high) value. It was an important priority to choose a scheme that offered protection against the possibility of having the final sample yield standard errors that represented an outlying case on the upper end of the distribution. The results indicated that the likelihood of this was much larger for 51 tracts compared to 66 or more tracts. It is also clear from Figure 2 that increasing the number of tracts in the sample beyond 66 provided very minor gains in terms of reducing the range of standard error estimates and lowering the likelihood of having high standard errors.

Thus, cost and design considerations all argue for choosing the smallest number of

clusters that meet the other needs reflected in the simulations. Based on the analysis described above, we decided to select a total of 65 tracts for the L.A.FANS sample from approximately 1,600 eligible tracts in Los Angeles County.

### **Stratification**

The L.A.FANS has a stratified sampling design. Stratification was adopted in order to obtain an oversample of poor and very poor tracts which, in turn, provides us with a relatively large number of respondents in poor households and of welfare recipients. A key feature of the L.A.FANS, however, is that it includes a sample of neighborhoods across the entire income range. This is important because neighborhoods are unlikely to exert an influence exclusively for children growing up in poor areas; rather, neighborhoods may also affect children growing up in middle-class or affluent areas. Nevertheless, the poorest neighborhoods are of particular scientific and policy interest and it is important to be able to conduct strata-specific analyses for the poorest neighborhoods and compare findings to results for other strata.

Stratification also reduces the variance of many estimates based on the full sample concerning the effects of poverty, welfare, and economic status. This reduction occurs because respondents are more similar to others within the same stratum, according to several important measures such as rates of welfare participation, but are quite different to respondents in other strata. It is straightforward to show that under these conditions, stratification will lead yield more precise estimates of these population parameters (see Kish, 1965).

Prior to sampling, census tracts were divided into three strata based on the percent of the tract's population in poverty. Tract-level estimates of percent in poverty in 1997 were developed by Los Angeles County's Urban Research Division using state and county administrative data. The sampling strata in the L.A.FANS design correspond to tracts that were very poor (those in

the top 10 percent of the poverty distribution), poor (tracts in the 60-89<sup>th</sup> percentiles), and non-poor (tracts in the bottom 60 percent of the distribution). The choice of three strata and the specific cut-offs were based on an analysis that examined the trade-off under different schemes between likely yield of welfare recipients and the concentration of the sample in a small number of high poverty areas. The chosen scheme represented the best balance between these two competing objectives.

A key decision regarding the stratification scheme concerns the allocation of the sample of tracts across strata. There are several considerations in determining this allocation. For example, holding the number of respondents per tract constant, different schemes will provide varying yields of poor families, welfare recipients, and respondents of different race and ethnic groups. To decide on the allocation of tracts across strata, we undertook an analysis of alternative schemes as part of the detailed simulation exercise.

The simulation analysis—as well as efficiency considerations—led to us to choose a scheme in which an approximately equal number of census tracts were allocated to each of the three strata (20 each from the very poor and poor strata and 25 from the larger non-poor stratum). The simulations also showed that oversampling of poor tracts was necessary to obtain a sufficient number of welfare recipients and minorities, particularly blacks (who are concentrated in poorer neighborhoods).

### **Sampling Tracts Within Strata**

By selecting an equal number of households (50) from each tract, we obtain a multistage design with a minimum design effect and tract-level estimates with roughly equal variances. Ignoring the intermediate step of selecting blocks within tracts, the probability of selecting the  $i$ th household in the  $j$ th tract in the  $k$ th stratum is given by:

$$\begin{aligned}\Pr(HH_{ijk}) &= n_k / N_k \times p_{jk} / P_k \times 50 / p_{jk} \\ &= n_k / N_k \times 50 / P_k ,\end{aligned}$$

where  $n_k$  is the number of tracts to be selected in the  $k$ th stratum,  $N_k$  is the total number of tracts in the  $k$ th stratum,  $p_{jk}$  is the population of the  $j$ th tract in the  $k$ th stratum, and  $P_k$  is the total population in the  $k$ th stratum. Because this expression does not depend on any characteristics of the tract, the sampling probabilities—and hence the sampling weights—are the same for all households in the stratum resulting in a self-weighted sample. As a consequence the weights are easily computed and the design is efficient because of less variability between the sampling weights. The probability proportional to size design was implemented using a systematic sampling algorithm.

