ABSTRACT

Title of dissertation: NEIGHBORHOOD CHARACTERISTICS AND
PARTICIPATION IN HOUSEHOLD SURVEYS

Carolina Casas-Cordero Valencia,
Doctor of Philosophy, 2010

Dissertation directed by: Professor Frauke Kreuter
Joint Program in Survey Methodology

Declining response rates in household surveys continue to demand not only a
better understanding of the mechanisms underlying nonresponse, but also the iden-
tification of auxiliary variables that can help assess, reduce, and hopefully correct
for this source of error in survey estimates. Using data from L.A. Family and Neigh-
borhood Study (L.A. FANS), this dissertation shows that observable characteristics
of the sampled neighborhoods have the potential to advance both survey research
topics.

Paper 1 of this dissertation advances our understanding of the role that local
neighborhood processes play in survey participation. The measures of social and
physical environments are shown to be significant predictors of household coopera-
tion in the L.A.FANS, even after controlling for the socio-economic composition of
households and neighborhoods.

A nice feature of the indicators of the physical environment is that they can
be observed without performing the actual interview. Thus they are available for

both respondents and nonrespondents. However, survey interviewers charged with this task might make errors that can limit the usability of these observations. Paper 2 uses a multilevel framework to examine 25 neighborhood items rated by survey interviewers. The results show that errors vary by type of item and that interviewer perceptions are largely driven by characteristics of the sampled areas – not by characteristics of the interviewers themselves.

If predictive of survey participation, neighborhood characteristics can be useful for survey fieldwork decisions aimed at increasing response rates. If neighborhood characteristics are also related to survey outcome variables they furthermore can be used to inform strategies aimed at reducing nonresponse bias. Paper 3 compares the effectiveness of several different neighborhood characteristics in nonresponse adjustments for the L.A.FANS, and shows that interviewer observations perform similar to Census variables when used for weighting key estimates of L.A. FANS.

Results of this dissertation can be relevant for those who want to increase response rates by tailoring efforts according to neighborhood characteristics. The most important contribution of this dissertation, however, lies in re-discovering intersections between survey methodology and urban sociology.

NEIGHBORHOOD CHARACTERISTICS AND PARTICIPATION IN HOUSEHOLD SURVEYS

by

Carolina Casas-Cordero Valencia

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:

Professor Frauke Kreuter, Chair/Advisor
Professor Mick Couper
Professor Judith Hellerstein
Dr. Steve Miller
Professor Narayan Sastry

Dedication

A mis padres,

Victor Casas-Cordero y Dina Valencia,

por haberme entregado herramientas y disciplina

para lograr mis metas.

A mi profesora guía,

Frauke Kreuter,

por ser la mejor mentora, colega y amiga

que habría imaginado encontrar.

Y a mi esposo,

Carlos Núñez Cataldo,

porque sólo gracias a tu paciencia, cariño y apoyo

es que tuve la posibilidad de volar tan alto.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. OVERVIEW

New theoretical insights from urban sociology and social epidemiology - and the development of instruments to measure neighborhood social processes and the physical environment - have revitalized the interest in researching the effects of neighborhood characteristics on individual-level outcomes across many fields (Sampson et al., 2002; Brooks-Gunn et al., 1997; Kawachi and Berkman, 2003). Survey methodology can also benefit from these insights and developments. Declining response rates in household surveys (Atrostic et al., 2001; De Leeuw and DeHeer, 2002; Curtin et al., 2005) increase the pressure to understand nonresponse mechanisms, and to identify auxiliary variables that can be routinely collected to assess the potential for nonresponse bias. To be successful for nonresponse bias investigations, however, neighborhood characteristics need first to comply with two conditions (Little, 1986; Kalton and Flores-Cervantes, 2003; Little and Vartivarian, 2003, 2005; Groves, 2006; Kreuter et al., 2010): (1) be available for respondents and non-respondents, and (2) be associated with both *individual outcomes* and *participation in household surveys*.

The few studies in the survey literature that included neighborhood-level indicators concluded that neighborhood-level effects on survey participation were rather small (Couper and Groves, 1996; Campanelli et al., 1997; Groves and Couper, 1998;

O'Muircheartaigh and Campanelli, 1999; Kennickell, 1999; Lynn et al., 2002; Kennickell, 2003; Johnson et al., 2006; Bates et al., 2008; Durrant and Steele, 2009). Nevertheless, there are reasons to revisit this topic. First, most of these studies used census demographic data *in-lieu* of direct measures of the neighborhood constructs of interest (Groves and Couper, 1998; Kennickell, 2003; Durrant and Steele, 2009). Second, most of the neighborhood constructs used for modeling survey participation were influenced by the early literature on social disorganization and crime. Since then, theory has evolved and a new understanding of neighborhood processes has led to constructs that might be much more suitable to explain the behavior of households and interviewers during the survey recruitment process. A careful test of these emerging ecological mechanisms has yet to be performed. Finally, some of the measures used to characterize neighborhoods are amenable to observation – which suggest this task could be passed on to regular survey interviewers or household listers. Taking these measures, however, would not be the primary job of interviewers or listers and thus might be more prone to measurement error. To date, only a few studies have provided evidence of the quality of observational data collected by survey interviewers (Pickery and Loosveldt, 2004a; Kreuter et al., 2007; Alwin, 2008; Carton, 2008).

For these reasons, three main research questions emerge: *Are theoretically-driven measures of neighborhood characteristics associated with participation in household surveys?; What are the measurement error properties of neighborhood measures collected by survey interviewers using direct observation methods?; How effective are nonresponse weighting adjustments based on neighborhood observations versus those*

based on Census records? Chapters 2, 3 and 4 of this dissertation aim to address these issues. Each one of these chapters is a stand-alone paper that provides motivation, theoretical background, analyses and a thorough discussion of the results. Chapter 5 summarizes the results across the three papers and identifies areas for future research. A consolidated appendix for the three papers is provided in Chapter 6.

As mentioned above, studies of neighborhood effects on survey participation are scarce. Paper 1 (Chapter 2) advances our understanding of the role that local neighborhood processes play in survey participation. It first extends the Groves-Couper model of survey participation to formally incorporate the influence of neighborhood physical conditions on household(er)s and interviewer. Hypotheses are derived from this extension and tested using multilevel analyses to account for individual and neighborhood-level factors known to influence the survey participation decision.

A nice feature of the indicators of the physical environment is that they can be observed without performing the actual interview. Thus they are available for both respondents and nonrespondents. However, survey interviewers charged with this task might make errors that can limit the usability of these observations. Paper 2 (Chapter 3) assess the measurement error properties of the neighborhood data collected by survey interviewers using direct observation methods. Estimates of the reliability of interviewer observations are obtained using different measures of consistency, like the percent agreement index, kappa, and intra-class correlation coefficients. Estimates of interviewer effects, sampling point effects, and the effects

of interviewer characteristics on perceptions of neighborhood characteristics are also provided for 25 different neighborhood items.

If predictive of survey participation, neighborhood characteristics can be useful for survey fieldwork decisions aimed at increasing response rates. If neighborhood characteristics are also related to survey outcome variables they furthermore can be used to inform strategies aimed at reducing nonresponse bias. Paper 3 (Chapter 4) assess the effectiveness of nonresponse adjustments based on neighborhood observations versus adjustments based on Census demographics. To address this issue, different response propensity models are fitted using different combinations of Census records and neighborhood observational data. Estimates of relative loss $(1 + L)$ and Kish's mean square error are used as criterion.

All analyses implemented here used data from the Los Angeles Family and Neighborhood Study (L.A. FANS). The L.A. FANS is a study of families in Los Angeles County and the neighborhoods in which they live (Sastry et al., 2003). Data available for research included the household interviews, Census demographic characteristics at the tract level, and interviewer observations about household composition and physical characteristics of the sampled blocks. The overall description of the study is provided in Paper 1. Papers 2 and 3 only provide descriptions on datasets or variables not introduced earlier.

2. PAPER 1: TESTING NEIGHBORHOOD MECHANISMS INFLUENCING PARTICIPATION IN HOUSEHOLD SURVEYS

2.1 Introduction

Social scientists realize that certain phenomena cluster spatially, and so *place* matters (Diez-Roux, 2001; Kawachi and Berkman, 2003; Morenoff, 2003). Survey researchers are no exception. Low response rates have been shown in urban areas across time and countries (Groves and Couper, 1998; Stoop, 2005) for both contact rates (Smith, 1983; Gfroerer et al., 1997; Groves and Couper, 1998) and cooperation rates (House and Wolf, 1978; DeMaio, 1980; Steeh, 1981; Smith, 1983; Goyder et al., 1992; Groves and Couper, 1998; Durrant and Steele, 2009). But what explains this urbanicity effect?

One of the urban correlates of survey participation that is most difficult to measure is the social environment. Typically, there are no readily available measures to study its effects on household cooperation. As a consequence, most research to date is based on measures of socio-economic composition derived from demographic variables that are compiled from Census records. These are then used as proxies for the social processes of interest. The current strategy to study the effect of area-level mechanisms on household cooperation presents three shortcomings. First, it cannot disentangle the area-level effect of socio-economic composition from that of the social environment. Second, the effect of the physical environment has not been formally included in theories of survey participation. And third, it has not clearly identified whether different levels of geography may be relevant for different area-level mechanisms influencing the survey response process.

Measures of the social and physical environment are increasingly being used in

the sociological literature to explain an array of individual-level phenomena (Diez-Roux, 2001; Sampson et al., 2002; Morenoff, 2003; Kawachi and Berkman, 2003). Given the increasing availability of neighborhood and residential data, this study now takes on the challenge of disentangling the effects of different area-level characteristics on survey participation. With this goal in mind this paper formulates a revised version of the well-known Groves and Couper model of survey participation and aims to answer the following research questions: (1) Does neighborhood socio-economic composition explain variability in cooperation rates? (2) Does the neighborhood social or physical environment explain additional variance? If no additional variance is explained, (3) Do they help explain how other neighborhood mechanisms work?

The Los Angeles Family and Neighborhood Study (L.A. FANS) is an ideal data set with which to test these ideas. L.A. FANS has neighborhood measures of the *physical* and *social environment*, and census measures of *socio-economic composition* traditionally used as proxies for social environments. In addition, the study collected a small set of household characteristics for both respondents and nonrespondents so household-level effects can be controlled for when analyzing area-level effects. The specific nature of the L.A. FANS data allows the examination of the process of contact and cooperation for different stages of the recruitment process (screening, rostering, and the survey interview). It also allows the assessment of the relative influence of different levels of geography on these corresponding response outcomes.

2.2   Background

Before looking at the L.A. FANS data it is important to review the literature of area-level effects on individual-level outcomes. I first review the study of area-level effects on the household cooperation decision. I then present several insights from urban sociology aimed at improving our understanding of contextual effects arising at the neighborhood level. Both bodies of literature will then be used to propose a revised version of the well-known Groves and Couper model of survey participation.

2.2.1   Contextual Effects on Household Cooperation

Groves and Couper (1998) hypothesize that four factors influence the householder's decision to cooperate with a cross-sectional survey request: the survey design, the interviewer, the household and the social environment (see Figure 2.1).[1] The authors conceptualized the influence of the social environment at two levels. The first are societal-level conditions that facilitate or mitigate survey participation (Groves and Couper, 1998; Johnson et al., 2002). The second are more local variations in context at the community or neighborhood level that shape the decision to participate or refuse (Groves and Couper, 1998, 155). Early nonresponse studies focused on the effects of different degrees of urbanization on survey response rates and found lower cooperation rates in urban areas and large cities (House and Wolf, 1978; DeMaio, 1980; Steeh, 1981; Smith, 1983; Goyder et al., 1992). Studies that followed focused on effects across smaller geographical areas like census tracts, zip

---

[1] See Lepkowski and Couper (2001) for a survey participation model under the context of a longitudinal survey request.

codes and census blocks (Goyder et al., 1992; Campanelli et al., 1997; Groves and Couper, 1998; Kennickell, 1999; Sastry and Pebley, 2003; Johnson et al., 2006).
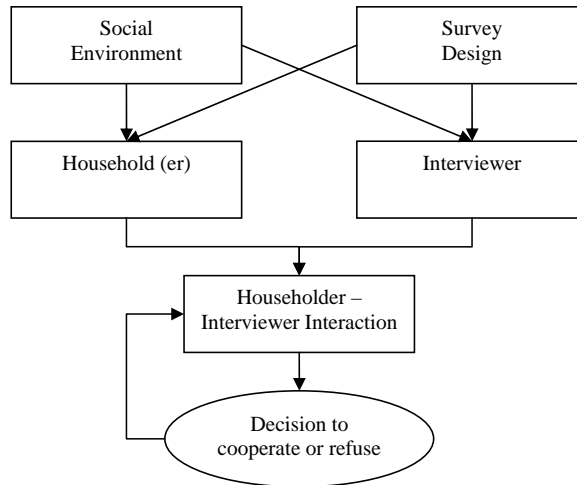
Fig. 2.1: Groves and Couper Model of participation in Household Surveys.

If we look across the literature, three area-level mechanisms have been hypothesized to explain this urbanicity effect on household cooperation: (1) population density, (2) crime, and (3) social disorganization. In the words of Groves and Couper (1998, 176) *"We believe that many of the effects of urbanicity found in the literature may be explained in terms of greater population density, higher crime rates, and social disorganization that are often associated with life in large urban areas"*. In the survey context, these three area-level mechanisms are hypothesized to reduce cooperation rates by decreasing residents' willingness to engage in activities (such as surveys) that are seen to benefit the community or the society at large (Groves and Couper, 1998; Abraham et al., 2009).

Living in areas with high *population density* has been associated with excessive social encounters and 'perceptions of crowding', which may lead to avoidance of

contact with strangers (Aellio and Baum, 1979). This, in turn, may lead to less helping behavior and greater distrust of strangers (Franck, 1980; Wilson, 1985). This reduction in 'helping behavior' is hypothesized to translate into reduced survey cooperation in crowded areas (Groves and Couper, 1998). Population density has been found to decrease cooperation in bivariate (House and Wolf, 1978; Goyder et al., 1992; Groves and Couper, 1998) and multivariate analyses (Smith, 1983; Brehm, 1993; Groves and Couper, 1998; Lynn et al., 2002).

Living in areas with high levels of *crime* has been shown to increase levels of 'fear of crime', 'risk perception', and 'distrust of strangers' among local residents (Taylor et al., 1995). The 'perception of potential harm' may cause the sample persons to react negatively to the survey request (Groves, 1989; Groves and Couper, 1998). Evidence from bivariate analyses suggest that crime reduces cooperation (House and Wolf, 1978; Goyder et al., 1992; Groves and Couper, 1998). However, crime is no longer a significant predictor of cooperation in multivariate analyses (Groves and Couper, 1998).

Social disorganization is an umbrella term that includes a variety of other concepts, among them sometimes *population density* and *crime* themselves. This framework emerged from work in urban sociology that proposed a link between neighborhood social processes and the emergence of crime and victimization in inner city areas (Shaw and McKay, 1942, 1969; Bursik, 1986, 1989). Areas with high *social disorganization* are characterized by the inability of the community to realize the 'common values' of its residents and maintain effective 'social controls' (Kornhauser, 1978). These areas are hypothesized to influence the attitudes and behavior of local

10

residents leading to fear of crime, distrust, and feelings of powerlessness (Taylor, 1997).

Groves and Couper (1998, 178) focused on the role of shared norms and values among potential respondents. The authors suggested that the lack of *cohesion* at the community level may have as its counterpart the isolation of individuals both from the local community and from the society at large. This relative lack of participation or involvement in the community may reduce the willingness to engage in activities such as surveys, that are seen to benefit the community of society at large (Groves and Couper, 1998, 177).

Unlike measures of *population density* and *crime*, which can be readily available from administrative records, measures of *social disorganization* are not readily available. In practice, most nonresponse studies compile indirect indicators of the social processes of interest from demographics derived from Census records.[2] In multivariate analyses, cooperation rates have shown to correlate positively with percentage of children (Groves and Couper, 1998), multi-unit structures (Bates et al., 2008), and college graduates (Kennickell, 1999). They have also shown to correlate negatively with percentage of whites (Kennickell, 1999), 65+ years (Kennickell, 1999), multi-unit structure (Goyder et al., 1992), boarded-up housing units (Gfroerer et al., 1997), working males (Kennickell, 1999), professional/managerial occupations (Johnson et al., 2006), non-movers in 5 years (Johnson et al., 2006) and high housing

---

[2] Groves and Couper (1998, 177), for example, identified the following as indirect indicators of *social cohesion*: the transience of the population, the presence of large apartment complexes, racial/ethnic heterogeneity, and the physical decay of a neighborhood.

11

values (Kennickell, 1999).

Few studies use data-reduction techniques to obtain more parsimonious measures from Census records. Goyder et al. (1992) and Van Goor et al. (2005) used factor analysis to derive area-level measures.[3] [4] Neither study found significant effect for the combined measures of social disorganization on household cooperation.[5] Both studies, however, included measures of crime into their social disorganization measure, thus the independent effect of *social disorganization* and *crime* could not be disentangled.

---

[3] The Goyder et al. (1992) study obtained a 2-factor solution for social disorganization using six census variables. The first factor, labeled 'social disorganization', had strong loadings on: (1) total non-traffic criminal offenses, (2) ratio of movers to non-movers during the past 5 years, (3) ratio of lone-parent families to husband-wife families, (4) multiple-story building dweller, and (5) proportion born in Asia, continental Europe, or Central-South America. The second factor, labeled factor 'density', loaded strongly on (6) population density and only moderately on (5) proportion born overseas.

[4] The Van Goor et al. (2005) study obtained a 4-factor solution for social disorganization. The first factor, labeled 'stable middle-class population', had strong loadings on: post-war neighborhoods with families with young children and with a stable population scored high on this component. The second factor, labeled 'socioeconomic deprivation and social marginality', had strong loadings on: neighborhoods with low average income, with relatively high numbers of singles, one-parent families, members of ethnic minority groups, and persons arrested for criminal activities scored high on this dimension. The third factor combined the 'physical decay (vacancies) of the neighborhood with criminal offences' to which local residents in particular fall victim. The fourth factor represented the 'mobility' of the neighborhood population.

[5] Using multilevel analysis Van Goor et al. (2005) found that the percentage of variance explained at the neighborhood level was a negligible 0.4% of the variance of overall nonresponse.

A different approach used by Schrapler et al. (2010) consists of acquiring variables from market research vendors. The authors linked additional (commercial) data on the immediate vicinity of the households from the MOSAIC data system to the German Socio-Economic Panel Study (SOEP).[6] The linked information is not necessarily in line with the specific reality of a particular household in the gross sample but it is used as an approximation for 'potential' respondent and neighborhood characteristics (Schrapler et al., 2010, 21).[7] Consistent with past research, the authors found high refusals in urban places. They also found higher cooperation in neighborhoods characterized by 'high earners', 'self-employed people' and 'households in high-quality new houses'.[8]

The research to date on the effect of area-level mechanisms on household cooperation presents three shortcomings. First, it cannot disentangle the area-level effect of socio-economic composition from that of the social environment. Second,

---

[6] The MOSAIC data system contains over 75 indicators with neighborhood characteristics. These data are normally used to analyze and describe 'customer databases' or 'markets'. This information is available at the address level and contains approximately 17.8 million buildings in Germany.

[7] Variables used as proxies for the 'potential respondent' in the Schrapler et al. (2010) study include: age, a lifestyle segmentation indicator, social status, purchasing power, and east-Germany indicator. Variables used as proxies for the 'neighborhood' include: city, number of households, household classification, size of buildings, percentage of foreigners, frequency of moves, affinity for gardens and desire for anonymity.

[8] Variables in the MOSAIC dataset are derived from cluster analyses and probability models. The probability model is developed on the basis of a 'calibration sample' by Sinus Sociovision (see http://www.sociovision.com for details).

the effect of the physical environment has not been formally included in theories of survey participation. And third, it has not clearly identified whether different levels of geography may be relevant for different area-level mechanisms influencing the survey response process. The next section introduces some insights from urban sociology that could help inform better studies in the survey literature.

2.2.2 Lessons from Urban Sociology

In the past, survey researchers were influenced by the social disorganization framework to understand part of the urbanicity effect on household cooperation. In the last decades, new insights from urban sociology have helped to improve the conceptualization and measurement of some of these complex social constructs better. I first briefly review here the concept of neighborhood. I then review two aspects of the neighborhood social and physical environment that could help inform area-level theories of survey participation.

Neighborhoods in the Urban Context

'Neighborhood' is a relatively flexible and amorphous concept which is generally defined spatially (Pebley and Sastry, 2003, 9). Early social ecologists saw urban neighborhoods as organic or 'natural' entities created as a result of the isolation of small geographic areas by physical barriers, such as railroads, rivers, and boulevards (Burgess, 1930) and/or through competition over land for residential and commercial use (Sampson et al., 2002). In practice, however, most social scientist in the United States rely on geographical boundaries defined by the U. S. Census Bureau

14

or other administrative agencies (Sampson et al., 2002). Alternatives to the use of administrative units include the use of resident's own definition of neighborhood boundaries using either maps (Coulton et al., 2001; Lee and Campbell, 1997; Guest and Lee, 1984) or average travel times (Newsome et al., 1998; Rindfuss et al., 2002; Crawford, 2002).

So what definition of neighborhood should be used to study area-level effects on individual outcomes? Pebley and Sastry (2003, 14) conclude that it is unlikely that a single set of neighborhood boundaries will be adequate to describe the individual's experience of neighborhood life. Not only do definitions of boundaries vary among individuals living in the same block (Lee et al., 1991; Guest and Lee, 1984; Logan and Collver, 1983; Coulton et al., 2001), they also may vary for a single household or individual over time (e.g., as children age) and may depend on context.

For the purpose of my research I will define neighborhoods by their Census tract boundaries, since these are the most ubiquitous definition in the urban sociology literature. Census tracts are also meaningful entities in survey sampling research. In the United States, Census tracts, blocks, and block groups are the units most often used in sampling (within primary sampling units) for household surveys (Dever et al., 2011).

Census tracts are small, relatively permanent statistical subdivisions of a county (see Figure 2.2). Tracts are delineated by a local committee of Census data users for the purpose of presenting data. Census tract boundaries normally follow visible features, but may follow governmental unit boundaries and other non-visible features in some instances; they always nest within counties. They are designed

to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions. Census tracts average about $4,000$ inhabitants (http://ask.census.gov/).
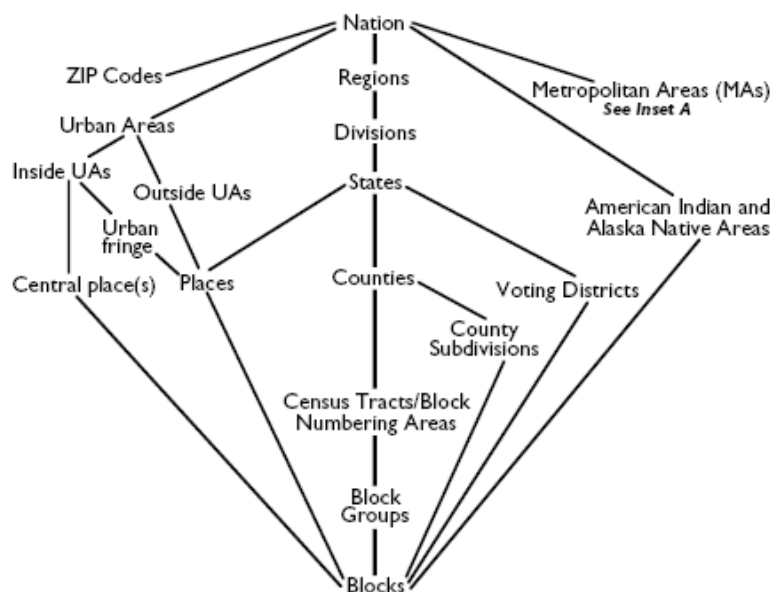
Fig. 2.2: Geographic hierarchy of units defined by the U.S. Census Bureau

Renewed Interest in Neighborhood Effects

Most of the early literature of neighborhood effects on individual outcomes relied almost exclusively on census demographic data to approximate the ecological constructs of interest. This approach was heavily criticized at the beginning of the 1990's (Jencks and Mayer, 1990; Mayer and Jencks, 1989). At the center of the problem was the realization that, if truly ecological mechanisms exist, these should be dynamic processes hypothesized to affect individual outcomes beyond the influence of structural factors like neighborhood socio-demographic composition (Jencks

and Mayer, 1990; Sampson et al., 1999). As Mayer and Jencks (1989) elaborates for child-related outcomes *"Why, for example, should concentrated poverty (which is, after all, the concentration of poor 'people') matter?. If neighborhood effects on child outcomes exists, presumably they are constituted from social processes that involve collective aspects of community life"*.

Following the advice of Jencks and Mayer, a new generation of social scientists has made enormous progress by (1) developing measurements of the intervening social processes evoked by ecological theories, and (2) by moving the scope of inquiry beyond crime and delinquency to outcomes relating to health and well-being. As a result of these developments, it is now possible to find specific survey instruments (e.g., walkability assessments) developed for particular outcomes (e.g., cardiovascular problems – see Gauvin et al. (2005)).

The downside of this renewed interest in the topic is the increasing number of constructs and indicators of neighborhood social processes that have emerged. Many of these indicators are highly correlated, raising the question of how many independent and valid constructs there really are (Cook et al., 1997; Furstenberg et al., 1999; Sampson et al., 1999). In other words, is there only one higher-order social process (e.g. social disorganization), or are there multiple subdimensions (e.g. social ties, social control)?

Neighborhood Social Environment – Collective Efficacy

A meta-analysis by Sampson et al. (2002) identified four classes of neighborhood social mechanisms, that, while related, appeared to have independent validity:

social ties/interaction, norms and collective efficacy, institutional resources, and *routine activities.* I focus here only on collective efficacy because I believe it is the process most likely to influence the household participation decision.

Collective efficacy has been defined as *'social cohesion among neighbors combined with their willingness to intervene on behalf of the common good'* (Sampson et al., 1997). Neighborhoods are high in collective efficacy when residents trust each other, share common values, and are willing to intervene on each other's behalf – for example, in supervising children and protecting public order. Sampson et al. (1997) developed measures of neighborhood collective efficacy based on reports from local residents on the capacity for 'informal social control' and 'social cohesion and trust' among neighbors.

Informal social control refers to the capacity of a group to regulate its members according to desired principles. One central goal would be the desire of community residents to live in safe and orderly environments that are free of predatory crime, especially interpersonal violence (Sampson et al., 1997). Examples of informal social control include the monitoring of spontaneous play groups among children, a willingness to intervene to prevent acts such as truancy and street corner loitering by teenage peer groups, and the confrontation of persons who are exploiting or disturbing public space.

At the neighborhood level, the willingness of local residents to intervene for the common good depends in large part on conditions of mutual trust and solidarity among neighbors. Sampson and colleagues notice that one is unlikely to intervene in a neighborhood context in which the rules are unclear and people distrust or fear

one another. They suggest that *socially cohesive* neighborhoods will prove the most fertile contexts for the realization of *informal social control.* In sum, the linkage of mutual trust and the willingness to intervene for the common good is what defines the neighborhood context of collective efficacy.

Collective efficacy has been defined as existing relative to the tasks of supervising children and maintaining the public order, but it generalizes to broader issues of importance to the well-being of neighborhoods. Collective efficacy has been found to predict partner violence, low rates of early sexual initiation, obesity in adolescents, mental health, and all-cause mortality and mortality from cardiovascular disease (Browning, 2002; Cohen et al., 003a; Lochner et al., 2003; Browning et al., 2005; Xue et al., 2005; Cohen et al., 2006; Araya et al., 2006). Section 2.2.3 elaborates on how collective efficacy could influence the survey participation decision.

Neighborhood Physical Environment – Disorder and Decay

A large body of literature has also focused in the influence of the physical (built) environment on crime and victimization outcomes such as 'fear of crime' and 'risk perception'. Research on neighborhoods antecedents of fear of crime have focused mostly around the 'Broken-Windows Theory' – a theoretical framework linking neighborhood *disorder* and residential *decay* to individual-level fear of crime and risk perception (Garofalo and Laub, 1978; Wilson and Kelling, 1982; Skogan, 1990). Public signs of disorder and decay are hypothesized to become important symbols that residents and others cannot or will not protect their neighborhood from *crime* and *fear.* It is further hypothesized that residents react to the symbolism of

19

these incivilities by withdrawing from social activity in the neighborhood.

Well known versions of the broken-windows thesis assert that areas with visible signs of *disorder* and *physical decay*, in the long run, will suffer from serious crime and a downward spiral of urban decay (Wilson and Kelling, 1982; Skogan, 1990). A somewhat different perspective, put forward by Sampson and Raudenbush (1999, 608), suggests that public *disorder* and predatory *crimes* are manifestations of the same explanatory process, albeit at different ends of a 'seriousness' continuum. For example, even those elements of disorder not obviously criminal in nature (e.g., garbage, vacant housing) are either violations of an ordinance (as in littering, slumlord abandonment) or may be conceptualized as sharing a similar causal structure and thus predicted by similar mechanisms (Hunter, 1985).

Sampson and Raudenbush (1999) suggest that what makes this conceptual move significant is that it provides the opportunity to observe - and hence systematically measure - important manifestations of crime-related processes. Muggings, assaults, and rapes might be impossible to observe reliably, but vandalism, prostitution, gang congregation, and evidence of drug use can, in principle, be observed by all, whether residents, business people, visitors, possible investors, local activists, or potential offenders.

Neighborhood incivilities have been found to correlate with individual-level outcomes like depression, psychological distress, perceived powerlessness, child and adolescent mental health, physical function in the elderly, psychological well-being, physical activity and smoking, and mortality among other health related outcomes (Cohen et al., 2000; Ross et al., 2000; Ross, 2000; Sampson et al., 2002; Caughy

et al., 2003). The next section elaborates on how incivilities could be linked to the survey participation decision.

2.2.3 The Conceptual Model

After reviewing both the survey literature and the urban sociology literature, a revision of the Groves and Couper model of survey participation seems warranted. First, to make explicit the relationship between the concepts of neighborhood socioeconomic composition and neighborhood social environment. And second, to formally integrate hypotheses about how the characteristics of the physical environment may influence the household cooperation decision.

Hypotheses about the Social Environment

The Groves and Couper model of survey participation emphasizes the role of *social cohesion* on the household cooperation decision (Groves and Couper, 1998, 177). The role of *informal social control*, however has not been similarly acknowledged. *Informal social control* is the component of agency in the collective efficacy construct. Since participation in social surveys has also been interpreted as a form of community involvement (Couper et al., 1998; Abraham et al., 2006), I hypothesize that the neighborhood *informal social control* will play an important role in explaining the householder decision to cooperate in a household survey.

When promoting the survey request, for example, survey interviewers usually advertise the survey as an opportunity 'to inform policy making', 'to let your authorities know about the issues your community cares about', 'to make your voice

21

heard' and the like. To the extent that surveys – especially those conducted by government agencies or academic institutions – are perceived as a vehicle towards achieving a 'common good' or 'benefit the society at large', I hypothesize that:

Hypothesis 1 : Neighborhood *informal social control* will increase the likelihood of cooperation among the contacted households.

Hypothesis 2 : Neighborhood *social cohesion* will increase the likelihood of cooperation among the contacted households.

It is also important to acknowledge that the effect of the social environment does not occur in a vacuum. Thus we cannot ignore the effect of the neighborhood social fabric. As Sampson and Raudenbush (1999, 613) point out *'a theory of collective efficacy does not render neighborhood structural constrains irrelevant, rather it proposes mediating mechanisms while at the same time insisting on an independent role for agency in all corners of the social structure'*. In the context of the survey request, I thus hypothesize that:

Hypothesis 3 : Neighborhood *informal social control* and *social cohesion* will mediate the effect of neighborhood socio-economic composition on household cooperation.

Hypotheses about the Physical Environment

Characteristics of the physical environment have not been formally incorporated in the Groves and Couper model of survey participation. However, insights from urban sociology suggest they should, since measures of *disorder* and *decay* at

the neighborhood level have been shown to influence residents' fear of crime and risk perception – two household level mechanisms that are expected to reduce cooperation with the survey request.

Figure 2.3 presents a revised version of the Groves and Couper model of survey participation shown in Figure 2.1. The revised version formally incorporates the influence of the physical environment on both householders and interviewers. For the purpose of this paper, however, I only elaborate on the mechanisms underlying the relationships on the left side of the diagram – the neighborhoods and the household(ers).
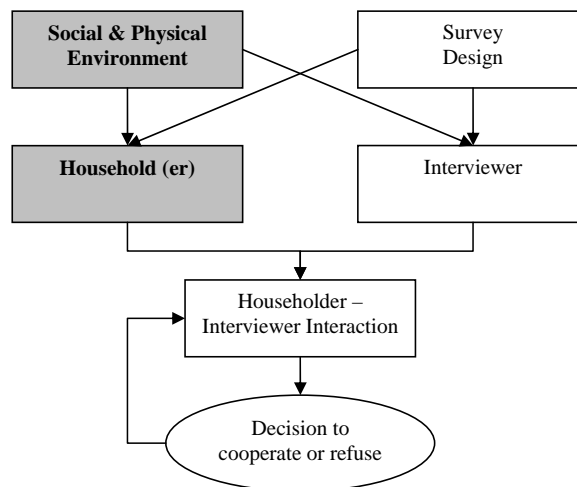
Fig. 2.3: Adapted Groves-Couper model of Participation in Household Surveys.

In the survey context, respondents living in a neighborhood with signs of incivilities may try to avoid long interactions with strangers, which may result in higher refusals (nonresponse) or break-offs (partial response). Interviewers too may try to reduce the amount of time spent in a dangerous neighborhood and thus will be more likely to accept refusals in such neighborhoods. These influences on

respondents' and interviewers' behaviors will likely reduce the cooperation with the survey request. Hypotheses about the influence of the physical environment on the household(er) are two:

Hypothesis 4 : Neighborhood *disorder* will reduce the likelihood of cooperation among the contacted households

Hypothesis 5 : Neighborhood *residential decay* will reduce the likelihood of cooperation among the contacted households

Figure 2.4 provides a 'close-up' on the mechanisms, within the urban context, linking neighborhood characteristics and the survey participation decision.

The box at the bottom depicts the household-level influences on cooperation. The participation decision is hypothesized to be mostly influenced by individual psychosocial characteristics, however household demographic characteristics are also hypothesized to play a role (Groves and Couper, 1998, 32).

The box at the top depicts the neighborhood-level influences on person-level psychosocial behavior. The measures of *population density, crime* and neighborhood *socio-economic composition* correspond to the three area-level predictors traditionally used in response propensity models of household cooperation. The measures of the social environment are now explicitly included in the model for the purpose of disentangling the effect of neighborhood social composition on household cooperation. Finally, the measures of the physical environment are added to the model to test the possibility of direct effect of these correlates of crime on household cooperation.
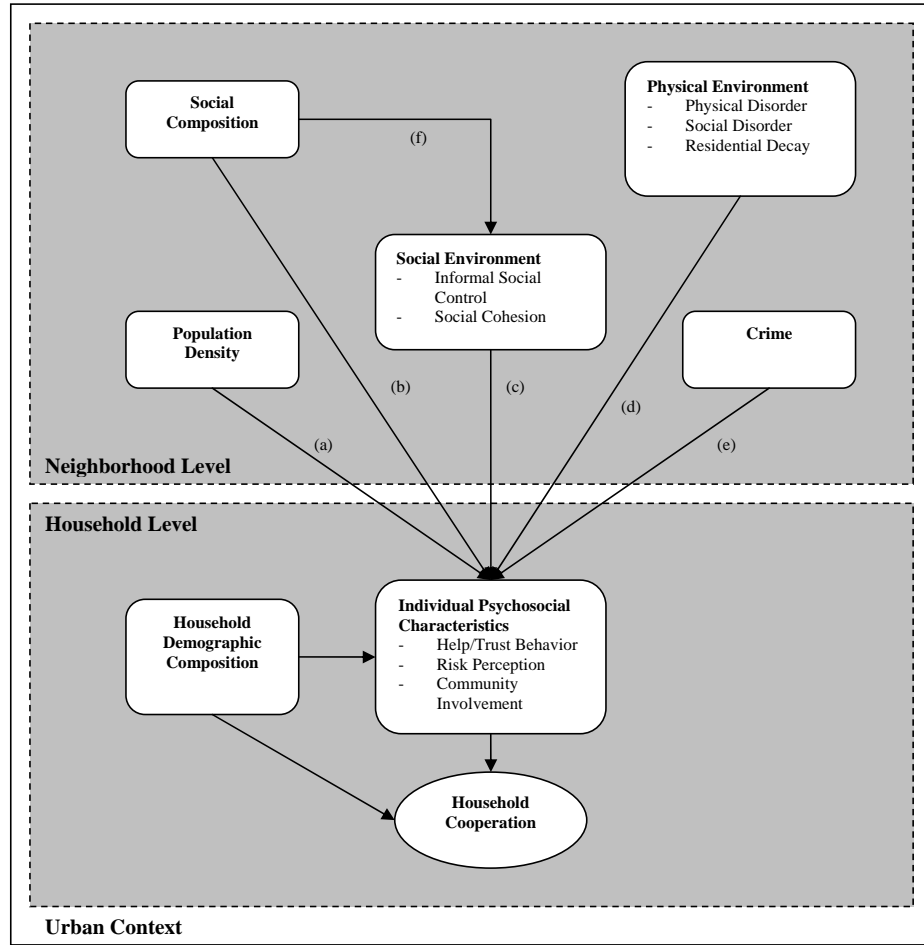
Fig. 2.4: Conceptual Model for Household Cooperation in the Urban Context.

To conclude, the purpose of this paper is to disentangle the direct effect of the different neighborhood mechanisms affecting the household cooperation decision, i.e. to provide estimates of the effects illustrated by arrows (a)-(e).

In this paper I will use multilevel analyses to develop a model of survey participation with predictors at the two levels illustrated in Figure 2.4 – neighborhoods and households. For the purpose of my analysis, neighborhoods will be defined by the boundaries given by the U.S. Census tracts because they are the most ubiquitous area-level unit used in studies of neighborhood effects on individual-level outcomes.

2.3 Data and Methods

The Los Angeles Family and Neighborhood Study (L.A. FANS) is an ideal data set with which to test hypotheses about neighborhood mechanisms. The study collected theoretically driven measures of neighborhood *social* and *physical* environments. In addition, it collected a small set of household characteristics for both respondents and nonrespondents. The specific nature of the L.A. FANS data allows the examination of these processes not just for a simple survey participation request, but for different stages of such a process (screening, rostering, and the survey interview). It also allows the assessment of the relative influence of different levels of geography (e.g. blocks and tracts) on these corresponding response outcomes. A brief description of the study design, the data sources and the analyses methods are presented in this section.

2.3.1 Study Design

The L.A. FANS study is a study of families in Los Angeles County and the neighborhoods in which they live. The study presents a complex design based on a multi-stage probability sample of tracts, households, and individuals in Los Angeles County. The sample was selected in two-phases (Kish, 1965). The first phase involved the selection of 65 Census tracts within three poverty strata.[9] Tracts in

---

[9] Prior to sampling, census tracts were divided into three strata based on the percent of the tract's population in poverty. The sampling strata in the L.A. FANS design correspond to tracts that were very poor (those in the top 10% of the poverty distribution), poor (tracts in the 60-89th percentiles), and non-poor (tracts in the bottom 60% of the distribution). Tract-level estimates

the 'poor' and 'very poor' stratum were oversampled at this first phase.

Within the selected tracts, a preliminary sample of approximately $9,400$ addresses was drawn to complete a screener interview, which consisted on answering a single question about the presence of children in the household. Among those successfully screened, a sub-sample of approximately $4,100$ households was selected for the L.A. FANS sample. Households with children were oversampled at this second phase.

Households selected for the sample were first asked to complete a roster interview. The roster was administered to an adult resident and collected information about all household members, including how they were related to each other as well as the basic demographic and social characteristics for each person. Household rosters were completed with $3,083$ households and individual interviews were completed with approximately 85% of the selected respondents. Figure 2.5 presents an overview of the outcomes of the sampling and the data collection strategy at the household-level. See Sastry and Pebley (2003) for details on the outcomes for the different types of L.A. FANS respondents.