### **Sample of Blocks within Selected Tracts**

In the second sampling stage, we selected census blocks within each sampled census tract. We then sampled households from these blocks (rather than from the tract as a whole) in order to simplify the fieldwork associated with listing addresses, interviewing households, and monitoring operations—and hence to reduce the costs. However, a sufficient number of blocks was selected in order to retain the diversity within each tract.

We determined the number of blocks to be sampled in each tract by dividing the target number of listings per tract by the tract's weighted average block size. Because the weighted average block size varies by tract, so too does the number of sampled blocks. In total, 439 blocks were selected (of which 11 turned out to have no households), for an average of 6.6 non-zero blocks per tract and with a range between 2 and 14.

We sampled census blocks with probability proportional to block population size. First, however, we calculated the distribution of block size for Los Angeles County and set the block

sizes in the lowest five percentiles to the fifth percentile in order to put a floor on the sampling probabilities. This results in a ceiling on the weights and protects against an unusual block having a very large weight, although it does result in some loss of efficiency. In addition, very large blocks were selected with certainty to avoid being sampled more than once in the systematic sampling algorithm that we employed.

### **Sample of Households**

The goal of the third sampling stage was to select 50 households within each tract. This number of households per tract was set by our decision to sample 65 tracts and our desire to have a balanced design, in which the same number of households per tract was interviewed. A balanced design was required for certain modeling approaches and minimizes the cluster effect and the variance. The 50 households were allocated evenly across the sampled blocks in each tract. This minimizes the harmonic mean and, therefore, both the block- and tract-level cluster effect. Thus within each tract, we sampled the same number of households per block (although households per block varied across tracts).

The 50 households were selected at random from a listing of all dwelling units within the sampled blocks. In total we listed about 41,000 addresses. Households with children under 18 years of age were oversampled so that they make up 70 percent of the sample, compared to an average of 35 percent they would otherwise comprise. Households that were unable to complete the interviews in one of the two survey languages—English and Spanish—were excluded from the sample.

### **Household Survey Respondents**

In the final sampling stage, one adult respondent was sampled at random in each selected household (designated the RSA or randomly sampled adult). In households with children, one

child respondent was also selected at random and designated the RSC (randomly sampled child). These two respondents will be followed throughout the longitudinal survey. The reason for selecting two primary respondents in each household, one adult and one child, is that it makes the rules straightforward for tracking respondents in subsequent waves, especially if the original household splits. Other adults and children were selected for the sample based on their relationship with the two primary respondents.

Table 2 summarizes which respondents were interviewed in households with and without children. In households with children, the mother of the randomly selected child was selected as a respondent and termed the Primary Caregiver (PCG). In many cases the RSA and the PCG were the same person, since the RSA was chosen at random from among all the adults in the household. If the RSC's mother did not live in the household or was unable to answer questions about the child, the child's actual primary caregiver was selected as the respondent to provide information on the selected child. Note that the selection of the PCG did not depend on his or her age—if necessary, we selected and interviewed a mother or other primary giver who was under 18 years of age. If the RSC had one or more siblings under 18 years of age who shared the same biological or adoptive mother and the same PCG, we randomly selected one of them for interview (and designated this child as the SIB).

### **L.A.FANS-1 Fieldwork Results**

The fieldwork for Wave 1 of L.A.FANS was completed between April 2001 and January 2002. A total of 3,090 households in 65 census tracts were interviewed, with 30% of households in tracts in the very poor stratum, 31% in tracts in the poor stratum, and 39% in tracts in the nonpoor stratum. The final sample size was smaller than the planned sample of 3,250 households because fieldwork costs were higher than anticipated (and the budget was fixed).

However, because this problem was identified soon after fieldwork began, the smaller sample size was incorporated into the design and did not reflect lower response rates, poorer data quality, or any other fieldwork problem. In particular, the fieldwork plan called for cases to be released in three batches. The third batch of cases released to the field was scaled down to reflect the revised sample size by randomly subsampling cases. Thus, the smaller number of cases completed reflects a deliberate design decision.

Screeners to determine eligibility were completed in 91.5% of occupied housing units selected for the sample. Among respondents selected for interview, interviews were completed with 85 percent of RSAs, 89 percent of PCGs, 87 percent of RSCs, and 86 percent of SIBs (for details see Sastry and Pebley, 2003). These response rates are similar to or better than other major surveys, such as the baseline NLSY-97 (comparably-defined response rate of 92 percent), the 1994-95 baseline wave of AddHealth (79 percent), the 1997 PSID Child Supplement (88 percent), and the 1999 baseline wave of the Welfare, Children, and Families Study (83 percent).<sup>10</sup>

Our final sample size of 3,090 was below our target sample size of 3,250. However, given that the original sample size was quite generous—and the gap between the planned and actual sample sizes was quite small—we expected this to have minimal effects on the statistical power available for most analyses using these data. To confirm this, we redid all of the power calculations described above to reflect the actual sample size. We also examined our

---

<sup>10</sup> These response rates were obtained, respectively, from the *NLSY-97 User's Guide* (Bureau of Labor Statistics, 2001), the ADDHealth web site (<http://www.cpc.unc.edu/projects/addhealth/>), the *1997 User's Guide* to the PSID Child Development Supplement (Hofferth et al., 1997), and supplemental materials to a published study using data from the Welfare, Children, and Families Study (Chase-Lansdale et al., 2003).

assumptions regarding intra-community and intra-household correlation by estimating these parameters from the data. We found that our assumed values for these correlations and for the design effects were very close to the actual values. In no case did any of our conclusions change about the adequacy of the sample size.

## **5. Discussion and Future Directions**

The design of innovative new surveys is a crucial methodological ingredient for advancing research in many areas of social science, but especially in understanding the effects of neighborhood factors on the health, development, and well-being of children, adolescents, and families. The goal of this paper was to outline the central substantive and methodological issues that shaped the sampling design of the Los Angeles Family and Neighborhood Survey, a new survey to support multilevel analyses on these and other topics. We discussed the process through which we arrived at decisions regarding the definition of neighborhoods, sample size, the number of neighborhoods in the sample, stratification, the sampling of tracts, blocks and households, and the selection of household respondents. Most of the design issues that we faced were generic—at least to the U.S.—rather than specific to the Los Angeles area. Hence our approach is broadly applicable to studies in different settings or focusing on other outcomes and processes. For example, similar multilevel designs are likely to be useful for disentangling causal effects of teachers and schools on children’s educational outcomes.

To help understand the causal effects of neighborhoods on children and families it is necessary to collect longitudinal data at all three of these levels. We plan to conduct a first L.A.FANS follow-up wave in 2005-2006. This second wave will allow researchers to capture changes in family composition and neighborhood characteristics over time and to model the process of neighborhood choice. Figure 3 illustrates the main ideas guiding the design of the

study in subsequent waves. In Wave 2, interviews will be conducted with sampled respondents who remain in the neighborhood as well as those who have left. We will also select a sample of “new entrants” into the neighborhood—that is, people who have moved into the neighborhood between the baseline survey and Wave 2. Thus, at each wave we will have a representative sample of all neighborhood residents. The new entrants will become part of the panel sample and will be followed in subsequent waves.

The RSCs and RSAs are considered our primary respondents, and once they join the sample they will be followed throughout the study whether they live together or apart. We plan to interview these sampled individuals wherever they move. For this reason, detailed contact information was collected for each respondent at Wave 1 and we are maintaining contact with respondents between waves. Respondents who remain in Los Angeles County will be interviewed in person in subsequent waves, regardless of the neighborhood in which they live. Those who move out of Los Angeles County will be interviewed by telephone, even if they leave the country. Other adults and children will be interviewed depending on whether or not they live with the RSCs. We expect most moves to be within Los Angeles County and to other locations in California, although there will be some moves to elsewhere in the U.S. and overseas.

In the second and subsequent waves, new entrants to sampled neighborhoods will be selected using several different methods. In particular, new household members in study households that remain in the 65 neighborhoods and new entrant households will be identified and sampled. The general principle guiding the selection of new entrants is that everyone moving into or being born into the tract will have a positive probability of being selected as an RSA or RSC.

The L.A.FANS design has several features that are important from an analytical

perspective. The first is the inclusion of dynamic measures of neighborhood characteristics that are collected prospectively. In particular, this survey will include not just longitudinal data on individuals, which are crucial, but also repeated cross-sectional samples of communities. By contrast, virtually all existing research is based on neighborhood data collected at just a single point in time. At best, this other research would be able to look at changes in neighborhood characteristics over a 10-year period, using data from two decennial censuses. An important design trade-off concerns the spacing of waves, where closer spacing may be necessary to capture changes in individual level behavior and outcomes—especially for children—while neighborhood-level change may unfold over a somewhat longer interval. One way to reconcile these issues would be to collect certain neighborhood data less frequently and perhaps to incorporate neighborhood history calendars (Axinn, Barber, and Ghimire, 1997) that would parallel the individual event history calendars that are completed by survey respondents. This would provide a way to characterize certain neighborhood attributes during the period between survey waves, as well as the period prior to the baseline. Through either repeated cross-sections or neighborhood history calendars, data can be collected on the precise timing and ordering of changes. This will allow researchers to incorporate better measures of neighborhood attributes and of the duration of exposure to neighborhood conditions, which is useful information for disentangling causal effects.