2.3.2 Data and Variables

Variables used to develop the analytic variables in this paper were assembled from different data sets available in the L.A. FANS study. I briefly describe here the development of the dependent and independent variable used in the response
_____

of percent in poverty in 1997 were developed by Los Angeles County's Urban Research Division using state and county administrative data. See Sastry et al. (2003) for more details.
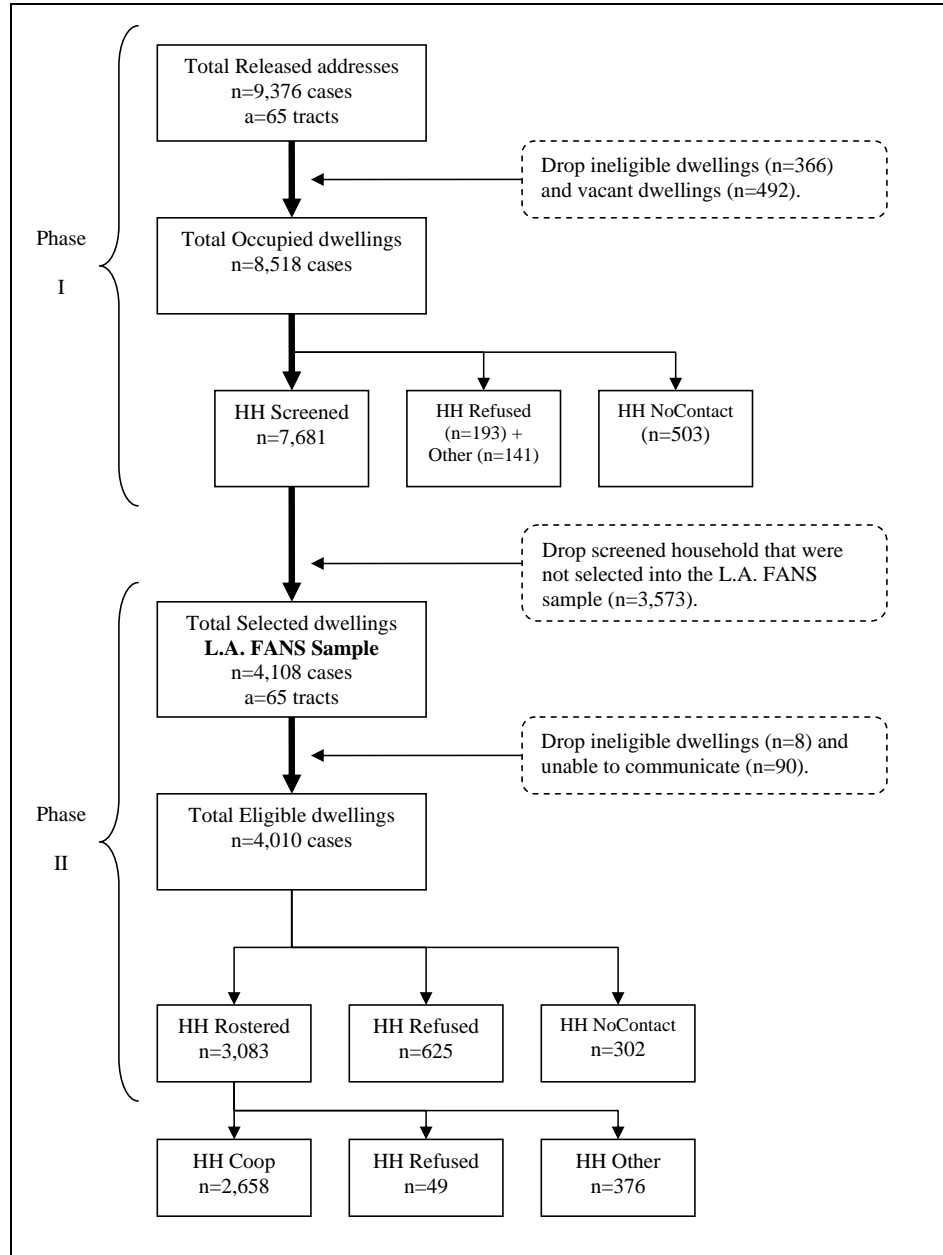
Fig. 2.5: Final disposition of L.A. FANS sampled cases for the two main phases of data collection.

propensity models used here.

Indicator of Household Cooperation

The staged design of the L.A. FANS recruitment process allows for the development of different types of dependent variables reflecting: (a) the three different stages of the survey request (screener, roster, and household interview) and (b) the two types of response processes under study: contactability and cooperation (conditional on contact). Overall, response rates for the L.A. FANS study were high and match up favorably to other major surveys of similar populations (Sastry et al., 2003).[10] However, recruitment efforts yielded somewhat variable results across the 65 neighborhoods in the sample. Table 2.1 displays descriptive statistics for the tract-level response rates for these six outcomes.

To test the hypotheses about household cooperation, however, the most appropriate dependent variable is cooperation at the roster interview stage. The binary variable reflecting **roster cooperation** takes on a value of 1 if the household completed the *roster interview* and 0 if not, and it is available for the $3,708$ households contacted for the roster interview. There are two reasons to focus on roster cooperation. First, it is the response outcome with larger variability across the 65 census tracts – with the lowest cooperation rate of 76% and the largest cooperation rate of 98%. The second and most important reason is that – in the context of the L.A. FANS study – roster cooperation resembles better a traditional survey request in terms of 'burden' and 'commitment'.

---

[10] Sastry et al. (2003) report comparably-defined response rates for the NLSY-97 (92%), the 1994-95 baseline wave of ADDHealth (79%), the 1997 PSID Child Supplement (88%), and the 1999 baseline wave of the Welfare, Children and Families Study (74%).

The relatively low burden of the screener interview - which consists on a single question about the presence of children – clearly disqualifies the cooperation at the screener interview as the focal dependent variable for this study. The relatively high level of commitment at the household interview stage, on the other hand, clearly disqualifies the household interview too. The request for the household interview happens after both the screener and the roster interview have been granted. At this stage, I argue, the mechanisms underlying cooperation may more closely resemble the request of a panel survey rather than a cross sectional survey.

In this study roster cooperation (conditional on contact) will be the focal dependent variable used to test the hypotheses derived from the conceptual model of survey cooperation presented in section 2.2.3. Table 2.1 displays the tract level response rates for the different stages of interviewing.

Tab. 2.1: L.A. FANS Tract Level Response Rates (n=65). Unweighted estimates.

Rates	Average	Min	Max
Screener Interview			
Contactability Rate	96.0	67.5	100.0
Cooperation Rate	96.1	85.5	100.0
Roster Interview			
Contactability Rate	92.6	68.2	100.0
Cooperation Rate	83.6	66.7	98.0
Household Interview			
Contactability Rate	88.0	75.5	100.0
Cooperation Rate	98.2	87.2	100.0

L.A. FANS interviewers collected observations on a few household characteristics for both respondents and nonrespondents to the screener interview. Household observations included *type of housing unit, estimate of rent* and *presence of children.* Respondent observations included *race, age* and *gender.* Interviewers also recorded the *language of the interview.*

Data based on interviewer observations, however, was missing for up to 30% of the variables in the cases attempted for the roster interview. This high percentage of missing data is most likely not missing at random, thus analyzing only the cases with complete data on these covariates is not a viable option. To overcome this problem the L.A. FANS team developed multiple-imputed data sets for the screener interview.[11] I used data from one of these files for the analysis of response propensity, in practice, treating the imputed values as if they were true values. I do not expect that using imputed data at the household level will affect the inferences regarding the neighborhood level mechanisms. Household-level variables are used as merely 'controls' in the response propensity models, where the hypotheses of interest involve the main effects of different types of neighborhood-level characteristics.

Neighborhood level predictors used in this paper were developed based on data from Census records, respondents' judgements, and interviewer observation of the physical environment. A detailed description on the construction of these measures

---

[11] Five imputed data sets were developed by sequential regression modeling (Raghunathan et al., 2001) using IVEware. Access to these multiple imputed files was granted via a special request to the L.A. FANS team.

is presented next.

Neighborhood Social Environment

Data on the neighborhood social environment was derived from questions included in the Adult Module of the L.A. FANS study. The ADULT1 file contains $3,557$ cases which correspond to the two types of adults responding the survey: the Randomly Selected Adult (RSA) and the Primary Care Giver (PCG). For my analyses I only used the records from the $2,619$ randomly selected adults in the sample.[12] I followed a four-step procedure to develop tract-level measures from the person-level responses. First, cases with missing values on all items in a given scale were dropped. Second, cases with missing values on some items were imputed using the mean response for the missing item from the entire sample. Third, mean scales were calculated for each individual from his/her responses to all the items in the corresponding scale. And fourth, individual-level responses were averaged within each census tract to create neighborhood-level responses.

The **social cohesion** scale consists of five statements about perceived neighborhood cohesion (how close knit is neighborhood; willing to help neighbors; neighbors get along; neighbors share values; and neighbors can be trusted). The tract-level measure developed from the respondents' reports is a continuous variable from

---

[12] In households with children, the Adult Module could have potentially be completed for up to two adults in cases when the randomly selected adult (RSA) and the primary care giver (PCG) were not the same person. To avoid duplicate reports from these households with children, I only used the response from the RSA. These cases include those cases where the RSA and the PCG are the same person.

1 to 5 where values closer to 1 mean 'low social cohesion' and values closer to 5 mean 'high social cohesion'. The **informal social control** scale consists of three statements about how likely it is for neighbors to do something if kids are hanging out; kids are painting graffiti; or kids disrespect adults. The tract-level measure is a continuous variable from 1 to 5 where values closer to 1 mean 'low informal social control' and values closer to 5 mean 'high informal social control'. Descriptive statistics for the two neighborhood measures are displayed in Table 2.3. Descriptive statistics for the eight person-level items are available in the Appendix.

Neighborhood Physical Environment

Assessments of the physical environment were completed by a small team of L.A. FANS trained interviewers, who provided multiple independent ratings on observable characteristics of the 65 L.A. FANS neighborhoods. For the analyses of household cooperation I only used the ratings recorded by the first interviewer completing observations in each Census block. I followed this strategy because it more closely resembles the situation of a typical data collection effort where it is unlikely to send multiple interviewers to collect multiple ratings per block. The file containing multiple observations per block had $5,966$ cases. The file containing only the assessments made by the first interviewer had $2,040$ cases, which correspond to 422 blocks rated by 30 interviewers across the 65 census tracts. Each interviewer worked in 1-22 tracts and each tract was assessed by 1-5 interviewers.

To develop tract-level measures from the block face-level responses I followed a similar four-step procedure described for the measures of the social environment.

An additional step involved dichotomizing the original Likert score before averaging within tracts. This step is typically done in studies of neighborhood disorder since the prevalence of some indicators of disorder is very low (Sampson et al., 1999).

The **physical disorder** scale consists of eight statements about perceived signs of physical disorder on the streets, alleys, lots, walls or sidewalks (abandoned cars; trash or junk; garbage or litter; needles or drug-paraphernalia; empty beer/liquor bottles; discarded cigars buts; graffiti on walls or buildings; and painted-over graffiti). The tract-level measure represents the percentage of 'physical disorder' observed in the tract. The **social disorder** scale consists of seven statements about perceived signs of social disorder on the streets (gangs; prostitutes; people selling drugs; people drinking; homeless people; and adults loitering). The tract-level measure represents the percentage of 'social disorder' observed in the tract. The **residential decay** scale consists of five statements about perceived signs of residential decay on the street, alleys, lots, walls or sidewalks (residential buildings bad condition; houses/appts boarded up; vacant lots; houses/appts damaged walls; and houses/appts without well tended yards). The tract-level measure represents the percentage of 'residential decay' observed in the tract. Descriptive statistics for the three neighborhood measures are displayed in Table 2.3. Descriptive statistics for the twenty block face-level items are available in the Appendix.

Neighborhood Socio-Economic Composition

The L.A. FANS assembled a wealth of aggregate-level data from the 1990 and 2000 Decennial Census, including variables constructed by the L.A. FANS team

34

(Peterson et al., 2007). In the spirit of Goyder et al. (1992) I used factor analysis to help reduce the dimensionality of the data and to develop measures that better reflect the social structure of the Los Angeles neighborhoods.[13] I included sixteen variables from the 2000 Census that have been used in contemporary neighborhood research on child-developmental outcomes (Sampson et al., 1999).

Analyses of screeplots and Eigenvalues (> 1) suggested that only the first three factors merited retention. I decided to retain the first five factors to try to parallel the factors recovered by Sampson et al. (1999). Table 2.2 displays the factor loadings greater than 0.6 and the uniqueness for the five dimensions representing the neighborhood socio-economic composition in the L.A. FANS sample.

The first factor extracted had an eigenvalue greater than 9 and was dominated by high positive loadings for *Spanish speakers* and *Hispanic origin* and high negative loadings for *higher education*, *high income*, and *executive/professional occupation*. After recoding the scores by -1, the interpretation of this factor seems to revolve around **concentrated affluence**. With eigenvalue greater than 2 and high positive loadings for percentage of *foreign born*, and *non-citizens*, and negative loadings for *owner-occupied*, the second dimension clearly captures the degree of **immigrant concentration**. The predominant interpretation for the third factor

---

[13] I used the principal-factor method to analyze the correlation matrix. Under this method the factor loadings are computed using the squared multiple correlations as estimates of the communality. I rotated the factors using the Varimax (orthogonal) rotation method. I used the regression method to predict the factor scores. I implemented the factor analysis using the *factor* procedure in Stata 10. I implemented the factor analysis on the 65 census tracts in the sample.

is **concentrated disadvantage**. This factor had an eigenvalue larger than 1.5 and loaded primarily on four variables: percentage *under poverty*, on *public assistance*, *female head-of-household*, and *black residents*. The fourth factor had an eigenvalue of 0.43 and was interpreted as **family structure**. The two variables with high loadings were *households with children* (positive loading) and *non-family households* (negative loadings). The fifth factor extracted had an eigenvalue of 0.36 and was dominated by high positive loadings on *percent unemployed* and negative loadings on *percentage same occupancy since 1995*. After recoding the scores by -1 the interpretation seems to revolve around the concept of **residential stability**. Descriptive statistics for the five neighborhood measures are displayed in Table 2.3. Descriptive statistics for the sixteen tract-level items are available in the Appendix.

Population Density and Crime

The tract-level measure of **population density** was readily available in the L.A. FANS dataset. It was computed as the *total number of persons* per tract divided by *total number of square miles* in each tract. I developed a tract-level measure of **crime** based on the following question from the Adult Module: *'While you have lived in this neighborhood, have you or anyone in your household had anything stolen or damaged inside or outside your home, including your cars or vehicles parked on the street?'*. The tract-level measure represents the percentage of households victimized in the tract. Descriptive statistics for the neighborhood measures are displayed in Table 2.3.

Tab. 2.2: Factor Loadings and Uniqueness for Exploratory Factor Analysis on Census Variables (n=2,619). Unweighted estimates.

Variable	Factor Loadings	Uniqueness
Concentrated Affluence (before reverse coding)		
Perc. residents 25 years or older with a college education	-0.8255	0.0091
Perc. households with a high income	-0.7139	0.0285
Perc. residents who are executives or professionals	-0.7778	0.0225
Perc. of adults Spanish speakers	0.8868	0.0064
Perc. Hispanic	0.9148	0.0021
Immigrant Concentration		
Perc. of population foreign-born	0.8977	0.0337
Perc. population non-citizens	0.8085	0.0146
Perc. of occupied housing owner-occupied	-0.6640	0.0411
Concentrated Disadvantage		
Perc. residents living below the poverty line	0.5748	0.0579
Perc. households on public assistance	0.7896	0.0950
Perc. female-headed households	0.6710	0.1334
Perc. black residents	0.6672	0.3411
Family Structure		
Perc. non-family households	-0.8435	0.0446
Perc. households with children	0.6425	0.0366
Residential Stability (before reverse coding)		
Perc. occupying same dwelling as in 1995	-0.7289	0.2546
Perc. individuals on unemployment	0.6087	0.3590

Tab. 2.3: Descriptive Statistics for Neighborhood Characteristics (n=65). Unweighted estimates.

Variable	Mean	Std. Dev.	N
Social Environment			
Social Cohesion	3.38	0.31	65
Informal Social Control	3.56	0.36	65
Physical Environment[†]			
Physical Disorder	0.42	0.18	65
Social Disorder	0.03	0.03	65
Residential Decay	0.52	0.16	65
Social Composition			
Concentrated Affluence	0.00	0.99	65
Residential Stability	0.00	0.88	65
Family Structure	0.00	0.96	65
Concentrated Disadvantage	0.00	0.95	65
Immigrant Concentration	0.00	0.98	65
Other Mechanisms			
Crime	0.43	0.11	65
Population Density	14,836	10,462	65

(†):Estimates based on the ratings of a single interviewer per block ($n = 2,040$).

2.3.3 Statistical Methods

Preliminary evidence of association between neighborhood characteristics and the survey response process was analyzed using correlational analyses at the tract level. A stronger test of these associations was developed using a multilevel model described in detail in this section. Additional sensitivity analyses test the robustness of the main results are also described here.

Multilevel Analyses of Survey Participation

The goal of the multivariate analyses is to test the validity of the conceptual model depicted in Figure 2.4. This model can be described as a two-level logistic regression model, where the first level corresponds to households and the second level corresponds to neighborhoods. Let y_{ik} be the indicator of *roster cooperation* status taking on a value of 1 if household i in neighborhood k cooperated with the roster interview, and $y_{ik} = 0$ if it refused; and let μ_{ik} denote the probability $y_{ik} = 1$, that is,

$$y_{ik} \mid \mu_{ik} \sim Bernoulli \tag{2.1}$$

$$E(y_{ik} \mid \mu_{ik}) = \mu_{ik}$$

$$Var(y_{ik} \mid \mu_{ik}) = \mu_{ik}(1 - \mu_{ik})$$

As is standard in logistic regression, we define η_{ik} as the log-odds of the probability of having completed the roster interview,

$$\eta_{ik} \equiv \log\left(\frac{\mu_{ik}}{1 - \mu_{ik}}\right) \tag{2.2}$$

The structural model at the lower level accounts for predictable variation within neighborhood across households,

$$\eta_{ik} = \pi_k + \sum_{p=1}^{P} \alpha_p X_{pik} \qquad (2.3)$$

Where π_k reflects the specific (constant) effect of neighborhood k on roster co-operation; X_{pik} are variables measuring P household-level characteristics observed during the screener interview; and α_p are the associated regression coefficients of the household-level variables. It is important to control for household-level covariates since they are hypothesized to be strong correlates of household cooperation.

The second level of the model describes variation between neighborhoods around the grand mean of roster cooperation status:

$$\pi_k = \theta + \sum_{q=1}^{Q} \beta_q W_{qk} + v_k \qquad (2.4)$$

$$v_k \sim IID(0, \psi_v^2)$$

Where θ is the grand mean level of neighborhood roster cooperation in the sample; W_{qk} correspond to the Q neighborhood-level variables measuring neighborhood characteristics; and β_q are the associated regression coefficients of the neighborhood variables. Random variability among tracts is modeled by v_k, assumed to be Independent and Identically Distributed (IID) with mean zero and variance ψ_v^2.[14]

Equation 2.4 implies that only the intercept varies randomly between neigh-

---

[14] Some nonresponse studies have modeled area effects as fixed effects (Groves and Couper, 1998). This modeling strategy is not appropriate to test area-level effects since, by definition, the fixed effects model 'explains' all differences between areas and there is no unexplained between-area variability that could further be explained by area-level variables (Snijders and Bosker, 1999, 43).

borhoods. I did not include random coefficients in my response propensity model because I have no evidence that suggests the need for random slopes. That is, I did not expect households in different tracts to react differently to the same neighborhood influences. Combining equations 2.3-2.4 the goal is to estimate the full *two-level logistic random intercept* model (Raudenbush and Bryk, 2002):

$$\eta_{ik} = \theta + \sum_{p=1}^{P} \alpha_p X_{pik} + \sum_{q=1}^{Q} \beta_q W_{qk} + v_k \tag{2.5}$$

$$v_k \sim IID(0, \psi_v^2)$$

This full model was developed in stages to assess the contribution of different types of predictors. The first model will contained only an indicator for the tract-level random effect (v_k) and no fixed effects. The purpose of this unconditional model was to assess the magnitude of tract-level effects on household cooperation by assessing the percentage of variance explained at the tract level. The second model added household-level predictors. The following models successively added the traditional Census derived predictors and the predictors from the social and physical environments. To facilitate comparison across the different models all neighborhood level variables were standardized.[15] Measures of model fit like the AIC and the BIC criteria were used to compare model performance across different (nested) models.

Accounting for the Complex Survey Design

Under a model-based framework, the effects of the sampling design variables on the estimation of the standard errors can be directly incorporated into the sub-

---

[15] Each variable was first centered around its mean and then divided by its standard deviation. The transformed variable had a mean of 0 and a standard deviation of 1.

stantive model (Pfeffermann et al., 1998). Indicators of stratification variables for both phases of the sampling design were included in all the conditional models. Dummy variables were included to reflect the 'poverty' strata used to select tracts in Phase I, and the 'children status' strata used to select households in Phase II. The clustering effect due to the sampling of primary selection unites (PSUs) was modeled using tract-level random effects.

The use of standard multilevel modeling to estimate model (2.5) could typically lead to biased parameter estimates if the units under analyses are selected with unequal probabilities (Rabe-Hesketh and Skrondal, 2006). Several authors have discussed the use of sampling weights to rectify this problem in the context of two-level linear (or linear mixed) models, particularly random-intercept models (Pfeffermann et al., 1998). Thus, given that households were selected with unequal probabilities into the L.A. FANS sample, the estimation of the two-level model in (2.5) required the use of sampling probability weights.

In this paper I followed the approach by Rabe-Hesketh and Skrondal (2008) which incorporates the sample probability weights directly into the pseudo log-likelihood formula. This full pseudo-maximum-likelihood estimation procedure is available for generalized linear models with any number of levels via adaptive quadrature in the *gllamm* procedure in Stata (Rabe-Hesketh et al., 2002, 2004; Rabe-Hesketh and Skrondal, 2008), where the standard errors for the weighted models are estimated using a type of sandwich estimator (Skinner, 1989). To properly estimate the model in (2.5), however, the probability weights need to be available for both levels – tracts and households. In addition, it is recommended to re-scale the

level-1 weights to reduce bias in the estimates of the variance components.

At the time of first writing this paper, the L.A. FANS did not have separate weights for the tract and household levels. For estimation purposes, I developed an estimate of the two set of weights following the suggestions in Rabe-Hesketh and Skrondal (2006, 825). I assumed that the tract-level weights were equal to 1 ($wgt\_tr^* = 1$) and the household-level weights were equal to the L.A. FANS household sampling weight ($wgt\_hh^* = wgthh$). This näive approximation ignores the unequal selection probability at the tract-level and assumes that all the variability arises at the household-level. Given the high variability of the selection weights across the household ('children status') strata, this assumption seemed reasonable.

$$wgthh = wgt\_tr^* * wgt\_hh^* \tag{2.6}$$

These estimates of the tract-level and household-level probabilities were used in the estimation of the multilevel model in (2.5).[16] Analyses using alternative set of weights and alternative set of covariates are discussed below as part of the sensitivity analyses.[17]

---

[16] Following Rabe-Hesketh and Skrondal (2006, 813), the household-level weights were re-scaled before estimation using scaling method #1.

[17] Thanks to a special request, the L.A. FANS team was able to give me access to the separate set of weights associated with the tract-level probabilities (wgt_tr) and the household-level probabilities (wgt_hh). These two set of weights, when combined, can reproduce the single set of household-level weights available for the L.A. FANS study ($wgthh^{NTNS} = wgt\_tr * wgt\_hh$). For more details on the L.A. FANS sampling weights, see the description of the development of the L.A. FANS household weights in Paper 3.

The stability of the results from the multilevel analyses was tested by re-running the models under different specification of the sampling weights, using alternative definition of the Census predictors, and dropping the single tract that was found potentially influential. Using the same set of covariates as the main analysis in the paper, the full model was replicated:

- using an alternative scaling method for the household weights (ALT),[18]

- using the original household weights without scaling (UNS), and

- without weights (UNW).

Using the same set of weights as the main analysis in the paper, the full model was also replicated:

- using an alternative set of Census variables closely resembling those used by Groves and Couper (1998): *perc. owner*, *perc. pop. less 18yrs*, *perc. non-white*, *perc. multi-unit*, *crime* and *population density.*,

- using a combined measure of *collective efficacy* in place of the individual indicators of *social cohesion* and *informal social control*, and

- after dropping the tract found influential by univariate test for outliers.

A table with descriptive statistics for the different set of weights used in the sensitivity analyses is available in the Appendix.

---

[18] See Rabe-Hesketh and Skrondal (2006, 813) for details on scaling method #2.

2.4 Analysis and Results

This paper tests the conceptual model in Figure 2.4, which proposes several mechanisms that connect neighborhood characteristics and household cooperation. Preliminary evidence of associations used correlational analyses. Stronger tests of these associations used multilevel analyses. Results of the sensitivity analyses are finally discussed to evaluate the robustness of the findings.

2.4.1 Correlational Analyses

Each data point in Figures 2.6 and 2.7 corresponds to the estimate of the Pearson correlation coefficient between the corresponding tract-level response rate and each neighborhood characteristic under study. The gray area in the middle of the figures helps identify those estimates that were not significantly different from zero at the 5% level.

To facilitate the comparisons across the different types of neighborhood characteristics I assigned them different symbols. Variables used in traditional nonresponse studies are displayed using diamonds, while the indicators of the social and physical environment are displayed using circles. Tables with the point estimates and 95% confidence intervals are available in the Appendix.[19]

---

[19] Confidence intervals for the estimated correlations were approximated using Fisher's Transformation (Rosner, 2000, 461,462).

Analyses of Roster Cooperation Rates

Figure 2.6 shows that, as expected from the conceptual model, tract level cooperation rates are influenced by several neighborhood characteristics. Neighborhood cooperation rates were negatively associated with the indicators of *affluence* ($r = -0.28, p = 0.0236$), *social cohesion* ($r = -0.45, p = 0.0002$) and *owner-occupied* ($r = -0.42, p = 0.0006$). In contrast, cooperation rates were positively associated with the indicators of *immigrant concentration* ($r = 0.43, p = 0.0004$), *physical disorder* ($r = 0.49, p = 0.0000$), *social disorder* ($r = 0.41, p = 0.0009$), and *residential decay* ($r = 0.46, p = 0.0001$). To a lesser extent, cooperation rates were also positively associated with *population density* ($r = 0.34, p = 0.0068$), *multi-unit structure* ($r = 0.30, p = 0.0171$), *non-white* ($r = 0.30, p = 0.0173$) and *pop.< 18 yrs* ($r = 0.32, p = 0.0114$).

Focusing on the pattern of the association of the environmental variables two trends seem to emerge. The indicators of the social environment cluster together with a negative correlation to roster cooperation rates. In contrast, indicators of the physical environment cluster with a positive correlation. These bivariate results contradict the expectations of the conceptual model, which hypothesizes a positive effect of the social environment (Hypothesis #1 and #2) and a negative effect of neighborhood incivilities (Hypothesis #4 and #5). These results, however, should be taken with caution. Given the complexity and interconnectedness of the social phenomenon under study, it is possible that some of these results will change in the multivariate setting.
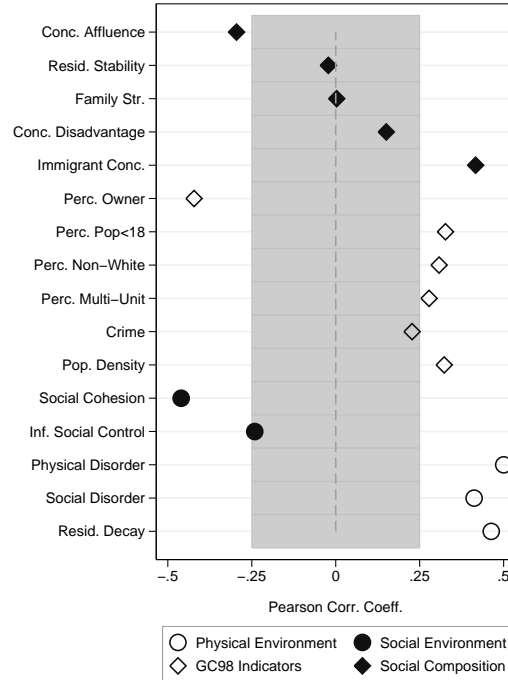
Fig. 2.6: Bivariate Pearson Correlation between Roster Cooperation Rates and Neighborhood Characteristics (n=65). Unweighted estimates.

Analyses of All Contact and Cooperation Rates

Figure 2.7 displays the correlational results for tract-level contact and cooperation rates at the three stages of recruitment. The graphs in the first row show that no tract-level effect seems to significantly influence the process of household contactability, although a few neighborhood characteristics are borderline significant at the screener stage. The graphs in the second row show that household cooperation is strongly influenced by neighborhood characteristics at the roster stage, but only marginally at the screener and household interview stage.

These results are consistent with our understanding of the L.A. FANS recruitment process. The screener interview represents the first opportunity of contact with

the selected household. It makes sense that some contextual effects would arise at this stage of the contact process and not at the later stages – when the availability of household-provided contact information is likely to render uninformative the 'influence' of the neighborhood context. Along the same lines, it is not surprising that the strongest effect of the neighborhood context on the cooperation process arises at the roster stage. Since, as argued before, the other interview stages do not reflect the typical cooperation request in terms of 'burden' and 'commitment'. Result from this bivariate analyses support the decision of using the roster stage for the multilevel analyses of household cooperation presented next.

2.4.2   Multilevel Analyses

Analyses presented in this section aimed to test the validity of the conceptual model for household cooperation depicted in Figure 2.4. Table 2.4 presents five multilevel logistic regressions testing the effects of different sets of predictors. Models (0)-(2) test whether neighborhoods influence the household cooperation decision and whether this influence can be explained by variables traditionally used in response propensity models. Models (3)-(4) test whether these influences can be explained by mechanisms arising from the social and physical environments.

Testing the Traditional Mechanisms

Model (0) in Table 2.4 includes only the tract-level random effect and no household or tract-level covariates. The estimate of the standard deviation of the tract-level variance was positive and statistically significant ($\hat{\psi} = 0.47$, $\text{SE}(\hat{\psi}) =$
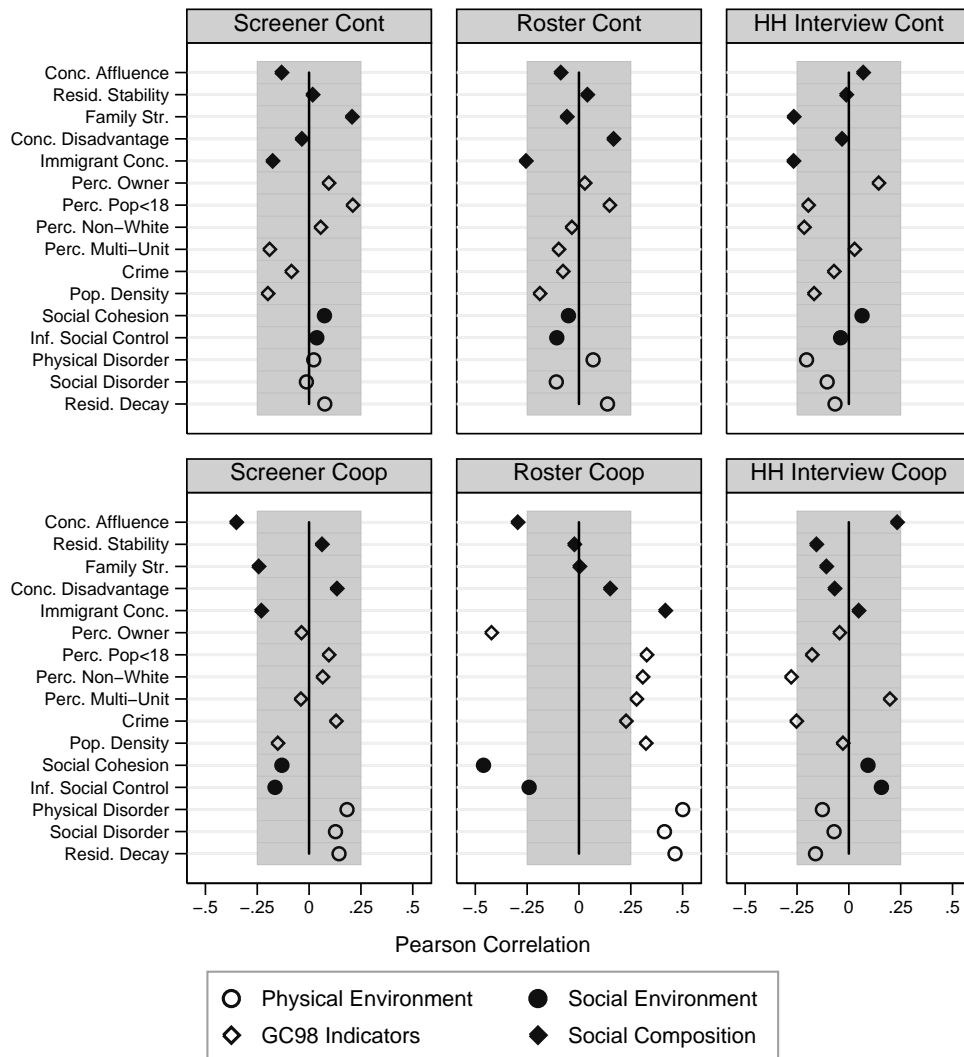
Fig. 2.7: Bivariate Pearson Correlation between All Response Rates and Neighborhood Characteristics (n=65). Unweighted estimates.

0.08), which is evidence of significant tract-level effects on household cooperation. In effect, an estimate of the intraclass correlation (ICC) shows that approximately 6.2% of the unexplained (random) variation in household cooperation is located at

the tract level.[20] This result differ from Van Goor et al. (2005) who finds negligible area-level random effects for household cooperation ($ICC = 0.4\%$). This percentage of variance explained is expected to decrease as neighborhood level predictors are included into the model.

Model (1) incorporates the household level predictors and the strata indicators.[21] Results indicate that households were more likely to cooperate when they *had children*, when the screener respondent was *Latino* (versus *white*), and when the screener respondent was *younger than 34 years old* (versus *35-54 yrs*). In contrast, households were less likely to cooperate when their estimated *rent was below $3000/month*. Adding households covariates significantly improved the fit of the model, based on the AIC and BIC criteria. In addition, the estimate of the ICC dropped from 6.2% to 4.8% from model (0) to model (1), which is an indication that household level characteristics also seem to vary across neighborhoods. See model fit statistics at the bottom of Table 2.4.

Model (2) incorporates the traditional neighborhood covariates into the model. Results shows that cooperation is less likely in *affluent* neighborhoods, and it is more likely in neighborhoods with high *immigrant* concentration. These multivariate results replicate the contradictory findings from the bivariate analyses. Adding these neighborhood covariates did not improve the fit of the model as evidenced

---

[20] An estimate of the percentage explained variance was obtained using $ICC = \frac{\hat{\psi}^2}{\hat{\psi}^2 + \pi^2/3}$ (Snijders and Bosker, 1999).

[21] Due to space constrains these results are suppressed from Table 2.4, however a table with all the regression coefficients is available in the Appendix.

by the increase in both AIC and BIC criteria. In addition, they only marginally reduced the percentage of tract-level unexplained variance from 4.8% to 4%.

Testing the Neighborhood Social Environment

Model (3) incorporates the measures of the social environment into the model. Results show that *informal social control* had a strong and positive effect on household cooperation, whereas *social cohesion* had a strong and negative effect. The direction of the association for *social cohesion* was consistent with the bivariate results, however the direction was reversed for *informal social control*. Although the result for *informal social control* is now consistent with the conceptual model (Hypothesis #1), it is possible that this change in sign is an artifact of the multivariate setting. Results not shown reveal a strong correlation between *social cohesion* and *informal social control* ($r = 0.857, p = 0.0000$), thus multicolinearity is a potential threat for the validity of these results.

To rule out the possibility of multicolinearity effects, model (3) was re-estimated using (a) only one of the social scales at a time, and (b) a combined measure of the two social scales, i.e., a measure of *collective efficacy* (Sampson et al., 1997).[22] When only *social cohesion* was added to model (2), its effect was negative but not significant ($\hat{\beta} = -0.10, SE(\hat{\beta}) = 0.19$). When only *informal social control* was added to model (2) its effect was positive and significant ($\hat{\beta} = 0.27, SE(\hat{\beta}) = 0.13$). Finally, when only *collective efficacy* was added to model (2) its effect was positive but not

---

[22] The tract level measure of collective efficacy was developed averaging responses within tracts using the items of social cohesion (5 items) and informal social control (3 items).

significant ($\hat{\beta} = 0.10, SE(\hat{\beta}) = 0.18$). These results are consistent with the findings in model (3) and strengthen the confidence that the positive effect of *informal social control* on household cooperation is not due to a multicolinearity effect. The results for *social cohesion*, however, remains at odds with the conceptual model in Figure 2.4.

Model (3) can also be used to test Hypothesis #3, which suggests a mediating role of the social environment between neighborhood socio-economic composition and household cooperation. When *social cohesion* and *informal social control* are added to the model, the coefficients of both *affluence* and *immigrant* concentration reduce their size and loose statistical significance. This results is consistent with the definition of a 'mediator effect' (Aneshensel, 2002). However, when *collective efficacy* is added in place of *social cohesion* and *informal social control*, the coefficients of both *affluence* and *immigrant* concentration increase their size and maintain their statistical significance. This mixed set of results precludes a conclusive assessment regarding the nature of the intervening role of the neighborhood social environment on household cooperation.[23]

As expected, adding the measures of the social environment helped reduce the percentage of unexplained variance from 4% to 2.5%. However it did not help

---

[23] I also estimated model (3) using either social cohesion or informal social control in the model. When only social cohesion was added, the coefficients of both *affluence* and *immigrant* concentration got reduced. When only informal social control was added, both coefficients increased. As shown in Table 2.4, the net effect when both social variables are included is the reduction of both of these coefficients.

improve the fit of the model as evidenced by the increase in both AIC and BIC criteria.

Testing the Neighborhood Physical Environment

Model (4) finally incorporates the measures of the physical environment into the model. Results show no significant effect of neighborhood *disorder* on cooperation, but a positive effect of *residential decay* ($\hat{\beta} = 0.24, SE(\hat{\beta}) = 0.13$). This last result echoes the findings from the bivariate analyses.[24]

When measures of disorder and decay were added, the coefficients of the two indicators of socio-economic composition dramatically reduced their size. The coefficient of *affluence* dropped from -0.20 to 0.05 and the coefficient for *immigrant* concentration dropped from 0.20 to 0.01.[25] This result suggests the possibility of a mediating role of the neighborhood physical environment between neighborhood socio-economic composition and household cooperation. This result was not anticipated by the theoretical model in Figure 2.4.

As expected, adding the measures of the physical environment helped reduce even further the percentage of unexplained variance from 2.5% to 1.7%. It did not help improve the fit of the model as evidenced by the increase in both AIC and BIC criteria.

[24] This result holds when the measure of *collective efficacy* was used instead of the measures of *social cohesion* and *informal social control*.

[25] Similar drops were evidenced when I added the measure of *collective efficacy* instead of the measures of *social cohesion* and *informal social control*. The coefficient of *affluence* dropped from -0.33 to -0.11 and the coefficient of *immigrant* concentration dropped from 0.32 to 0.06.

Tab. 2.4: Multilevel Logistic Regression predicting the Probability of Roster Coopera-
tion (n=3,708). Estimates obtained using the single set of household weights
$(wgthh^{NTNS})$.

Variables	Model (0) $\hat{\beta}$	Model (0) $SE(\hat{\beta})$	Model (1) $\hat{\beta}$	Model (1) $SE(\hat{\beta})$	Model (2) $\hat{\beta}$	Model (2) $SE(\hat{\beta})$	Model (3) $\hat{\beta}$	Model (3) $SE(\hat{\beta})$	Model (4) $\hat{\beta}$	Model (4) $SE(\hat{\beta})$
Poverty Strata			–	–	–	–	–	–	–	–
Household Predictors			–	–	–	–	–	–	–	–
Traditional Neighborhood Predictors										
Pop. Density					-0.05	(0.10)	-0.03	(0.10)	0.00	(0.10)
Crime					0.06	(0.09)	0.13	(0.09)	0.11	(0.08)
Conc. Affluence					-0.26*	(0.12)	-0.20	(0.16)	0.05	(0.21)
Resid. Stability					-0.02	(0.09)	-0.03	(0.09)	0.02	(0.10)
Family Structure					0.01	(0.09)	-0.04	(0.09)	-0.04	(0.09)
Conc. Disadvantage					0.07	(0.10)	0.10	(0.10)	-0.07	(0.14)
Immigrant Conc.					0.27*	(0.14)	0.20	(0.14)	0.01	(0.18)
Neighborhood Environment Predictors										
Social Cohesion							-0.45*	(0.21)	-0.47*	(0.22)
Informal Social Control							0.49***	(0.14)	0.47**	(0.14)
Physical Disorder									0.08	(0.18)
Social Disorder									0.10	(0.09)
Resid. Decay									0.24	(0.13)
Intercept	1.50***	(0.09)	1.06***	(0.18)	1.54***	(0.31)	1.46***	(0.31)	1.44***	(0.32)
Tract Level Std.Dev.$(\hat{\psi})$	0.47***	(0.08)	0.39***	(0.09)	0.37***	(0.09)	0.29***	(0.08)	0.24**	(0.09)
Variance Components										
Household Level $(\hat{\pi}^2/3)$	3.2899		3.2899		3.2899		3.2899		3.2899	
Tract Level (ψ^2)	0.2177		0.1652		0.1384		0.0842		0.0571	
ICC	6.21%		4.78%		4.04%		2.49%		1.71%	
Fit Statistics										
Log Likelihood	-1051.42		-999.01		-995.94		-990.57		-987.61	
df	2		16		25		27		30	
AIC	2106.84		2030.02		2041.88		2035.14		2035.21	
BIC (n=65)	2111.19		2064.81		2096.24		2093.85		2100.45	
Observations	3708		3708		3708		3708		3708	

A few strong results emerge from the multilevel analyses of household cooperation. As expected, household-level predictors explained most of the variability on household cooperation. Evidence of area-level effects, however, was strong and remained significant even after controlling for the household covariates. Area-level effects on cooperation arose at the neighborhood level and they were mostly explained by characteristics of the social and physical environment. After controlling for the social and physical environment, the independent effect of neighborhood socio-economic composition on household cooperation was negligible.

Results in Table 2.4 can more easily be interpreted in terms of odds ratios. Odds ratio convey by what multiplicative factor the odds of the predicted event increase per one unit change in the predictor variable. Since all the neighborhood predictors in the multilevel model were standardized, we can directly interpret a unit change in any of the predictors as a change in one standard deviation (SD) unit.

When only measures derived from Census records were included in the model (model 2), results showed that living in a neighborhood with higher *immigrant* concentration (1 SD higher) increased the odds of cooperation by a factor of 1.32, whereas living in a neighborhood with higher *affluence* decreased cooperation by a factor of 0.77. These apparent effects of the neighborhood socio-economic composition were, however, explained by characteristics of the social and physical environment. When measures of the social and physical environment were added to the

model, results showed that living in a neighborhood where residents were perceived to 'act on behalf of the common good' (high *informal social control*) increased the odds of cooperation by a factor of 1.59, however living in a neighborhood where residents shared 'norms and values' (high *social control*) decreased the odds of co-operation by a factor of 0.63.

These models were re-estimated after dropping from the analysis the single tract that was identified as 'potentially influential'.[26] The same substantive results were found using the reduced dataset (n=3,650), regarding the magnitude, sign and statistical significance of the estimates of the neighborhood β coefficients and the tract-level random effect ($\hat{\psi}$). A few differences were observed on the magnitude of the β coefficients for the household covariates. Results from additional sensitivity analyses are presented in the next section.

2.4.3   Sensitivity Analyses

The robustness of the findings from the multilevel analyses was assessed here by re-estimating the models in Table 2.4 using alternative weighting schemes and alternative measures of socio-economic composition. Table with the results from the sensitivity analyses are available in the Appendix.

---

[26] Analyses of univariate outliers identified a single tract as potentially influential due to its relatively low contact rate and large sample size. This tract had the lowest contact rate at the screener stage (67%) and at the roster stage (68%). It also had more than 700 households when the average number of household per tract was approximately 230. This tract was located in a high socio-economic status neighborhood which had many high rise buildings.

Alternative Weighting Schemes

A first test of the stability of the multilevel analyses consisted on assessing how results would change under different specification of the sampling weights. The analyses in Table 2.4 were first replicated using the original weights with an alternative scaling method (ALT), the original weights without scaling (UNS), and without sampling weights (UNW).

Models using different specifications of the weighting scheme presented a few differences across the tract level β coefficients, however the main results hold: (a) no significant effect of neighborhood socio-economic composition, (b) strong positive effect of *informal social control*, strong negative effect of *social cohesion*, and weak positive effect of *physical decay*.

The methods differed somewhat on the estimate of the tract level random effect ($\hat{\psi}$). The estimate using the alternative scaling method showed the same pattern of decreasing size for $\hat{\psi}$ as tract-level covariates were included into the model with a final estimate close to that presented for model (4) in Table 2.4 ($\hat{\psi}_{ALT}$=0.3, SE(ψ_{ALT})=0.08). The estimate using the unscaled weights however did not show the decreasing pattern for $\hat{\psi}$ and the final estimate (model 4) was larger in magnitude than that in Table 2.4 ($\hat{\psi}_{UNS}$=0.46, SE($\hat{\psi}_{UNS}$)=0.08).

As expected, the unweighted results yielded somewhat different results for the tract-level β coefficients. The effect of *immigrant* and *crime* was strong across all models and remained marginally significant in the final model. The effect of the environmental variables, however remained substantively the same: strong positive

effect of *informal social control*, weaker negative effect of *social cohesion*, and weak positive effect of *physical decay*. The estimate of the random effect followed the same pattern of decrease and magnitude to those presented in Table 2.4 and the estimates for model (4) were very close in size ($\hat{\psi}_{UNW}$=0.27, SE($\hat{\psi}_{UNW}$)=0.06).

Analysis for model (4) were re-run a last time after receiving the two set of sampling weights corresponding to the tract and the household level (TWO). Estimating the multilevel model with the two set of weights yielded the same substantive results as model (4) with the single set of weights. The estimate of the random effect for the full model was very close in size ($\hat{\psi}_{TWO}$=0.29, SE($\hat{\psi}_{TWO}$)=0.10) to those presented in Table 2.4.

Alternative Measures of Neighborhood Socio-Economic Composition

A potential critique of the comparability of the results presented here is that the measures of socio-economic composition derived from factor analysis would likely yield different results across different estimation samples. To address this concern, the analyses in Table 2.4 were replicated using Census derived variables closely resembling those used by Groves and Couper (1998): *percentage owner occupied, multi-unit structure, non-white* and *population under 18 yrs*. The substantive results were the same as those presented earlier: (a) strong tract-level random effects, (b) no significant effect of neighborhood social composition, (c) strong positive effect of *informal social control*, strong negative effect of *social cohesion*, and weak positive effect of *physical decay*. In effect, none of the indicators of socio-economic composition reached statistical significance when added to the model.

2.5 Discussion

The purpose of this study was to extend the Groves and Couper model of survey participation and test the validity of the propositions derived from the revised model. At the core of the research inquiry was the question of the ability of the conceptual model to reveal the mechanisms underlying the influence of the urban context on the household cooperation decision. Evidence from this study supports this claim. A somewhat unexpected result, however, was the direction of some of the associations found.

In the bivariate analyses, living in a neighborhood with signs of *disorder* and *residential decay* was found to increase the likelihood of cooperation. In the multivariate analyses, this effect remained significant only for *residential decay*. This result contradicts expectations under the *'Broken-Windows Theory'* (Wilson and Kelling, 1982). A plausible *post-hoc* explanation for this finding can be elaborated for the survey context under the 'social needs' hypothesis (Suttles, 1972). Under this framework, residents in neighborhoods high in crime and physical dilapidation could be motivated to participate in formal and informal initiatives that help them address these problems. Given the particular focus of the L.A. FANS survey on family and neighborhood issues, it is possible that the request for survey participation was most likely perceived as 'an opportunity to improve neighborhood conditions'. This could explain the higher cooperation rates observed in tracts with high levels of *crime*, signs of *disorder* and signs of *residential decay*. Given the periodical monitoring of tract-level response rates during data collection (Groves and Heeringa, 2006),

survey practitioners could benefit from a better understanding of this perspective.

Another result that contradicted theoretical expectations was the negative association between measures of the social environment and household cooperation rates. Both *social cohesion* and *informal social control* presented this pattern in the bivariate analyses. This effect was only observed for *social cohesion* in the multivariate analyses. This result contradicts the expectations under the *'Collective Efficacy Theory'* (Sampson et al., 1997). A plausible *post-hoc* explanation for this phenomena is an 'upward bias' in the estimates of the social environment, which were developed from interviews with local residents in the L.A. FANS study. The idea of an upward bias on neighborhood-level estimates was inspired by the work of Abraham et al. (2009) on biases on individual-level estimates of pro-social behavior. Their study finds a strong connection between the causes of volunteering and the causes of survey participation. The authors demonstrate an upward bias in individual-level estimates of volunteering derived from the American Time Use Survey (ATUS) surveys and they conclude that estimates of pro-social behavior derived from household surveys will tend to overestimate the prevalence of these type of activities.[27]

Measures of neighborhood *social cohesion* and *informal social control* are, arguably, measures of collective pro-social behavior. Under this context, I hypothesize that a similar mechanism to that described by Abraham et al. (2009) at the national-

---

[27] The authors analyzed respondents and nonrespondents to the American Time Use Survey ATUS. Data on volunteering for both groups was available from the Current Population Survey (CPS).

level could be at play for small-area estimates from the Los Angeles neighborhoods. In effect, one should probably expect lower estimates of the social environment when more households respond in a given tract than when fewer households do. In the extreme if we only get responses from the few 'socially engaged' neighbors – and thus observe low tract-level response rates – then we probably should expect higher (*upward biased*) tract-level estimates of the social environment. This could explain the negative relationship observed, at the tract level, between the roster cooperation rates and the measures of the social environment.

Another unexpected result was the positive effect of ethnic and racial composition on household cooperation when only Census-derived variables were used in response propensity models. Both correlates of cooperation – neighborhood *affluence* and *immigrant concentration* – loaded on Census indicators of *Hispanic origin*, *Spanish speaker*, *foreign born* and *non-citizen*. Given the important presence of Hispanic immigrants in Los Angeles County, this results should be revisited under a framework that incorporates 'cultural' effects and not only 'area-level' effects on household cooperation.

This study also attempted to open the discussion about the appropriate levels of geography to assess area-level effects on survey participation: Is there a single macro-level influence on household cooperation or are there multiple (e.g. blocks, tracts, counties)? Is it different by type of response process (e.g. contact, cooperation)? Is it different by stage of the recruitment process (e.g. screener, roster, interviewing)? Evidence from this study suggest that different response processes are influenced by different levels of geography. Block-level effects were somewhat

stronger on contactability, whereas tract-level effects were stronger on cooperation. This pattern hold for both the roster and the household interview stage. Blocks and tracts exercised roughly the same degree of influence on both response processes at the screener interview stage. Future studies of area-level effects should attempt to state clearly the scope of the inferences based on the kind of aggregate data used for investigation.

Observational studies present many challenges and limitations that threaten the validity and generalizability of their results. I attempted to address some of them through the modeling strategy (e.g. multilevel modeling, complex survey design, imputation), and others by replicating the results under alternative conditions (e.g. weighting schemes, neighborhood predictors, outliers). A clear limitation of this study is, however, the generalizability of its results. The sample used is only representative of the Los Angeles County, which is an ethnically diverse urban enclave where, for example, non-hispanic whites represent only 29% of the population.[28] As a result, it is not possible to generalize the findings from this study to the population of counties in the United States. Emerging evidence suggest, however, that neighborhood effects could, indeed, be very local (Carroll-Scott, 2008).[29] The exis-

---

[28] http://quickfacts.census.gov/qfd/states/06/06037.html

[29] Carroll-Scott (2008) studied the effect of neighborhoods on child developmental outcomes using data from Chicago and Los Angeles neighborhoods. She concluded that the importance of *affluence* as a predictor of neighborhood social processes may differ across regional contexts rather than being universal. The data used in her analyses for Los Angeles was available from the L.A. FANS study (Sastry et al., 2003). The data for Chicago was available from the Project on Human Development in Chicago Neighborhoods (PHDCN) (Earls et al., 1997).

tence of a universal set of contextual effects on survey cooperation is an empirical question that would require the replication of this kind of study across different type of communities.

In spite of its limitations, the most important contribution of this paper lies in re-discovering important intersections between survey methods and urban sociology. The review of the urban sociology literature evidenced the simplicity of current ecological theories of survey participation. Furthermore, results from this study helped to highlight the complexities involved in the study of neighborhood effects on household participation. The benefit of the cross examination of the two bodies of literature extends beyond the theoretical model developed and tested in this study. Given the increase threat of nonresponse bias due to declining response rates in household surveys, the evaluation of non-traditional neighborhood variables for nonresponse adjustments seems like a natural follow-up from this study. In addition, testing for interviewer effects on ratings of neighborhood characteristics surges as another area suited for further research in the survey methods literature.

3. PAPER 2: ASSESSING THE MEASUREMENT ERROR PROPERTIES OF INTERVIEWER OBSERVATIONS OF NEIGHBORHOOD CHARACTERISTICS

3.1 Introduction

Interviewer observations play an increasing role in survey research. As part of face-to-face data collection efforts, interviewers are often charged with making observations of respondents, housing units, and neighborhood characteristics. Those interviewer observations are currently used to inform responsive design decisions, to expand the set of covariates for nonresponse adjustments, to explain participation in surveys, and to assess nonresponse bias.

While the use of interviewer observations is growing, little effort has been made to assess the quality of these assessments. Notable exceptions are Pickering et al. (2003) and Alwin (2008). Inspired by the recent interest in paradata and the use of interviewer observations mentioned above, a special session was held at AAPOR 2010 to bring new attention to measurement error in interviewer observations. There, observations of respondent characteristics (West, 2010; McCulloch et al., 2010) and housing units (Sinibaldi, 2010) were evaluated and first attempts were made to explain errors in such observations across interviewers, subgroups, or time.

Quality assessments of neighborhood observations, however, have not been performed in the survey literature. Urban sociologists, on the other hand, have looked in detail at the measurement error properties of neighborhood measures (Raudenbush and Sampson, 1999; Caughy et al., 2001). However, it is an empirical question whether measurement error properties derived from such specialized studies can serve as benchmarks for studies that use interviewers to collect neighborhood

observations in typical survey settings.

This paper addresses this issue by providing quality assessments based on neighborhood data from the Los Angeles Family and Neighborhood Study (Sastry and Pebley, 2003). Two features make this study particulary relevant for survey research. First, the study trained a single group of interviewers to collect both survey data and neighborhood data – which mirrors how regular surveys collect observational data. Second, the study collected multiple independent observations on each sampled block – which allows the estimation of variance components associated with interviewers, tracts, and blocks. Furthermore, data on interviewers is available to test whether certain interviewer characteristics influence their perceptions of neighborhood characteristics.

The rest of this chapter is organized as follows. Section 3.2 provides the background material from the fields of survey methodology and urban sociology which helps to set the research questions that this paper aims to answer. Section 3.3 presents the data and methods used to answer these questions, and section 3.4 presents the results. Section 3.5 provides a discussion of the results, limitations of the analyses, and ideas for future research.

3.2   Background

Assessing the quality of neighborhood observations collected by survey interviewers presents a particular set of challenges. I review here two sets of literature to provide some context on the different issues involved. Section 3.2.1 presents findings from the survey methodology literature on the errors introduced by interviewers on the data they collect. Section 3.2.2 presents findings from the urban sociology literature on the quality of neighborhood observations derived from specialized studies of neighborhood effects. A summary of the contributions of both bodies of literature and the outstanding research questions that this dissertation seeks to answer are presented in section 3.2.3.

3.2.1   Interviewers and Measurement Error

Interviewers' idiosyncratic behavior usually leaves a *trace* on the data they collect. An important topic of research in survey methods deals with the measurement of 'interviewer effects' on survey responses. The literature is scarce, however, on the assessment of error in the data interviewers collect by direct observation. The first part of this section provides evidence on the magnitude of interviewer and sampling point clustering effects derived from analyses of survey responses. The second part present recent evidence on the quality of the data interviewers collect using direct observation methods.

Survey researchers generally conceptualize the influence of interviewers as random. The most commonly used statistic for estimating interviewer related error is the intraclass correlation, typically referred as ρ_{int} (Fowler, 1991, 261), and the most widely used approach to calculate ρ_{int} is Kish's ANOVA method (Groves, 1989; O'Muircheartaigh and Campanelli, 1998; Biemer and Lyberg, 2003). Mathematically, ρ_{int} represents the correlation between two different responses collected by the same interviewer – thus ρ_{int} is a measure of how similar (*homogeneous*) are the responses collected by interviewers.

A value of ρ_{int} greater than zero means that answers to a particular question are susceptible to interviewer influence, whereas values equal to zero indicate no interviewer influence.[1] Groves (1989, 365) reported values of ρ_{int} for survey responses averaging 0.03 in face-to-face surveys, with the majority of the values tending to be between 0.01 and 0.02. The average values of ρ_{int} are much smaller, around 0.01, in the context of telephone surveys (Groves, 1989, 367).

Many face-to-face surveys suffer from the fact that their design confounds two distinct sources of homogeneity – the sampling point (ρ_{sp}) and the interviewer (ρ_{int}) clustering effect (Hox et al., 1991). The first factor yields *spatial homogeneity*, which reflects the fact that respondents are more alike within a geographical area than in the population as a whole. The second factor yields interviewer homogeneity, which suggests that interviewers' idiosyncratic behavior makes the data they collect from

---

[1] The values of ρ_{int} can range from $\frac{-1}{(m-1)}$ to $+1$ (Groves, 1989, 318). In practice researchers set negative values to zero.

their assigned cases more similar. Sampling point homogeneity is a characteristic of the *true values* of elements in the population (O'Muircheartaigh and Campanelli, 1998, 67). Interviewer homogeneity, on the other hand, represent features of the method of data collection and thus can be regarded as *error of measurement.*

The development of multilevel modeling techniques (Goldstein, 1995) and the use of careful designs has allowed survey researchers to take on several measurement challenges: the joint estimation of multiple sources of random variation (*e.g.* O'Muircheartaigh and Campanelli (1998); Schnell and Kreuter (2005); Durrant and Steele (2009)); the joint estimation of fixed and random effects (*e.g.* O'Muircheartaigh and Campanelli (1998); Durrant and Steele (2009)), and the simultaneous analysis of multiple items (*e.g.* Pickery and Loosveldt (2004b)). In terms of content, most of these studies look into the effects of interviewer characteristics on survey participation (Groves and Couper, 1998; Pickery and Loosveldt, 2002; Durrant and Steele, 2009; Durrant et al., 2010) and the quality of survey responses (Pickery and Loosveldt, 2001, 2004b). I only review here selected results from the two studies that separate spatial clustering ρ_{sp} from interviewer clustering ρ_{int} in the analysis of survey responses.

O'Muircheartaigh and Campanelli (1998) reported ρ_{sp} estimates significantly greater than zero for four out of ten items and ρ_{int} estimates greater than zero for three out of ten. When considering only items collected by interviewer observation, however, the percentage of significant $\hat{\rho}_{int}$ rose to three out of four. Overall, the authors found the highest values of $\hat{\rho}_{int}$ for tenure and ethnicity, and the lowest values for gender and marital status. Schnell and Kreuter (2005), on the other

hand, reported the largest effects $(0.49 < \hat{\rho}_{int} < 0.58)$ for respondents' reports of neighborhood conditions such as 'graffiti on the walls'.

O'Muircheartaigh and Campanelli (1998) also reported similar order of magnitude for $\hat{\rho}_{sp}$ and $\hat{\rho}_{int}$, both ranging from -0.02 to 0.18. Prior studies had found both larger interviewer variance (Hansen et al., 1961) and larger sampling point variance (Bailey et al., 1978).

Before closing this section of the review it is important to acknowledge that there is another large body of the survey literature that conceptualizes interviewer effects as *systematic* rather than *random*. Here too, studies focusing on the influence of interviewers on the survey responses themselves are rare. Two landmark studies in this area investigated the effects of interviewer's race and gender on survey responses. Schuman and Converse (1971) found that questions dealing with militant protest and hostility to whites showed the greatest sensitivity to interviewer's race. A study by Kane and Macaulay (1993) found that male and female respondents expressed more egalitarian gender-related attitudes or greater criticism of existing gender inequalities to female interviewers. In most surveys in which they have been studied, however, interviewer demographic characteristics are found to be unrelated to the data that results (Fowler, 1991).

Interviewer Effects on Direct Observation Items

A few large survey programs currently charge interviewers with the responsibility of collecting observational data on *respondents'* characteristics, attitudes and behaviors; *housing units'* physical characteristics; and *neighborhoods'* physical char-

acteristics. In the United States, surveys collecting these data include the Current Population Survey (CPS), the National Health Interview Survey (NHIS) and National Survey of Family Growth (NSFG). In Europe, countries participating in the European Social Survey (ESS) also collect similar observations.

Awareness about the completeness and quality of this type of observational data has been discussed in conferences and technical reports (Kreuter et al., 2007; Carton, 2008; Blom et al., 2008; Maitland et al., 2009; Casas-Cordero, 2010). To date, however, only a few studies have attempted to assess the quality of observational data collected by interviewers.

The work by Alwin (2008) represents the most comprehensive effort to date to provide estimates on one crucial aspect of measurement – the *reliability* of survey data. Alwin (2008) found high reliability for interviewers' reports of 'facts' pertaining the households, but low reliability for interviewers' 'beliefs' about respondent characteristics or reactions to the interview. Low reliability of subjective ('belief') items, however, occurred not only among interviewers' reports but also among respondents' self-reports (Alwin, 2008, 159). The paper by West (2010) presents the first attempt to measure the *accuracy* of interviewer observations of respondents' characteristics in the context of the National Survey of Family Growth (NSFG). In this study, agreement between the interviewer observations and the respondents' report was above 70% – and the Kappa statistic around 0.3 – for observations of children in the household and judgement of whether the respondent is sexually active.

Pickering et al. (2003) provides evidence of the *accuracy* of interviewer's as-

sessment of the *age* and *tenure* of housing units. Their report compares interviewers' initial guesses to householders' survey responses. Agreement between the two tenure sources varied considerably for different types of units $(89\% - 46\%)$ and, in general, were higher in rural areas than in urban areas. Finally, the paper by Sinibaldi (2010) looked at agreement rates for three housing unit characteristics and one neighborhood characteristic.[2] These variables are routinely collected as part of the contact protocol of surveys conducted by NatCen in the United Kingdom. Overall, agreement between the raters was higher for the housing unit items $(78\% - 91\%)$ than the area item (63%). Area-level characteristics were strong predictors of agreement, but interviewer characteristics were not.[3]

To my knowledge, no other studies in the survey literature provide estimates of the quality of neighborhood observational data collected by interviewers. Urban sociologists, on the other hand, have looked in detail at the measurement error properties of neighborhood measures. The next section describes a framework that helped to guide the research questions in this paper.

---

[2] Observations collected for the housing units (HU) were: *barriers to entry*, *type of dwelling*, and *condition of the HU compared to others in the area*. The item on the area surrounding the sampled HU was *'Condition of residential properties in the area'*.

[3] Area-level predictors used by Sinibaldi (2010) included: region of the country, safety of the area, and a composite measure of socioeconomic deprivation culled from Census records. Interviewer level predictors included: age, sex and pay grade.

3.2.2 Measuring Neighborhood Environments

Drawing on findings in urban sociology, this section discusses the challenges encountered in measuring neighborhood environments and provides benchmarks for the assessment of the measurement error properties of neighborhood ratings. The first part describes data collection protocols used in the urban sociology literature to collect data on physical characteristics of the neighborhood environments. The second and third parts provide estimates of the measurement error properties of neighborhood observations derived under such specialized studies. The last section provides insights into the factors that influence perceptions of disorder and decay.

Systematic Social Observations

Urban sociologists increasingly collect data on neighborhood physical characteristics and resources. They sometimes asks survey respondents, or their neighbors, about the conditions of the surrounding area. Increasingly, they use trained observers to collect such assessments using special neighborhood questionnaires. Examples of observations collected about the neighborhood physical environment are presented in Table 3.1.

Research by Taylor and colleagues pioneered the collection of physical signs of disorder and decay in urban neighborhoods (Taylor et al., 1985; Perkins et al., 1992; Taylor, 2001). The work by Sampson and Raudenbush, on the other hand, sparked a renewed interest in the approach by providing a framework for the analysis of the measurement error properties of neighborhood measures (Raudenbush and Sampson,

1999; Sampson and Raudenbush, 1999). As a result, many neighborhood inventories have been developed in the last 20 years inspired by these studies. I only review here studies that include observations on physical disorder and decay collected by trained observers.

Studies collecting observational data share a strong focus on the training and selection of observers. Some studies recruit observers based on their expertise in the subject matter (Weich et al., 2001; Laraia et al., 2006), while others prefer to use residents of the target areas (Taylor et al., 1995; Mujahid et al., 2007). Most studies train observer using written definitions and pictures of neighborhood features and answer categories. They also implement field practice sessions, individual evaluations and group debriefings. Training sessions last on average 30 hours and span multiple days (Caughy et al., 2001; Sastry and Pebley, 2003; Dunstan et al., 2005; Laraia et al., 2006; Zenk et al., 2007; Furr-Holden et al., 2008). Observers are sometimes hired based on their performance on reliability targets achieved during the training sessions (Caughy et al., 2001; Zenk et al., 2007). Most studies hire observers for the sole purpose of collecting neighborhood data (Caughy et al., 2001; Weich et al., 2001; Dunstan et al., 2005; Zenk et al., 2007). A few studies train the same crew of interviewers to collect both the survey data and the neighborhood data (Sastry and Pebley, 2003; Andresen et al., 2006).

The implementation of the data collection protocol also differs markedly across studies. Most studies collect data on paper and pencil (Raudenbush and Sampson, 1999; Caughy et al., 2001; Sastry and Pebley, 2003; Dunstan et al., 2005), but others use hand held devices (Zenk et al., 2007; Furr-Holden et al., 2008). A more expensive

approach records all or part of the observations using videotapes, which are coded later by trained observers (Raudenbush and Sampson, 1999; Cohen et al., 2000). Most studies collect data over a few months (Caughy et al., 2001; Weich et al., 2001; Sastry and Pebley, 2003; Laraia et al., 2006; Zenk et al., 2007), but some take only a few days (Dunstan et al., 2005).

Studies typically collect data using pairs of observers, and ratings for each block are usually arrived at independently (Taylor et al., 1995; Raudenbush and Sampson, 1999; Sastry and Pebley, 2003; Dunstan et al., 2005; Zenk et al., 2007; Furr-Holden et al., 2008), but some are developed by consensus (Caughy et al., 2001; Franzini et al., 2008). Some pairs complete their observations simultaneously (Taylor et al., 1995; Raudenbush and Sampson, 1999; Caughy et al., 2001; Dunstan et al., 2005; Laraia et al., 2006; Andresen et al., 2006; Furr-Holden et al., 2008; Franzini et al., 2008), while other work on separate occasions (Sastry and Pebley, 2003; Zenk et al., 2007).

As evident from this review, there is not such a thing as a 'universal' approach to collect neighborhood data using trained observers. What is common to all these studies, however, is the strong desire to reduce the errors associated with individual observers by carefully training, selecting and supervising them. The next section provides some results derived from the analysis of the measurement error properties of the neighborhood data collected in these specialized studies.

Evaluation of Measurement Error using Classical Psychometric Methods

Studies of neighborhood effects have traditionally used well know psychometric techniques to assess the validity and reliability of neighborhood constructs (Raudenbush and Sampson, 1999). Most studies assess the validity of neighborhood measures by correlating the ratings provided by trained observers with the ratings provided by residents of the same areas (McGuire, 1997; Raudenbush and Sampson, 1999; Caughy et al., 2001; Dunstan et al., 2005; Furr-Holden et al., 2008). Raudenbush and Sampson (1999), for example, reported strong correlations between raters and adult residents on the *physical disorder* scale ($r = 0.71$) and the *social disorder* scale ($r = 0.65$). Furr-Holden et al. (2008) reported significant correlations between raters' and teenagers' reports on items related to *violence* ($r = 0.17$), *alcohol* ($r = 0.24$) and *drug use* ($r = 0.29$). The validity of neighborhood measures is also assessed by correlating reports from trained observers with Census variables (Laraia et al., 2006) or composites derived from Census variables such as measures of socio-economic status, ethnic heterogeneity and crime (Raudenbush and Sampson, 1999; Caughy et al., 2001; Dunstan et al., 2005).

Another key criterion used to evaluate the quality of observational data is the reliability of the neighborhood ratings. Reliability assessments are implemented on both neighborhood scales and individual items. Weich et al. (2001), for example, reported interobserver agreement as low as 0.36 for the item *evidence of vandalism*, and as high as 0.90 for the item *evidence of graffiti*. Overall, most estimates of percentage agreement between pairs of trained observers range from 0.65 to 0.85

(Taylor et al., 1995; Raudenbush and Sampson, 1999; Weich et al., 2001; Caughy et al., 2001; Craig et al., 2002; Brown et al., 2004; Laraia et al., 2006; Andresen et al., 2006; Zenk et al., 2007; Furr-Holden et al., 2008).

Estimates of reliability are also reported using the intraclass correlation coefficient (ICC) (Shrout and Fleiss, 1979) and the Kappa statistic (Cohen, 1960). Furr-Holden et al. (2008, 253) reported high ICC's across the items in scales of *physical layout* (0.61−0.98), *residential decay* (0.71−0.94), *physical disorder* (0.60−0.99), *social disorder* (0.70 − 0.82), *adult activity* (0.69 − 0.85), *youth activity* (0.62 − 0.82), and *drug use* (0.67 − 0.79). Estimates of the Kappa statistics, which control for agreement due to chance, are mostly in the range from 0.50 to 0.85 for items on disorder and decay (Weich et al., 2001; Dunstan et al., 2005; Andresen et al., 2006).

Estimates of reliability based on classical techniques, such as those reported here, do not reflect the complex structure of neighborhood observational data. The next section introduces a more general framework which, among other features, incorporates multiple sources of random and systematic variation in the analysis of neighborhood data.

Evaluation of Measurement Error Properties using a Multilevel Framework

A second generation of neighborhood studies uses a framework known as *ecometrics* (Raudenbush and Sampson, 1999) to assess the measurement error properties of neighborhood characteristics. The framework borrows, integrates and adapts three analytic strategies that are prominent in psychometrics – *item response modeling, generalizability theory*, and *factor analysis* (Raudenbush and Sampson, 1999).

Under the ecometric framework, data collected by trained observers is modeled using a three-level hierarchical level model, where neighborhood items are at the lowest level, block faces are at the second level, and neighborhoods are at the highest level. Figure 3.1 illustrates this structure. The model can be viewed as an item response model (level-1) embedded within a hierarchical structure in which the secondary units of measurement, the block faces (level-2), are nested within the units of primary interest, the neighborhoods (level-3). The model can be extended by allowing for multiple characteristics (*factors*) to be measured simultaneously, for example *physical disorder* and *social disorder*, rather than a single, unidimensional trait.
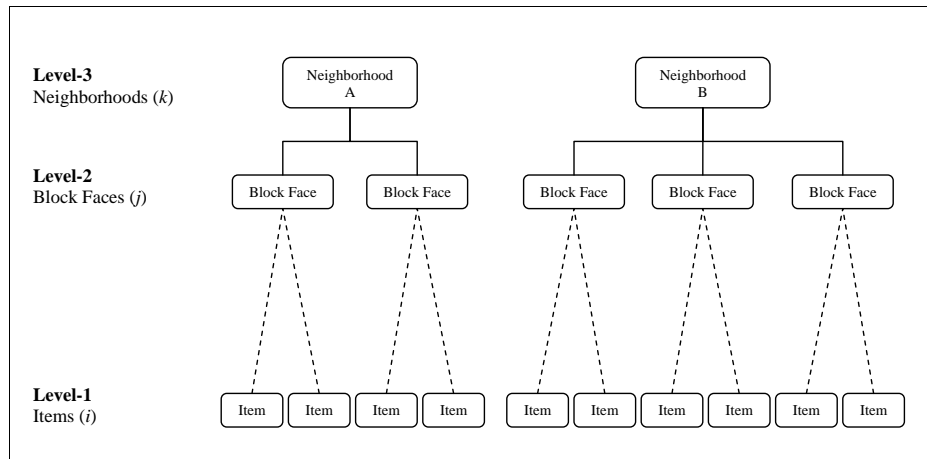
Fig. 3.1: Data Structure for Ecometric Analysis

Fitting the model in Figure 3.1 produces several estimates of interest for assessing the quality of neighborhood measures. Analyses of the lowest level has shown that items in the *physical disorder* scale have three desirable qualities, first they show a high degree of variation, which is one of the characteristics of a *good* scale;

second they are ordered in a way which is consistent with what one would expect from the frequency distributions of the items; and third, they show 'face validity' – the items which are observed more frequently are considered less 'severe' or 'difficult' than the items observed less frequently (Raudenbush and Sampson, 1999; Sampson and Raudenbush, 1999; Caughy et al., 2001). Items in the *social disorder* scale, on the other hand, have not shown the same desirable properties (Raudenbush and Sampson, 1999; Sampson and Raudenbush, 1999).

Analyses at the second level consists of assessing how much agreement exists between observations at the level of the small areas (*e.g.* block faces) used to characterize the neighborhoods (*e.g.* census tracts). Variability at the block face level has been partially explained by one characteristic of the occasion of measurement – the *time of day* in which the block face was rated. Results show that 'time of day' is a significant predictor of the likelihood of seeing *social disorder* on the block face, but not *physical disorder* (Raudenbush and Sampson, 1999; Sampson and Raudenbush, 1999). Effects of other systematic sources of measurement error have not been incorporated in this multilevel framework.

What influences perceptions of disorder?

This review has focused on the measurement error properties of neighborhood observations reported by trained observers. Our understanding of the factors that influence observers' perceptions, however, is rather limited. Andresen et al. (2006) used a single-level regression model to assess the effect of the observers' 'work experience' and 'place of residence' on observer's perceptions of disorder. The authors

report that experienced local observers were more likely to perceive signs of disorder than inexperienced and nonlocal observers. Sampson and Raudenbush (1999) found that measures of neighborhood *concentrated disadvantage*, *immigrant concentration*, *collective efficacy* and *mixed land use* showed the strongest associations with tract-level measures of physical and social disorder reported by trained observers. Measures of *residential stability* and *population density* were not influential.

As mentioned at the beginning, two types of informants have been used to provide assessments of neighborhood disorder – local residents (*subjective perceptions*) and trained observers (*objective perceptions*). Most of our understanding of what shapes perceptions of disorder comes from studies based on residents' reports of disorder and decay. I briefly review here findings from studies that have used multilevel models to assess these influences.

Recent studies have provided evidence of both person-level and neighborhood-level characteristics influencing residents' perceptions of disorder and decay (Sampson and Raudenbush, 2004; Mujahid et al., 2007; Franzini et al., 2008). At the neighborhood level, the studies by Sampson and Franzini also controlled for *objective* measures of disorder and decay provided by specially trained observers.

At the person-level, the only consistent result indicates that – compared to other residents in the same neighborhood – black and minority residents are less likely to report signs of disorder (Sampson and Raudenbush, 2004; Mujahid et al., 2007; Franzini et al., 2008). Sampson and Raudenbush (2004) argue that increased past exposure to disorder increases the threshold at which disorder is being perceived as a problem. Given the history of racial segregation in the United States, it is pos-

sible that blacks have been exposed to more disorder than whites in the past and therefore it is possible that blacks and whites judge disorder differently. As Sampson and Raudenbush (2004, 329) point out *"the basic psychological mechanism [in perceiving disorder] involves the perception of discrepancies based on expectations, underscoring the fact that perceived disorder reflects more than meets the eye"*. Older residents are less likely to perceive disorder than younger residents (Sampson and Raudenbush, 2004; Franzini et al., 2008). Other person level characteristics, however, show mixed results. Franzini et al. (2008) found that residents that were ever married, who move frequently, or have more education tend to perceive less disorder. Sampson and Raudenbush (2004) found higher reports of disorder among females residents and those divorced, but lower reports among residents embedded in networks of reciprocal exchange. Employment status, socioeconomic status, mobility and home ownership, however, were unrelated to perceptions of disorder. In contrast, Mujahid et al. (2007) found perceptions of disorder uncorrelated to gender, but lower among high SES residents.

At the neighborhood level, two correlates of perceived disorder arise after controlling for both residents' characteristics and objective measures of disorder and decay – neighborhood 'socioeconomic status' and 'ethnic composition'. Perceived disorder is higher in neighborhoods with high percentage of the population *under the poverty level* (Sampson and Raudenbush, 2004; Franzini et al., 2008), a high percentage of *black* residents (Sampson and Raudenbush, 2004), and high percentage of *latino* residents (Sampson and Raudenbush, 2004). Measures of *population size* (Sampson and Raudenbush, 2004; Franzini et al., 2008) and *population density*

(Sampson and Raudenbush, 2004), however are uncorrelated to perceived disorder.

The above discussion provided context, from two different disciplines, on the challenges to collecting observational data using trained observers. The review also reveled important gaps in the assessment of the quality of neighborhood measures. The next section presents the research questions this dissertation aims to answer.

Tab. 3.1: Neighborhood characteristics and related variables. Classification by Nicotera (2007).

Social Composition	Social Processes
Age	Organizational participation
Race/ethnicity	Unsupervised teens
Nativity	Neighboring
Residential mobility	Crime
Density of children	Value consensus
Percent of female headed households	Community monitoring
Percent of elderly	Social capital/social networks
Percent of single parents	Civic participation
Economic Composition	**Physical Composition/Resources**
Percent affluent neighbors	Condition of housing
Poverty	Trash/litter
Employment	Graffiti
Percent white collar workers	Traffic, street, and parking conditions
Percent managerial/professional workers	Play grounds/parks
Education	Proximity to employment and public transportation
Public housing	Community centers
Home ownership	Schools
Proximity to affluent neighborhoods	Bars, grocery stores, retail shops, cafes
	Libraries
	Abandoned homes and Vacant lots
	Crowding
	Architecture

3.2.3 Research Questions and Hypotheses

The review of the urban sociology literature above showed that there are many ways in which specialized studies collect observational data. Different aspects of the data collection protocols are especially sensitive to cost and quality considerations: the need for repeated assessments, how many observations of each unit to collect, the time elapsed between repeated observations, and whether the personnel in charge of collecting the observations is solely dedicated to that task or is given other tasks (*e.g.* listing, interviewing).

Learning about the effects of these features of the protocol is important to inform cost-quality trade-offs. Unfortunately, not enough studies are available to estimate the measurement error properties of neighborhood observational data collected under different protocols. One of the aims of this dissertation is to contribute to fill this gap.

One of the interesting contrasts between the urban sociology literature and the survey methods literature has to do with the conceptualization and treatment of the effects associated with the personnel in charge of data collection.

Urban sociologists use sophisticated techniques to train and monitor observers with the aim of *reducing* observer effects on the data they collect. Most studies, however, use classical methods to assess one dimension of measurement – the reliability of the data collected by two different observers. Assessments of other dimensions of measurement, such as the degree of homogeneity of the data collected by the same interviewer (ρ_{int}), is not addressed in the studies.

Survey methodologists, on the other hand, have a long tradition studying the influences of interviewers on the data they collect, conceptualizing these influences as both systematic and random. Most of the research in this field, however, has focused on the influence of interviewers on survey responses and not on the observational data they collect.

Thus, three research questions emerge and guide the analyses in this paper:

1. How large are the estimates of *interobserver agreement* for neighborhood items collected by trained observers?

2. How large are the estimates of *observer clustering effects* for neighborhood items? How large are these estimates relative to *sampling point clustering effects*?

3. What characteristics of the observers are associated with the values they assign to neighborhood observations? In other words, are there any *systematic effects* of observers on the data they collect?

To investigate these issues, I propose two modifications to the measurement error model of Raudenbush and Sampson (1999) illustrated in Figure 3.1. First, I incorporate the effect of observers as another source of random variation in the measurement error model of neighborhood observations. Second, I focus on individual items rather than on scales. Below I discuss the rationale for this extended model and I present three hypotheses to be tested within this model.

As discussed earlier, the standard measurement error model used for neighborhood constructs has been conceptualized as having three hierarchical levels given by *neighborhoods, block faces* and *items* (Figure 3.1). I propose to extend this model by conceptualizing the data as having at least three levels given by *neighborhoods (k), blocks (j)* and *observers (r)*. An additional level, associated with *items (i)*, could be added when analyzing scales containing multiple items or suppressed when only analyzing individual items. Figure 3.2 illustrates the extended model.

The revised model combines the hierarchical structure given by the geography of small areas, and the cross-classification with observers, which is given by the distribution of their work assignments. Each block and its block faces are nested in a single tract. At the same time, however, each block could also be nested in two different observers if there is a partial interpenetration of their work assignments. In Figure 3.2, for example, Interviewer #1 collects observations in neighborhoods A and B, whereas Interviewer #3 only collected data on neighborhood B. In this setting, 'interviewers and blocks' and 'interviewers and tracts' are *crossed* levels because none of them is completely nested within the other.

Estimates derived from this extended model can be used to develop estimates of interobserver agreement and observer clustering effect, which can be used to answer the first two research questions. To answer the third question different sets of covariates can be added to the model to test hypotheses about the influence of different factors on the neighborhood ratings. These hypothesis are motivated
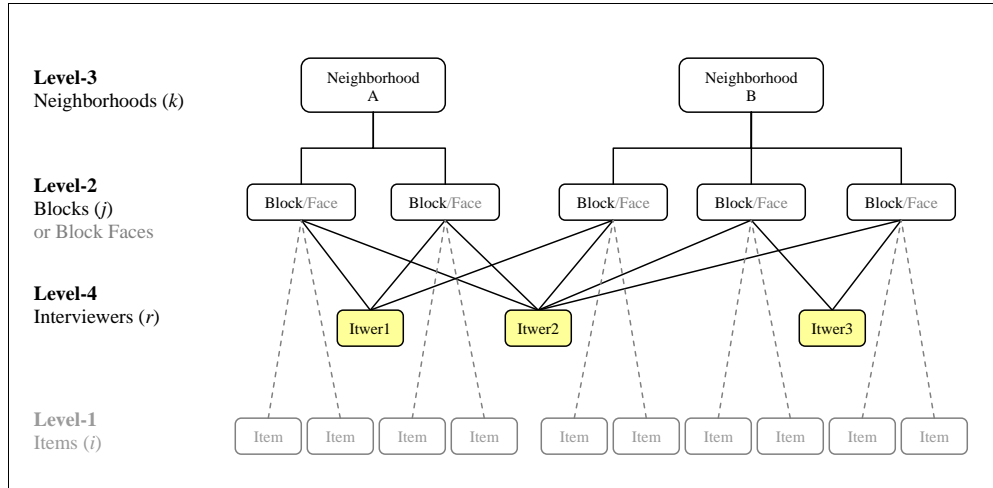
Fig. 3.2: Revised Measurement Error Model for Neighborhood Observations.

below.

Hypotheses about Influences on Neighborhood Ratings

Our understanding of the factors that influence observers' perceptions of disorder is rather limited. Urban sociologists have only documented two influences on neighborhood data reported by trained observers – characteristics of the neighborhoods (*e.g.* ethnic composition) and characteristics of the occasion of measurement (*e.g.* time of day). Most studies about what shapes perceptions of disorder are based on residents' perceptions.

It is an empirical question whether the evidence of 'observer biases' found in studies of residents' perceptions can help inform hypotheses about the biases of non-residents. In particular when those non-residents are subject to training specially designed to reduce observer biases and promote a common understanding of concepts and classification categories. The hypotheses elaborated here will be discussed in

terms of the two mechanisms influencing the perceptions of trained observers – (1) individual biases and (2) specialized training.

Sampson and Raudenbush (2004) found that black and latino residents report less disorder than white residents. The authors argue that the basic psychological mechanisms underlying this phenomena involves the perception of discrepancies based on expectations. Given the history of racial segregation in the United States, it is possible that blacks have been exposed to more disorder than whites in the past, and therefore it may be that blacks and whites judge disorder differently. I would expect the same psychological mechanism to influence the perception of non-residents, and thus would expect lower reports of disorder from non-white observers. Given specialized training, however, I would expect no difference between the reports of white and non-white observers. The following indicator of observers' 'potential exposure to disorder' could be used to test the this hypothesis:

Hypothesis 1 Estimates of disorder and decay collected by 'white' observers will be no different than those collected by 'non-white' observers.