A second important design feature is to collect detailed retrospective and prospective information on migration and residential changes of respondents, which will allow researchers to address the problem of residential mobility. This issue has plagued research to date on the causal effects of neighborhood characteristics. By studying residential moves, researchers will gain insights into one important selection process through which neighborhood characteristics

may be endogenous. Another way the survey will allow researchers to address the problem of residential mobility is by assembling contextual data for all of Los Angeles County. Because the vast majority of migration occurs within the county—and we will have detailed measures for certain key characteristics of all neighborhoods in Los Angeles—this will mean that fewer respondents are omitted in analyses based on these data. For research that requires the detailed neighborhood measures collected for the 65 communities in the sample, residential mobility creates a sample selection problem. Several approaches are possible to tackle this problem, including imputation (to address the issue as a problem of missing data) and the estimation of sample selection models. Nevertheless, being able to investigate the determinants of residential mobility will provide useful insights into possible selection effects.

Finally, the design will provide researchers with an opportunity to control for the endogeneity of neighborhood characteristics. Three main factors underlie the potential endogeneity of neighborhood effects and make it difficult to identify causal effects of neighborhood characteristics. First, there may be systematic processes through which resources and programs are targeted towards certain neighborhoods and not others. This would result in differences in the availability and quality of neighborhood institutions that reflect the priorities of decisionmakers responsible for providing services. This may result, for instance, in health care facilities being placed in areas with greatest demand. In this situation, studies that did not correctly account for the purposive program placement would find the greater availability of these facilities associated with *worse* health outcomes. Second, neighborhood characteristics, including the presence and quality of local institutions, may reflect the ability of residents to mobilize and demand these facilities and services. Finally, people may select the neighborhoods in which they live on the basis of their own needs and preferences, as well as the incentives

created by macro-level social and economic processes such as historical and contemporary racial discrimination (Massey and Denton, 1993). Unmeasured parent characteristics may lead them both to choose certain kinds of neighborhoods and to make other investments in their children (Tienda, 1991; Evans Oates and Schwab, 1992; Duncan and Raudenbush, 1998). Thus, a common set of parent/family characteristics determines both children's development *and* neighborhood choice. Some of these characteristics, such as household income and parents' education, are measurable, and can be controlled in models of children's development. However, other characteristics are unobserved and hence are picked up in the random component of the statistical model, where correlation with included regressors—i.e., neighborhood characteristics—leads to biased and inconsistent estimates of all model parameter.

The statistical methods to disentangle the causal effects of neighborhood characteristics on key outcomes of interest are established and generally well-known. They comprise of instrumental variables methods—or, more generally, simultaneous equation models—as well as fixed effects and correlated random effects models. As Duncan and others have explained (e.g., Duncan, Connell, and Klebanov, 1997), the problem of neighborhood selection is really one of omitted variables. Specific unobserved (omitted) parent/family factors are the parents' cognitive ability and family motivation and aspirations, which may influence the degree to which a family values their children's development (as well as their choice of place of residence). An instrumental variable approach requires researchers to identify a new variable that is correlated with each endogenous neighborhood variable in question but is uncorrelated with the error components (including the unobserved family effects). However, a serious problem with this approach is the difficulty in finding *any* credible and viable instruments (Aaronson, 1997, 1998); moreover, finding *multiple* instruments for several different neighborhood variables of interest is

likely to be virtually impossible. Fixed effects models assume that the unobserved effects are time invariant, as are the correlations between the unobserved effects and the regressors.

Through differencing, the fixed effects models eliminate all time-invariant effects (both observed and unobserved) and hence the problem of omitted variables. Finally, the unobserved effects can be captured in a model as family-specific random effects within a multilevel modeling framework. By allowing for correlation between the random effects and neighborhood variables in the model, a correlated random effects approach (Chamberlain, 1980) provides researchers with an approach to assess the nature and strength of this association and hence to test for the potential endogeneity of neighborhood characteristics. Correlated random effects models have the advantage of being able to incorporate time-vary unobserved effects; the correlation between the unobserved effects and the regressors can also be allowed to change over time. Moreover, this approach nests the fixed effects model—that is, by imposing a set of restrictions in the correlated random effects model, researchers can constrain the unobserved effects to be time invariant and test this restriction.