Research shows that residents embedded in networks of social support and greater reciprocal exchange perceive less disorder than those residents that are more socially isolated (Sampson and Raudenbush, 2004). The same mechanisms may not apply to non-residents. Observers who are more 'involved in their communities', for example, could be more critical when rating signs of disorder and decay in other communities. Given specialized training, however, I would expect no difference between the reports of observers with different degrees of community involvement. Three indicators of

'community involvement' could be used to test this hypothesis:

Hypothesis 2 Estimates of disorder and decay collected by observers that reported doing 'some type of community work', having 'children', and being 'married' will be no different than those from other observers.

The study by Andresen et al. (2006) found that, among experienced interviewers, those who were 'local' reported more signs of disorder than the 'non-local'. These results are contrary to the expectations under the literature on fear of crime, which shows that the 'lack of familiarity' with a place is correlated with a heightened sense of insecurity and risk perception (Taylor et al., 1984). I expect this mechanism to apply similarly to residents and non-residents. In effect, I would expect that observers working in 'unfamiliar places' would be more likely to perceive signs of disorder. However, specialized training should reduce this type of bias. Two indicators of 'familiarity with the area' could be used to test this hypothesis:

Hypothesis 3 Estimates of disorder and decay will be no different whether observers are working on 'areas close to their own neighborhoods' or not. Estimates of disorder and decay will be no different whether observers are working on areas in which 'they had prior experiences' or not.

Another causal mechanisms of fear of crime studied at the individual level is known as the 'vulnerability perspective'. This perspective emphasizes individual demographics to explain fear and is based on the assumption that fear is greatest when individuals perceive that they are at a physical disadvantage against potential assaults and/or when individuals believe that they are particularly vulnerable to being

victims of crime (Wyant, 2008). Early research found that women (Clemente and Kleiman, 1977) and the elderly (Lee, 1983) were more fearful of crime – despite the fact that they were less likely to be victimized (Garofalo and Laub, 1978). I expect this mechanism to apply similarly to residents and non-residents. In effect, I would expect that 'vulnerable' observers would be more likely to perceive signs of disorder. Again, though, training may reduce this type of bias. Two indicators of 'vulnerability' could be used to test this hypothesis:

Hypothesis 4 Estimates of disorder and decay collected by 'older' observers will be no different than those of younger observers. Estimates collected by 'female' observers will be no different than those of male observers.

This dissertation aims to assess an array of measurement error properties associated with neighborhood observational data. The next section describes the data and the methods used for analyses.

3.3   Data and Methods

The Los Angeles Family and Neighborhood Study (L.A. FANS) is an ideal dataset with which to estimates parameters of the measurement error model in Figure 3.2. Three features make this study particularly relevant for survey research. First, the study trained a single group of interviewers to collect both survey data and neighborhood data. Second, the study collected multiple independent observations on each sampled block. And third, data collected on the observers themselves allows for testing the research hypotheses associated with observer characteristics.

Earlier chapters of this dissertation have already given a general introduction to the L.A. FANS study. I only describe here the aspects concerning the assessment of measurement error in the neighborhood observational data. Section 3.3.1 describes the dataset used for analyses. Section 3.3.2 describes the protocol used to collect neighborhood observations and it also provides descriptive statistics for the items under analyses. Section 3.3.3 presents the variables used for hypothesis testing and those used as controls. Finally, section 3.3.4 presents an overview of the statistical methods used.

3.3.1   Analytic Datasets

One of the special features of the L.A. FANS data collection effort is that several interviewers completed the neighborhood observations on the same areas *independently* of each other. In addition, interviewers completed their observations on their own work schedule – without having to coordinate with the other observers

collecting observations on the same areas. Because of this flexibility in their work schedule, however, ratings on the same block were often completed at different times of the day, days of the week, or months of the fieldwork period. This feature has consequences for the analysis of interobserver agreement, which will be discussed later.

The dataset containing the neighborhood observations had a total of $5,966$ records, which includes between two and six independent assessments per block. The dataset used for analyses, however, contains $3,998$ records which includes only two ratings per block. I retained only two ratings per block to create a dataset that supported the computation of classical and multilevel estimates of agreement. This reduced dataset is consistent with most neighborhood studies that collect only two ratings per block face.

The development of the analytic dataset proceeded as follows. For all blocks with more than two ratings, I analyzed only the pair of ratings that was closest in time, using the number of elapsed days between ratings as criterion. To begin, 267 block faces were dropped because of mismatches between the two set of ratings.[4] Then, 27 block faces were dropped because of missing data in all neighborhood observations, and 145 block faces were dropped due to the exclusion of interviewers with too few ratings.[5] Through these manipulations, 16 block faces were left with

---

[4] An example of a mismatch is when one block had 5 block faces rated by the first rater but only 4 block faces rated by the second rater. In that case, only the block faces with matching identification numbers were retained (*i.e.* block faces #1 through #4) and the unmatched block face (#5) was dropped from the dataset.

[5] Seven of the 35 interviewers rated fewer than 30 block faces, compared to a median number

only a single rating per block and thus had to be dropped from the analytic dataset too. Finally, $1,513$ block faces were dropped which corresponded to the 3rd, 4th, 5th and 6th rating for those blocks with multiple ratings. Table 3.2 presents a comparison of the original and the final data sets.

Tab. 3.2: Number of observations in the original and analytic datasets.

	Original Data	Analytic Data
Number of data records	5,966	3,998
Number of repeated observations	2-6	2
Number of unique Interviewers	35	28
Number of unique Tracts	65	65
Number of unique Blocks	422	419
Number of unique Block Faces	2,029	1,999

The final data set consists of $3,998$ block faces, where 22% of the block faces were observed on the same day, 32.1% within $1-3$ days, 18.6% within $4-7$ days, 13.4% within $8-14$ days, 11.0% within $15-90$ days and 2.2% within more than 90 days.[6]

3.3.2 Neighborhood Observations

L.A.FANS interviewers collected observations of the physical environments for 419 Census blocks in the 65 Census tracts that were included in the analytic sample. A special data collection instrument and training protocol was designed, and survey interviewers were trained to observe and collect data from all 'faces' of the selected

of block face of 108 ratings per interviewer. To ensure enough variability those seven interviewers, and the 45 block faces they rated, were dropped from the analytic dataset.

[6] Note that the data set used here differs from the data set used in Paper 1, which retained only the first interviewer that rated each block and only dropped the 27 records without any observational data.

blocks in a *systematic and standardized manner.* This section describes the L.A. FANS data collection protocol and the neighborhood items used in the analyses.

Neighborhood Observation Protocol

It is important to note that, in the L.A. FANS study, the tasks of collecting neighborhood observations was assigned to interviewers who were also in charge of conducting the household interviews. This approach is similar to how regular surveys collect neighborhood observations, however it is unique among neighborhood effects studies and very relevant for survey research methods.

Interviewers were trained to carry out their observations systematically, but fairly quickly. The observation protocol involved driving around the entire block, and walking down each block face and recording the characteristics of that block face at the end of the walk. Interviewers were instructed to complete these observations the first time they visited the sampled block. Neighborhood observations were conducted between April 2000 and July 2001, with one third (27%) being done in April and May of 2000, and the remainder (63%) in April and May of 2001.

For their assessments, the interviewers used the *Neighborhood Observation Forms* (NOF). The NOF is a paper and pencil instrument consisting of 4 forms, the *Neighborhood Observations Cover Page*, the *Block Face Observation Form*, a *Social Observation Form*, and the *Alley Observation Form.* The four forms are available in the Appendix. Figure 3.3 helps illustrates which information was collected in what way in a typical Census block.

The Neighborhood Observations Cover Page contained 9 items, and interview-
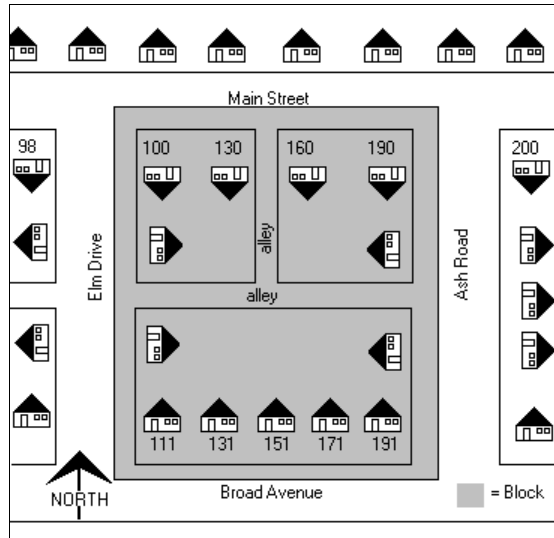
Fig. 3.3: Illustration of a typical Census Block in an urban settings.

ers were instructed to complete it at the beginning of the rating task. Looking at Figure 3.3, the information recorded on the NOF Cover Page is common to all four block faces in the grey shaded Census block. The Alley Observation Form collected observations about any alleys present in a given Census block.

In contrast, items collected on the Block Face and Social observation forms had to be completed for each block face in the block. Thus the values recorded here differ for the four sides shown in Figure 3.3. The Block Face Observation Form contained 42 items in total, spanning topics like physical disorder, residential decay, condition of the streets and presence of businesses and institutions on the block face. Those observations were supplemented by information collected on the *Social Observation Form*, which contained 23 social disorder items such as presence of gangs or prostitutes on the block face.

95

The Neighborhood Items

L.A. FANS collected a total of 64 block face attributes. In this paper I analyze the 25 items that correspond to neighborhood observations of *physical disorder* (i=8), *social disorder* (i=7), *residential decay* (i=5) and *residential security* (i=5). Signs of disorder and decay, also know an *'incivilities'*, are the focus of the analyses because they are theoretically linked to the mechanisms explaining cooperation in household surveys (see Paper 1). Items on residential security are included in the analyses because similar variables are currently being collected by large survey projects.

Most of the items were collected using 4 and 5-category Likert-type questions, but a few were collected using yes/no questions. All Likert-type items were dichotomized so that 1 means 'presence' and 0 means 'absence' of the characteristic being rated. Exact wording for each item and their original scales are displayed in Tables 3.5 and 3.6. Estimates of the prevalence of each item based on all $3,998$ records is also provided.

Items in the *physical disorder* scale captured a wide spectrum of the disorder phenomena. Consistent with the literature, items considered 'less severe' (*e.g.* litter) were reported much more frequently than 'more severe' items (*e.g.* drugs). Items in the *social disorder* scale showed lower prevalence, which is consistent with indications of the higher severity of the observations covered.

This section introduces the variables used to operationalize the hypotheses about factors influencing interviewer's perceptions of disorder and decay in urban neighborhoods (see section 3.2.3).

Characteristics of Interviewers

Interviewer characteristics were recorded in the *Interviewer Background Questionnaire.*[7] In addition to the usual demographic characteristics (*e.g.* age, education, race) the interviewer questionnaire also captured information about the interviewers' own neighborhoods (*e.g.* city of residence, how satisfied, how long lived there). The interviewer questionnaire is available in the Appendix. Descriptive statistics for the variables used in this paper are displayed in Table 3.3.

The first variable in the table is *race*, which was used here as an indicator of 'potential exposure to disorder'. The second variable is *age* and it represents a correlate of 'vulnerability'. The questionnaire did not ask about interviewer's *gender*, so I was not able to test this additional indicator of vulnerability. The third set of variables aimed to capture interviewer's 'community involvement' and was derived from questions on *marital status*, *presence of children*, and *community activities.*[8]

---

[7] This data set is not part of the public or the restricted L.A. FANS datasets. I am grateful to the L.A. FANS team for allowing me to use those data here.

[8] The exact wording of the question on community activities was "*How active are you in community activities or organizations or in volunteer activities?*". The variable *community activities* was coded 1=(very active, somewhat active) and 0=(not active).

Table 3.3 shows that, while white interviewers represented 35% of the interviewer crew, they collected 51% of the observations. Interviewers that were married, on the other hand, represented 58% of the interviewers but only collected 36% of the observations. When looking at the analytic dataset ($n = 3,998$) it is important to remember that this data reflects 'workloads' and not the distribution of the characteristics of the crew of interviewers.

Tab. 3.3: Descriptive Statistics for Interviewer Characteristics. Unweighted estimates.

Indicators of		Interviewer Data (n=28)	Block Face Data (n=3,998)
Potential Exposure			
White	No	65.4%	48.6%
	Yes	34.6%	51.4%
Vulnerability			
55+ yrs	No	92.0%	84.4%
	Yes	8.0%	15.6%
Community Involvement			
Ever Married	No	57.7%	36.0%
	Yes	42.3%	64.0%
Has Kids	No	65.4%	53.4%
	Yes	34.6%	46.6%
Comm. Activities	No	46.2%	54.9%
	Yes	53.8%	45.1%

The characteristics described here are considered 'fixed' for each interviewer, *i.e.* they do not change as the field work progresses. The next section describes a different set of variables that do not correspond to particular characteristics of interviewers or the blocks they are rating, but to the *interaction* of both.

The *Neighborhood Observations Cover Page* records features that cannot be classified as being 'fixed' properties of either blocks or interviewers, because they only arise though the interaction among both. The 'time of day' at which a block was rated, for example, is not a property of the block itself – the same block could have been rated at a different time if a different interviewer was assigned to it. The 'time of day' is also not a property of the interviewer herself since – the same interviewer can choose a different time to rate another block. I this decided to label these variables 'characteristics of the occasion of measurement'.

Descriptive statistics for the variables used in this paper are displayed in Table 3.4. In principle, the distribution of the variables based on the block-level data ($n = 419$) and that based on the block face level data ($n = 3,998$) should be exactly the same. As explained in section 3.3.1, small discrepancies arise due to the process of dropping unmatched block faces for analyses.

The first two variables were used to test hypotheses about the influence of interviewers' 'familiarity with the area' on the ratings. The variable *Neighborhood Close* indicates if the distance between the interviewer's neighborhood and the tract rated by her is less or equal than 5 miles.[9] The variable *Experience with Block*

---

[9] L.A. FANS contains geographic positioning data for each household in the sample. For this analysis, however, I used longitude and latitude associated with the centroid of each Census tracts in the sample. Information on the interviewer's place of residence was available in the Interviewer Background Questionnaire. I identified 9 different cities of residence among the 28 interviewers completing the neighborhood ratings. I assigned the longitude and latitude coordinates of the

Tab. 3.4: Descriptive Statistics for Characteristics of the Measurement Occasion. Un-weighted estimates.

Indicators of		Block Data (n=419)	Block Face Data (n=3,998)
Familiarity with Area			
Neighborhood Close	No	87.3%	87.1%
	Yes	12.7%	12.9%
Experience with Block	No	22.9%	23.4%
	Yes	77.1%	76.6%
Temporal Variability			
Rated after 5pm	No	88.7%	90.9%
	Yes	11.3%	9.1%
Rated on Weekend	No	74.2%	72.7%
	Yes	25.8%	27.3%

indicates whether the interviewer had any type of previous experience with the block (e.g. listing, interviewing) or not.

The last two variables attempt to capture the effect of 'temporal variability' (Raudenbush, 2003) on the ratings. For the purpose of this paper, I used an indicator of time of day (*Rated After 5pm*) and one indicator of day of week (*Rated on Weekend*) to control for the influence of temporal variability on perceptions of disorder.

centroid of each city to each interviewer. I used the *geodist* procedure in Stata 10 to calculate ellipsoidal distances between two geo-referenced points – tracts' location and interviewers' residence location. The procedure uses Vincenty's equations to approximate the distance between two points on the earth's surface (Vincenty, 1975).

Neighborhood Structural Characteristics

The final set of correlates correspond to neighborhood level attributes associated with neighborhood socio-economic composition, such measures are typically derived from Census records. Following Sampson and Raudenbush (1999) I used three measures of socio-economic composition: *immigrant concentration, concentrated disadvantage* and *concentrated affluence*, plus an indicator of *population density*.

The indicator of population density is already available in the L.A. FANS data. The measures of socio-economic composition were already available from the factor analysis developed for the analyses of response propensity. Details about the factor analysis and the definition of the factors are available in Paper 1. All neighborhood variables were standardized to have a mean of zero and a standard deviation of one.

Tab. 3.5: Wording and Descriptive Statistics for the Physical Disorder and Social Disorder Items.

Neighborhood Items	Label	Perc.	n	Scale[†]	Question[††]
Physical Disorder Items (n=8)					
Are there abandoned cars on the street or in alleys or lots?	cars	9.8%	3,998	1-4	10
Is there trash or junk on the street or sidewalks, in yards/lots?	trash	52.7%	3,998	1-4	11
Is there garbage, litter, or broken glass on the street or sidewalk, in yards, or vacant lots?	litter	73.6%	3,998	1-4	12
Are there needles, syringes, condoms, or drug-related paraphernalia on the street or sidewalk, in yards/lots?	drugs	3.4%	3,998	1-4	13
Are there empty beer containers or liquor bottles on the street or sidewalks, in yards, or vacant lots?	bottles	21.0%	3,998	1-4	14
Are there cigarettes or cigar butts or discarded cigarette packages on the street or sidewalks, in yards/lots or gutters?	cigars	59.7%	3,998	1-4	15
Is there graffiti on buildings, sidewalks, walls, or signs?	graffiti	53.5%	3,998	1-4	16
Is there painted-over graffiti on buildings, sidewalks, walls, or signs?	pograff	36.3%	3,998	1-4	17
Social Disorder Items (n=7)					
Did any of the groups of teens you saw appear to be a gang?	gang	1.2%	3,962	1-0	8, 9
Did you see any adults on the block face loitering, congregating or hanging out?	loitering	8.6%	3,982	1-0	11
Did you see any prostitutes on the block face?	prostit	0.3%	3,986	1-0	12, 13
Did you see any homeless people or people begging on the block face?	homeless	2.0%	3,996	1-0	14, 15
Did you see people who were selling illegal drugs on the block face?	selling	0.5%	3,996	1-0	16, 17
Did you see any people drinking alcohol openly on the block face?	drinking	2.3%	3,996	1-0	18
Did you see any drunken or otherwise intoxicated people on the block face?	intox	1.3%	3,996	1-0	19, 20

Notes: (†) The original scale for the physical disorder items was: 1=None; 2=A little; 3=Some; 4=A lot. The original scale for the social disorder items was: 1=Yes; 0=No.

(††) Question number correspond to the original numbers in the *Block Face Observation Form* (all physical disorder items) and the *Social Observation Form* (all social disorder items).

Tab. 3.6: Wording and Descriptive Statistics for the Residential Decay and Residential Security Items.

Neighborhood Items	Label	Perc.	n[†††]	Scale[†]	Question[††]
Residential Decay Items (n=5)					
What is the overall condition of the residential buildings?	bldgs	84.6%	3,627	1-5	21
How many houses/apartments are burned out, boarded	boarded	10.0%	3,627	1-5	22
How many vacant lots are there on this block?	vacant	16.0%	3,627	1-5	23
How many houses/apartments have peeling paint or damaged exterior walls?	walls	66.7%	3,627	1-5	24
How many houses/apartments have well-tended yards or gardens?	yards	77.1%	3,627	1-5	25
Residential Security Items (n=5)					
How many houses/apartments have window bars or gratings on doors or windows?	barswin	63.8%	3,627	1-5	26
How many houses/apartments have signs indicating they are protected by private security services?	secsign	52.2%	3,627	1-5	27
How many houses/apartments have signs indicating they are protected by dogs?	dogsign	32.4%	3,627	1-5	28
How many houses/apartments have security gates or security fences?	gates	59.2%	3,627	1-5	29
Are there signs indicating there is a neighborhood watch on this block?	ngwatch	17.7%	3,618	1-0	30

Notes: (†) The original scale for most of the residential decay and residential security items was: 1=None; 2=Very few; 3=Some; 4=Many; 5=All. The item *yards* had to be reverse coded before analysis. The item *bldgs* had a different scale (1=Very poor; 2=Poor; 3=Fair; 4=Very good; 5=Excellent) and also had to be reverse coded before analysis.

(††) Question number correspond to the original numbers in the *Block Face Observation Form* for all residential decay and residential security items.

(†††) Sample sizes are smaller for the residential decay and residential security items because they were not collected for block faces rated as 'nonresidential'.

3.3.4 Statistical Methods

Analyses of Random Influences

Initial estimates of interobserver agreement were based on classical measures like Percentage Agreement Index $(\hat{\delta}_1)$ and the Kappa statistic $(\hat{\delta}_2)$ (Hintze, 2005). The next set of results were based on a form of the Intraclass Correlation Coefficient $(\hat{\delta}_3)$. This estimate relaxed the classic assumption of a single source of random variation in the variable under analysis. Following Raudenbush and Sampson (1999) I derived a measure of the intraclass correlation coefficient using multilevel modeling techniques.

The measurement error model proposed in this paper (Figure 3.2) can be adapted to analyze individual items or sets of items. In this paper I focused on the analysis of individual items, thus I conceptualized the observational data as having three levels of clustering: *tracts (k), blocks (b)* and *interviewers (r)* (Figure 3.4). The labels O_m and O_p in the first block face are a reminder of the fact that each pair of ratings could have been completed on two different occasions for each block face. The structure of the data combines the nesting of the geography levels and the cross-classification of 'interviewers and blocks' and 'interviewers and tracts'. For simplicity, I only modeled the crossing of interviewers and tracts in this paper.

The multilevel structure of the neighborhood data is represented in model (3.1), where y_{ijkr} is an indicator taking on a value of 1 if a neighborhood characteristic, *e.g. graffiti*, is observed in block face i of block j of tract k, rated by interviewer r, with $y_{ijkr} = 0$ if not. The model includes three random effects to take

104
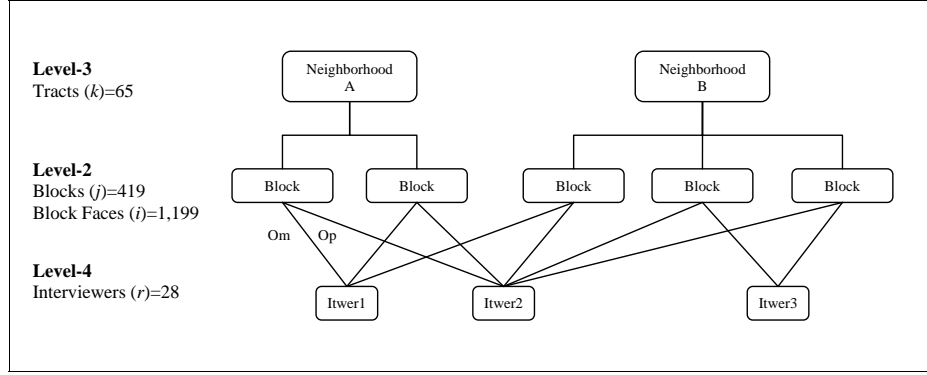
Fig. 3.4: Measurement Error Model for Neighborhood Items.

into account the dependency among the observations within tracts (t_k), blocks (b_{jk}) and interviewers (i_r). The term e_{ijkr} reflects residual variability associated with the 'block faces' and the interaction of 'tracts and interviewers'. The random effects follow the distributional assumptions in equation (3.1), where IID stands for Independent and Identically Distributed. To be consistent with the data structure, the random effects of tracts (γ_t^2) and interviewers (γ_i^2) are modeled as crossed factors and the random effect of blocks (γ_b^2) is modeled as nested within tracts.

$$y_{ijkr} = \beta + t_k + b_{jk} + i_r + e_{ijkr} \tag{3.1}$$

$$t_k \sim IID(0, \gamma_t^2)$$

$$b_{jk}|t_k \sim IID(0, \gamma_b^2)$$

$$i_r \sim IID(0, \gamma_i^2)$$

$$e_{ijkr} \sim IID(0, \gamma_e^2)$$

The model in (3.1) corresponds to an 'unconditional' Linear Probability Model because it fits the probability of observing *graffiti* as a function of an overall mean (β) without covariates, and it uses a linear model to estimate this probability. The

model was fitted using the *xtmixed* command in Stata 10. This command fits linear mixed models, which are characterized as containing both fixed effects and random effects. The overall error distribution of the linear mixed model is assumed to be Gaussian, but heteroskedasticity and correlations within lowest-level groups may also be modeled. By default, *xtmixed* specifies that the model be fit using restricted maximum likelihood (REML). For my analyses, I specified that the model be fit using maximum likelihood (ML). This specification makes the additional assumption of normality beyond that of IID in (3.1). The assumption of normality is not crucial to the properties of the resulting estimators of the variance components derived below. Those estimators are consistent for distributional assumptions other than normal, and in particular include the general class described in (3.1).[10]

Estimates of variance components derived from model (3.1) produce a measure of reliability that corresponds to the correlation between two observations made on the *same block face* by *different interviewers*. This form of the intraclass correlation, defined by expression (3.2), attempts to measure the degree of agreement between the two observers rating the same block face. One important limitation of the L.A. FANS data, for the purpose of estimating interobserver agreement, is that most of the time the two observation of each block face were made at two different time points. Among the 1,999 pairs of observations under analysis, 77.3% were observed on different days. And among those observed on the same day, 35.5% were observed at different times of the day. Thus, an estimate of agreement between the two

---

[10] See Stata Manual [XT] for more details on *xtmixed* (StataCorp, 2007, 282). See the Appendix for more details on the precise parametrization of the model implemented here.

observations made on each block face cannot completely disentangle interobserver variability from temporal variability. To improve our understanding of the variability associated with using a different observer, estimates of agreement will be provided based on the full sample ($n = 3,998$), the 'same day' sample ($n = 908$), and the 'same day and time' sample ($n = 586$).

Model (3.1) can also be used to derive estimates of interviewer effects ($\hat{\rho}_{int}$) and sampling point effects ($\hat{\rho}_{sp}$). These estimates represent, respectively, the correlation between two observations made by the *same interviewer* in *different block faces* (eq. 3.3), and the correlation between two observations made on the *same tract* by *different interviewers* on *different blocks* and *block faces* (eq. 3.4). The derivation of these expressions are given in the Appendix.

$$\hat{\delta}_3 = Corr(y_{ijkr}, y_{ijkr'}) = \frac{\hat{\gamma}_t^2 + \hat{\gamma}_b^2}{\hat{\gamma}_t^2 + \hat{\gamma}_b^2 + \hat{\gamma}_i^2 + \hat{\gamma}_e^2} \tag{3.2}$$

$$\hat{\rho}_{int} = Corr(y_{ijkr}, y_{i'j'k'r}) = \frac{\hat{\gamma}_i^2}{\hat{\gamma}_t^2 + \hat{\gamma}_b^2 + \hat{\gamma}_i^2 + \hat{\gamma}_e^2} \tag{3.3}$$

$$\hat{\rho}_{sp} = Corr(y_{ijkr}, y_{i'j'kr'}) = \frac{\hat{\gamma}_t^2}{\hat{\gamma}_t^2 + \hat{\gamma}_b^2 + \hat{\gamma}_i^2 + \hat{\gamma}_e^2} \tag{3.4}$$

Estimates of $\hat{\rho}_{int}$ and $\hat{\rho}_{sp}$ will be used to compare the extent of interviewer effects across different neighborhood items and to gauge the magnitude of interviewer effects. The latter is challenging. On one side, no estimates of $\hat{\rho}_{int}$ for neighborhood items are available from the urban sociology literature. On the other side, comparing these estimates to those previously found in the survey literature for survey responses is tricky, because an additional variance component, associated with the respondent, is included in the survey items which does not arise in the neighborhood items.

Due to the different nature of the neighborhood items I suggest to use a relative

measure, rather than the direct estimate $\hat{\rho}_{int}$, as an indicator of the error associated with the interviewers. To accomplish this goal, I used the ratio $\frac{\hat{\rho}_{sp}}{\hat{\rho}_{int}}$ as a measure of the relative size of true score variance $(\hat{\rho}_{sp})$ to measurement error variance $(\hat{\rho}_{int})$ in the neighborhood ratings.

<div align="center">Analyses of Systematic Influences</div>

Analyses described above were used to assess different measurement properties conceptualized as random variation. Analyses described in this section assessed systematic sources of variation associated with the interviewers, the occasion of measurement, and the neighborhoods.

Initial analyses produced estimates of the probability of seeing disorder as a function each one of the covariates hypothesized to influence the neighborhood observations. These probabilities were estimated by $\bar{y}_c = \left(\frac{\sum_{i \in c} y_i}{n_c} \right)$ where, for each item y, the numerator represents the sum of all the binary ratings (y_i) in category c of the covariate under analysis, and the denominator (n_c) is the total number of observations in that category.

The next set of analyses tested the significance of the relationships in a setting that takes into account the multilevel structure of the data. For this purpose, I extended model (3.1) to including the seven covariates hypothesized to influence perceptions of disorder. I also added controls for the six variables that had already been found influential on perceptions of disorder. The extended model is given by

the equation below,

$$y_{ijkr} = \beta + \sum_{p=1}^{5} \alpha_p \mathbf{I}_{pr} + \sum_{q=1}^{6} \theta_q \mathbf{B}_{qjr} + \sum_{s=1}^{2} \tau_s \mathbf{T}_{st} + t_k + b_{jk} + i_r + e_{ijkr} \qquad (3.5)$$

$$t_k | \mathbf{I}_{pr}, \mathbf{B}_{qjr}, \mathbf{T}_{st} \sim IID(0, \gamma_t^{2*})$$

$$b_{jk} | t_k, \mathbf{I}_{pr}, \mathbf{B}_{qjr}, \mathbf{T}_{st} \sim IID(0, \gamma_f^{2*})$$

$$i_r | \mathbf{I}_{pr}, \mathbf{B}_{qjr}, \mathbf{T}_{st} \sim IID(0, \gamma_i^{2*})$$

$$e_{ijkr} | \mathbf{I}_{pr}, \mathbf{B}_{qjr}, \mathbf{T}_{st} \sim IID(0, \gamma_e^{2*})$$

where $\mathbf{I}_{pr}$ is a vector containing five interviewer variables ($p = 5$), $\mathbf{B}_{qjr}$ is a vector containing four variables derived from the interaction of interviewers and blocks ($q = 4$), and $\mathbf{T}_{ts}$ contains four tract-level variables ($s = 4$). The regression coefficients associated with these covariates are α_p, θ_q and τ_s. The same procedures used to estimate the unconditional model in equation (3.1) were used to estimate the conditional model in equation (3.5).

3.4 Analyses and Results

3.4.1 How much do trained observers agree about their ratings?

The results presented here addressed the question of interobserver agreement associated with neighborhood items. In this study, this question amounts to: how large are the discrepancies between the block face observations collected by two independent observers on two separate occasions? Figure 3.5 displays the 65 tract level estimates of the prevalence for ten items using data from the first occasion of measurement ($Perc.\#1, n_1 = 1,999$), and the second occasion of measurement ($Perc.\#2, n_2 = 1,999$). The dispersion of the points in the graph illustrates the extent of disagreement based on estimates derived on two different occasions of measurement.
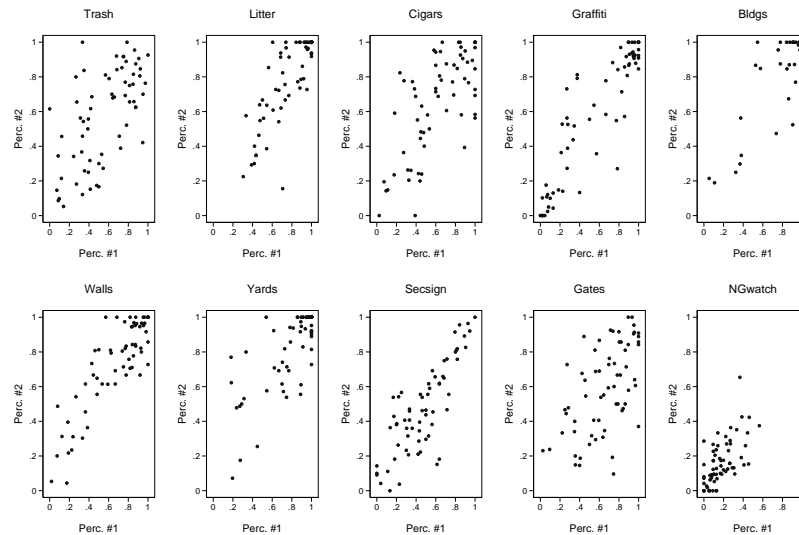
Fig. 3.5: Tract level Estimates of Disorder on Two Different Occasions.

Classical estimates of the extent of disagreement between the two binary outcomes – the percentage agreement ($\hat{\delta}_1$) and the Kappa statistic ($\hat{\delta}_2$) – are provided in Figure 3.6 using black and gray symbols respectivey. Percentage agreement was relatively high across all items ($min = 0.66, max = 0.99$). The social disorder items, however, achieved the highest scores ($min = 0.87, max = 0.99$). This result is not surprising since, given the 'severity' of the disorder items, most of the time the two independent observers agreed on 'not having seen' signs of social disorder.

Once agreement due to chance is taken into account by the Kappa statistic, the performance of the disorder items decreased ($min = 0.00, max = 0.17$). The Kappa statistics for the remaining items varied considerably ($min = 0.12, max = 0.62$).

Estimates of interobserver agreement derived from the *intraclass correlation coefficient* $\hat{\delta}_3$ showed the same pattern as the kappa statistic ($min = 0.05, max = 0.56$). These results are displayed in Figure 3.7a. Tables with the point estimates for all three estimates of agreement are available in the Appendix.
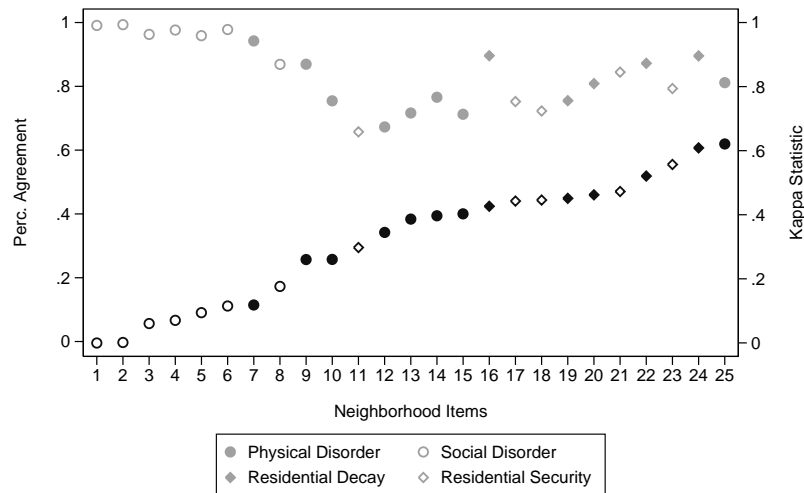
Fig. 3.6: Estimate of Percentage Agreement ($\hat{\delta}_1$) and Kappa ($\hat{\delta}_2$).

One of the purposes of this study was to provide benchmarks for surveys collecting observational data on neighborhood characteristics. Three items currently collected by some major survey projects include 'condition of buildings in the area', presence of 'graffiti' and presence of 'trash or litter'. In this study, estimates of interobserver agreement for *graffiti* ($\hat{\delta}_3 = 0.56$) and *condition of buildings* were moderate ($\hat{\delta}_3 = 0.45$), and those for *litter* ($\hat{\delta}_3 = 0.29$) and *trash* ($\hat{\delta}_3 = 0.26$) were low.

In the survey research context, the comparison of these estimates with other types of observational data collected by interviewers is not straight forward. Previous reports of agreement between interviewers and respondents' reports are not directly comparable (e.g. Pickering et al. (2003); West (2010)). The only comparable results come from the British Crime Survey (BCS), where estimates of percentage agreement between pairs of observers on 'condition of buildings in the area' achieved $\hat{\delta}_1 = 0.63$ (Sinibaldi, 2010). This estimate is relatively low compared with to the L.A. FANS estimate $\hat{\delta}_1 = 0.87$ for 'condition of residential buildings' in the block face. The BCS estimate, however, was derived based on observations taken a year apart.

Overall, estimates of agreement for the L.A. FANS are lower than most agreement rates reported in the urban sociology literature, which most likely reflects the fact that most of the paired observations (77%) were collected on separate days. Thus, the low estimates of agreement reported here not only reflect observer variability but also temporal variability.

To improve our understanding of the variability associated with using different

observers, estimates of agreement were derived using the full sample ($n = 3,998$), the 'same day' sample ($n = 908$), and the 'same day and time' sample ($n = 586$). A table with these results is available in the Appendix. Estimates of agreement increased from an average of $\hat{\bar{\delta}}_3 = 0.25$ (full sample) to $\hat{\bar{\delta}}_3 = 0.26$ (same day) to $\hat{\bar{\delta}}_3 = 0.30$ (same day and time) across all items. For items such as *trash*, agreement increased from $\hat{\delta}_3 = 0.26$ to $\hat{\delta}_3 = 0.28$ and to $\hat{\delta}_3 = 0.41$. In contrast, agreement decreased for other items such as *security gates* ($\hat{\delta}_3 = 0.28$; $\hat{\delta}_3 = 0.19$; $\hat{\delta}_3 = 0.17$). These last results is puzzling.

3.4.2 How large are interviewer clustering effects?

This paper's second research question was concerned with the magnitude of interviewer ($\hat{\rho}_{int}$) and sampling point ($\hat{\rho}_{sp}$) clustering effects. Figures 3.7b and 3.7c displays these results graphically. A table with the point estimates is available in the Appendix.

Estimates of ρ_{int} ranged from 0.01 to 0.22, with the average at $\hat{\bar{\rho}}_{int} = 0.07$. Estimates $\hat{\rho}_{sp}$ ranged from 0.01 to 0.48, with the average at $\hat{\bar{\rho}}_{sp} = 0.17$. Two results are interesting to discuss here: (1) the relative size of interviewer effects across different items, and (2) the relative size of interviewer effects to sampling point effects.

Items such as *vacant lots* ($\hat{\rho}_{int} = 0.01$), *boarded up* housing ($\hat{\rho}_{int} = 0.02$) and presence of *gangs* ($\hat{\rho}_{int} = 0.02$) showed relatively low interviewer effects, suggesting that observers have a relatively clear – and common – idea about how to rate these features. In the context of this study, this 'common ground' most likely comes from

the special training they received prior to conducting the observations. For other

items, the influence of interviewer 'individual judgments' was probably stronger.

Examples of items with large interviewer effects include *cigarettes* ($\hat{\rho}_{int} = 0.22$),

litter ($\hat{\rho}_{int} = 0.17$) and *security gates* ($\hat{\rho}_{int} = 0.16$).

The estimates of $\hat{\rho}_{int}$ derived from L.A. FANS for neighborhood items are

much larger than those reported in the survey literature for survey responses. The

direct comparison of $\hat{\rho}_{int}$ for survey items and neighborhood items, however, is not

appropriate. A better indicator of the relative size of the interviewer-related error

in the neighborhood data is $\frac{\hat{\rho}_{sp}}{\hat{\rho}_{int}}$.

The item that performed best, *vacant lots*, showed 21 times more true score

variance ($\hat{\rho}_{sp}$) than measurement error variance ($\hat{\rho}_{int}$). The item with the poorest

performance, presence of *gangs*, showed up to 2 times more measurement error

variance than true score variance. The ratio $\frac{\hat{\rho}_{sp}}{\hat{\rho}_{int}}$ averaged 8.2 for the *residential*

security items, 4.05 for the *residential decay* items, 2.33 for the *physical disorder*

items, and 1.81 for the *social disorder* items. I interpret these results as an indication

that the extent of measurement error variance compared to the true score variance

is low for neighborhood items. To my knowledge, there is no benchmark for the

ratio $\frac{\hat{\rho}_{sp}}{\hat{\rho}_{int}}$ for survey items. A table with these results is available in the Appendix.

3.4.3   What factors have a systematic influence on perceptions of disorder?

Results in the previous section provide evidence of the magnitude of inter-

viewer random effects on neighborhood items. Analyses in this section test whether

part of this variability can be explained by fixed characteristics of the interviewers,

Fig. 3.7: Estimates of Intraclass Correlation from the Linear Unconditional Models

(*a*) Interobserver Agreement ($\hat{\delta}_3$)

(*b*) Interviewer Effect ($\hat{\rho}_{int}$)

(*c*) Sampling Point Effect ($\hat{\rho}_{sp}$)
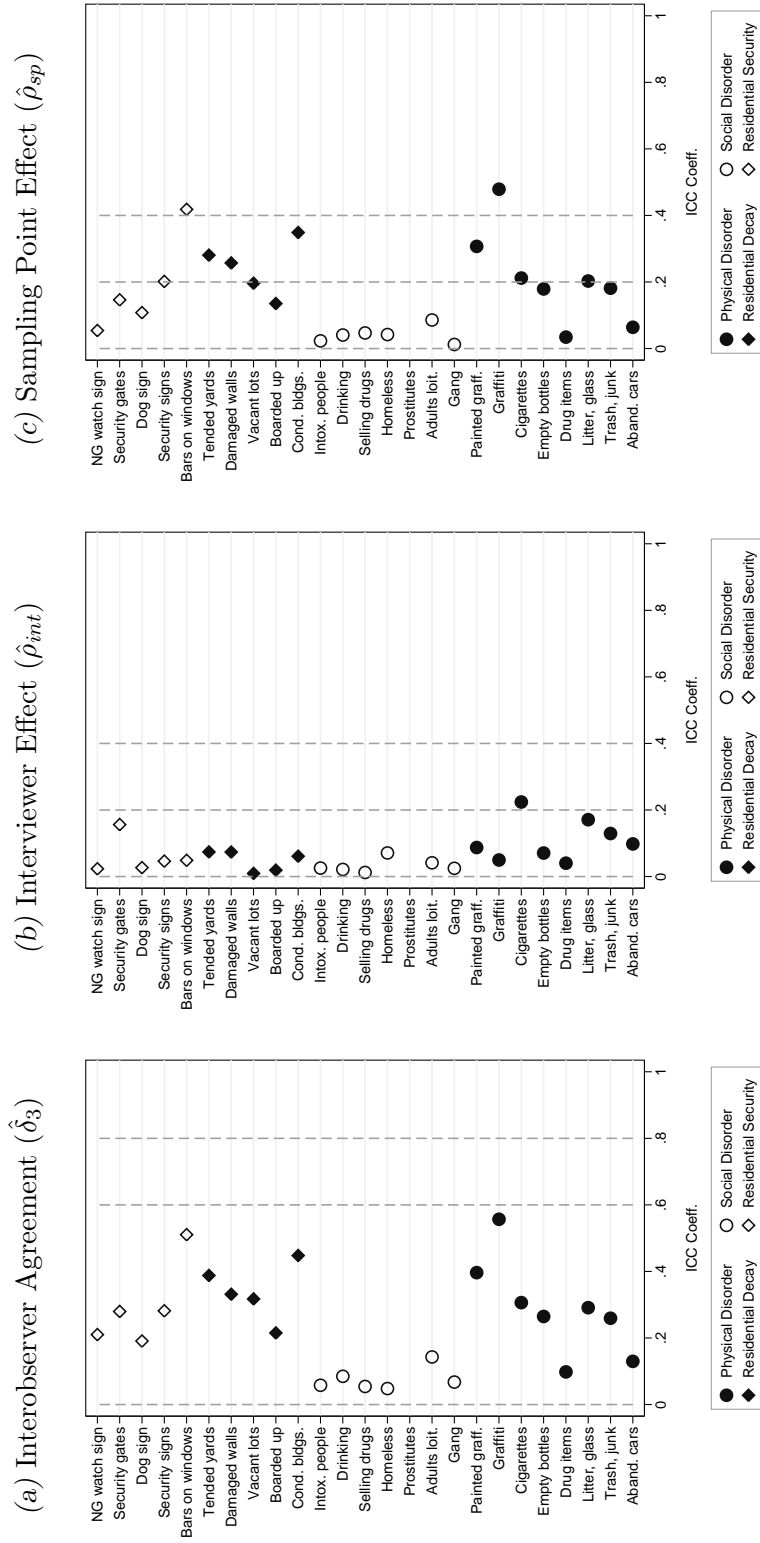
the occasion of measurement, or the social composition of the neighborhoods.

Results of Bivariate Analyses

Initial bivariate analyses provided estimates of the probability of seeing disorder as a function of a single covariate at a time. Results showed mixed support for the hypotheses formulated in section 3.2.3. Table 3.7 illustrates the patterns of the results: a "+" sign indicates higher probability of seeing disorder and a "-" sign indicates lower probability of seeing disorder. Only statistically significant results are displayed in the table. Tables with the estimated probabilities are available in the Appendix.

Results in the first panel were used to test whether observers influenced by 'prior exposure to disorder' (*e.g. non-whites*) perceived fewer signs of disorder than those not exposed (*e.g. whites*). The pattern of results was mixed. *White* interviewers were more likely to see some signs of disorder (*e.g. trash, cigars, gates*), but they were also less likely to see others (*e.g. graffiti, walls, yards*). The indicator of 'vulnerability' also showed an inconsistent pattern. *Older* interviewer were more likely to see cigarettes (72%) than younger interviewers (60%), but they were also less likely to see damaged walls ($43\%vs.69\%$).

The indicators of 'community involvement', on the other hand, showed a very consistent pattern. Interviewers with children, for example, were more likely to see *cigarettes* ($73\%vs.51\%$), deteriorated *buildings* ($88\%vs.79\%$) and deteriorated *walls* ($71\%vs.59\%$) than those without children. The same pattern hold for interviewer that were 'ever married' and interviewers that were involved in 'community activi-

ties'. These results provided preliminary support for the mechanism which suggests that interviewers who are more engaged in their own communities will report more signs of disorder.

Finally, the indicators of 'familiarity with the area' showed consistent, but contradictory results. As hypothesized, interviewers that lived close to the areas they were rating (*neighborhood close*) were less likely to report signs of disorder and decay. However, interviewers were also more likely to report disorder in areas where they had some previous experience.

Results from Multilevel Multivariate Analyses

Results from these bivariate analyses provide only preliminary tests of the hypotheses under study. Among the limitations of these analyses are the lack of control for confounding variables and the inadequate modeling of the multilevel structure of the data. Multilevel analyses in this section addressed these shortcomings.

Table 3.8 displays the estimated coefficients for the conditional multilevel model in equation (3.5) for two representative items: *trash*, which showed poor measurement properties ($\hat{\delta}_3 = 0.26, \hat{\rho}_{int} = 0.13, \hat{\rho}_{sp} = 0.18$) and *graffiti*, which showed the best measurement properties ($\hat{\delta}_3 = 0.56, \hat{\rho}_{int} = 0.05, \hat{\rho}_{sp} = 0.48$).

The panel in the bottom displays the estimates of the random effects coefficients for interviewers (γ_i^2), tracts (γ_t^2), blocks (γ_b^2), and the residual (γ_e^2) – which incorporates the variability associated with the 'block faces' and the interaction between 'tracts and interviewers'. Stars next to each coefficient represent the significance level for the test of the corresponding variance component equal to zero.

117

Probability of Perceiving signs of Disorder and Decay.

Indicators of		Physical Disorder Items				Res. Decay Items			Res. Security Items		
		trash	litter	cigars	graffiti	bldgs	walls	yards	secsign	gates	ngwatch
Potential Exposure											
Race White	No	-		-	+		+	+		-	
	Yes	+		+	-		-	-		+	
Vulnerability											
Age $55 + yrs$	No		+	-	+		+	+		-	
	Yes		-	+	-		-	-		+	
Community Involvement											
Comm. Activities	No		-		-	-	-	-	+	-	
	Yes		+		+	+	+	+	-	+	
Has Kids	No	-	-	-	-	-	-	-		-	
	Yes	+	+	+	+	+	+	+		+	
Ever Married	No	-	-	-		-	-				
	Yes	+	+	+		+	+				
Familiarity with Area											
Experience with Block	No		-	-	-	-	-	-	+		+
	Yes		+	+	+	+	+	+	-		-
Neighborhood Close	No	+	+	+	+		+		-		
	Yes	-	-	-	-		-		+		

Note: Only statistically significant results are displayed in the table. (+) indicates higher probability
of perceiving disorder, and (-) indicates lower probability of perceiving disorder.

The panel on top displays the coefficients of the fixed effects associated with the co-variates used for hypothesis testing and those used as controls in the measurement error model of equation (3.5).

Model (1) is used as a reference to compare the performance of the conditional

models that incorporate the fixed effects of interviewers (Model 2), the occasions of measurement (Model 3) and the neighborhoods (Model 4).

The interviewer variables in model (2) did not reach statistical significance for either of the two items. However, the effect size of the three variables associated with interviewer 'community involvement' are larger for *trash* than for *graffiti*. I interpret this result as an indication that an item like *trash* is more vulnerable to this type of interviewer effect than an item like *graffiti*. Adding the interviewer covariates improved the overall fit of the model. This is evidenced by the decrease in both the AIC and the BIC criteria from model (1) to model (2).

Model (3) added the covariates associated with the occasion of measurement. Again, none of the covariates reached statistical significance. The overall fit of the model increased for both items, based on the decrease in the AIC/BIC criteria from model (2) to model (3).

Model (4) incorporated the last set of covariates, which were derived from census records and represented features of the socio-economic composition of the neighborhoods being rated. Consistent with prior research, indicators of *concentrated disadvantage*, *concentrated affluence* and *immigrant concentration* were significant predictors of disorder. As expected, the indicator of *population density* was not significant.

The probability of perceiving signs of *graffiti* was higher in neighborhoods with higher *disadvantage* ($\hat{\beta} = 0.14, SE(\hat{\beta}) = 0.02$) and *immigrant concentration* ($\hat{\beta} = 0.14, SE(\hat{\beta}) = 0.02$). As expected, the probability of observing *graffiti* in *affluent* neighborhoods was lower ($\hat{\beta} = -0.23, SE(\hat{\beta}) = 0.02$). The same pattern

was observed for the perceptions of *trash*.

Two other findings are worth noting. First, the effect size of the coefficients of neighborhood covariates was larger for *graffiti* than for *trash*. This result is consistent with all prior results that suggest that *graffiti*, among all items, is very much driven by neighborhood (tract-level) phenomena. The second finding is the dramatic reduction in the coefficient associated with the tract-level random effect, which occurred for both *trash* and *graffiti*. This phenomenon suggests that the neighborhood covariates were successful in explaining the random variability associated with tracts (neighborhoods). The increase in the overall fit of the model is evidenced by the decrease in the AIC/BIC criteria between model (3) and model (4).

The results discussed for *trash* and *graffiti* were replicated, to a large extent, across the other items of physical disorder, residential decay and residential security. These results are displayed in Table 3.9.

The lack of predictive power of some of the covariates in the model is evident in the panel with the fixed effects. None of the variables involved in my hypotheses showed a consistent pattern – neither effect sizes nor statistical significance of the results. Interviewers with children, for example, were not more likely to see signs of *litter* on the block face ($\hat{\beta} = -0.03; SE(\hat{\beta}) = 0.09$). And interviewers living close to the area they were rating were just as likely to see signs of deteriorated *buildings* ($\hat{\beta} = 0.01; SE(\hat{\beta}) = 0.03$) as those living further away. The strongest influence on the ratings was, by far, neighborhood socio-economic composition.

Tab. 3.8: Multilevel Multivariate Regression of Trash and Graffiti.

	Trash				Graffiti			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Fixed Effects								
Interviewer Characteristics								
Race White		0.10	0.13	0.13		0.10	0.13	0.13
Age 55+ yrs		-0.13	-0.16	-0.14		-0.13	-0.17	-0.16
Has Kids		0.15	0.18	0.15		0.03	0.06	0.05
Comm. Involvement		-0.10	-0.08	-0.08		0.03	0.03	0.04
Ever Married		-0.13	-0.19	-0.16		-0.04	-0.05	-0.04
Occasion of Measurement								
Rated after 5pm			0.00	-0.01			-0.02	-0.03
Rated on Weekend			0.03	0.04			-0.03	-0.00
Prior Experience			0.00	0.01			0.01	0.03
Neighborhood close			0.02	0.00			0.01	-0.01
Neighborhood Characteristics								
Immigrant				0.07***				0.14***
Disadvantage				0.10***				0.14***
Affluence				-0.13***				-0.23***
Pop. Density				-0.01				0.02
Intercept	0.58***	0.60***	0.61***	0.58***	0.57***	0.54***	0.54***	0.49***
Random Effects[†]								
Interviewers	-1.73***	-1.71***	-1.68***	-1.72***	-2.21***	-2.26***	-2.30***	-2.32***
Tracts	-1.56***	-1.56***	-1.60***	-2.44***	-1.08***	-1.09***	-1.10***	-2.51***
Blocks	-1.98***	-2.03***	-1.96***	-1.96***	-1.99***	-1.98***	-1.93***	-1.93***
Residual	-0.95***	-0.96***	-0.98***	-0.98***	-1.18***	-1.16***	-1.17***	-1.17***
Summary Statistics								
n	3998	3621	2892	2892	3998	3621	2892	2892
LL	-2133.9	-1900.8	-1495.2	-1460.1	-1318.1	-1244.1	-1005.0	-937.4
df	5	10	14	18	5	10	14	18
AIC	4277.8	3821.6	3018.3	2956.1	2646.2	2508.2	2038.0	1910.8
BIC	4309.3	3883.5	3101.9	3063.6	2677.6	2570.1	2121.6	2018.3

$(*) = p < 0.05; (**) = p < 0.01; (***) = p < 0.001$

(†): Estimates of the random effects are on the logarithmic scale.

Tab. 3.9: Multilevel Multivariate Regression of selected Neighborhood Items.

	Physical Disorder Items				Res. Decay Items			Res. Security Items		
	trash	litter	cigars	graffiti	bldgs	walls	yards	secsign	gates	ngwatch
Fixed Effects										
Interviewer Characteristics										
Race White	0.13	0.00	0.01	0.13	0.06	0.04	0.03	-0.17*	0.05	-0.02
Age 55+ yrs	-0.14	-0.08	0.09	-0.16	-0.01	-0.25***	-0.12	-0.05	0.23	-0.07
Has Kids	0.15	-0.03	0.04	0.05	0.09	0.11*	0.09	-0.01	-0.14	0.06
Comm. Involvement	-0.08	0.09	0.05	0.04	0.05	0.04	0.01	-0.00	0.16*	0.02
Ever Married	-0.16	0.12	0.10	-0.04	-0.05	0.05	-0.06	0.16	-0.01	-0.01
Occasion of Measurement										
Prior Experience	-0.01	0.03	-0.01	-0.03	0.02	-0.02	0.05*	0.01	-0.03	-0.03
Neighborhood close	0.04	-0.00	-0.04	-0.00	0.01	-0.01	0.12***	0.05	-0.02	0.01
Rated after 5pm	0.01	0.01	0.05	0.03	0.07**	0.02	-0.03	-0.03	-0.00	-0.01
Rated on Weekend	0.00	0.03	0.01	-0.01	-0.04**	0.03	0.01	0.01	0.01	-0.01
Neighborhood Composition										
Immigrant Conc.	0.07***	0.09***	0.09***	0.14***	0.03	0.07**	0.10***	-0.07**	0.08**	-0.04*
Conc. Disadvantage	0.10***	0.06***	0.11***	0.14***	0.07***	0.09***	0.09***	-0.07***	0.07**	-0.03*
Conc. Affluence	-0.13***	-0.16***	-0.13***	-0.23***	-0.15***	-0.18***	-0.16***	0.18***	-0.11***	0.01
Pop. Density	-0.01	0.00	0.02	0.02	-0.01	-0.00	-0.02	-0.00	0.01	-0.00
Intercept	0.58***	0.60***	0.47***	0.49***	0.82***	0.56***	0.72***	0.47***	0.60***	0.21***
Random Effects[†]										
Interviewers	-1.72***	-1.97***	-1.49***	-2.32***	-2.53***	-3.02***	-2.44***	-2.25***	-1.78***	-2.52***
Tracts	-2.44***	-2.54***	-2.61***	-2.51***	-2.09***	-2.10***	-2.22***	-2.23***	-2.05***	-2.51***
Blocks	-1.96***	-2.09***	-1.90***	-1.93***	-2.16***	-2.12***	-1.92***	-1.93***	-1.68***	-1.98***
Residual	-0.98***	-1.09***	-1.04***	-1.17***	-1.44***	-1.05***	-1.19***	-0.90***	-1.04***	-1.06***
Summary Statistics										
n	2892	2892	2892	2892	2631	2631	2631	2631	2631	2623
LL	-1460.05	-1140.88	-1299.52	-937.42	-192.00	-1114.95	-816.78	-1516.90	-1245.89	-1110.94

$(*) = p < 0.05; (**) = p < 0.01; (***) = p < 0.001$

(†): Estimates of the random effects are on the logarithmic scale.

Summary of Results

Estimates of interobserver agreement ($\hat{\delta}_3$) derived from the L.A. FANS data were smaller than those in the urban sociology literature. This result remained even after restricting the analysis to the pairs of observations collected on the same day ($n = 908$) and those collected on the same day and time of day ($n = 586$).

The magnitude of the interviewer clustering effect ($\hat{\rho}_{int}$) for neighborhood items was much larger than previously reported for survey items. When compared to the size of the sampling point clustering effects ($\hat{\rho}_{sp}$), however, the relative influence of interviewers in the variance of the neighborhood items was rather small. From the analyses of random influences, five items can be identified as performing relatively better: *graffiti*, condition of *buildings*, condition of *yards*, condition of *walls*, and *bars on windows*.

The final set of analyses focused on testing whether the variability in the ratings could be explained by fixed characteristics of the interviewers, the occasions of measurement, or the neighborhoods. Bivariate analyses showed that interviewers engaged in their communities were more likely to report signs of disorder, and interviewers completing ratings close to their homes were less likely to report disorder. These results did not hold in the multivariate analyses. Neighborhood measures of *immigrant concentration*, *concentrated disadvantage* and *concentrated affluence* emerged as the only strong and significant predictors of perceived disorder by trained observers. In a specialized neighborhood studies such as L.A. FANS, the minimization of observer idiosyncratic behavior could have been achieved though specialized

training and supervision during the fieldwork.

3.5 Discussion

This dissertation brings new attention to the potential of neighborhood observational data to inform survey fieldwork operations and methodological studies about survey participation. To be useful, however, neighborhood data first needs to be reliable and not full of measurement error.

A review of the survey methodology and the urban sociology literature helped identify important dimensions of measurement for neighborhood observational data collected by survey interviewers. As a result, a revised version of the Raudenbush and Sampson model of measurement error for neighborhood data was developed and hypotheses about the influence of interviewer characteristics on the ratings were elaborated.

The revised model had three important features. First, it set up the analysis based on individual items rather than on group of items (scales). Second, it incorporated interviewers as an additional level of clustering, which enabled the derivation of estimators for interobserver agreement and interviewer clustering effects. The model finally incorporated different sets of covariates which were used to test hypotheses about the systematic influence of interviewers, neighborhoods and occasions of measurement.

This paper showed that neighborhood items perform differently along the different dimensions of measurement. There was limited evidence of the systematic influence of interviewer characteristics on the neighborhood ratings. In addition, interviewer clustering effects were rather small when compared to sampling point

clustering effects. These two results suggest that idiosyncratic characteristics of interviewers do not influence their perceptions of disorder to a large extent. In specialized neighborhood studies like the L.A. FANS, the minimization of observer idiosyncratic behavior is usually achieved though specialized training, supervision and re-calibration during the fieldwork.

Agreement among pairs of ratings in the same block face, however, was rather low. It is important to remember that, in the L.A. FANS study, interviewers were permitted to complete observations on their own schedule during their first visit to the sampled bock. This flexibility in the fieldwork resulted in most of the block faces having repeated observations collected on different days. Thus, except for a few cases, one cannot disentangle the effect of using different interviewers (*inter-observer variability*) from that of change associated with the passage of time (*time variability*). Results based on the non-random subset of the cases that were rated on the same day (n=908), revealed that estimates of interobserver agreement increase for some items, but decreased for others. The same phenomenon occurred when restricting the analyses to those cases rated at the same time of the day ($n = 586$). These results are puzzling and await further research.

A critical review of this paper identifies three important limitations regarding the design, analysis and generalizability of these results: (1) lack of random assignment of interviewers to areas; (2) focus on the analysis of individual items rather than scales; and (3) limits on the generalizability of the results from this study. I will briefly discuss each issue here.

The lack of randomization in investigations of interviewer variability could

lead to overestimation of the interviewer effect ($\hat{\rho}_{int}$) (Kish, 1962). In the context of our study, convenience or preference could led interviewers to select different areas to work on, thus a component due to the average difference among these factors is confounded with the component for interviewer variability. As Kish's points out, the overestimation of $\hat{\rho}_{int}$ could be great in those sampling operations where the interviewer has wide latitude in choosing his workload, but it might be small in surveys carried out at one limited site, where an approximation to randomization occurs automatically. The L.A. FANS study is much closer to the latter case, because all observations were completed in a single county in the United States.

This paper purposefully avoids joint analysis of multiple items for two reasons: (a) to explore the properties of neighborhood items that are currently being collected by large survey projects, and (b) to inform future studies that need to evaluate the cost/quality trade-offs of collecting a few items versus full neighborhood scales. As a result of using this approach, estimates of measurement error derived from the current analyses are most likely larger (*i.e.* provide an upper bound) than those derived from analysis of multiple items or composite scores. Researchers should consider these implications when evaluating whether to use measures disorder derived from single items or composites scores for their substantive analysis.

One of the key feature of the L.A. FANS study is that it combined (a) a state of the art questionnaire and training protocols for the collection of observational data, and (b) a regular crew of survey interviewers to collect those observations. This study, however, was conducted in a single-city in the Unites States, thus results presented here cannot be generalized to a broader setting. It is an empirical

question, however, whether these results can be replicated in a multi-site study or studies where training and other conditions in the field would vary greatly. This is particularly relevant for the National Children study (NCS), which will collect data on neighborhood environments throughout the U.S. using the same items in the L.A. FANS questionnaire.

Survey researchers are currently starting to assess the potential for observational data for methodological and practical purposes. Many surveys routinely require survey interviewers to collect observations on survey respondents and the selected housing units. In this setting, adding neighborhood observations may seem relatively easy to add to current call record forms. The proper implementation of training and the supervision of these collections, however, can be challenging and costly.

A very practical question that may arise from the survey methods audience is whether, based on the results from this paper, we can identify a minimum set of items that can be used for nonresponse investigations. Results from this paper identified five items performing relatively better on the measurement error properties evaluated: *graffiti*, condition of *buildings*, condition of *yards*, condition of *walls*, and *bars on windows*. The worse performing items were *litter* and *security gates*.

4. PAPER 3: ASSESSING THE EFFECTIVENESS OF NEIGHBORHOOD CHARACTERISTICS IN NONRESPONSE WEIGHTING ADJUSTMENTS

4.1 Introduction

Decreasing response rates continues to be the overwhelming concern of survey researchers and methods to address this problem are being pursued around the globe. In recent years, however, the focus has shifted to address nonresponse bias rather than nonresponse rates.

Any strategy that aims to reduce nonresponse bias will require auxiliary information from both respondents and nonrespondents. Data available from sampling frames and census records are typically used for this purpose. Observational data collected by survey interviewers during fieldwork operations could also serve this purpose.

The collection of observational data, however, comes at additional cost. Without evidence that these observational data have the potential for nonresponse bias reduction, cost spent on collecting them is not justifiable.

Paper 1 showed that interviewer observations of neighborhood characteristics can be strong predictors of cooperation in household surveys. Paper 2 showed, however, that certain neighborhood variables are more prone to measurement errors than others. Analyses in this paper aim to assess the potential of neighborhood observational data to be used in post-survey adjustments for nonresponse.

4.2 Background

One of the primary areas of focus in the nonresponse literature has been on the potential biasing effects of nonresponse. Nonresponse errors can be of two types - random or systematic. Random errors increase the variance of survey estimates, but they decrease as sample sizes increases. Systematic errors, on the other hand, bias the survey estimates and are not reduced with increases in sample size (Kish, 1992).

Strategies used to reduce the risk of nonresponse bias are implemented before, during, and after data collection. Before data collection, careful training of survey interviewers in survey research techniques can help reduce nonresponse errors (Lessler and Kalsbeek, 1992; Groves and McGonagle, 2001). Strategies employed during the data collection involve the subsampling of nonrespondents for further follow up in two-phase sampling designs (Hansen and Hurwitz, 1946; Deming, 1953) and the manipulation of different aspects of the data collection protocol to foster parcipation (Groves and Heeringa, 2006; Lepkowski et al., 2010).[1] In the U.S., researchers at the National Survey of Family Growth (NSFG) monitor indicators of interviewer recruitment efforts and interviewer observations of the sampled respondents in an effort to increase response rates and decrease nonresponse bias (Lepkowski et al., 2010, 17). In Statistics Netherlands, survey researchers monitor indicators of sam-

---

[1] Large survey projects using the two-phase, or double sampling approach to sub sample nonrespondents include the American Community Survey (BUREAU, 2009), the General Social Survey (Davis and Smith, 2005, 2108), and the National Survey of Family Growth (Lepkowski et al., 2010, 16).

ple composition (R-indicators) to increase the balance of the resulting sample in the distribution of selected covariates (Schouten et al., 2009; Särndal and Lundström, 2008).

After data collection, however, the most ubiquitous method used to correct for nonresponse biased is weighting. Weights are commonly assigned to respondents records in a survey data file in order to make the weighted records represent the population of inference as closely as possible. The weights are usually developed in a series of stages to compensate for unequal selection probabilities, nonresponse, noncoverage, and sampling fluctuations from known population values (Brick and Kalton, 1996).

There are different strategies used to develop weights that adjust for sample nonresponse, such as cell weighting (Oh and Scheuren, 1983; Kalton, 1983; Little, 1986; Brick and Kalton, 1996), raking (Oh and Scheuren, 1983; Deville and Särndal, 1992), GREG weighting (Bethlehem, 1988; Deville and Särndal, 1992), and logistic regression weighting (Little, 1986; Little and Rubin, 2002). The methods differ in the flexibility for modeling and the ability to handle large number of auxiliary variables. A review of these methods is given in Kalton and Flores-Cervantes (2003).

The success of non-response weighting depends on the variables that are used in constructing non-response weights. To be successful at reducing nonresponse, variables used in nonresponse adjustments (Z-variables) should have two properties (Little, 1986; Kalton and Flores-Cervantes, 2003; Little and Vartivarian, 2003, 2005; Groves, 2006): variables should be be predictive of the sampled persons probability of responding to a survey request (P-variable) and be predictive of the survey

outcome variables of interest (Y-variables).

A simulation study by Little and Vartivarian (2005) showed that error reductions are possible with high Z-P correlations and Z-Y correlations in the range of $.48 - .80$. Whether such correlations can be found in practice is an empirical question. Kreuter et al. (2010) addressed this question. The authors found magnitudes of the Z-P and the Z-Y correlations lower than those suggested by Little and Vartivarian (2005). The authors also assessed the effect of adding new auxiliary variable(s) to the current weighting scheme across five large survey project. The reported change in the weighted estimates were also modest. The authors raised concerns about correlations possibly being attenuated by measurement error in the auxiliary variables, but they were unable to assess this effect with the data at hand.

The meta-analysis by Groves and Peytcheva (2008) reminded the field to focus on reducing nonresponse bias rather than nonresponse rates. Towards that end, survey researchers should search for adjustment variables with high correlation with the survey outcomes. In recent years a handful of studies have provided evidence of the empirical magnitudes of the Z-Y relationship (Kreuter et al., 2010; Peytchev et al., 2010). This study aims to contribute to the literature by assessing the potential of neighborhood data collected by survey interviewers to be used in nonresponse adjustments. The research strategy is described in the next section.

4.2.1   Research Questions

A big challenge for the field of survey methodology is the identification of auxiliary variables that are strongly related to survey outcomes. Observational data

collected by interviewers seems like a natural candidate for this purpose. Interviewers are currently charged with the responsibility to collect observational data on the sampled members, their housing units and the neighborhoods where they live. In most cases, this data would be available for respondents and nonrespondents. Without evidence that these observational data have the potential for nonresponse bias reduction, however, cost spent on collecting them is not justifiable. One of the aims of this paper is to assess whether neighborhood observations collected by survey interviewers can be used for this purpose.

Paper 1 provided evidence that physical characteristics of the sampled neighborhoods could influence the household(er) survey participation decision in face to face surveys. In fact, the effect of the physical environment went beyond the influence of the social environment and the socio-economic composition of the neighborhoods.

Collecting data on the physical look and appearance of local areas, however, adds additional burden to interviewers and costs to the data collection process. Given the theoretical relevance of these neighborhood measures, an empirical research question is whether they make a significant difference when used in nonresponse adjustments.

To answer this question I developed a small set of nonresponse adjustments that differed only on the type and number of area-level variables they are based on. Building on the results of Paper 1, the first adjustment is based on a 'full' model that includes area-level measures of *socio-economic composition* and the *physical environment*. The performance of this adjustment is then compared with that of

alternative sets of adjustments based on 'reduced' models that only include, for example, measures of socio-economic composition. The performance of different nonresponse adjustments is thus evaluated based on three criteria:

1. How strongly related are the different sets of area-level variables to both, survey outcomes and household cooperation.

2. How large of an effect do nonresponse adjustments, based on different sets of area-level variables, have on the variance of the weighted mean.

3. How large of an effect do nonresponse adjustments, based on different sets of area-level variables, have on the estimate of the weighted means.

The Los Angeles Family and Neighborhood Study (L.A. FANS) provides an interesting opportunity to perform these analyses because it has a wealth of auxiliary variables available at the area-level and observational data is also available at the household-level for both respondents and nonrespondents. Analyses in Paper 1 showed that household-level covariates were significant predictors of cooperation in the L.A. FANS survey. Given these results, these variables were included (controlled for) in all the models under analysis here.

4.3 Data and Methods

Analyses in this paper were implemented using data from the Los Angeles Family and Neighborhood Study (L.A. FANS). Earlier chapters of this dissertation have already given a general introduction to the L.A. FANS study. I only present here variables and methods that are particular to this paper. A brief description of the correlational analyses and the methods used to develop nonresponse adjustments are presented in sections 4.3.1 and 4.3.2. Strategies used to assess the effect of alternative nonresponse adjustments on estimated means and variances are discussed in section 4.3.3. Variables used here are presented in section 4.3.4.

4.3.1 Correlational Analyses

Correlational analyses were implemented to assess the strength of the association between different types of area-level variables (Z's), the response indicator (P) and selected survey outcomes (Y's). Estimates of the bivariate Pearson correlation are provided for each area-level variable and the response indicator (Z-P correlations) and different survey outcomes (Z-Y correlations). Confidence intervals for the estimated correlations were approximated using Fisher's Transformation.[2]

4.3.2 Development of Nonresponse Adjustments

Before describing the methods used to develop the new nonresponse adjustments it is important to understand the structure of the current weighting scheme

---

[2] I used the Stata command *ci2* to compute 95% confidence intervals for the bivariate Pearson correlations. For a reference on Fisher's Transformation see Rosner (2000, 461-462).

available for the L.A. FANS data. The development of the L.A. FANS weights involved a series of step which include the development of appropriate sampling selection probabilities and corrections for nonresponse. I only discuss here the household weights and the weights developed for the randomly selected adult (RSA). A complete description of the L.A. FANS weights and weighting procedures is available in Sastry et al. (2003).

L.A. FANS Current Weighting Scheme

The L.A. FANS household-level weight is calculated as the inverse of three different terms, which reflect the probability of sampling tracts ($Prob1$), the probability of sampling households ($Prob2$), and an overall adjustment for household nonresponse and the over-sampling of households with children ($Prob3$).

The probability of selecting tracts is given in expression (4.1). The first term in the expression reflects the average rate at which all tracts were selected. The second term reflects the differential rates at which tracts were sampled across the p poverty strata ($p = 1, 2, 3$). This probability differs across the strata because tracts were over-sampled in the 'very poor' strata and under-sampled in the 'non-poor' strata.

$$Prob1_p = \left(\frac{Num.HH.in.65.sampled.tracts}{Num.HH.in.LA.County} \right) * \left(\frac{Prop.HH.in.sampled.tracts}{Prop.HH.in.LA.County} \right)_p$$

(4.1)

The probability of selecting a household within a given tract t ($t = 1, \ldots, 65$) is a product of two factors – the average rate at which households were selected in each

tract and the rate of under/over sampling of households with children. The expression below shows the first factor. The second factor is accounted for in expression (4.3).

$$Prob2_t = \left(\frac{Num.HH.sampled}{Total.HH.Census2000} \right)_t \qquad (4.2)$$

Two additional factors contribute to expression (4.3): an adjustment for household nonresponse within each tract t, and an adjustment for the over-sampling of households with children – which takes a different form depending on the household's children status c ($c = children, nochildren$) within each tract t. More details on the first term are given in the next section.

$$Prob3_{tc} = \left(\frac{Num.HH.Rostered}{Num.HH.Sampled} \right)_t * Adj_{tc} \qquad (4.3)$$

In households with children, the factor Adj_{tc} reflects the proportion of households with children that completed the roster interview in tract t ($phhk_t$) to the proportion of households with children in the 2000 Census in tract t ($PHHK_t$). When this ratio is greater than one, due to over-sampling of households with children, the inverse of this factor will down weight households with children. The opposite effect applies for households without children.

$$Adj_{tc} = \begin{cases} \frac{phhk_t}{PHHK_t} & \text{for HHs with children} \\[2ex] \frac{(1-phhk_t)}{(1-PHHK_t)} & \text{for HHs without children} \end{cases}$$

Note that this last adjustment is different from calculating selection probabilities

at each stage of sampling and is meant to avoid some extreme weights that would result if the more standard procedures were used. The figure below illustrates the main sampling phases and outcomes of the L.A. FANS recruitment process.

Phase I	**Phase II** – Household Roster Interview			Adult Interview	
Screened	Sampled	Contacted	Interviewed	Interviewed	...... etc.
7,681	4,010	3,708	3,083	~ 2,600	

Refused

NonContacted
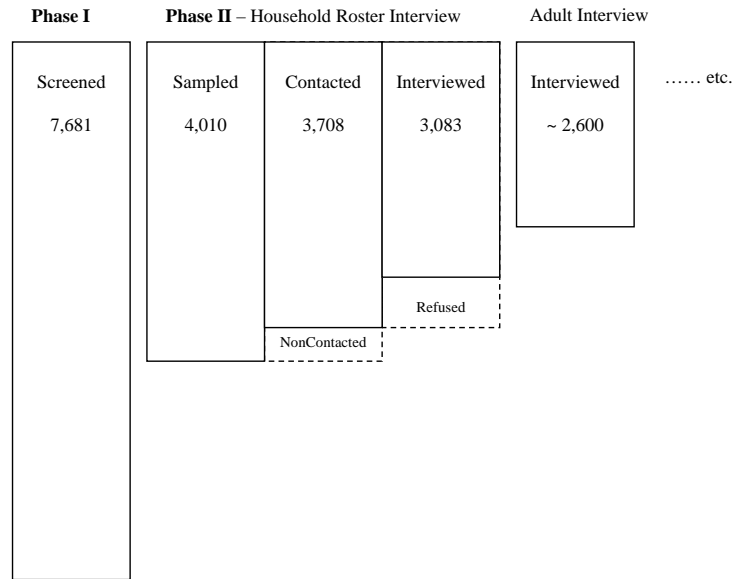
Fig. 4.1: Illustration of the outcomes of the L.A. FANS sample recruitment efforts.

The L.A. FANS household-level weight is computed as the inverse of the product of expressions (4.1)-(4.3). To reduce the variance of the weights, the upper tail of expression (4.4) was trimmed by setting all weights beyond the $95\%^{th}$ percentile to the $95\%^{th}$ percentile value. Household weights were then normalized to have a

mean of one.[3]

$$wgthh^{NTNS} = (Prob1_p * Prob2_t * Prob3_{tc})^{-1} \qquad (4.4)$$

The weight for the randomly selected adult (RSA) is computed by multiplying expression (4.4) by the inverse of the probability of selecting a random adult in the household ($ProbAdult$).[4]

$$wgtrsa^{NTNS} = wgthh^{NTNS} * (ProbAdult)_i^{-1} \qquad (4.5)$$

The last step in the development of the RSA weights involved raking (Oh and Scheuren, 1983). Raking has been widely used for many years for benchmarking sample distributions to external distributions (Kalton and Flores-Cervantes, 2003). In raking the RSA weights, the full set of two-way crossclassification of age, race/ethnicity, and sex was used to match the marginal distributions of the sample to the same distributions for all Los Angeles County residents from the 2000 Census. This procedure produced RSA weights which correct for potential coverage and nonresponse errors in the L.A. FANS survey.

To reduce the variance of the weights the raked weights were also trimmed

---

[3] The final set of household-level weights sum up to the total number of households responding the roster interview (n=3,083). The variable '$wgthh$', available in the restricted files, identifies the household-level weights in the L.A. FANS datasets. The upper script 'NTNS' in expression (4.4) stands for 'Not-Trimmed Not-Standardized' and it is used here to highlight the fact that this expression corresponds to the weights before the process of trimming and standardization.

[4] The probability of selecting an adult is a function of the number of eligible adults in each household. See Sastry et al. (2003) for details on the adjustment for selection of persons in the L.A. FANS survey.

beyond the $95\%^{th}$ percentile and normalized to have a mean of one.[5] Descriptive statistics for selected L.A. FANS weights are displayed in Table 4.2.

Tab. 4.1: Descriptive Statistics for Different L.A. FANS Tract Level Estimates.

Variable	Label	Mean	Std. Dev.	n
Subsampling Rate HH with Children	–	0.98	0.02	65
Subsampling Rate HH without Children	–	0.44	0.26	65
Roster Response Rate	$\hat{RR}_t$	0.76	0.09	65
Prop. HH with Children in Roster	$phhk_t$	0.74	0.07	65
Prop. HH with Children in L.A. County	$PHHK_t$	0.45	0.14	65

Tab. 4.2: Descriptive statistics for factors used in the development of the L.A. FANS household weights

Variable	Label	Mean	Std. Dev.	n
Household NTNS weight	$wgthh^{NTNS}$	1020.62	1292.98	3083
Household weight	$wgthh$	1	1.11	3083
RSA weight	$wgtrsa$	1	1.17	2609

Removing the Current Nonresponse Adjustment

The purpose of this study was to assess the effect of nonresponse adjustments based on different sets of area-level variables. As described in the previous section, the L.A. FANS household-level weight contains two different types of nonresponse adjustments embedded in its structure.

The first component corresponds to the inverse of the *household response rate* within each Census tract (see first term in expression (4.3)). This sample-

---

[5] The final set of RSA weights sum up to the total number of adults eligible for the adult interview (n=2,619). The variable '*wgtrsa*' identifies the RSA weights in the L.A. FANS datasets.

based adjustment, which corresponds to a 'weighting class adjustment' (Kalton and Kasprzyk, 1986), aims to correct for the potential biasing effects of noncontact and refusals to the roster interview. The second component corresponds to the ratio of the proportion of children in the sample to the proportion of children in Los Angeles County (see second term in expression 4.3). This adjustment aims to correct for the over/under sampling of households with children. At the same time, however, it is exercising a type of nonresponse adjustment for the potential under/over coverage of households with children.

In addition to these nonresponse adjustments at the household-level, the RSA weight includes a raking adjustment. This adjustment aims to correct for errors of nonresponse and coverage for the distribution of age, gender and race/etchnicity of the adults in the sample.

In this study I attempted to remove only the first type of nonresponse adjustment from the L.A. FANS household-level weight. For that purpose, I developed the following estimate of the roster response rate within each tract using data from the recruitment efforts available in the MODSTAT1 file:[6]

$$\hat{RR}_t = \left(\frac{Num.HH.Rostered}{Num.HH.Sampled} \right)_t \tag{4.6}$$

Multiplying the estimate of the roster response rate within each tract $(\hat{RR}_t)$ by ex-

---

[6] In expression (4.6), the term in the numerator corresponds to the total number of households that completed the roster interview in each tract t. The term in the denominator corresponds to the total number of households that were selected for the roster interview in each tract t. Descriptive statistics for this estimate of the tract-level response rate for the roster interview are available in Table 4.1.

pression (4.4) 'removes' the effect of the weighting-cell adjustment from the household weights. This adjusted set of weights, defined in expression (4.7) below, is used later in the development of the new set of weights being compared in this paper.[7]

$$wgthh\_r^{NTNS} = wgthh^{NTNS} * \hat{RR}_t \qquad (4.7)$$

Producing new sets of RSA weights requires, in addition, the incorporation of additional adjustment factors that account for the probability of selecting an adult in the household ($ProbAdult$) and raking to population controls ($Adjust_{Raking}$). Unfortunately, I did not have the means to replicate these procedures. For the purpose of this paper, I used the ratio of the publicly available RSA weights ($wgtrsa$) to the household weights ($wgthh$) as an approximation for this adjustment factor:

$$Adjust_{Adult} = \frac{wgtrsa}{wgthh} \qquad (4.8)$$

$$\frac{(Prob1 * Prob2 * Prob3)^{-1} * (ProbAdult)^{-1} * Adjust_{Raking}}{(Prob1 * Prob2 * Prob3)^{-1}} =$$

I expect this approximation to be reasonable for most of the cases, except perhaps for those cases with weights in the high end of the distribution. Overall, I do not expect that using this approximation in the development of the new RSA weights affects the comparisons across the new sets weights under analyses.

Development of New Nonresponse Adjustments

There are different ways of developing weights for nonresponse adjustments (Kalton and Flores-Cervantes, 2003). In this paper I used stratification on the

---

[7] The household-level weight in expression (4.4) was provided under spacial request to the LA FANS team.

propensity score (Rosenbaum and Rubin, 1983; Little, 1986). I used five general steps to create nonresponse adjustments: (1) fit a logistic regression of household and area-level variables on the binary indicator of roster response; (2) calculate predicted propensities for each unit in the sample used for modeling; (3) sort cases by their predicted response propensities, from lowest to highest; (4) form classes with about the same number of units each; and (5) decide on what type of nonresponse adjustment to use within each class.

The propensity score method requires the development of a model that predicts the probability that a household responds to the roster interview.[8] In this paper I developed six response propensity models which included different sets of area-level variables (Z's) available for both respondents and nonrespondents.[9] The equation below illustrates the structure of the models fitted:

$$log\left(\frac{Pr(y_{ik} = 1 | Z^H, Z^T)}{1 - Pr(y_{ik} = 1 | Z^H, Z^T)}\right) = \theta + \sum_{p=1}^{P} \alpha_p Z_{pik}^H + \sum_{q=1}^{Q} \beta_q Z_{qk}^T \qquad (4.9)$$

where y_{ik} is the binary indicator of response during the roster interview for household i in Census tract k; Z_{ik}^H is a set of household-level characteristics recorded by the interviewers for respondents and nonrespondents to the roster interview; and Z_k^T is a set of area-level characteristics available from Census records and interviewer observations. Table 4.3 illustrates the structure of the six response propensity models fitted.[10] Section 4.3.4 provides the description of the covariates used.

At the time of estimation, a choice must be made whether to use the survey

---

[8] The binary indicator of roster cooperation y_{ik} takes on a value of 1 if the household completed the roster interview and 0 if it was not contacted or refused.

[9] All area-level variables are used in their original measurement scales. See descriptive statistics

Tab. 4.3: Alternative Models Used to Develop Nonresponse Weighting Adjustments.

Variables in each model	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Census Records						
Census Scales (3 variables)		x				x
Census Items (5 variables)			x			
Interviewer Observations						
Neighborhood Scales (2 variables)				x		x
Neighborhood Items (7 variables)					x	
Household Observations (6 variables)	x	x	x	x	x	x

base weights to estimate the model parameters in (4.9). Little (1986) suggests that, since *probabilities conditional on being selected for the sample* are desired, it implies that unweighted regression should be fit. Another argument for using unweighed analysis, illustrated by Little and Vartivarian (2003), is that using variable base weights can yield estimators with higher variances in the context of cell nonresponse adjustments. In this paper I used unweighted logistic regression to estimate the model in (4.9).