Very few previous analyses of neighborhood effects on children's outcomes have actually used these methods to control for the potential endogeneity of neighborhood factors. Several studies have used instrumental variables as a means of eliminating the correlation between unobserved parent attributes and neighborhood variables (Evans, Oates, and Schwab, 1992; Case and Katz, 1991; Foster and McLanahan, 1996). Aaronson (1997, 1998) and Plotnick and Hoffman (1996) used sibling fixed effects models in analyses of educational attainment, adult economic status, and teen pregnancy. While Aaronson found significant neighborhood effects once unobserved family characteristics were controlled, Plotnick and Hoffman (1996) did not. No studies to date have used correlated random effects models to study the effects of

neighborhood characteristics on child outcomes, although this approach clearly has some important advantages.

In conclusion, the Los Angeles Family and Neighborhood Survey is an important new resource for policymakers with an interest in Los Angeles and researchers in the social and behavioral sciences. It will be especially useful for researchers seeking to understand the effects of neighborhoods on child development, migration and residential mobility, and welfare reform. Finally, by tackling several key sample design issues in the L.A.FANS, we have laid out a general approach for designing multilevel models that can be applied elsewhere.

## References

- Aaronson, D. 1997. "Sibling estimates of neighborhood effects," in J. Brooks-Gunn, G.J. Duncan, and J.L. Aber (eds.) *Neighborhood Poverty*, Volume II. New York: Russell Sage Foundation.
- Aaronson, D. 1998. "Using sibling data to estimate the impact of neighborhoods on children's educational outcomes," *Journal of Human Resources* 33: 915-946.
- Afshartous, David. 1995. "Determination of sample size for multilevel model design," paper presented at the AERA meeting in San Francisco, CA, April 1995.
- Aneshensel, C.S., and Sucoff, C.A. 1996. "The neighborhood context of adolescent mental health," *Journal of Health and Social Behavior* 37: 293-310.
- Axinn, William G., Jennifer S. Barber, and Dirgha J. Ghimire. 1997. "The neighborhood history calendar: A data collection method designed for dynamic multilevel modeling," *Sociological Methodology* 27: 355-392.
- Bassiri, D. 1988. "Large and small sample properties of maximum likelihood estimates for the hierarchical linear model." Unpublished doctoral dissertation, Michigan State University, East Lansing, MI. (Cited in Kreft and de Leeuw, 1998.)
- Billy, John O.G., and David E. Moore. 1992. "A multilevel analysis of marital and nonmarital fertility in the United States," *Social Forces* 70: 977-1011.
- Bureau of Labor Statistics. 2001 *NLSY-97 User's Guide*. Prepared by Center for Human Resources, The Ohio State University, Columbus, OH.
- Case, A.C., and L.F. Katz. 1991. "The company you keep: The effects of family and neighborhood on disadvantaged youth" NBER Manuscript.
- Chamberlain, G. 1980. "Analysis of covariance with qualitative data," *Review of Economic Studies* 47: 225-238.
- Chase-Lansdale, P. Lindsay, Robert A. Moffitt, Brenda J. Lohman, Andrew J. Cherlin, Rebekah Levine Coley, Laura D. Pittman, Jennifer Roff, and Elizabeth Votruba-Drzal. 2003. "Mothers' Transitions from Welfare to Work and the Well-Being of Preschoolers and Adolescents," *Science* 299: 1548-1552.

- Cochran, W.G. 1977. *Sampling techniques*. Third edition. New York: Wiley.
- Cohen, Michael P. 1998. "Determining sample sizes for surveys with data analyzed by hierarchical linear models," *Journal of Official Statistics* 14: 267-275.
- Donner, A., and N. Klar. 2000. *Design and analysis of cluster randomization trials in health research*. London: Arnold.
- Duncan, G.J., J.P. Connell, and P.K. Klebanov. 1997. "Conceptual and methodological issues in estimating causal effects of neighborhoods and family conditions on individual development," in J. Brooks-Gunn, G.J.Duncan, and J.L. Aber (eds.) *Neighborhood Poverty*, Volume I. New York: Russell Sage Foundation.
- Duncan, G.J., and S.W. Raudenbush. 1998. "Neighborhoods and adolescent development: How can we determine the links," Joint Center for Poverty Research Working Paper Num. 3.
- Duncan, G.J., J. Brooks-Gunn, and P. Klebanov. 1994. "Economic deprivation and early childhood development," *Child Development* 65: 296-318.
- Evans, W.N., W.E. Oates, and R.M. Schwab. 1992. "Measuring peer group effects: A study of teenage behavior," *Journal of Political Economy* 100: 966-991.
- Foster, E.M., and S. McLanahan. 1996. "An illustration of the use of instrumental variables: Do neighborhood conditions affect a young person's chance of finishing high school?" Manuscript.
- Furstenberg, Frank F., and Mary E. Hughes. 1997. "The influence of neighborhoods on children's development: A theoretical perspective and a research agenda," in Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber (eds.), *Neighborhood poverty: Policy implications in studying neighborhoods*, Volume II. New York: Russell Sage.
- Garner, C., and S. Raudenbush. 1991. "Neighborhood effects on educational attainment: A multilevel analysis," *Sociology of Education* 64: 251-262.
- Gephard, Martha A. 1997. "Neighborhoods and communities as contexts for development," in Jeanne Brooks-Gunn, Greg J. Duncan, and J. Lawrence Aber (eds.), *Neighborhood poverty: Policy implications in studying neighborhoods*, Volume I. New York: Russell Sage.
- Ginther, Donna, Robert Haveman, and Barbara Wolfe. 2000. "Neighborhood attributes as

- determinants of children's outcomes: How robust are the relationships?" *Journal of Human Resources* 35: 603-642.
- Hedeker, D., R.D. Gibbons, and C. Waternaux. 1999. "Sample size estimation for longitudinal designs with attrition: Comparing time-related contrasts between two groups," *Journal of Educational and Behavioral Statistics* 24:70-93.
- Hofferth, S., Davis-Kean, P.E., Davis, J., Finkelstein, J. 1997. *The Child Development Supplement of the Panel Study of Income Dynamics, 1997 User Guide*. Survey Research Center, Institute for Social Research, The University of Michigan. Ann Arbor, MI.
- Hox, J. 1998. "Multilevel modeling: When and how," in I. Balderjahn, R. Mathar, and M. Schader (eds.), *Classification, data analysts, and data highways*. New York: Springer Verlag.
- Jencks, C., and S. Mayer. 1990. "The social consequences of growing up in a poor neighborhood," in L.E. Lynn and M.F.H. McGeary (eds.), *Inner-city poverty in the United States*. Washington: National Academy Press.
- Kish, L. 1965. *Survey Sampling*. New York: John Wiley & Sons, Inc.
- Kreft, Ita, and Jan de Leeuw. 1998. *Introducing multilevel modeling*. Thousand Oaks: Sage.
- Lee, B.A., R.S. Oropesa, and J.W. Kanan. 1994. "Neighborhood context and residential mobility," *Demography* 31: 249-270.
- Leventhal, T., and Brooks-Gunn, J. 2000 "The neighborhoods they live in: The effects of neighborhood residence upon child and adolescent outcomes," *Psychological Bulletin* 126: 309-337.
- Massey, D.S., and N.A. Denton. 1993. *American apartheid: Segregation and the making of the underclass*. Cambridge: Harvard University Press.
- Mok, M. 1995. "Sample size requirements for 2-level designs in educational research," *Multilevel Modeling Newsletter* 7(2): 11-15.
- Moerbeek, M., G.J.P. van Breukelen, and M.P.F. Berger. 2000. "Design issues for multilevel experiments," *Journal of Educational and Behavioral Statistics*.