For each model in Table 4.3 decisions have to be made regarding (a) the number of adjustment cells to use and (b) the method of estimating the nonresponse adjustment within each cell. Following Dever et al. (2011), these decisions were made based on the inspection of the variability of the predicted response propensities ($\hat{p}_z$) within an 'initial' number of ten adjustment cells. After these steps, the new

---

in Tables 4.4 and 4.5.

[10] The full model (6) includes 5 household-level covariates, 3 scales measuring the physical environment and 3 scales measuring the socio-economic composition of the neighborhood. Reduced model (5), on the other hand, includes only the household-level covariates and 7 items measuring the social environment.

nonresponse adjustment were computed as the inverse of the response propensities in each weighting cell ($w_z = \frac{1}{\hat{p}_z}$).

The new sets of RSA weights are obtained by multiplying the household-level weights in expression (4.4) $wgthh\_r^{NTNS}$, the new nonresponse adjustment w_z, and the adjustment factor $Adjust_{Adult}$ estimated in (4.8). To be consistent with the procedures implemented in the L.A. FANS survey, the weights resulting from expression (4.10) were trimmed at the $95\%^{th}$ percentile to conform the final RSA weights ($wgtrsa\_r$) used for analyses here.

$$wgtrsa\_r^{NTNS} = wgthh\_r^{NTNS} * w_z * Adjust_{Adult} \qquad (4.10)$$

4.3.3   Kish's Variance Inflation Factor and MSE

The use of weighing methods for nonresponse bias reduction comes at the expense of increased variances (Kalton and Kasprzyk, 1986; Kish, 1992). As Kish (1992) notes, these increases in variance tend to persist undiminished for all statistics as if they come to increase the element variance from σ^2 to $(1+L)\sigma^2$, or to decrease the effective number of elements from n to $n/(1+L)$, where L denotes "relative loss". Kish provided the following expression as a summary factor of this potential increase in variance, where w_i is the weight associated with element i, and CV_w^2 is the square of the coefficient of variation of the weights for element i.

$$(1 + L) = \frac{n \sum w_i^2}{(\sum w_i)^2} = 1 + CV_w^2 \qquad (4.11)$$

The factor $(1 + L)$ is an approximate potential relative increase in the variance of the estimated means that can be attributed to the distribution of the weights. This

146

formula is widely used to gauge the impact of alternative weighting schemes (*e.g.* Lepkowski et al. (2010)). It has been noted, however, that the factor $(1 + L)$ is only a good approximation when the adjustment cell is weakly associated with the survey outcome (Little and Vartivarian, 2005).

It is important to remember that nonresponse adjustments are added to the weights with the expectation that they will increase the variance of the weighted estimates, but with the hope that they will help reduce the bias due to nonresponse. The expression below, also in Kish (1992), provides a means to gauge this variance-bias trade-offs for the weighted mean $(\bar{y}_w)$:

$$MSE(\bar{y}_w) = S_y^2 \left(\frac{1}{n_d} + \frac{CV_w^2}{n_d} \right) = \frac{S_y^2}{n_d} (1 + L) \tag{4.12}$$

In this expression S_y^2 is the population parameter for the element variance of the variable y; n_d is the effective sample size; and $(1 + L)$ is the variance inflation factor defined in expression (4.11). The effective sample size can be estimated using $n_d = \frac{n}{Deff}$, where $Deff$ corresponds to the design effect and is computed as the ratio of the variance under a complex design to the variance under simple random sample $\left(\frac{Var(\bar{y})}{Var(\bar{y}_{srs})} \right)$. I computed an estimate of the population element variance S_y^2 using:

$$\hat{S}_y^2 = \frac{n}{n-1} * \frac{1}{\sum w_i} * \left(\sum w_i y_i^2 - \frac{(\sum w_i y_i)}{\sum w_i} \right) \tag{4.13}$$

In this paper, estimates of $(1 + L)$ and $MSE(\bar{y}_w)$ were computed for different survey variables for the new different sets of RSA weights. All else equal, one would prefer the weighting scheme with the lowest variance inflation factor $(1 + L)$ and

mean square error (MSE). The next section introduces the survey variables that were used in these analyses.

4.3.4 Datasets and Variables

This section briefly introduces the variables used in the response propensity models and the analyses of weighted means. Detailed description of these variables have already been provided in Paper 1 and Paper 2.

Variables used in Response Propensity Models (Z's)

The response propensity models illustrated in Table 4.3 used as covariates neighborhood measures derived from interviewer observations and Census records. In addition, for each data source, two types of variables were derived – *items* and composite scores (*scales*). Descriptive statistics for the final set of variables used are displayed in Tables 4.4 and 4.5.

Census variables are readily available in the L.A. FANS NSC_STF3 file.[11] Census scales used here correspond to the three measures of socio-economic composition that showed stronger association with cooperation in the analyses of Paper 1. Details about the factor analysis and the interpretation of the factors extracted is available in Paper 1. Five individual Census items were selected trying to match those used in the landmark nonresponse study by Groves and Couper (1998).

---

[11] Summary File 3, prepared by the U.S. Census Bureau and containing tabulations of selected demographic and socio-economic characteristics of the U.S. population by various geographic levels. See Peterson et al. (2007).

Tab. 4.4: Descriptive statistics for variables derived from Census Records. Unweighted estimates.

Area-level variables	Estimate	Std. Error	n
Census Scales			
Immigrant Concentration Score	0.000	0.122	65
Concentrated Disadvantage Score	0.000	0.118	65
Concentrated Affluence Score	0.000	0.123	65
Census Items			
Perc. *Pop* < 18 yrs.	30.19	1.06	65
Perc. Non-White	76.17	3.18	65
Perc. Owner Occupied	43.38	3.25	65
Perc. Multi-Unit	39.26	3.43	65
Population Density	14836.44	1297.641	65

L.A. FANS interviewers completed observations on several neighborhood characteristics that aimed to capture the level of disorder and decay of the sampled neighborhoods. Area-level estimates for composite scores (scales) of *physical disorder*, and *residential decay* were derived by aggregating block face-level data to the tract-level using the same procedures described in Paper 1. The *social disorder* scale was not considered for analysis due to the poor measurement error properties of the items associated with it (see Paper 2). Area-level estimates for individual items were produced for seven indicators following the same procedures described for the neighborhood scales.[12] All disorder and decay estimates were developed based on

---

[12] The seven items were selected based on two criteria. First, they were considered the most *representative* of the neighborhood physical environment, based on the results of an exploratory factor analysis on thirteen neighborhood items (loadings in the first factor higher than 0.6). These items were: *graffiti*, *condition of the buildings*, *damaged walls* and *damaged yards*. I added *litter*, *bars on windows* and *security gates* because similar items are currently collected in surveys such as the National Survey of Family Growth (NSFG) and the European Social Survey (ESS).

the same subset of 3,998 records used in Paper 2, where ratings from exactly two interviewers were retained for each block-face in the sample.

Tab. 4.5: Descriptive statistics for variables derived from Interviewer Observations.

Area-level variables	Estimate	Std. Error	n
Neighborhood Scales			
Perc. Physical Disorder	42.10	2.38	65
Perc. Residential Decay	53.54	1.94	65
Neighborhood Items			
Perc. Litter	78.59	2.70	65
Perc. Graffiti	59.39	4.39	65
Perc. Condition Buildings	88.01	2.63	65
Perc. Damaged Walls	70.19	3.27	65
Perc. Damaged Yards	81.22	2.78	65
Perc. Bars on Windows	66.82	3.91	65
Perc. Security Gates	62.42	2.76	65

The L.A. FANS study also collected data on interviewers perceptions on a few household level characteristics for both respondents and nonrespondents to the screener interview. Household observations included *type of housing unit*, estimate of *rent* and *presence of children*. Respondent observations included *race, age* and *gender*.[13] Household-level variables were significant predictors of roster cooperation in Paper 1, and thus they are also included as predictors in all six response propensity models used here. Descriptive statistics for these variables are available in the Appendix.[14]

---

[13] Interviewers also recorded the *language* of the interview. This variable was not included in the analyses of Paper 1 and it is not included here.

[14] Household-level observations were missing for up to 30% of the households contacted for the roster interview. For the purpose of this analysis I used the same multiple imputed dataset of

Analyses of the change in the weighted means requires the use of responses provided by the survey respondents. Analyses are restricted here to eight outcomes provided by the Randomly Selected Adult (RSA). These variables were expected to show high, medium, and low association with the neighborhood physical environment. Table 4.6 provides descriptive statistics for the variables used.

Outcomes expected to have high Z-Y correlation included two reports on the attitudes and behavior of their neighbors. The concept of *reciprocated exchange* attempts to measure how often people in the neighborhood do certain things for each other, such as: asking favors to each other, watching over each other's property, and asking for advice to each other. Answer categories were coded from low to high, so 1 means 'never' and 4 means 'often'. The concept of *intergenerational closure* attempted to measure perceived closeness between adults and children in the neighborhood, such as: kids have adults to look up to, adults watch out for kids, adults know child's friends, adult knows local children, and adults know other parents. Answer categories were coded from low to high, so 1 means 'strongly disagree' and 5 means 'strongly agree'.

Outcomes expected to have medium sized Z-Y correlations include self report of *health status*, *unsafe neighborhood*, *church member* and participation in *social groups.* I expected the first two variables to be positively associated, and the last

household observations used in the analyses of Paper 1. Since these variables are included in all six models under comparison, I do not expect that the use of this imputed data set will affect the comparisons across the models in this paper.

two negatively associated, with neighborhood disorder and residential decay. In the survey literature, indicators such as church membership and social groups are typically expected to correlate positively with participation in household surveys.[15]

Finally, I selected two fertility-related variables that I expected to have Z-Y correlation 'close-to-zero': a variable measuring the *use of contraception method*, and a variable reflecting *disapproval towards having children without being married*. The first variable was measured with the yes/no question *'Are you and your partner currently using any of these types of contraception or any method of preventing pregnancy?'*. For the second variable, adults were asked their opinions on three situations: a teenage girl has a baby without being married, a woman in her twenties has a baby without being married; man in his twenties fathers a child without being married to the baby's mother. The original answer categories were 1 'strongly approves' 2 'approves' 3 'neither approves nor disapproves' 4 'disapproves' 5 'strongly disapproves'. I average the scores across the three questions and coded as 0 the scores lower than 4, and 1 the scores greater or equal than 4.

---

[15] All these variables were dichotomized for analysis: *Poor Health* (1=poor, fair; 0=good, very good, excellent), *Unsafe NG* (1=extremely dangerous, somewhat dangerous; 0=fairly safe, completely safe), *Church Member* (1=yes; 0=no); and *Social Groups* (1=1-9 groups; 0=none).

Tab. 4.6: Descriptive statistics for Survey Outcomes (Y's). Unweighted estimates. RSA only ($n = 2,619$).

Survey variables	Estimate	Std. Error	n
Neighborhood Outcomes			
Reciprocated Exchange Score	2.641	0.016	2582
Intergenerational Closure Score	3.530	0.013	2585
Individual Outcomes			
Perc. Church Member	39.27	0.96	2572
Perc. Participates 1+ Social Groups	31.23	0.91	2581
Perc. Poor Health	22.51	0.83	2532
Perc. Unsafe Neighborhood	33.06	0.93	2568
Perc. Uses Contraception Method	57.33	1.06	2184
Perc. Disapproves Babies without being Married	54.54	0.98	2609

4.4 Analyses and Results

Inspection of the Z-P and the Z-Y correlations showed a very consistent pattern in terms of the size and the direction of the correlations. In general, *items* did as well as *scales* for both interviewer observations and for Census records. One clear advantage of the scales, however, was the interpretability of the results.

4.4.1 Strength of Z-P Correlations

Figure 4.2 displays the bivariate Pearson correlation between the binary indicator of roster response and each one of the thirteen area-level variables under analysis. For each variable, the figure displays the estimate of the correlation and an approximation for its 95% confidence interval.

The first noticeable result in Figure 4.2 is that the direction of the Z-P correlations did not meet the theoretical expectations laid out in theories of survey participation in household surveys. In the figure, response to the roster interview is positively associated with area-level signs of *disorder* and *decay*. The association is also positive with *immigrant* concentration and concentrated *disadvantage* in the area. The correlation is negative with the measure of concentrated *affluence*. The same trends are visible among items covering similar constructs.[16]

These household-level results are consistent with those presented in Paper 1

---

[16] Under the current paradigm, households in areas characterized by social disorganization are less likely to participate in household surveys. These areas are characterized by low levels of social cohesion, high levels of ethnic heterogeneity, boarded up houses, multiunit structures, population density, etc. See Paper 1 for a review of the literature.

for the tract-level correlation of roster cooperation and some of the same area-level measures used here. A detailed discussion is provided there about the interpretation of these unexpected results.

The second interesting result is that the size of the Z-P correlations for scales and items were within the same range for indicators of related constructs. For the *disorder* and *decay* measures, for example, the Z-P correlation for the scales ranged $r = (.09 - .12)$ and for the items $r = (.06 - .11)$. For measures of *disadvantage* and *immigrant* concentration the scales ranged $r = (.04 - .05)$ and the items $r = (.03 - .09)$. The Z-P correlation for concentrated *affluence* was $r = (-.07)$ and it was $r = (-.07)$ for percent *owner* occupied.

Overall, however, the size of the Z-P correlations were modest. This result is consistent with Kreuter et al. (2010) which reported most of the Z-P correlations below 0.1 (in absolute value).

4.4.2   Strength of Z-Y Correlations

The next set of figures display the bivariate Pearson correlation between eight different survey outcomes and the thirteen area-level variables under analysis. For each variable, the figure displays the estimate of the correlation and an approximation for its 95% confidence interval.

In general, the direction of the association between the different survey outcomes and the indicators of *disorder and decay* met most of the theoretical expectations. Positive Z-Y correlations were observed with self-reports of *poor health* (Fig. 4.4c) and *unsafe neighborhood* (Fig. 4.4d). Negative correlations were ob-
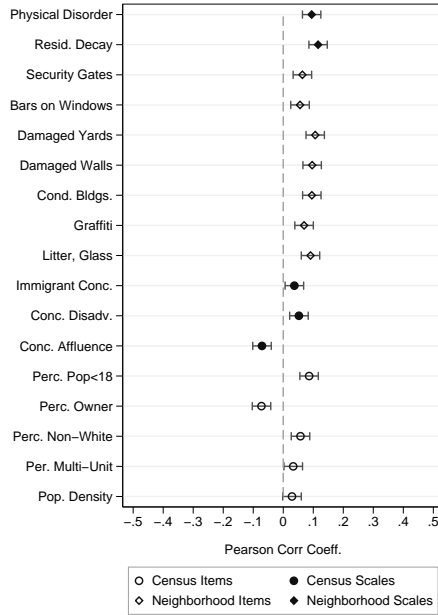
Fig. 4.2: Bivariate Correlation between the Household Response Indicator (P) and each Area-level variable (Z's).

served with self-reports of participation in *church* (Fig. 4.4a) and *social groups* (Fig. 4.4b), as well as reports on neighbors' perceived degree of *intergenerational closure* (Fig. 4.5b) and *reciprocated exchange* (Fig. 4.5a). For the items on use of *contraceptive methods* (Fig. 4.3a) and disapproval of *babies out of marriage* (Fig. 4.3b) the correlations were close to zero.

Another result that emerged from these figures is that there is a great deal of overlap between the estimates derived from scales and items. The 95% confidence intervals (CI) for the scales of disorder and decay, for example, overlapped with the CI of all seven neighborhood items most of the time. A similar phenomenon occurred between the CI of scales and items derived from Census records. The CI of concentrated *affluence* and *owner* occupied overlapped across all twelve survey outcomes. The CI for concentrated *disadvantage* and *immigrant* concentration

156

overlapped with percentage *multi-unit*, percentage *pop < 18* and *population density* across all survey outcomes. The item percentage *non-white*, however, showed stronger correlations for adult reports of *poor health*, participation in *social groups* and neighbors' *intergenerational closure.*

Looking at the patterns across the figures, we could conclude that there is no consistent evidence of large differences between the magnitude of the Z-Y correlations observed for the different sets of area-level measures. For some survey outcomes, the correlations with Census derived or interviewer derived measures were indistinguishable (*e.g.* Fig. 4.4a). For other variables, the indicators derived from interviewer observations were much stronger (*e.g.* Fig. 4.4b). For most variables, however, the confidence interval of the correlations with variables derived from both sources overlapped considerably.

Looking at the actual size of the Z-Y correlation shows, on the other hand, that they were somewhat higher than previously reported. The average size of the correlations, in absolute terms, was larger for variables such as unsafe neighborhood ($|\bar{r}| = .33$), *intergenerational closure* ($|\bar{r}| = .24$), and *participation in social groups* ($|\bar{r}| = .24$).[17] The area-level variable with the strongest correlation, across all survey outcomes, was the *physical disorder* scale. Excluding the variable on use of *contraceptive methods*, the range of the correlation for *physical disorder* ranged from .13 for disapproval of *babies out of marriage* to .44 for *unsafe neighborhood.*

In comparison, most of the Z-Y correlations reported by Kreuter et al. (2010)

---

[17] The estimator $|\bar{r}|$ corresponds to the unweighted average of the correlation for each survey outcome across the thirteen area-level variables under analysis.

Fig. 4.3: Bivariate Correlation between ADULT Y's and Area Z's.

(a) Use Contraception   *(b)* Babies Out of Marriage
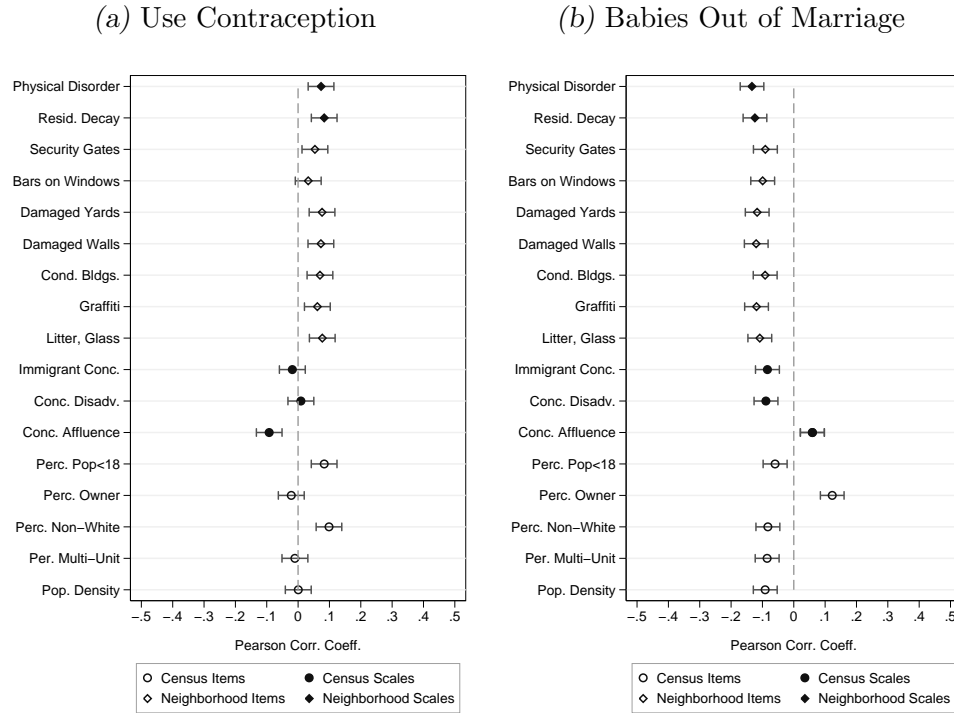
were below $|r| = 0.2$. A correlation of $|r| <= 0.1$, for example, was reported for the interviewer observations of *litter* and fifteen different survey outcomes on trust in people and government, general health and victimization derived from the European Social Survey (ESS). In this study the range of the Z-Y correlations for the interviewer observations of *litter* was $|r| = (0.11 - 0.38)$ and for *graffiti* was $|r| = (0.12 - 0.43)$.[18] This result is remarkable for two reasons. First, it shows that the strength of the association between the 'potential' auxiliary variables and the survey outcomes are higher than previously reported. Most interesting, however, is the fact that even the lowest performing item in term of measurement error (*litter*),

[18] Both of these estimates exclude the item on use of *contraceptive methods*.

can be as strong of a predictor as the item with the best measurement properties (*graffiti*). This pattern was observed across all eight items under study here.

It is important to notice, however, that even though some of the Z-Y correlations are much higher than previously reported, they are still considered 'low' under de Little and Vartivarian (2005) simulations. Under their framework, the magnitude of the Z-P correlations observed in the previous section and the Z-Y correlations observed here suggest that neither census variables nor neighborhood observations might have the potential for bias (or variance) reduction.

Results in this section suggest that there is not a clear advantage of interviewer derived variables over Census derived variables in terms of their 'potential' for bias reduction. Assessing the strength of the Z-P and the Z-Y correlations, however, was only the first step in the assessment of the potential for bias reduction. The next section shows the results from the two other criteria used here.

Fig. 4.4: Bivariate Correlation between ADULT Y's and Area Z's.

(a) Church Particip.
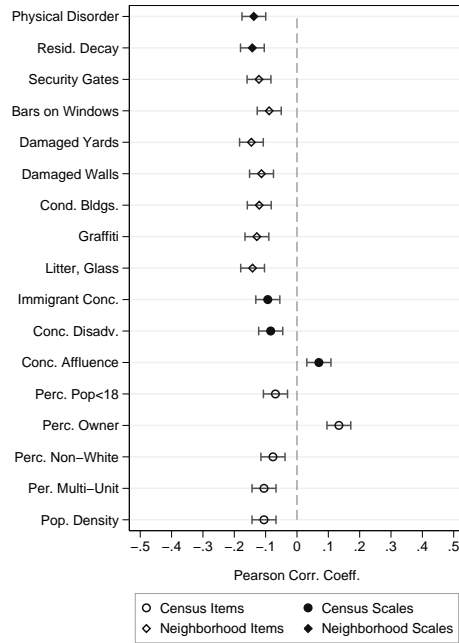
(b) Social Particip.
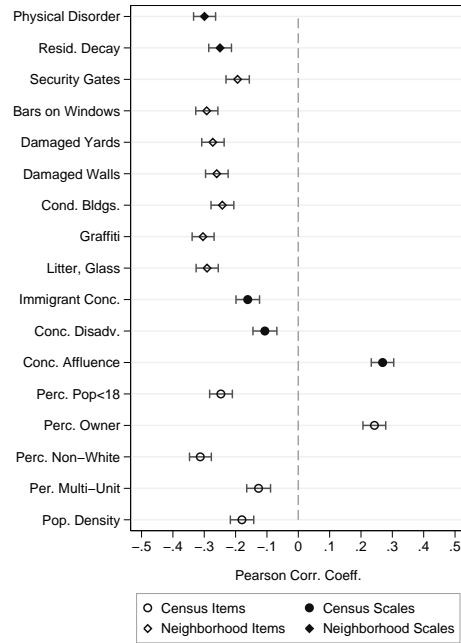
(c) Poor Health
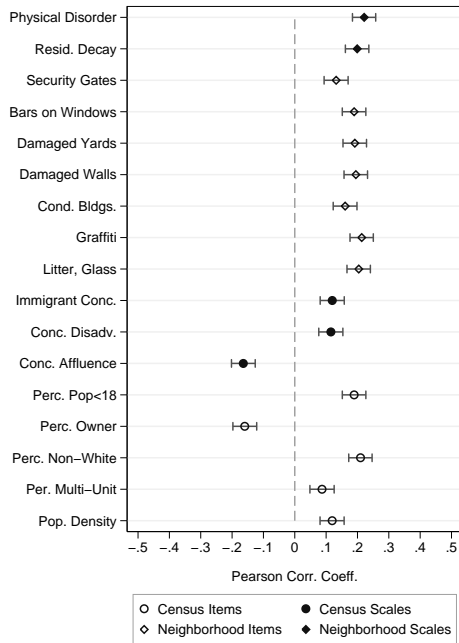
(d) Unsafe Neighborhood
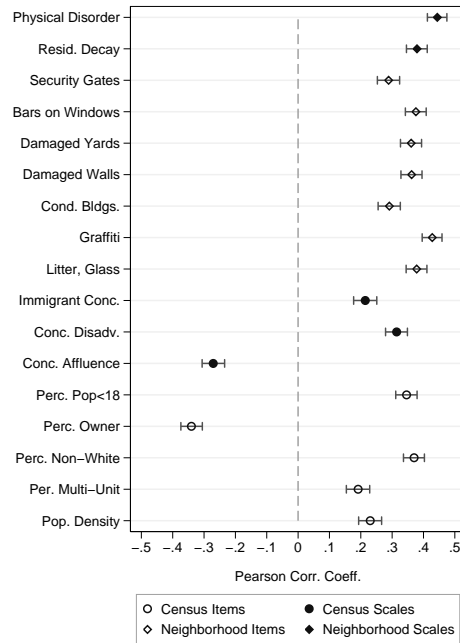
Fig. 4.5: Bivariate Correlation between NEIGHBORHOOD Y's and Area Z's.

(a) Reciprocated Exchange

(b) Intergenerational Closure
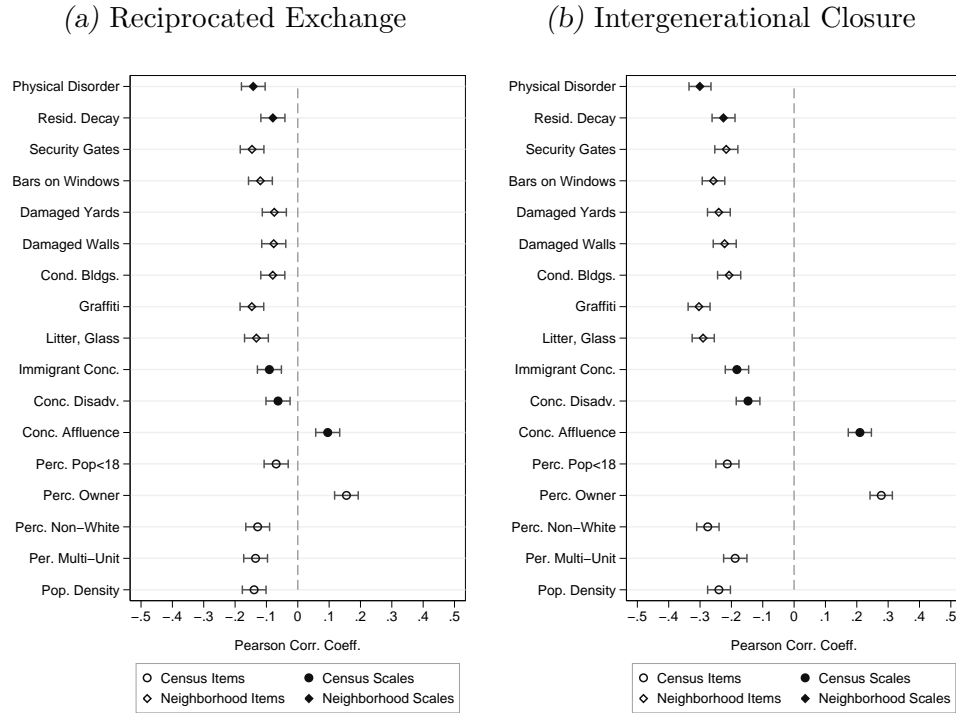
4.4.3 Development of New RSA Weights

The top two panels in Figure 4.6 display the distribution of the estimated response propensities across the ten deciles used as the starting point in the development of the new nonresponse adjustments. The white line in the middle of each box corresponds to the median response propensity and the end points of the box correspond to the interquartile range.

The propensities displayed in Figure 4.6a correspond to the reduced model that includes only the household-level items (Model 1). Figure 4.6b corresponds to the full model that includes the household-level items, the Census scales, and

the physical environment scales (Model 6). Both figures reveal high variability in the propensities of the households in the 10th decile – which present the lowest (predicted) response propensities. This pattern, a long tail in the lowest decile, emerged across all six models fitted (see figures in the Appendix).

This result can be problematic for theoretical and practical reasons. On one hand, it violates the assumption of homogeneity in the response probabilities within the adjustment classes. This is one of the conditions that nonresponse adjustment cells have to meet to reduce nonresponse bias (Lessler and Kalsbeek, 1992). On the other hand, it can yield high variability on the resulting weights, which has the potential to increase the variance of the weighted estimates.

Having large variability within the classes influences the decision regarding the type of nonresponse adjustment to use. Dever et al. (2011) provides a good discussion of the options available. Given the low variability in deciles (classes) $2 - 9$, I decided to use the inverse of the predicted response propensities as the adjustment in those classes. In the first decile, however, I decided to used the median of the response propensities as the nonresponse adjustment. This strategy should help reduce the variability of the resulting weights. Lepkowski et al. (2010, 19) describes a similar protocol as part of the nonresponse weighting scheme in the National Survey of Family Growth (NSFG) .

After creating the nonresponse adjustments, the next step consists of multiplying the new nonresponse adjustments derived by the household weights derived in expression (4.7). The two panels in the middle of Figure 4.6 display the resulting distributions for models 1 and 6. Model 1 shows large variability in the three lowest

deciles and model 6 in the bottom two. The median size of the household weights was close to $2,000$ for model 6, but almost $3,000$ for model 1. It can be inferred from these figures that the variability in the household-level weights is more driven by the variability in the households' selection probabilities than the variability in the households' (estimated) response propensities.

The last step in the development of the RSA weights was the multiplication of the newly derived household-level weights by the adult adjustment factor approximated by expression (4.8). The panels at the bottom of Figure 4.6 display the distributions of the final RSA weights after trimming the upper $95\%^{th}$ percentile to the $95\%^{th}$ value. As it is evident by looking at the distributions in Figures (4.6c)-(4.6f), trimming the weights reduced the size of the largest weights in the first decile by almost half. Descriptive statistics for the six new RSA weights are available in Table 4.7. A table with descriptive statistics for the new RSA weights before and after trimming is available in the Appendix.

One of the consequences of high variability in the final weights is the potential increase in the variance of the weighted estimates. The next section present results that estimate this effect for the new sets of RSA weights.

Fig. 4.6: Distribution of predicted propensities and final RSA weights within 'preliminary' propensity strata.

(a) Propensities M#1
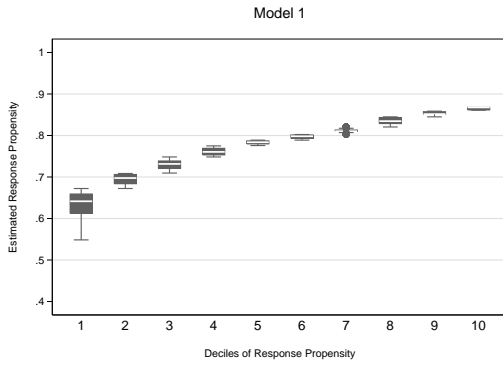
(b) Propensities M#6
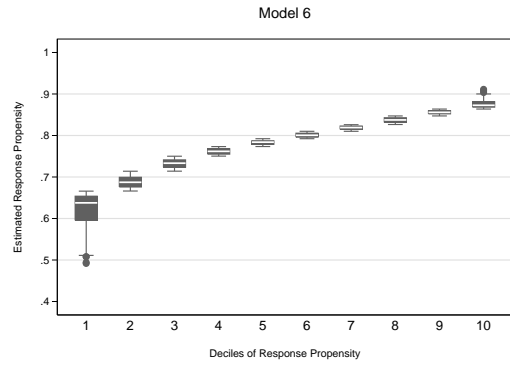
(c) HH Weights M#1
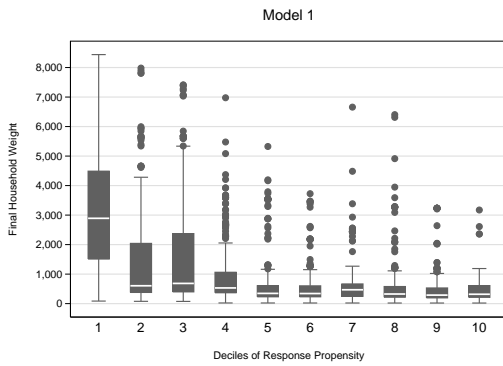
(d) HH Weights M#6
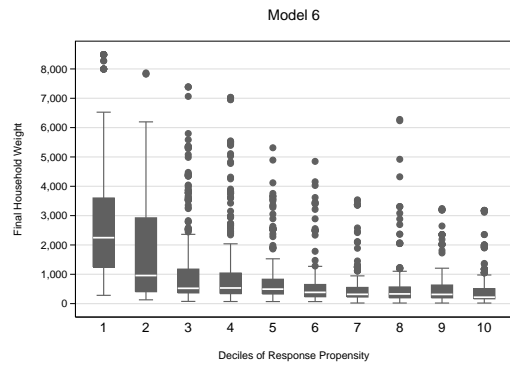
(e) RSA Weights M#1
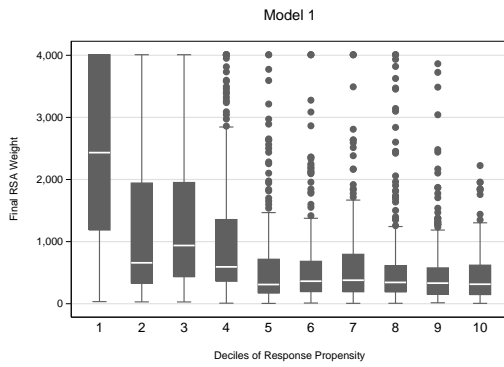
(f) RSA Weights M#6
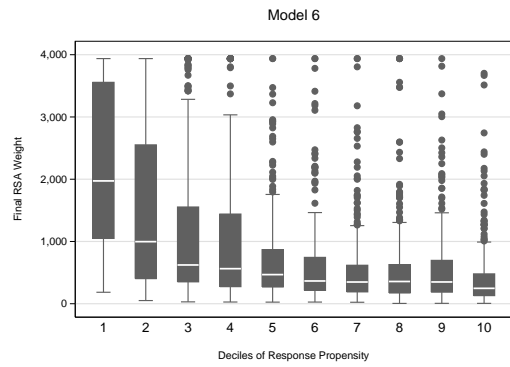
Tab. 4.7: Descriptive Statistics for New RSA Weights.

Nonresponse Adjustment Model	Param	Mean	Std. Dev.	Min	Max	n
Model 1: HHobs	7	957.4	1093.9	5.3	4009.2	2609
Model 2: HHobs, Census Scales	10	964.1	1108.4	5.2	4063.6	2609
Model 3: HHobs, Census Items	12	954.5	1085.8	5.1	3970.2	2609
Model 4: HHobs, NGobs Scales	9	950.4	1076.7	5.1	3924.2	2609
Model 5: HHobs, NGobs Items	14	945.0	1066.1	5.4	3897.0	2609
Model 6: HHobs, Census Scales, Ngobs Scales	12	950.9	1078.3	5.1	3937.8	2609

4.4.4   Kish's Variance Inflation Factor and MSE

It is important to remember that nonresponse adjustments are added to the weights with the expectation that they will increase the variance of the weighted estimates, but with the hope that they will help reduce the bias due to nonresponse. An estimate of the potential variance increase (1+L) in the weighted estimates is given in the first column of Table 4.8. Comparing across weighting schemes, the set of weights with the lowest potential to increase the variance of the weighted estimates were those including neighborhood observations in the corresponding response propensity models. Thus, based on $(1+L)$, the best performing set of weights are those associated with model 5 followed by models 4 and 6.

An estimate of the combined effect of variances and biases, the mean square error MSE for the weighted estimate, is also provided in Table 4.8. Estimates of MSE for the eight different survey outcomes yields consistent results. Overall, the best performing set of weights corresponds to model 5 followed by models 4 and 6. Difference in the MSE estimates derived from different weights is more noticeable

for the variables related to neighbor's behaviors – *reciprocated exchange* and *inter-generational closure.* And to a lesser extent for the variables on participation in *social groups* and *unsafe neighborhood.*[19]

Tab. 4.8: Estimates of Kish's Variance Inflation Factor $(1 + L)$ and MSE

Nonresponse Adjustment Model	(1+L)	Mean Square Error (MSE)*1,000							
		intclos	rexch	church	social	health	unsafe	ctruse	babies
Model 1: HHobs	2.3056	3.2586	2.5848	0.6258	1.2884	0.2614	0.7906	1.0575	0.5135
Model 2: HHobs, Census Scales	2.3216	3.2885	2.6002	0.6325	1.3048	0.2653	0.7898	1.0677	0.5203
Model 3: HHobs, Census Items	2.2941	3.3169	2.6155	0.6327	1.2855	0.2543	0.7848	1.0352	0.5101
Model 4: HHobs, Ngobs Scales	2.2835	3.1381	2.5252	0.6373	1.2710	0.2520	0.7732	**1.0247**	0.5127
Model 5: HHobs, Ngobs Items	**2.2727**	**3.0481**	**2.4963**	**0.6163**	**1.2460**	**0.2460**	**0.7548**	1.0283	**0.5073**
Model 6: HHobs, Census Scales, Ngobs Scales	2.2858	3.1324	2.5228	0.6391	1.2737	0.2510	0.7738	1.0296	0.5123

Based on these results, it is tempting to conclude that weighting schemes 4-6 performed better and would be preferred over weighting schemes 1-3. Before making such statements, however, it is important to put these results in perspective - the survey analysts' perspective. How large are these differences when estimating the weighted means and their standard errors?

---

[19] Kish's estimate of MSE is computed under the assumption that the adjustment variables are uncorrelated with the survey outcomes (Little and Vartivarian, 2005; Lepkowski et al., 2010). Simulations by Little and Vartivarian (2005) showed that Kish's RMSE is a good approximation of the empirical MSE with Z-Y correlations below .48. In this study, the highest Z-Y correlation was 0.44, thus I expect that Kish's approximation is appropriate here.

4.4.5 Estimate of the Weighted Means

Tables 4.9 and 4.10 display the estimates of the weighted means for the eight variables under analysis and their corresponding standard errors. Estimates of the standard errors were computed using Taylor Linearization and took into account the clustering and stratification of the L.A. FANS data.

Results in Table 4.9 show virtually no difference between the estimates derived under the six weighting schemes. For the estimates of percentages, not a single one of the estimates departed beyond one percentage point from each other. For the two continuous measures – *reciprocated exchange* and *intergenerational closure* – the differences across the models were not noticeable at the second decimal place.

The relative better performance of the weights based on model 5 is somewhat visible again in Table 4.10. This result provides some support for the claim that, if using covariates with high Z-Y correlation, one could obtain variance reductions (Little and Vartivarian, 2005).

One possible explanation for this lack of differentiation between the models could be the trimming of the upper 95th percentile. I re-computed the revised RSA weights without trimming and I re-estimated the results in Tables 4.9 and 4.10. The non-trimmed RSA weights yielded the same substantive results - no differentiation between the estimates derived using weights based on the six different models. As expected, the standard errors of the non-trimmed estimates of the weighted mean were somewhat higher than those from the trimmed weights. Tables with these results are available in the Appendix.

Tab. 4.9: Estimates of the Weighted Mean – $\hat{y}_z$.

Nonresponse Adjustment Model	rexch	intclos	church	social	health	unsafe	ctruse	babies
Model 1: HHobs	2.6511	3.6098	37.71	37.15	18.02	22.18	52.05	57.69
Model 2: HHobs, Census Scales	2.6521	3.6110	37.77	37.20	17.97	22.00	52.03	57.72
Model 3: HHobs, Census Items	2.6507	3.6102	37.84	37.12	17.91	22.08	51.99	57.70
Model 4: HHobs, Ngobs Scales	2.6510	3.6100	37.92	37.08	17.88	22.05	52.00	57.82
Model 5: HHobs, Ngobs Items	2.6499	3.6092	37.97	37.10	17.86	22.12	51.94	57.88
Model 6: HHobs, Census Scales, Ngobs Scales	2.6506	3.6099	37.94	37.09	17.87	22.04	51.99	57.83

Tab. 4.10: Estimates of the Standard Error of the Weighted Mean – $SE(\hat{y}_z)$.

Nonresponse Adjustment Model	rexch	intclos	church	social	health	unsafe	ctruse	babies
Model 1: HHobs	0.0374	0.0327	**1.66**	2.27	1.17	2.11	2.14	**1.50**
Model 2: HHobs, Census Scales	0.0375	0.0327	1.67	2.28	1.17	2.10	2.15	1.51
Model 3: HHobs, Census Items	0.0378	0.0329	1.68	2.27	1.16	2.11	2.13	**1.50**
Model 4: HHobs, Ngobs Scales	0.0369	0.0324	1.69	2.27	1.16	2.10	**2.12**	1.51
Model 5: HHobs, Ngobs Items	**0.0364**	**0.0323**	**1.66**	**2.25**	**1.15**	**2.08**	2.13	1.51
Model 6: HHobs, Census Scales, Ngobs Scales	0.0368	0.0324	1.69	2.27	**1.15**	2.10	2.13	1.51

4.5 Discussion

This paper aimed at contrasting the potential for bias reduction of area-level variables derived from Census records and interviewer observations of neighborhood characteristics. Correlational analyses showed modest Z-P correlations for both neighborhood observations and Census variables. Analyses of the Z-Y correlations showed somewhat higher correlations for survey variables such as participation in social groups, neighborhood safety and intergenerational closure.

Another consistent result was the high overlap between the correlations achieved using items and composite scores (scales). Except for a few variables, the confidence interval for both items and scales overlapped most of the time for both, interviewer observations and Census records. A clear advantage of the scales, however, was the interpretability of their results.

Results from the correlational analyses did not provide evidence to support the use of interviewer observations over Census records. Both performed consistently in terms of their patterns, and very close in terms of the magnitude of their Z-P and Z-Y correlations.