- Moerbeek, M., and W.K. Wong. 2002. "Multiple-objective optimal designs for the hierarchical linear model," *Journal of Official Statistics* 18: 291-303.
- Normand, S.-L.T., and K.H. Zou. 2002. "Sample size considerations in observational health care quality studies," *Statistics in Medicine* 21:331-345.
- O'Campo, P., A.C. Gielen, R.R. Faden, X. Xue, N. Kass, and M. Wang. 1995. "Violence by male partners against women during the childbearing year: A contextual analysis," *American Journal of Public Health* 85: 1092-1097.
- Pebley, A.R., N. Goldman, and G. Rodriguez. 1996. "Prenatal and delivery care and childhood immunization in Guatemala: Do family and community matter?" *Demography* 33: 197-210.
- Plotnick, R.D., and S.D. Hoffman. 1996. "The effect of neighborhood characteristics on young adult outcomes: Alternative estimates," Institute for Research on Poverty Discussion Paper No. 1106-96.
- Raudenbush, S.W. 1997. "Statistical analysis and optimal design for cluster randomized trials," *Psychological Methods* 2: 173-185.
- Raudenbush, S.W., and X. Liu. 2000. "Statistical power and optimal design for multisite randomized trials," *Psychological Methods* 5: 199-213.
- Raudenbush, S.W., and X. Liu. 2001. "Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change," *Psychological Methods* 5: 199-213.
- Robert, SA (1999) "Socioeconomic Position and Health: the Independent Contribution of Community Socioeconomic Context" *Annual Review of Sociology*, 25: 489-516.
- Sampson, R.J., and W.B. Groves. 1989. "Community structure and crime: Testing social disorganization theory," *American Journal of Sociology* 94: 774-802.
- Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and violent crime: A multilevel study of collective efficacy," *Science* 277: 918-924.
- Sampson, Robert J., Jeffrey D. Morenoff and Thomas Gannon-Rowley. 2002. "Assessing neighborhood effects: Social processes and new directions in research, *Annual Review of Sociology* 28: 443-478.

- Sastry, Narayan. 1996. "Community characteristics, individual and household attributes, and child survival in Brazil," *Demography* 33: 211-229.
- Sastry, N., and A.R. Pebley. 2003. "Non-response in the Los Angeles Family and Neighborhood Survey," RAND Labor and Population Program Working Paper DRU-2400/7-LAFANS.
- Sastry, N., A.R. Pebley, and M. Zonta. 2001. "Neighborhood Definitions and the Spatial Dimensions of Daily Life in Los Angeles." RAND Labor and Population Program Working Paper DRU-2400/8-LAFANS.
- Snijders, T.A.B., and R.J. Bosker. 1993. "Standard errors and sample sizes for two-level research," *Journal of Educational Statistics* 18: 237-259.
- Snijders, Tom A.B., and Roel J. Bosker. 1999. *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks: Sage.
- South, Scott J., and Eric P. Baumer 2000. "Deciphering community and race effects on adolescent premarital childbearing," *Social Forces* 78: 1379-1407.
- Tienda, M. 1991. "Poor people, poor places: Deciphering neighborhood effects on poverty outcomes," in Joan Huber (ed.) *Macro-micro linkages in sociology*. Newbury Park, CA: Sage.
- Van der Leeden, R., and F.M.T.A. Busing. 1994. "First iteration versus IGLS/RIGLS estimates in two-level models: A Monte Carlo study with ML3," Working paper PRM 94-03, Psychometrics and Research Methodology, Leiden, Netherlands. (Cited in Kreft and de Leeuw, 1998.)
- Zou, K.H., and S-L.T. Normand. 2001. "On determination of sample size in hierarchical binomial models," *Statistics in Medicine* 20:2163-2182.

**Table 1.** Standard Errors for Logistic Regression Coefficients from a Simulation Analysis Based on a Model of Unemployment Status

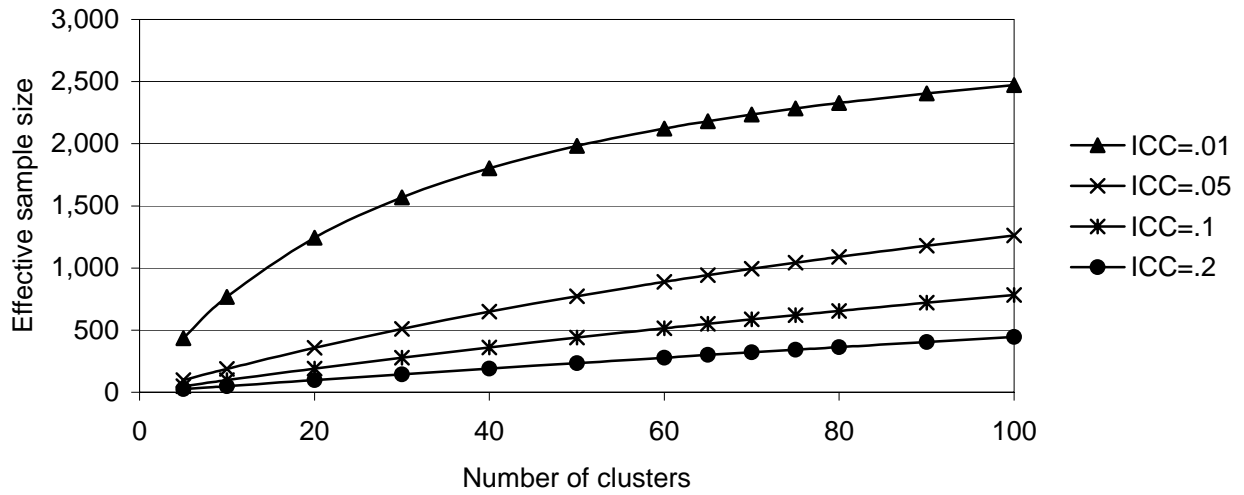
Number of census tracts	Coefficient (Level of measurement)			
	Sex is female (Individual)	Race is black (Individual)	Percent Hispanic (Tract)	Unemployment rate (Block group)
51	0.140	0.191	0.196	1.81
66	0.136	0.175	0.188	1.75
75	0.134	0.173	0.189	1.70
81	0.130	0.170	0.190	1.73

Note: Tracts were allocated equally across three strata. See text for a description of the simulations.

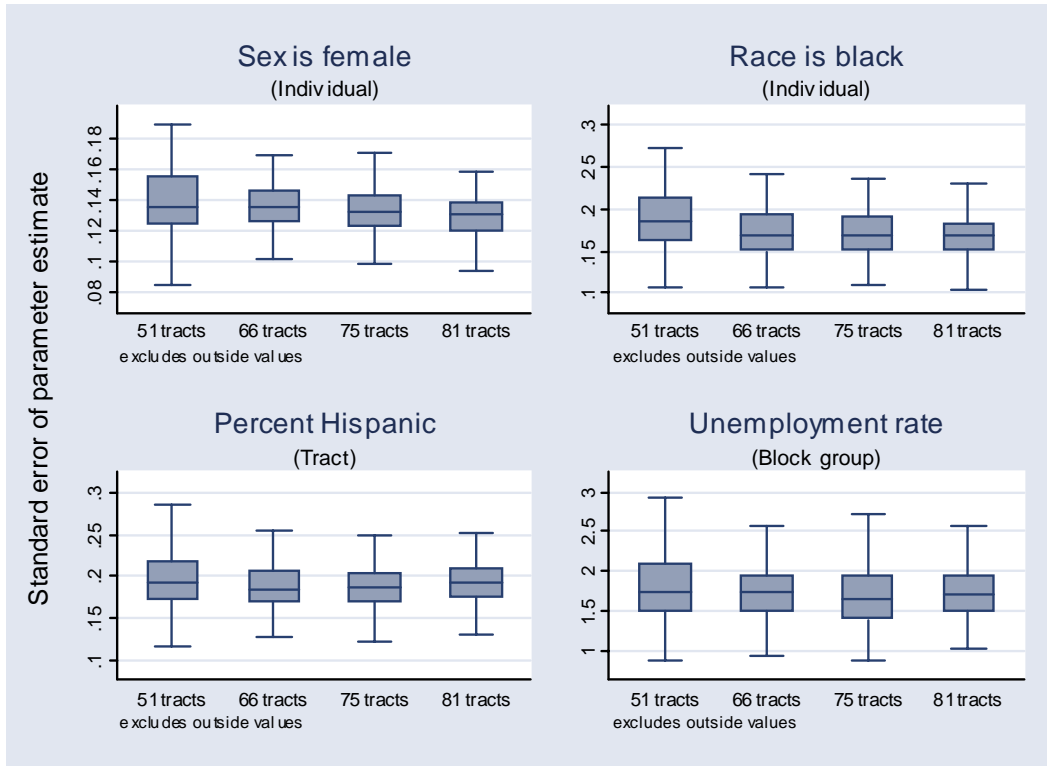
**Table 2.** L.A.FANS Respondents by Household Type

<b>HHs with Children &lt;18</b>	<b>HHs without Children &lt;18</b>
Randomly selected adult (RSA) Randomly selected child (RSC) Primary Caregiver of RSC (PCG) Sibling of RSC (SIB)	Randomly selected adult (RSA)

**Figure 1.** Effective sample size by number of clusters, holding actual sample size at 3,250



**Figure 2.** Box plots showing distribution of standard error estimates for logistic regression coefficients from simulation analysis based on a model of unemployment status



**Figure 3.** Design of L.A.FANS Follow-Up Waves