The second set of analyses aimed at testing directly the performance of the area-level variables when used in nonreponse weighting adjustments. For this purpose, a single weighting method – the propensity score method – was used to developed alternative sets of weights based on the different sets of area-level variables.

Analyses of the influence of the different wights on the estimate of the variance inflation factor and Kish's estimate of the mean square error yielded one consistent

result – the best performing set of weights were those that included interviewer observations. In particular, the model that included neighborhood items (Model 5). However, when looking at the actual estimates of the weighted means and their standard errors the differences between the six alternative adjustments were almost negligible.

This last result is surprising, considering that interviewer observations were much stronger predictors of cooperation than the Census derived variables in Paper 1. There were, however, some differences between the response propensity models implemented in Paper 1 and those implemented here that are worth noting.

The propensity models in Paper 1 modeled cooperation in the roster interview conditional on contact. The models used here, however, model cooperation against overall nonresponse, *i.e.*, the model here includes noncontacts ($n = 302$) among the nonrespondents. Evidence from Paper 1 suggests that the process of contactability was not influenced by tract-level variability during the roster interview. It is possible, then, that the stronger effects observed for the process of cooperation in Paper 1 could have been 'attenuated' here because of the inclusion of noncontacts in the response propensity model estimated here.

Another important difference with the analyses of Paper 1 was at the estimation stage. The complex structure of the L.A. FANS data was completely taken into account for the estimation of the response propensity models in Paper 1: random effects accounted for clustering of the primary sampling units, fixed effects accounted for stratification, and sampling weights were incorporated at the two levels of the multilevel model structure – Census tracts and households. In this paper, however,

I followed the recommendation by Little (1986) and estimated unweighted response propensity models for the development of the nonresponse adjustments. Evidence from the sensitivity analyses in Paper 1 suggest that, when modeling cooperation, the sampling weights were informative and ignoring them in the estimation affected the magnitude of both the standard errors and the regression coefficients. It is an empirical question how these effects could have influenced differently the predicted response propensities from the six alternative models compared here.

Results from this paper suggest that there is no gain, in term of potential for bias reduction, in using neighborhood observational data to develop nonresponse weighting adjustments if Census records are readily available. The last 'if' in the paragraph is important. A final post-hoc explanation for the consistency of the results across both sets of variables could be the close gap between the collection of the L.A. FANS data and the 2000 Census. Neighborhood observations were conducted between April 2000 and July 2001, which is somewhat close to the fieldwork of the 2000 Census. An extension of this study could be to assess whether this high degree of consistency between neighborhood observations and Census derived variables still holds as Census data gets out-dated for small areas such as Census tracts.

5. CONCLUSIONS AND FUTURE RESEARCH

Theories of survey participation reflect on societal-level mechanisms, acknowledge important variations in urbanicity, and highlight potential personal-level mechanisms when explaining survey participation. What has been neglected in the past is that processes on a lower level - the neighborhood that respondents live in - might influence the survey participation decision. In this dissertation I used the interviewer observations of neighborhood characteristics collected as part of the L.A. FANS data to fill this gap in the survey methodology literature. In summary, I found that characteristics of the social and physical environment are strong predictors of household cooperation. Some measures of the physical environment presented strong correlations with key survey outcomes. Not surprisingly, however, some measures show better measurement error properties than others. Overall, the results of this dissertation can be relevant for those who want to increase response rates by tailoring efforts to neighborhood characteristics.

Looking into more detail, in Paper 1 I extended the survey participation model developed by Groves and Couper (1998) to include neighborhood physical conditions and to disentangle the relationship between neighborhood social processes and neighborhood socio-economic composition. From there I developed theoretically driven measures of neighborhood social processes and the physical environment us-

ing data from both the L.A. FANS interview and from interviewer observations of the sampled neighborhoods. Factor analysis using Census demographics was used to extract a reduced set of indicators of the socio-economic composition of the Los Angeles neighborhoods. I tested the influence of the different measures of neighborhood characteristics on contactability and cooperation, controlling for household-level demographic characteristics (available from the screener interview). Neighborhood social processes and physical conditions were significant predictors of household cooperation (after controlling for household predictors). They explained about 6% of the total variance in cooperation. Interestingly, the effects found were contrary to the expectations one would derive from the 'collective efficacy theory' or the 'broken-windows theory'. Persons living in neighborhoods with more signs of physical disorder and residential decay were more likely to cooperate in the L.A.FANS survey than those living in well maintained neighborhoods. A post-hoc explanation for these results could be that the L.A. FANS is a survey about the families and the neighborhoods in which they live. The interviewers might have pitched possible improvements for the neighborhoods as reasons to participate in the survey. While the advance mailings and the news briefs only talked about research insights that will be gained from this study, the interviewer manual did point out that the information gathered in the survey "will be used [. . .] to determine how to improve neighborhoods [. . .] in Los Angeles". In this case survey methodologists would argue, with 'leverage salience theory' (Groves et al., 2000), that the topic is particularly salient in the more dilapidated neighborhoods. However, this study was not set out as a test of these competing theories. Future studies that do investigate the influence of

neighborhood physical conditions on participation should collect such observational variables on a larger range of topical surveys than the L.A. FANS.

It is also important to note that the L.A.FANS is a survey conducted in two phases. First there is a screener stage at which the interviewers only collect basic information about the households. This screener information is then used for selecting cases for the main survey. Explaining cooperation to the main survey is therefore restricted to cases that already have been successfully contacted. If the research interest is overall nonresponse which includes contactability as well as cooperation, other mechanisms might be present that have not been studied here.

There are some shortcomings in Paper 1 that deserve to be discussed. There is an additional variance component that one can tease apart from the components discussed here. Interviewers also vary in their ability to gain cooperation. In nation-wide surveys there is usually a single interviewer per sampled area. However, given the small area of the L.A. FANS survey, interviewers worked in several different neighborhoods, and each neighborhood included interviews done by different interviewers. Because of this design, I do not expect my results to change substantially when extracting interviewer variance components. Researchers of interviewer effects might want to use the L.A. FANS data set, given that it is one of the few that have a partially interpenetrated design.

The measures used to contrast the effects of social and physical environment as predictors for survey participation came from two different sources. Survey respondents reported on social activities in the neighborhood and the like, from which *social cohesion* and *informal social control* scales were formed. Interviewers ob-

served signs of *physical disorder*, *social disorder* and *residential decay.* This strategy matches data collection efforts known from other substantive studies on neighborhood effects. Survey methodologists are well aware of possible measurement errors in reports given by survey respondents and questionnaire developers try hard to reduce them. The measurement error properties of interviewer observations are much less studied in the survey literature. This has not discouraged survey researchers from collecting such data. Interviewers are often charged with making some basic neighborhood and household observations during the recruitment process – data that are usually part of contact protocols. Thus Paper 2, while illuminating measurement errors for variables used in the participation models for L.A.FANS, also speaks to a larger class of surveys that collect such data.

The measurement error analyses performed in Paper 2 differ from what has been done to study interviewer observations in the sociological literature in so far as they add interviewer effects into the measurement error models. Taking the multilevel nature of the data into account, the models I used to study the reliability of the neighborhood variables go beyond 'classical test theory' approaches. Not surprisingly, different items showed different measurement error properties. The best performing items were those that reflect more or less stable characteristics of the structures themselves. Interviewers differed most in their perception of litter in the neighborhood, and also in their assessments of gates as being security gates.

In my attempt to explain interviewers' perceptions of disorder I did find that characteristics of the interviewers' that I suspected would shape their perceptions did not have the hypothesized effects. This could be due to a successful uniform

175

training all interviewers received. Instead, interviewer perceptions were driven to a large extent by the characteristics of the neighborhoods – in particular the socio-economic composition.

As mentioned before, interviewer observations are increasingly common variables on contact protocols, such as the NSFG and the ESS. They are sometimes used to guide field efforts, or to control the interviewers themselves. Other surveys, like the U.S. Consumer Expenditure Survey, uses these observations for nonresponse adjustments.

In Paper 3 I assessed the potential of neighborhood observations to be used in nonresponse adjustments. In particular, the indicators of physical disorder and residential decay seemed to be good candidates, given their predictive power in the models of survey participation. If the underlying mechanisms that lead to participation in the survey itself also affect how respondents answer survey questions then nonresponse bias will appear. Capturing proxy variables for this participation mechanism and subsequently using them in nonresponse adjustments should reduce a possible nonresponse bias. Which proxy variables serve best to capture the mechanism is an empirical question. The interviewer observations of disorder and decay are one option. Several Census variables are strong correlates of the observations made by the interviewers and thus could also be suited for nonresponse adjustments.

As a preliminary assessment, I estimated the strength of the correlations between the neighborhood variables considered for adjustment and the response indicator, as well as the correlation between the adjustment candidate variables and the survey outcomes. I used thirteen neighborhood variables and eight survey items in

176

these analyses. The neighborhood variables considered for adjustment did not correlate highly with the response indicator. This is not too surprising given what was learned from Paper 1: that neighborhood characteristics did influence participation but explained only a small fraction of the variance compared to household-level characteristics. However, the correlations with the survey outcome variables were larger than previously reported in the literature. This is good news for L.A.FANS – having variables in the adjustment models that correlate highly with the survey outcome variable(s), have the potential to reduce the variance of the weighted estimates.

Consistent with these findings, the different sets of weighting adjustments did not lead to noticeable differences in the weighted means, but a slight reduction in the variances was achieved. Of course means are only one statistic of interest, the effects could be different for totals or regression coefficients. Extensions of this research should examine those more closely. Another interesting extension of the models in Paper 3 would be the inclusion of Census based variables using data from the 1990 and 1980 Census records instead of those from 2000, to test the post-hoc hypothesis that adjustments based on Census data from out-dated Censuses will not perform similarly to those adjustments based on interviewer observations.

One important difference between the response propensity models that feed into the nonresponse adjustments in Paper 3 and those used in Paper 1 is that the former predicted overall nonresponse including non-contact, not just refusal. The efficiency of the neighborhood variables in the nonresponse adjustment may have been attenuated due to the fact that the mechanisms these represent do not influence contact. Paper 1 gave an indication for this as well. Thus while the use

of the neighborhood variables for overall nonresponse adjustment is negligible, they might nevertheless be useful to guide fieldwork decisions for example in responsive designs, as in two-phase sampling for nonrespondents. Those would be interesting empirical questions future research can tackle.

In working with the L.A. FANS data and thinking about the link between neighborhood effects, interviewer observations and improvements of survey estimates, several open questions emerged. I summarize them here into three additional lines of research that I hope will inspire readers of this dissertation.

The first line of research would focus on the unit of observation and the level at which neighborhood effects are likely to affect such units. Starting with the latter, one question that I believe is important to address is what geographical area is relevant to explain different nonresponse processes for different stages of interviewing. Several preliminary analyses that I performed leading up to what is discussed in Paper 1 showed much larger variations in the noncontact rates at the block than on the track level. The opposite was the case for refusal (conditional on contact). I observed this pattern for the roster interview and the household interview. Developing theories and exploring the differences of block-level mechanisms versus tract-level mechanisms for different response outcomes and stages of the recruitment process could be useful. If certain block-level covariates are strong predictors of contactability for example, fieldwork decisions on when or how to contact certain blocks could be informed by such covariates. This brings me to another aspect that is relevant not only for this particular nonresponse study performed here, but for non-experimental research of face-to-face survey participation in general.

Most nonresponse research is household-oriented. On the surface this is a reasonable approach, given that the survey participation decision is household based. However, households are not approached to participate in a survey at random at all possible times and days. The decision to work on a particular case at a particular time relies, in face-to-face surveys, usually in the hands of the interviewer. Thus, in order to plan optimal interventions to increase response rates, a shift in focus seems appropriate. When interviewers drive by the sampled segments they most likely attempt to contact all unresolved sample units within the segment. Recent methodological work as part of the National Survey of Family Growth supports this view. Here, attempts to force interviewers to act on particular cases and not others (within a segment) were not successful. Thus relevant questions are - what influences interviewers decisions to spent more time in some areas (and not in other areas)? Which areas will interviewers visit more or less often? Do interviewers use different recruitment strategies in some areas (and not in other areas)? If area level characteristics affect the effort interviewers put into the approach and recruitment of respondent cases then those need to be taken into account when managers re-organize fieldwork procedures throughout data collection.

This discussion of neighborhood effects on the decisions interviewers make leads to the second line of research I see needed. The revised survey participation model I introduced in Paper 1 postulates that the neighborhood physical environment will have an influence on both – householders living in the areas and the interviewers working in those areas. Theoretical thinking about contextual influences on interviewers' scheduling attempts and other aspects of their work (*e.g.* how timidly

179

or forcefully they approach a household or their expectations on their likelihood of success in a certain neighborhood) should be extended. This effort would shed light on the variations in response rates across interviewers. So far those variations are either attributed to interviewers, or to characteristics of the sample units in a given area. The area itself has largely been left untouched in such discussions (probably with the exception of gated communities). Knowing more about the source of this variation can inform interviewer training.

A third and somewhat different line of research is related to auxiliary variables in general. This includes taking a closer look at administrative data and other external sources that provide information about the neighborhoods. For substantive reasons L.A.FANS collected a wide variety of data from external sources including information on neighborhood services, population characteristics, housing characteristics, family and household socioeconomic status, education, employment and earnings, and health care services and facilities. Some of those are likely to be close proxy variables for substantive outcomes in the L.A.FANS and could be candidates for nonresponse adjustment just like the interviewer observations. Depending on the survey variables of interest ecological observations made by the interviewer such as the types of land use, presence of institutions (churches, schools, banks), type of commercial outlets (fast foods, liquor stores) might also be considered. Paper 3 already showed reasonably sized correlations of interviewer observations on physical characteristics of local areas with L.A. FANS survey outcomes. Learning about these relationships could assist survey managers to assess whether survey outcomes could be affected by nonresponse bias as the data collection progresses. In effect,

an interesting possibility would be the development of neighborhood observations that are hypothesized to be related to particular survey outcomes. These developments should be informed by substantive theories, such as those in environmental psychology, urban sociology and social epidemiology.

6. APPENDICES

6.1 Appendix to Paper 1

Tab. 6.1: Descriptive Statistics for Household Observed Characteristics (Part 1/2)

Variable	All Eligible Cases			Only Contacted Cases		
	Screener	Roster	HH Int.	Screener	Roster	HH Int.
	(8,518)	(4,010)	(3,083)	(8,015)	(3,708)	(2,707)
Type of Housing						
Apartment and Others	53.8%	47.4%	48.6%	53.3%	46.8%	49.1%
Single Family Home	46.2%	52.6%	51.4%	46.7%	53.2%	50.9%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Estimate of Rent						
Less than $500/month	11.6%	14.9%	15.1%	11.6%	14.9%	14.9%
$500-$999/month	40.4%	50.0%	51.5%	41.3%	49.9%	51.8%
$1000-$1999/month	29.1%	23.7%	23.2%	28.5%	24.0%	23.1%
$2000-$2999/month	8.9%	5.6%	5.1%	8.3%	5.5%	5.1%
$3000+/month	10.0%	5.8%	5.1%	10.2%	5.6%	5.1%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Children Status						
HH without Children	56.7%	26.2%	23.2%	59.0%	25.6%	23.8%
HH with Children	43.3%	73.8%	76.8%	41.0%	74.4%	76.2%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Tab. 6.2: Descriptive Statistics for Respondent Observed Characteristics (Part 2/2)

Variable	All Eligible Cases			Only Contacted Cases		
	Screener	Roster	HH Int.	Screener	Roster	HH Int.
	(8,518)	(4,010)	(3,083)	(8,015)	(3,708)	(2,707)
Race of Respondent						
Resp. Race Latino	34.4%	52.4%	54.8%	34.8%	52.3%	54.0%
Resp. Race White	43.0%	27.3%	25.5%	43.1%	27.5%	26.4%
Resp. Race Black	8.2%	9.8%	9.9%	8.4%	9.9%	10.2%
Resp. Race Other	14.4%	10.4%	9.8%	13.8%	10.3%	9.4%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Age of Respondent						
18-24 yrs	7.3%	8.0%	8.4%	7.3%	8.0%	8.1%
25-34 yrs	26.5%	30.6%	33.1%	26.8%	31.0%	33.8%
35-54 yrs	43.1%	47.6%	46.1%	42.5%	47.1%	45.8%
55-69 yrs	16.6%	10.5%	9.7%	16.9%	10.5%	9.5%
70+ yrs	6.4%	3.3%	2.8%	6.5%	3.4%	2.9%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Gender of Respondent						
Male	43.1%	37.5%	36.6%	43.4%	37.0%	36.2%
Female	56.9%	62.5%	63.4%	56.6%	63.0%	63.8%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Tab. 6.3: Descriptive Statistics for Census variables included in Factor Analysis (n=65).

Variable	Mean	Std. Dev.	N
perc. adults 25+ yrs with 13+ yrs school	0.35	0.24	65
perc. of families with income $75k and over	0.22	0.20	65
perc. workers in exec/prof occupations	0.27	0.18	65
perc. of adults who are spanish speakers in tract	0.47	0.28	65
perc. pop hispanic-latino in tract	0.55	0.30	65
perc. pop foreign-born in tract	0.40	0.15	65
perc. pop non-citizens in population in tract	0.27	0.15	65
perc. of occupied housing that is owner-occupied	0.43	0.26	65
perc. of individuals in poverty in tract	0.23	0.14	65
perc. hh receiving public assistance	0.09	0.07	65
perc. hhs headed by females with children	0.11	0.06	65
perc. pop non-hispanic black in tract	0.08	0.10	65
perc. non-family households in tract	0.26	0.14	65
perc. hhs with children	0.45	0.14	65
perc. occupying same dwelling as in 1995 in tract	0.51	0.10	65
perc. unemployed in civilian labor force	0.06	0.03	65

Tab. 6.4: Factor Loadings and Uniqueness from Factor Analysis

Variable	Factor1 C. Affluence	Factor2 Immigrant C.	Factor3 C. Disadvantage	Factor4 Family Str.	Factor5 Resid. Stab.	Uniqueness
perc. of families with income $75k and over	**-0.7139**	-0.4138	-0.3641	-0.1070	-0.2235	0.0285
perc. adults 25+ yrs with 13+ yrs school	**-0.8255**	-0.3067	-0.3451	-0.2781	-0.0335	0.0091
perc. workers in exec/prof occupations	**-0.7778**	-0.3015	-0.3014	-0.3200	-0.1467	0.0225
perc. pop hispanic-latino in tract	**0.9148**	0.3132	0.1141	0.1876	0.0125	0.0021
perc. of adults who are spanish speakers in tract	**0.8868**	0.3651	0.1413	0.1760	0.0075	0.0064
perc. pop foreign-born in tract	0.3856	**0.8977**	0.0301	0.0750	0.0518	0.0337
perc. pop non-citizens in population in tract	0.5121	**0.8085**	0.1859	0.0710	0.1491	0.0146
perc. of occupied housing that is owner-occupied	-0.3417	**-0.6640**	-0.3187	0.3499	-0.3792	0.0411
perc. of individuals in poverty in tract	0.5139	0.4806	**0.5748**	0.0737	0.3179	0.0579
perc. hh receiving public assistance	0.3521	0.3015	**0.7896**	0.1950	0.1587	0.0950
perc. hhs headed by females with children	0.5375	0.2388	**0.6710**	0.1762	0.1217	0.1334
perc. pop non-hispanic black in tract	0.0370	-0.2227	**0.6672**	0.0176	0.0515	0.3411
perc. non-family households in tract	-0.4355	0.0405	-0.0982	**-0.8435**	0.1986	0.0446
perc. hhs with children	0.6297	0.1478	0.3287	**0.6425**	0.0404	0.0366
perc. occupying same dwelling as in 1995 in tract	0.0699	-0.3392	-0.1064	0.2679	**-0.7289**	0.2546
perc. unemployed in civilian labor force	0.2604	0.0387	0.3870	-0.0049	**0.6087**	0.3590

Tab. 6.5: Descriptive Statistics for Neighborhood Census Characteristics

Variable	Mean	Std. Dev.	N
Factor Scores Indicators			
Concentrated Affluence	0.00	0.99	65
Residential Stability	0.00	0.88	65
Family Structure	0.00	0.96	65
Concentrated Disadvantage	0.00	0.95	65
Immigrant Concentration	0.00	0.98	65
Groves and Couper Indicators			
Perc. Multi-Unit	0.39	0.28	65
Perc. Owner	0.43	0.26	65
Race Non-White	0.76	0.26	65
Perc. Pop < 18yrs	0.30	0.09	65
Crime	0.43	0.11	65
Population Density	14,836	10,462	65

Tab. 6.6: Number of Adult (RSA) Respondents per Tract used to derive measures of the Neighborhood Social Environment.

Variable	Mean	Std. Dev.	Min.	Max.	N
Social Cohesion	40.65	6.19	27	56	2,585
Informal Social Control	40.64	6.2	27	56	2,584

Tab. 6.7: Descriptive statistics for items used to derive measures of the Neighborhood Social Environment.

Variable	Mean	Std. Dev.	Min.	Max.	N
Social Cohesion					
close knit neighborhood	3.15	1.14	1	5	2585
willing to help neighbors	3.67	0.94	1	5	2585
neighbors get along	3.56	0.96	1	5	2585
neighbors share values	3.09	1.05	1	5	2585
neighbors can be trusted	3.40	1.02	1	5	2585
Informal Social Control					
NG act if kids hanging out	3.45	1.31	1	5	2584
NG act if kids paint grafitti	3.87	1.26	1	5	2584
NG act if kids disrespect adults	3.34	1.24	1	5	2584

Tab. 6.8: Number of Cases used to derive measures of the Neighborhood Physical Environment

Variable	Mean	Std. Dev.	Min.	Max.	N
Cases per Tract					
Blocks	7.48	3.11	2	14	2040
Block Faces	37.14	13.48	8	62	2040
Interviewers	2.96	1.05	1	5	2040
Cases per Block					
Block Faces	6.5	5.06	2	28	2040
Interviewers	1	0	1	1	2040

Tab. 6.9: Descriptive statistics for original items used to derive measures of the Neighborhood Physical Environment.

Variable	Mean	Std. Dev.	Min.	Max.	N
Physical Disorder					
abandoned cars on street	1.10	0.37	1	4	2040
trash or junk on street	1.91	0.99	1	4	2040
garbage, litter or broken glass on street	2.28	1.02	1	4	2040
needles, syringes, condoms or drug re-items on street	1.04	0.24	1	4	2040
empty beer containers or liquor bottles on street	1.29	0.62	1	4	2040
cigarettes or cigar butts or discarded packages on street	1.98	0.99	1	4	2040
graffiti on buildings, sidewalks, walls or signs	1.87	0.98	1	4	2040
painted-over graffiti on buildings, sidewalks, etc	1.50	0.79	1	4	2040
Residential Decay					
condition of residential buildings (rev coded)	2.39	0.89	1	5	1819
# houses/appts burned out, boarded up, or abandoned	1.10	0.34	1	5	1819
# vacant lots on the block	1.22	0.57	1	5	1819
# houses/appts w/peeling paint or damaged exterior walls	2.15	1.03	1	5	1819
# houses/appts well tended yards or gardens (rev coded)	2.62	1.29	1	5	1819
Residential Security					
# houses/appts w/window bars or gratings on doors/windows	2.49	1.35	1	5	1819
# houses/appts w/sign private security	1.84	1.06	1	5	1819
# houses/appts w/sign protected by dog	1.44	0.70	1	5	1819
# houses/appts w/security gates or security fences	2.18	1.26	1	5	1819

Tab. 6.10: Descriptive statistics for dichotomized items used to derive measures of the Neighborhood Physical Environment.

Variable	Mean	Std. Dev.	Min.	Max.	N
Social Disorder					
saw group appear to be gang on the block	0.01	0.11	0	1	2023
saw adults loitering, or hanging out on block	0.08	0.27	0	1	2031
saw prostitutes on the block	0.00	0.07	0	1	2034
saw homeless people or people begging on the block	0.02	0.15	0	1	2038
saw people selling illegal drugs on the block	0.01	0.08	0	1	2039
saw people drinking alcohol openly on the block	0.02	0.14	0	1	2039
saw intoxicated people on the block	0.01	0.11	0	1	2037
Physical Disorder					
Abandoned cars	0.08	0.27	0	1	2040
Trash or junk	0.56	0.50	0	1	2040
Garbage, litter or broken glass	0.75	0.43	0	1	2040
Needles or drug re-items	0.04	0.19	0	1	2040
Empty beer/bottles	0.22	0.41	0	1	2040
Cigarettes/discarded packages	0.61	0.49	0	1	2040
Graffiti on walls/signs	0.54	0.50	0	1	2040
Painted-over graffiti	0.36	0.48	0	1	2040
Residential Decay					
Condition of resid. bldgs. (rev)	0.82	0.38	0	1	1819
Houses/appts burned out	0.09	0.29	0	1	1819
Vacant lots on the block	0.16	0.37	0	1	1819
Houses/appts w/damaged walls	0.69	0.46	0	1	1819
Houses/appts well tended yards (rev)	0.77	0.42	0	1	1819
Residential Security					
Houses/appts w/window bars	0.65	0.48	0	1	1819
Houses/appts w/security sign	0.50	0.50	0	1	1819
Houses/appts w/dog sign	0.34	0.48	0	1	1819
Houses/appts w/security gates	0.58	0.49	0	1	1819
Neighborhood watch sign	0.16	0.37	0	1	1816

Tab. 6.11: Screener Interview: Pearson Correlation of Screener Interview Rates and Neighborhood Characteristics (n=65)

Neighborhood	Contact Rate			Cooperation Rate			Response Rate		
Variable	Corr	Min95	Max95	Corr	Min95	Max95	Corr	Min95	Max95
Conc. Affluence	-0.1326	-0.3647	0.1150	-0.3507	-0.5477	-0.1167	-0.2884	-0.4973	-0.0478
Resid. Stability	0.0175	-0.2274	0.2603	0.0624	-0.1844	0.3017	0.0424	-0.2036	0.2833
Family Structure	0.2072	-0.0387	0.4294	-0.2431	-0.4597	0.0008	0.0435	-0.2025	0.2844
Conc. Disadvantage	-0.0348	-0.2763	0.2109	0.1343	-0.1133	0.3662	0.0391	-0.2068	0.2803
Immigrant Conc.	-0.1758	-0.4024	0.0712	-0.2301	-0.4488	0.0146	-0.2567	-0.4711	-0.0137
Perc. Owner	0.0950	-0.1524	0.3313	-0.0365	-0.2779	0.2093	0.0555	-0.1910	0.2954
Perc. Pop¡18	0.2108	-0.0349	0.4325	0.0954	-0.1520	0.3316	0.2205	-0.0247	0.4407
Perc. Non-White	0.0560	-0.1905	0.2958	0.0658	-0.1810	0.3048	0.0824	-0.1648	0.3198
Perc. Multi-Unit	-0.1905	-0.4151	0.0560	-0.0394	-0.2806	0.2065	-0.1709	-0.3982	0.0762
Crime	-0.0850	-0.3222	0.1623	0.1303	-0.1174	0.3626	0.0000	-0.2439	0.2439
Pop. Density	-0.1988	-0.4222	0.0474	-0.1509	-0.3808	0.0966	-0.2340	-0.4521	0.0105
Social Cohesion	0.0742	-0.1728	0.3125	-0.1311	-0.3634	0.1165	-0.0093	-0.2526	0.2352
Informal Social Control	0.0371	-0.2086	0.2785	-0.1636	-0.3919	0.0836	-0.0548	-0.2948	0.1916
Physical Disorder	0.0224	-0.2227	0.2649	0.1831	-0.0636	0.4088	0.1126	-0.1350	0.3469
Social Disorder	-0.0128	-0.2559	0.2319	0.1270	-0.1207	0.3597	0.0559	-0.1906	0.2957
Residential Decay	0.0760	-0.1711	0.3141	0.1448	-0.1028	0.3754	0.1347	-0.1129	0.3665

Tab. 6.12: Roster Interview: Pearson Correlation of Roster Interview Rates and Neighborhood Characteristics (n=65)

Neighborhood Variable	Contact Rate			Cooperation Rate			Response Rate		
	Corr	Min95	Max95	Corr	Min95	Max95	Corr	Min95	Max95
Conc. Affluence	-0.0879	-0.3249	0.1594	-0.2954	-0.5031	-0.0555	-0.2883	-0.4972	-0.0477
Resid. Stability	0.0406	-0.2053	0.2817	-0.0220	-0.2645	0.2231	-0.0024	-0.2461	0.2417
Family Structure	-0.0580	-0.2977	0.1886	0.0028	-0.2412	0.2466	-0.0345	-0.2761	0.2111
Conc. Disadvantage	0.1668	-0.0803	0.3947	0.1505	-0.0969	0.3805	0.2176	-0.0278	0.4382
Immigrant Conc.	-0.2557	-0.4702	-0.0125	0.4165	0.1921	0.5995	0.2043	-0.0417	0.4269
Perc. Owner	0.0286	-0.2168	0.2706	-0.4219	-0.6037	-0.1984	-0.3340	-0.5344	-0.0981
Perc. Pop¡18	0.1475	-0.1000	0.3778	0.3268	0.0901	0.5285	0.3466	0.1122	0.5445
Perc. Non-White	-0.0342	-0.2758	0.2115	0.3078	0.0691	0.5132	0.2322	-0.0123	0.4506
Perc. Multi-Unit	-0.0976	-0.3335	0.1499	0.2780	0.0365	0.4887	0.1823	-0.0645	0.4080
Crime	-0.0769	-0.3149	0.1702	0.2273	-0.0175	0.4465	0.1438	-0.1037	0.3746
Pop. Density	-0.1887	-0.4135	0.0579	0.3231	0.0859	0.5255	0.1628	-0.0844	0.3912
Social Cohesion	-0.0505	-0.2908	0.1958	-0.4601	-0.6330	-0.2436	-0.4086	-0.5933	-0.1829
Informal Social Control	-0.1070	-0.3420	0.1406	-0.2409	-0.4579	0.0032	-0.2588	-0.4729	-0.0159
Physical Disorder	0.0687	-0.1782	0.3075	0.5002	0.2919	0.6632	0.4481	0.2293	0.6239
Social Disorder	-0.1091	-0.3438	0.1385	0.4122	0.1871	0.5962	0.2701	0.0280	0.4822
Residential Decay	0.1378	-0.1098	0.3693	0.4631	0.2472	0.6353	0.4574	0.2403	0.6309

Tab. 6.13: Household Interview: Pearson Correlation of Household Interview Rates and Neighborhood Characteristics (n=65)

Neighborhood	Contact Rate			Cooperation Rate			Response Rate		
Variable	Corr	Min95	Max95	Corr	Min95	Max95	Corr	Min95	Max95
Conc. Affluence	0.0693	-0.1776	0.3080	0.2330	-0.0115	0.4513	0.1506	-0.0969	0.3805
Resid. Stability	-0.0122	-0.2553	0.2324	-0.1564	-0.3856	0.0910	-0.0669	-0.3058	0.1799
Family Structure	-0.2655	-0.4784	-0.0231	-0.1084	-0.3432	0.1392	-0.2792	-0.4898	-0.0379
Conc. Disadvantage	-0.0327	-0.2745	0.2129	-0.0683	-0.3071	0.1785	-0.0609	-0.3003	0.1858
Immigrant Conc.	-0.2668	-0.4795	-0.0245	0.0472	-0.1989	0.2878	-0.2257	-0.4451	0.0192
Perc. Owner	0.1439	-0.1037	0.3746	-0.0447	-0.2855	0.2014	0.1177	-0.1299	0.3515
Perc. Pop¡18	-0.1954	-0.4193	0.0509	-0.1776	-0.4040	0.0693	-0.2456	-0.4618	-0.0018
Perc. Non-White	-0.2153	-0.4363	0.0302	-0.2779	-0.4887	-0.0365	-0.2999	-0.5067	-0.0604
Perc. Multi-Unit	0.0279	-0.2175	0.2699	0.1982	-0.0480	0.4217	0.0951	-0.1524	0.3313
Crime	-0.0715	-0.3100	0.1755	-0.2524	-0.4675	-0.0090	-0.1603	-0.3890	0.0870
Pop. Density	-0.1675	-0.3952	0.0797	-0.0286	-0.2706	0.2168	-0.1676	-0.3954	0.0795
Social Cohesion	0.0640	-0.1828	0.3031	0.0919	-0.1555	0.3285	0.0957	-0.1517	0.3319
Informal Social Control	-0.0393	-0.2806	0.2065	0.1574	-0.0899	0.3865	0.0269	-0.2185	0.2690
Physical Disorder	-0.2036	-0.4263	0.0424	-0.1275	-0.3602	0.1202	-0.2368	-0.4545	0.0075
Social Disorder	-0.1044	-0.3396	0.1432	-0.0709	-0.3094	0.1761	-0.1263	-0.3592	0.1213
Residential Decay	-0.0665	-0.3054	0.1803	-0.1612	-0.3898	0.0860	-0.1233	-0.3565	0.1244
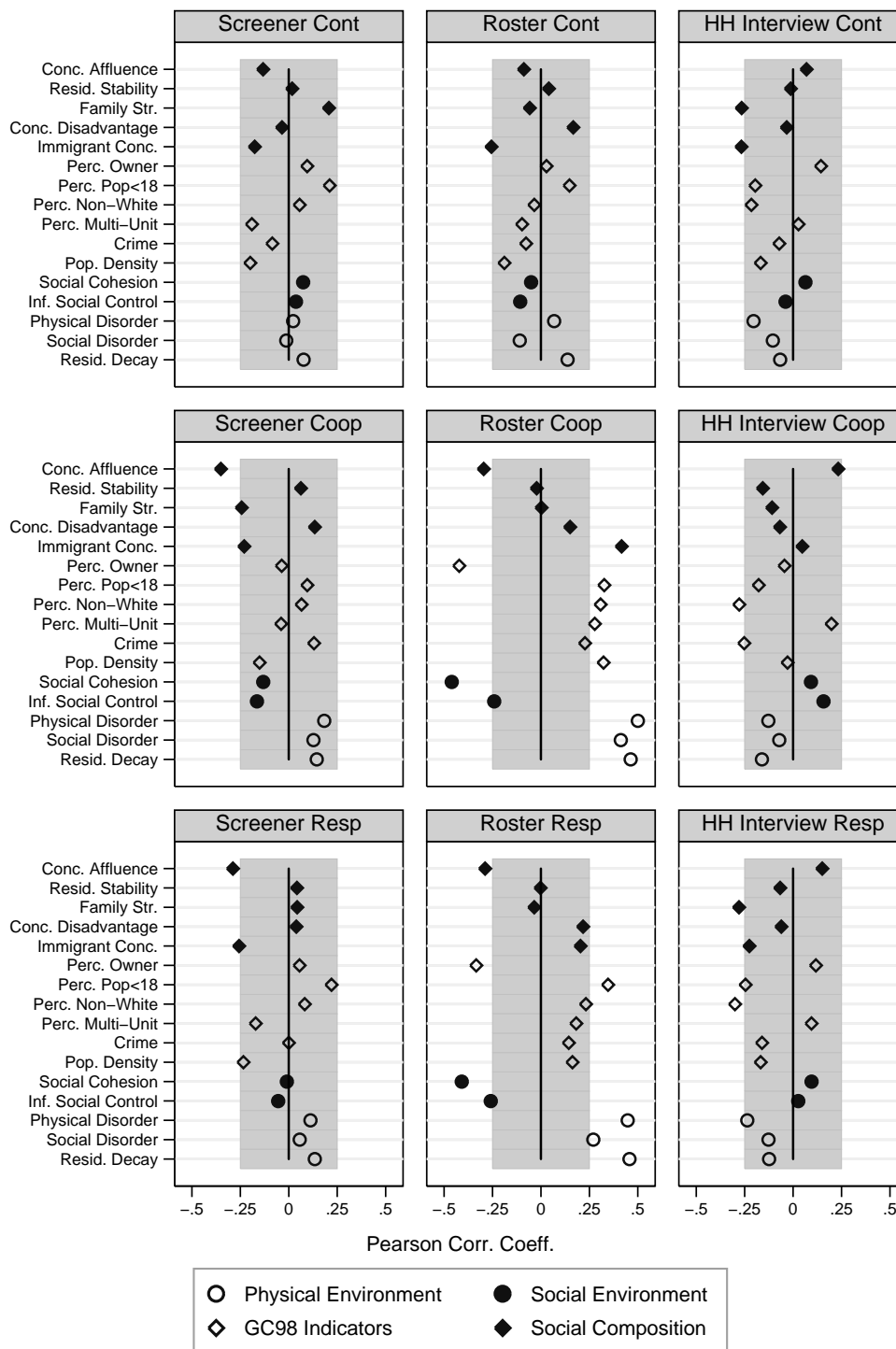
Fig. 6.1: Bivariate Pearson Correlations between Response Rates and Neighborhood Characteristics (n=65)

Tab. 6.14: Descriptive statistics for Different Sets of Household Weights used in the main analyses and the sensitivity analyses.[†]

Type	Description	Mean	Std. Dev.	Min	Max	Obs
One set of weights: L.A. FANS Official dataset						
A	*Unscaled Combined*					
	(Tract*Household) weight	1.0499	1.1416	0.0203	4.0221	3708
Two sets of weights: Approximation in Paper 1						
B	*Unscaled*					
	Tract weight	1.0000	0.0000	1.0000	1.0000	3708
	Household weight	1.0499	1.1416	0.0203	4.0221	3708
C	*Scaling method #1*					
	Tract weight	1.0000	0.0000	1.0000	1.0000	3708
	Household weight	0.5930	0.4168	0.0725	1.6576	3708
D	*Scaling method #2*					
	Tract weight	1.0000	0.0000	1.0000	1.0000	3708
	Household weight	0.7499	0.5296	0.1073	2.3510	3708
E	*Equally weighted*					
	Tract weight	1.0000	0.0000	1.0000	1.0000	3708
	Household weight	1.0000	0.0000	1.0000	1.0000	3708
Two sets of weights: L.A. FANS Special request						
F	*Unscaled*					
	Tract weight	0.6586	0.4742	0.0777	1.8329	3708
	Household weight	21.1331	9.8207	7.2489	31.2869	3708
G	*Unscaled Combined*					
	(Tract*Household) weight	1047.6580	1332.6200	19.2096	7858.2330	3708
H	*Scaling method #1*					
	Tract weight	21.1331	9.8207	7.2489	31.2869	3708
	Household weight	47.2400	46.1919	2.6500	259.3333	3708

(†): The main analysis discussed in Paper 1 used the "C" weights. These weights were also used in the sensitivity analyses labeled *Drop Outlier, Coll. Efficacy*, and *Groves-Couper*. The "B" weights were used in the analysis labeled *Unscaled*, the "D" weights for the analysis labeled *Alternative*, the "E" weights for the analysis labeled *Unweighted*, and the "H" weights for the analysis labeled *Two Weights*. Tables with results for the sensitivity analysis using alternative set of weights and covariates are displayed in the next set of tables.

Tab. 6.15: Sensitivity Analysis for Different Sets of Household Weights (Part 1/2). Multilevel Logistic Regression predicting Probability of Roster Cooperation. Displaying Neighborhood level predictors only.

Variables	Original		Alternative		Unscaled		Unweighted		Two Weights	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Fixed Effects										
Household Predictors	–	–	–	–	–	–	–	–	–	–
Poverty Strata										
Very Poor	-0.40	(0.35)	-0.35	(0.35)	-0.12	(0.46)	-0.21	(0.35)	-0.23	(0.42)
Poor	-0.50	(0.27)	-0.52*	(0.26)	-0.42	(0.32)	-0.51	(0.26)	-0.41	(0.27)
Non Poor	–	–	–	–	–	–	–	–	–	–
Traditional Neighborhood Predictors										
Pop. Density	0.00	(0.10)	0.02	(0.10)	-0.03	(0.13)	0.01	(0.10)	0.03	(0.13)
Crime	0.11	(0.08)	0.11	(0.08)	0.14	(0.10)	0.18*	(0.07)	0.17	(0.09)
Conc. Affluence	0.05	(0.21)	0.06	(0.21)	0.09	(0.25)	-0.03	(0.18)	0.07	(0.24)
Resid. Stability	0.02	(0.10)	0.03	(0.09)	0.08	(0.10)	0.03	(0.08)	0.08	(0.10)
Family Structure	-0.04	(0.09)	-0.06	(0.09)	-0.12	(0.11)	-0.03	(0.08)	-0.00	(0.10)
Conc. Disadvantage	-0.07	(0.14)	-0.10	(0.14)	-0.05	(0.17)	-0.03	(0.12)	-0.02	(0.17)
Immigrant Conc.	0.01	(0.18)	0.01	(0.18)	0.03	(0.21)	0.22	(0.17)	-0.07	(0.20)
Neighborhood Environmental Predictors										
Social Cohesion	-0.47*	(0.22)	-0.46*	(0.21)	-0.59*	(0.25)	-0.35*	(0.15)	-0.52*	(0.24)
Informal Social Control	0.47**	(0.14)	0.47**	(0.15)	0.56***	(0.16)	0.43***	(0.13)	0.59**	(0.18)
Physical Disorder	0.08	(0.18)	0.09	(0.18)	-0.01	(0.21)	-0.00	(0.17)	0.08	(0.18)
Social Disorder	0.10	(0.09)	0.10	(0.09)	0.03	(0.10)	0.02	(0.09)	0.02	(0.11)
Resid. Decay	0.24	(0.13)	0.24	(0.13)	0.29*	(0.14)	0.21	(0.11)	0.32*	(0.13)
Intercept	1.44***	(0.32)	1.46***	(0.31)	1.30***	(0.32)	1.15***	(0.31)	1.27***	(0.29)
Random Effects										
Tract Level Std.Dev.($\hat{\phi}$)	0.24**	(0.09)	0.30***	(0.08)	0.46***	(0.08)	0.27***	(0.07)	0.29**	(0.10)
Observations	3708		3708		3708		3708		3708	

Tab. 6.16: Sensitivity Analysis for Different Sets of Household Weights (Part 2/2). Multilevel Logistic Regression predicting Probability of Roster Cooperation. Displaying Household level predictors only.

Variables	Original $\hat{\beta}$	Original SE($\hat{\beta}$)	Alternative $\hat{\beta}$	Alternative SE($\hat{\beta}$)	Unscaled $\hat{\beta}$	Unscaled SE($\hat{\beta}$)	Unweighted $\hat{\beta}$	Unweighted SE($\hat{\beta}$)	Two Weights $\hat{\beta}$	Two Weights SE($\hat{\beta}$)
Fixed Effects										
Household Predictors										
Female	-0.09	(0.11)	-0.08	(0.11)	-0.02	(0.13)	-0.00	(0.10)	-0.04	(0.13)
Male	–	–	–	–	–	–	–	–	–	–
18-24 yrs	0.55*	(0.25)	0.53*	(0.25)	0.74	(0.38)	0.35	(0.19)	0.44	(0.31)
25-34 yrs	0.64***	(0.16)	0.64***	(0.17)	0.68**	(0.23)	0.45***	(0.12)	0.62**	(0.21)
35-54 yrs	–	–	–	–	–	–	–	–	–	–
55-69 yrs	0.08	(0.17)	0.08	(0.17)	0.08	(0.17)	0.02	(0.15)	0.11	(0.18)
70+ yrs	-0.25	(0.25)	-0.31	(0.25)	-0.36	(0.29)	-0.32	(0.22)	-0.40	(0.26)
Resp. Race White	–	–	–	–	–	–	–	–	–	–
Resp. Race Latino	0.43**	(0.15)	0.45**	(0.15)	0.55**	(0.17)	0.36*	(0.14)	0.44**	(0.15)
Resp. Race Black	0.21	(0.17)	0.22	(0.18)	0.28	(0.27)	0.29	(0.19)	0.25	(0.23)
Resp. Race Other	0.13	(0.22)	0.17	(0.22)	0.21	(0.24)	-0.00	(0.16)	0.06	(0.23)
Single Family Home	–	–	–	–	–	–	–	–	–	–
Apartment + Others	0.28	(0.18)	0.27	(0.18)	0.23	(0.24)	0.31*	(0.12)	0.33	(0.22)
Less than $500/month	-0.74*	(0.30)	-0.78**	(0.29)	-1.05**	(0.37)	-0.49	(0.31)	-0.81**	(0.31)
$500-$999/month	-0.54*	(0.27)	-0.58*	(0.26)	-0.42	(0.32)	-0.17	(0.27)	-0.42	(0.27)
$1000-$1999/month	-0.59*	(0.26)	-0.63*	(0.25)	-0.50	(0.26)	-0.14	(0.25)	-0.47	(0.24)
$2000-$2999/month	-0.31	(0.22)	-0.36	(0.22)	-0.37	(0.23)	-0.13	(0.26)	-0.15	(0.24)
$3000 or more/month	–	–	–	–	–	–	–	–	–	–
HH with Children	0.60***	(0.15)	0.64***	(0.14)	0.64**	(0.20)	0.54***	(0.11)	0.43*	(0.18)
HH without Children	–	–	–	–	–	–	–	–	–	–
Poverty Strata	–	–	–	–	–	–	–	–	–	–
Traditional Neighborhood Predictors	–	–	–	–	–	–	–	–	–	–
Neighborhood Environmental Predictors	–	–	–	–	–	–	–	–	–	–
Intercept	1.44***	(0.32)	1.46***	(0.31)	1.30***	(0.32)	1.15***	(0.31)	1.27***	(0.29)
Random Effects										
Tract Level Std.Dev.($\hat{\psi}$)	0.24**	(0.09)	0.30***	(0.08)	0.46***	(0.08)	0.27***	(0.07)	0.29**	(0.10)
Observations	3708		3708		3708		3708		3708	

Tab. 6.17: Sensitivity Analysis for Different Indicators of the Social Environment. Multilevel Logistic Regression predicting Probability of Roster Cooperation. Estimates obtained using the single set of household weights ($wgthh^{NTNS}$). Displaying Neighborhood level predictors only.

Variables	Original		Drop Outlier		Coll. Efficacy	
	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{\beta}$	$SE(\hat{\beta})$	$\hat{\beta}$	$SE(\hat{\beta})$
Fixed Effects						
Household Predictors	–	–	–	–	–	–
Poverty Strata						
Very Poor	-0.40	(0.35)	-0.52	(0.43)	-0.50	(0.44)
Poor	-0.50	(0.27)	-0.61	(0.34)	-0.60	(0.33)
Non Poor	–	–	–	–	–	–
Traditional Neighborhood Predictors						
Pop. Density	0.00	(0.10)	-0.01	(0.11)	-0.01	(0.10)
Crime	0.11	(0.08)	0.12	(0.08)	0.06	(0.08)
Conc. Affluence	0.05	(0.21)	0.04	(0.22)	-0.11	(0.23)
Resid. Stability	0.02	(0.10)	0.03	(0.09)	0.01	(0.10)
Family Structure	-0.04	(0.09)	-0.01	(0.09)	0.00	(0.09)
Conc. Disadvantage	-0.07	(0.14)	-0.07	(0.14)	-0.06	(0.15)
Immigrant Conc.	0.01	(0.18)	0.01	(0.19)	0.16	(0.22)
Neighborhood Environmental Predictors						
Collective Efficacy	–	–	–	–	0.06	(0.20)
Social Cohesion	-0.47*	(0.22)	-0.47*	(0.22)	–	–
Informal Social Control	0.47**	(0.14)	0.46**	(0.14)	–	–
Physical Disorder	0.08	(0.18)	0.11	(0.18)	0.00	(0.20)
Social Disorder	0.10	(0.09)	0.10	(0.09)	0.11	(0.10)
Resid. Decay	0.24	(0.13)	0.25	(0.13)	0.26	(0.14)
Intercept	1.44***	(0.32)	1.36***	(0.32)	1.56***	(0.32)
Random Effects						
Tract Level Std.Dev.($\hat{\psi}$)	0.24**	(0.09)	0.24**	(0.09)	0.33**	(0.11)
Observations	3708		3650		3708.00	

Tab. 6.18: Sensitivity Analysis for Different Indicators of Socio-Economic Composition. Multilevel Logistic Regression predicting Probability of Roster Cooperation. Estimates obtained using the single set of household weights ($wgthh^{NTNS}$). Displaying Neighborhood level predictors only.

Variables	Original		Groves-Couper	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Fixed Effects				
Household Predictors	–	–	–	–
Poverty Strata				
Very Poor	-0.40	(0.35)	-0.23	(0.30)
Poor	-0.50	(0.27)	-0.25	(0.22)
Non Poor	–	–	–	–
Traditional Neighborhood Predictors				
Pop. Density	0.00	(0.10)	-0.01	(0.11)
Crime	0.11	(0.08)	0.09	(0.08)
Conc. Affluence	0.05	(0.21)	–	–
Resid. Stability	0.02	(0.10)	–	–
Family Structure	-0.04	(0.09)	–	–
Conc. Disadvantage	-0.07	(0.14)	–	–
Immigrant Conc.	0.01	(0.18)	–	–
Perc. Owner	–	–	0.20	(0.19)
Perc. Non-White	–	–	-0.28	(0.16)
Perc. Multi-Unit	–	–	0.14	(0.19)
Perc. Pop < 18yrs	–	–	0.04	(0.18)
Neighborhood Environmental Predictors				
Social Cohesion	-0.47*	(0.22)	-0.52**	(0.19)
Informal Social Control	0.47**	(0.14)	0.47***	(0.13)
Physical Disorder	0.08	(0.18)	0.14	(0.15)
Social Disorder	0.10	(0.09)	0.11	(0.09)
Resid. Decay	0.24	(0.13)	0.21	(0.12)
Intercept	1.44***	(0.32)	1.08***	(0.30)
Random Effects				
Tract Level Std.Dev.($\hat{\psi}$)	0.24**	(0.09)	0.20*	(0.10)
Observations	3708		3708	
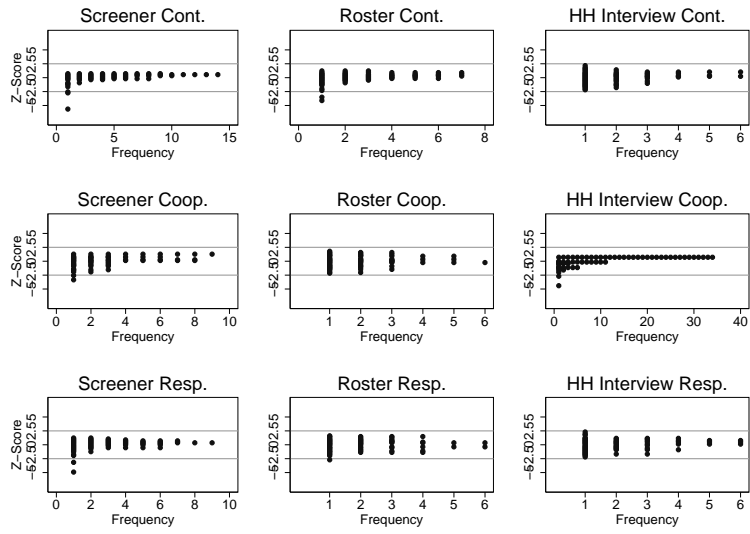
Fig. 6.2: Tract-level Outliers for Neighborhood Response Rates (n=65)
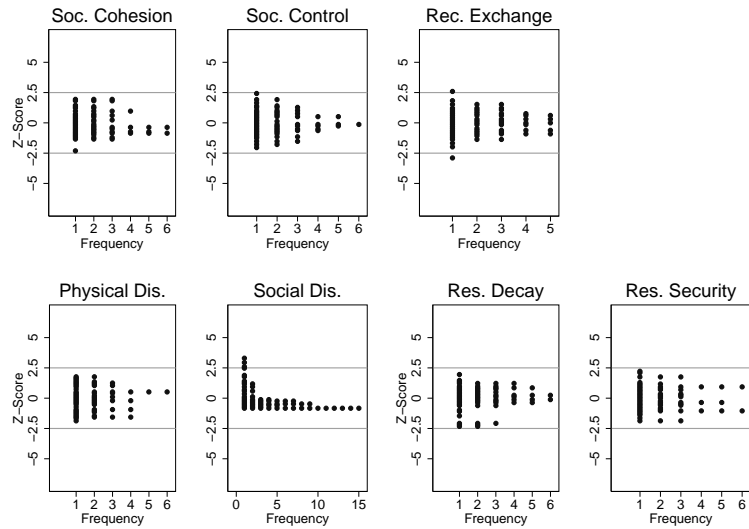
Fig. 6.3: Tract-level Outliers for Neighborhood Environmental Characteristics (n=65)

6.2   Appendix to Paper 2

Estimation of the Multilevel Model

The measurement error model for perceptions of disorder, introduced in expression (3.1) and reproduced below, incorporates the nesting of the geography levels and the cross-classification of 'interviewers and tracts':

$$y_{ijkr} = \beta + t_k + b_{jk} + i_r + e_{ijkr}$$

$$t_k \sim IID(0, \gamma_t^2)$$

$$b_{jk}|t_k \sim IID(0, \gamma_b^2)$$

$$i_r \sim IID(0, \gamma_i^2)$$

$$e_{ijkr} \sim IID(0, \gamma_e^2)$$

where y_{ijkr} is an indicator taking on a value of 1 if a neighborhood characteristic, e.g. *graffiti*, is observed in block face i of block j of tract k, rated by interviewer r, with $y_{ijkr} = 0$ if no graffiti was observed. The model includes three random effects to take into account the dependency among the observations within tracts (t_k), blocks (b_{jk}) and interviewers (i_r). The term e_{ijkr} reflects residual variability associated with the 'block faces' and the interaction of 'tracts and interviewers'. The random effects follow the distributional assumptions above, where IID stands for Independent and Identically Distributed.

This model was fitted using the *xtmixed* command in Stata 10. This command is primarily designed for multilevel models with nested random effects. To fit models with crossed random effects, I followed the approach by Goldstein (1987) described in Rabe-Hesketh and Skrondal (2008, 475):

- Consider the entire dataset as an artificial level-4 unit a within which both tracts and interviewers are nested (see Figure 6.4).

- Treat tracts or interviewers as the level-3 units and specify a random intercept for them. It is best to choose the factor with more levels, so I chose to model the random effects of tracts here (ut_{ka}).[1]

- For the other factor, here interviewers, specify a level-3 random intercept for each tract, ui_{pa}, (p=1,...,28). This can be constructed by treating u_{pa} as the random coefficient of the dummy variable d_{pijkr} for interviewer p, where

$$d_{pijkr} = \begin{cases} 1 & \text{if p=r} \\ 0 & \text{otherwise} \end{cases}$$

The 28 random coefficients associated with the interviewers are then specified as having equal variance γ_i^2 and being uncorrelated. Model (3.1) can then be written as:

$$y_{ijkra} = \beta + ut_{ka} + ub_{jka} + \sum_p u_{pa} d_{pijkr} + e_{ijkra}$$
$$= \beta + ut_{ka} + ub_{jka} + ui_{ra} + e_{ijkra}$$

where ut_{ka} is used to estimate the random effect of tracts (γ_t^2); ui_{ra} for the random effect of interviewers (γ_i^2); ub_{jka} for the random effect of blocks (γ_b^2); and e_{ijkra} for the residual variability (γ_e^2). The syntax to implement this procedure in Stata 10 is:

```
xtmixed y || _all: R.int_id || tr_id: || blk_id:, mle
```

---

[1] I added the letter u to the notation for the random effects to avoid confusion between the original and the new formulations described here.

where y corresponds to the binary neighborhood item *graffiti*, *int_id* is the identi-fication number for each interviewer ($n = 28$), *tr_id* is the identification number for each Census tract ($n = 65$), and *blk_id* is the identification number for each block ($n = 419$). The order of the variables in the syntax specifies the nesting structure. The *mle* option requests that the model is estimated using Maximum Likelihood estimation. For more on the modeling of crossed random effects in Stata see Rabe-Hesketh and Skrondal (2008, 473-508).
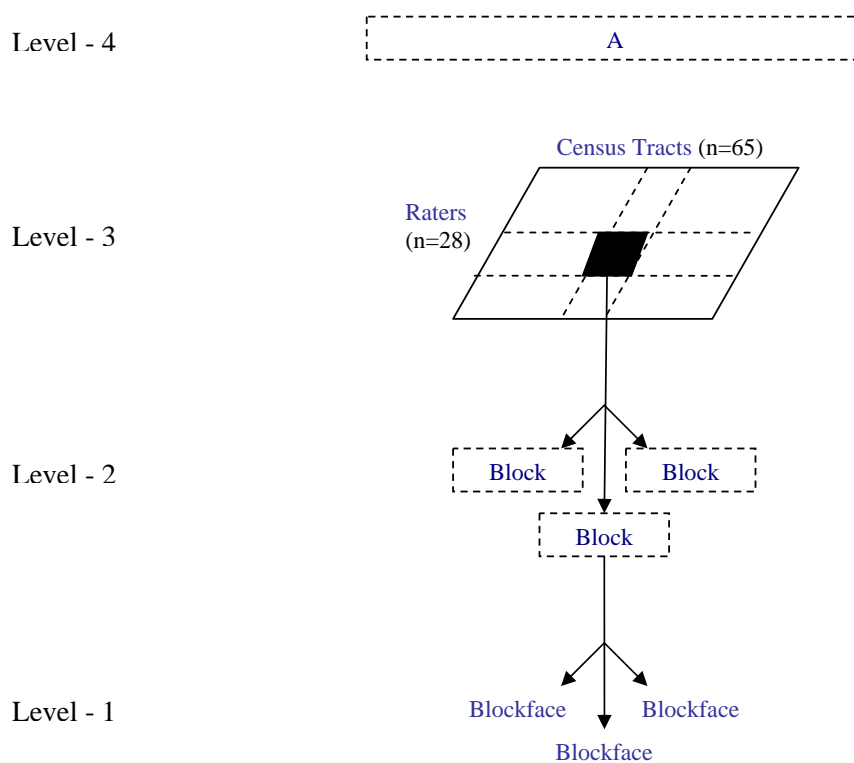
Fig. 6.4: Illustration of the model structure used to estimate the error-components model in expression (3.1).

Derivation of Intraclass Correlation Coefficients

The estimates of variance components derived from model (3.1) were used to derive three forms of the intra class correlation coefficient: *interobserver agreement* (δ_3), *interviewer clustering effects* (ρ_{int}), and *sampling point clustering effects* (ρ_{sp}).

$$\delta_3 = Corr(y_{ijkr}, y_{ijkr'}) = \frac{Cov(y_{ijkr}, y_{ijkr'})}{Var(y_{ijkr})}$$

$$\rho_{int} = Corr(y_{ijkr}, y_{i'j'k'r}) = \frac{Cov(y_{ijkr}, y_{i'j'k'r})}{Var(y_{ijkr})}$$

$$\rho_{sp} = Corr(y_{ijkr}, y_{i'j'kr'}) = \frac{Cov(y_{ijkr}, y_{i'j'kr'})}{Var(y_{ijkr})}$$

Given model (3.1), and assuming the usual assumption of uncorrelated error terms hold, the expression for the common variance ($Var(y_{ijkr})$) and the covariances are given by:

$$Var(y_{ijkr}) = Var(\beta + t_k + b_{jk} + i_r + e_{ijkr}) \tag{6.1}$$

$$= V(t_k) + V(b_{jk}) + V(i_r) + V(e_{ijkr})$$

$$= \gamma_t^2 + \gamma_b^2 + \gamma_i^2 + \gamma_e^2$$

$$Cov(y_{ijkr}, y_{ijkr'}) = Cov(\beta + t_k + b_{jk} + i_r + e_{ijkr}, \beta + t_k + b_{jk} + i_{r'} + e_{ijkr'}) \tag{6.2}$$

$$= C(t_k, t_k) + C(t_k, b_{jk}) + C(t_k, i_{r'}) + C(t_k, e_{ijkr'})$$

$$+ C(b_{jk}, t_k) + C(b_{jk}, b_{jk}) + C(b_{jk}, i_{r'}) + C(b_{jk}, e_{ijkr'})$$

$$+ C(i_r, t_k) + C(i_r, b_{jk}) + C(i_r, i_{r'}) + C(i_r, e_{ijkr'})$$

$$+ C(e_{ijkr}, t_k) + C(e_{ijkr}, b_{jk}) + C(e_{ijkr}, i_{r'}) + C(e_{ijkr}, e_{ijkr'})$$

$$= V(t_k) + V(b_{jk})$$

$$= \gamma_t^2 + \gamma_b^2$$

$$Cov(y_{ijkr}, y_{i'j'k'r}) = Cov(\beta + t_k + b_{jk} + i_r + e_{ijkr}, \beta + t_{k'} + b_{j'k'} + i_r + e_{i'j'k'r}) \ (6.3)$$

$$= C(t_k, t_{k'}) + C(t_k, b_{j'k'}) + C(t_k, i_r) + C(t_k, e_{i'j'k'r})$$

$$+C(b_{jk}, t_{k'}) + C(b_{jk}, b_{j'k'}) + C(b_{jk}, i_r) + C(b_{jk}, e_{i'j'k'r})$$

$$+C(i_r, t_{k'}) + C(i_r, b_{j'k'}) + C(i_r, i_r) + C(i_r, e_{i'j'k'r})$$

$$+C(e_{ijkr}, t_{k'}) + C(e_{ijkr}, b_{j'k'}) + C(e_{ijkr}, i_r) + C(e_{ijkr}, e_{i'j'k'r})$$

$$= V(i_r)$$

$$= \gamma_i^2$$

$$Cov(y_{ijkr}, y_{i'j'kr'}) = Cov(\beta + t_k + b_{jk} + i_r + e_{ijkr}, \beta + t_k + b_{j'k} + i_{r'} + e_{i'j'kr'}) \ (6.4)$$

$$= C(t_k, t_k) + C(t_k, b_{j'k}) + C(t_k, i_{r'}) + C(t_k, e_{i'j'kr'})$$

$$+C(b_{jk}, t_k) + C(b_{jk}, b_{j'k}) + C(b_{jk}, i_{r'}) + C(b_{jk}, e_{i'j'kr'})$$

$$+C(i_r, t_k) + C(i_r, b_{j'k}) + C(i_r, i_{r'}) + C(i_r, e_{i'j'kr'})$$

$$+C(e_{ijkr}, t_k) + C(e_{ijkr}, b_{j'k}) + C(e_{ijkr}, i_{r'}) + C(e_{i'j'kr'})$$

$$= V(t_k)$$

$$= \gamma_t^2$$

The final expressions for the estimates of the intraclass correlation coefficients, which correspond to expressions (3.2)-(3.4), are given below:

$$\delta_3 = Corr(y_{ijkr}, y_{ijkr'}) = \frac{Cov(y_{ijkr}, y_{ijkr'})}{Var(y_{ijkr})} = \frac{\gamma_t^2 + \gamma_b^2}{\gamma_t^2 + \gamma_b^2 + \gamma_i^2 + \gamma_e^2}$$

$$\rho_{int} = Corr(y_{ijkr}, y_{i'j'k'r}) = \frac{Cov(y_{ijkr}, y_{i'j'k'r})}{Var(y_{ijkr})} = \frac{\gamma_i^2}{\gamma_t^2 + \gamma_b^2 + \gamma_i^2 + \gamma_e^2}$$

$$\rho_{sp} = Corr(y_{ijkr}, y_{i'j'kr'}) = \frac{Cov(y_{ijkr}, y_{i'j'kr'})}{Var(y_{ijkr})} = \frac{\gamma_t^2}{\gamma_t^2 + \gamma_b^2 + \gamma_i^2 + \gamma_e^2}$$

Tab. 6.19: Unconditional Multilevel Model of Probability of Perceiving Physical Disorder.

	cars $\hat{\beta}$	SE($\hat{\beta}$)	trash $\hat{\beta}$	SE($\hat{\beta}$)	litter $\hat{\beta}$	SE($\hat{\beta}$)	drugs $\hat{\beta}$	SE($\hat{\beta}$)	bottles $\hat{\beta}$	SE($\hat{\beta}$)	cigars $\hat{\beta}$	SE($\hat{\beta}$)	graffiti $\hat{\beta}$	SE($\hat{\beta}$)	pograffiti $\hat{\beta}$	SE($\hat{\beta}$)
Intercept	0.11***	(0.02)	0.58***	(0.04)	0.74***	(0.04)	0.04***	(0.01)	0.21***	(0.03)	0.59***	(0.06)	0.57***	(0.05)	0.40***	(0.04)
Random Effects[†]																
Interviewers	-2.39***	(0.15)	-1.73***	(0.15)	-1.68***	(0.15)	-3.32***	(0.18)	-2.21***	(0.16)	-1.42***	(0.14)	-2.21***	(0.16)	-1.95***	(0.15)
Tracts	-2.60***	(0.13)	-1.56***	(0.10)	-1.59***	(0.10)	-3.39***	(0.17)	-1.75***	(0.11)	-1.45***	(0.10)	-1.08***	(0.09)	-1.32***	(0.10)
Blocks	-2.59***	(0.09)	-1.98***	(0.07)	-2.01***	(0.07)	-3.09***	(0.10)	-2.12***	(0.07)	-1.85***	(0.06)	-1.99***	(0.06)	-1.94***	(0.07)
Residual	-1.36***	(0.01)	-0.95***	(0.01)	-1.11***	(0.01)	-1.78***	(0.01)	-1.09***	(0.01)	-1.05***	(0.01)	-1.18***	(0.01)	-1.06***	(0.01)
Summary Statistics																
n	3998		3998		3998		3998		3998		3998		3998		3998	
LL	-443.45		-2133.90		-1554.51		1297.59		-1568.53		-1805.32		-1318.08		-1751.83	

(*) $= p < 0.05$; (**) $= p < 0.01$; (***) $= p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

Tab. 6.20: Unconditional Multilevel Model of Probability of Perceiving Social Disorder.

	gang		loitering		prostit		homeless		selling		drinking		intox	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Intercept	0.02***	(0.00)	0.10***	(0.02)	–	–	0.03***	(0.01)	0.01	(0.00)	0.02***	(0.01)	0.02**	(0.00)
Random Effects†														
Interviewers	-4.05***	(0.22)	-2.86***	(0.18)	–	–	-3.26***	(0.19)	-4.90***	(0.24)	-3.83***	(0.19)	-4.02***	(0.19)
Tracts	-4.40***	(0.33)	-2.49***	(0.12)	–	–	-3.52***	(0.14)	-4.23***	(0.14)	-3.50***	(0.15)	-4.07***	(0.21)
Blocks	-3.65***	(0.12)	-2.69***	(0.11)	–	–	-4.51***	(0.86)	-5.14***	(0.72)	-3.46***	(0.14)	-3.86***	(0.18)
Residual	-2.25***	(0.01)	-1.37***	(0.01)	–	–	-2.00***	(0.01)	-2.73***	(0.01)	-1.96***	(0.01)	-2.23***	(0.01)
Summary Statistics														
n	3962		3982		–		3996		3996		3996		3996	
LL	3169.84		-384.98		–		2251.92		5176.04		2025.29		3125.15	

(∗) $= p < 0.05$; (∗∗) $= p < 0.01$; (∗∗∗) $= p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

Tab. 6.21: Unconditional Multilevel Model of Probability of Perceiving Residential Decay.

	bldgs		boarded		vacant		walls		yards	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Intercept	0.90***	(0.03)	0.11***	(0.02)	0.18***	(0.02)	0.69***	(0.04)	0.83***	(0.04)
Random Effects†										
Interviewers	-2.51***	(0.15)	-3.17***	(0.25)	-3.32***	(0.28)	-2.10***	(0.16)	-2.20***	(0.16)
Tracts	-1.64***	(0.10)	-2.19***	(0.11)	-1.80***	(0.11)	-1.48***	(0.10)	-1.54***	(0.10)
Blocks	-2.27***	(0.06)	-2.46***	(0.09)	-2.04***	(0.07)	-2.10***	(0.08)	-2.02***	(0.06)
Residual	-1.47***	(0.01)	-1.33***	(0.01)	-1.18***	(0.01)	-1.06***	(0.01)	-1.21***	(0.01)
Summary Statistics										
n	3627		3627		3627		3627		3627	
LL	-129.23		-528.24		-1119.13		-1563.28		-1061.32	

(*) = $p < 0.05$; (**) = $p < 0.01$; (***) = $p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

Tab. 6.22: Unconditional Multilevel Model of Probability of Perceiving Residential Security.

	barswin		secsign		dogsign		gates		ngwatch	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Intercept	0.68***	(0.04)	0.46***	(0.04)	0.32***	(0.03)	0.68***	(0.05)	0.18***	(0.02)
Random Effects†										
Interviewers	-2.25***	(0.16)	-2.22***	(0.18)	-2.57***	(0.25)	-1.64***	(0.14)	-2.84***	(0.27)
Tracts	-1.17***	(0.09)	-1.48***	(0.10)	-1.87***	(0.12)	-1.67***	(0.11)	-2.42***	(0.15)
Blocks	-1.93***	(0.06)	-1.94***	(0.08)	-2.00***	(0.09)	-1.72***	(0.06)	-1.89***	(0.06)
Residual	-1.15***	(0.01)	-0.88***	(0.01)	-0.88***	(0.01)	-1.00***	(0.01)	-1.09***	(0.01)
Summary Statistics										
n	3627		3627		3627		3627		3618	
LL	-1309.04		-2182.48		-2146.44		-1838.22		-1420.49	

(*) $= p < 0.05$; (**) $= p < 0.01$; (***) $= p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

211

Tab. 6.23: Estimates of Intraclass Correlation Coefficients.

	Interobserver Agreement	Interviewer Clustering	Sampling Point Clustering	Ratio	
Items	$\hat{\delta}_3$	$\hat{\rho}_{int}$	$\hat{\rho}_{sp}$	$\frac{\hat{\rho}_{sp}}{\hat{\rho}_{int}}$	n
Physical Disorder					
cars	0.130	0.098	0.064	0.651	3998
trash	0.259	0.129	0.181	1.402	3998
litter	0.291	0.171	0.203	1.187	3998
drugs	0.098	0.040	0.034	0.858	3998
bottles	0.265	0.071	0.179	2.538	3998
cigars	0.306	0.224	0.212	0.944	3998
graffiti	0.557	0.050	0.479	9.629	3998
pograff	0.396	0.087	0.307	3.516	3998
gang	0.067	0.025	0.012	0.493	3962
loitering	0.143	0.041	0.086	2.078	3982
Social Disorder					
prostit					
homeless	0.048	0.071	0.042	0.594	3996
selling	0.054	0.012	0.047	3.810	3996
drinking	0.085	0.021	0.041	1.917	3996
intox	0.058	0.025	0.023	0.906	3996
Residential decay					
bldgs	0.448	0.061	0.349	5.710	3627
boarded	0.215	0.019	0.135	7.053	3627
vacant	0.317	0.009	0.196	20.948	3627
walls	0.331	0.074	0.257	3.497	3627
yards	0.388	0.074	0.281	3.796	3627
Residential Security					
barswin	0.510	0.049	0.418	8.595	3627
secsign	0.282	0.046	0.202	4.363	3627
dogsign	0.191	0.027	0.108	4.037	3627
gates	0.280	0.156	0.146	0.937	3627
ngwatch	0.210	0.023	0.054	2.323	3618

Tab. 6.24: Estimates of Interobserver Agreement derived from Different Samples.

	Any Day		Same Day		Same Day & Time	
Item	$\hat{\delta}_3$	n	$\hat{\delta}_3$	n	$\hat{\delta}_3$	n
Physical Disorder						
cars	0.130	3,998	0.184	908	0.219	586
trash	0.259	3,998	0.279	908	0.410	586
litter	0.291	3,998	0.316	908	0.255	586
drugs	0.098	3,998	0.200	908	0.256	586
bottles	0.265	3,998	0.216	908	0.216	586
cigars	0.306	3,998	0.277	908	0.258	586
graffiti	0.557	3,998	0.588	908	0.626	586
pograff	0.396	3,998	0.319	908	0.320	586
gang	0.067	3,962	0.079	901		
loitering	0.143	3,982	0.108	903	0.088	582
Social Disorder						
prostit						
homeless	0.048	3,996	0.038	908		
selling	0.054	3,996				
drinking	0.085	3,996				
intox	0.058	3,996	0.030	906		
Residential decay						
bldgs	0.448	3,627	0.350	824	0.352	540
boarded	0.215	3,627	0.318	824	0.307	540
vacant	0.317	3,627			0.318	540
walls	0.331	3,627	0.428	824	0.425	540
yards	0.388	3,627	0.358	824	0.381	540
Residential Security						
barswin	0.510	3,627	0.408	824	0.407	540
secsign	0.282	3,627	0.260	824	0.336	540
dogsign	0.191	3,627				
gates	0.280	3,627	0.190	824	0.167	540
ngwatch	0.210	3,618	0.155	823	0.102	539

Fig. 6.5: Estimates of Interobserver Agreement ($\hat{\delta}_3$) for Different Subsamples.

Tab. 6.25: Probability of Perceiving Disorder, by Characteristics of the Interviewers (Part 1/4). Unweighted estimates.

Covariates		cars	trash	litter	drugs	bottles	cigars	graffiti	pograff	gang	loitering	prostit	homeless	selling	drinking	intox
		Physical Disorder Items								Social Disorder Items						
Race White	No	13%	48%	73%	3%	19%	57%	55%	39%	2%	12%	0%	3%	1%	3%	1%
	Yes	4%	55%	72%	4%	23%	65%	51%	34%	0%	6%	0%	2%	0%	2%	1%
	χ^2	116.00	17.06	0.37	0.32	7.24	23.40	4.50	9.80	30.48	31.69	0.88	3.94	2.36	11.82	0.12
	pvalue (adj)	0.0000	0.0009	1.0000	1.0000	0.1785	0.0000	0.8494	0.0436	0.0000	0.0000	1.0000	1.0000	1.0000	0.0146	1.0000
Age 55+ yrs	No	9%	52%	76%	3%	21%	60%	57%	38%	2%	11%	0%	2%	1%	3%	1%
	Yes	6%	52%	58%	7%	24%	72%	37%	29%	0%	1%	0%	1%	0%	1%	0%
	χ^2	3.57	0.00	77.77	24.86	3.08	31.45	76.33	15.08	4.95	54.66	0.62	2.95	1.39	10.93	4.66
	pvalue (adj)	1.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0026	0.6526	0.0000	1.0000	1.0000	1.0000	0.0236	0.7711

Note: Table displays Pearson's χ^2 for the hypothesis that the rows and columns in each two-way table are independent.

Pvalue(adj) displays Bonferroni-adjusted significance level for k=25 simultaneous tests.

Tab. 6.26: Probability of Perceiving Disorder, by Characteristics of the Interviewers (Part 2/4). Unweighted estimates.

Covariates		Residential Decay Items					Residential Security Items				
		bldgs	boarded	vacant	walls	yards	barswin	secsign	dogsign	gates	ngwatch
Race White	No	84%	13%	14%	66%	79%	70%	53%	36%	55%	18%
	Yes	83%	7%	16%	63%	72%	58%	52%	28%	63%	18%
	χ^2	0.25	26.96	2.69	4.50	20.48	48.43	0.20	23.73	22.10	0.12
	pvalue (adj)	1.0000	0.0000	1.0000	0.8489	0.0002	0.0000	1.0000	0.0000	0.0001	1.0000
Age 55+ yrs	No	83%	11%	15%	69%	78%	64%	52%	34%	56%	19%
	Yes	82%	5%	20%	43%	61%	62%	51%	20%	75%	15%
	χ^2	0.66	16.25	8.32	131.53	65.90	0.61	0.18	37.54	61.74	3.66
	pvalue (adj)	1.0000	0.0014	0.0979	0.0000	0.0000	1.0000	1.0000	0.0000	0.0000	1.0000

Note: Table displays Pearson's χ^2 for the hypothesis that the rows and columns in each two-way table are independent.

Pvalue(adj) displays Bonferroni-adjusted significance level for k=25 simultaneous tests.

Tab. 6.27: Probability of Perceiving Disorder, by Characteristics of the Interviewers (Part 3/4). Unweighted estimates.

Covariates		Physical Disorder Items								Social Disorder Items						
		cars	trash	litter	drugs	bottles	cigars	graffiti	pograff	gang	loitering	prostit	homeless	selling	drinking	intox
Comm. Involvement	No	9%	53%	67%	3%	17%	60%	49%	30%	1%	9%	0%	1%	0%	1%	1%
	Yes	7%	51%	79%	4%	25%	62%	58%	43%	1%	9%	0%	3%	1%	4%	2%
	χ^2	6.98	1.59	55.56	4.93	38.47	2.01	26.93	64.68	0.07	0.27	0.41	16.26	5.37	20.88	8.17
	pvalue (adj)	0.2062	1.0000	0.0000	0.6599	0.0000	1.0000	0.0000	0.0000	1.0000	1.0000	1.0000	0.0014	0.5110	0.0001	0.1067
Has Kids	No	4%	47%	70%	2%	12%	51%	51%	31%	1%	10%	0%	1%	0%	1%	1%
	Yes	13%	58%	75%	6%	31%	73%	55%	42%	2%	9%	0%	3%	1%	4%	2%
	χ^2	81.45	39.87	10.41	44.38	183.93	182.61	5.66	48.98	12.34	1.34	0.35	7.92	2.93	21.89	1.45
	pvalue (adj)	0.0000	0.0000	0.0313	0.0000	0.0000	0.0000	0.4336	0.0000	0.0111	1.0000	1.0000	0.1224	1.0000	0.0001	1.0000
Ever Married	No	6%	46%	68%	2%	11%	50%	51%	31%	1%	10%	0%	2%	0%	1%	1%
	Yes	10%	55%	75%	5%	26%	67%	54%	39%	1%	8%	0%	2%	1%	3%	2%
	χ^2	14.12	30.70	21.66	25.87	116.50	101.84	2.38	22.06	0.24	3.70	0.04	0.93	7.27	15.86	5.23
	pvalue (adj)	0.0043	0.0000	0.0001	0.0000	0.0000	0.0000	1.0000	0.0001	1.0000	1.0000	1.0000	1.0000	0.1753	0.0017	0.5552

Note: Table displays Pearson's χ^2 for the hypothesis that the rows and columns in each two-way table are independent.

Pvalue(adj) displays Bonferroni-adjusted significance level for k=25 simultaneous tests.

217

Tab. 6.28: Probability of Perceiving Disorder, by Characteristics of the Interviewers (Part 4/4). Unweighted estimates.

Covariates		bldgs	boarded	vacant	walls	yards	barswin	secsign	dogsign	gates	ngwatch
		Residential Decay Items					Residential Security Items				
Comm. Involvement	No	79%	9%	15%	60%	71%	65%	54%	30%	54%	17%
	Yes	88%	11%	16%	71%	81%	63%	51%	34%	65%	19%
	χ^2	47.26	3.34	1.77	41.59	40.32	1.58	3.80	7.64	36.84	3.11
	pvalue (adj)	0.0000	1.0000	1.0000	0.0000	0.0000	1.0000	1.0000	0.1430	0.0000	1.0000
Has Kids	No	79%	8%	14%	59%	72%	65%	52%	30%	62%	17%
	Yes	88%	13%	17%	71%	79%	63%	54%	34%	55%	19%
	χ^2	48.80	25.52	3.66	50.61	17.92	1.65	1.60	7.53	16.74	1.88
	pvalue (adj)	0.0000	0.0000	1.0000	0.0000	0.0006	1.0000	1.0000	0.1516	0.0011	1.0000
Ever Married	No	79%	8%	12%	58%	74%	70%	51%	30%	59%	18%
	Yes	86%	11%	17%	69%	76%	60%	54%	33%	59%	18%
	χ^2	30.70	7.17	18.05	37.13	2.80	27.70	2.54	2.07	0.00	0.06
	pvalue (adj)	0.0000	0.1852	0.0005	0.0000	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000

Note: Table displays Pearson's χ^2 for the hypothesis that the rows and columns in each two-way table are independent.

Pvalue(adj) displays Bonferroni-adjusted significance level for k=25 simultaneous tests.

Tab. 6.29: Probability of Perceiving Disorder, by Characteristics of the Occasion of Measurement (1). Unweighted estimates.

Covariates		cars	trash	litter	drugs	bottles	cigars	graffiti	pograff	gang	loitering	prostit	homeless	selling	drinking	intox
				Physical Disorder Items								Social Disorder Items				
Experience with Block	No	10%	54%	66%	3%	18%	53%	43%	28%	1%	5%	0%	2%	0%	2%	1%
	Yes	10%	52%	76%	4%	22%	62%	57%	39%	1%	10%	0%	2%	0%	2%	1%
	χ^2	0.09	0.56	39.16	1.31	8.63	25.02	57.55	35.80	0.03	18.68	3.97	0.20	0.45	0.75	0.95
	pvalue (adj)	1.000	1.000	0.000	1.000	0.083	0.000	0.000	0.000	1.000	0.000	1.000	1.000	1.000	1.000	1.000
Neighborhood Close	No	9%	53%	74%	4%	22%	62%	55%	39%	1%	10%	0%	2%	1%	3%	2%
	Yes	4%	44%	64%	2%	12%	52%	37%	18%	0%	5%	0%	2%	0%	0%	0%
	χ^2	13.37	15.24	19.62	4.53	26.94	17.90	53.68	80.61	3.47	12.79	0.07	0.35	2.69	11.75	5.23
	pvalue (adj)	0.006	0.002	0.000	0.831	0.000	0.001	0.000	0.000	1.000	0.009	1.000	1.000	1.000	0.015	0.555
Rated after 5pm	No	10%	52%	73%	4%	21%	59%	52%	35%	1%	8%	0%	2%	0%	2%	1%
	Yes	5%	58%	71%	2%	17%	70%	58%	36%	0%	4%	0%	1%	0%	0%	0%
	χ^2	5.22	3.64	0.66	1.80	2.59	14.89	3.30	0.01	1.14	6.56	0.05	2.44	0.06	6.20	1.75
	pvalue (adj)	0.557	1.000	1.000	1.000	1.000	0.003	1.000	1.000	1.000	0.261	1.000	1.000	1.000	0.319	1.000
Rated on Weekend	No	9%	51%	71%	3%	18%	57%	50%	34%	1%	7%	0%	2%	0%	2%	1%
	Yes	13%	57%	81%	5%	29%	67%	64%	43%	1%	12%	0%	3%	1%	3%	1%
	χ^2	13.03	13.03	46.76	8.84	54.69	31.88	63.87	30.39	0.29	26.95	0.83	2.45	4.69	0.83	0.44
	pvalue (adj)	0.008	0.008	0.000	0.074	0.000	0.000	0.000	0.000	1.000	0.000	1.000	1.000	0.757	1.000	1.000

Note: Table displays Pearson's χ^2 for the hypothesis that the rows and columns in each two-way table are independent.

Pvalue(adj) displays Bonferroni-adjusted significance level for k=25 simultaneous tests.

Tab. 6.30: Probability of Perceiving Disorder, by Characteristics of the Occasion of Measurement (2). Unweighted estimates.

Covariates		Residential Decay Items					Residential Security Items				
		bldgs	boarded	vacant	walls	yards	barswin	secsign	dogsign	gates	ngwatch
Experience with Block	No	78%	9%	17%	61%	72%	53%	58%	30%	59%	20%
	Yes	86%	10%	16%	68%	79%	67%	51%	33%	59%	17%
	χ^2	31.67	2.12	1.05	16.70	14.59	58.92	12.82	2.75	0.00	5.47
	pvalue (adj)	0.000	1.000	1.000	0.001	0.003	0.000	0.009	1.000	1.000	0.484
Neighborhood Close	No	84%	11%	16%	66%	75%	64%	51%	32%	59%	18%
	Yes	82%	5%	14%	53%	76%	59%	67%	28%	59%	20%
	χ^2	0.44	12.60	0.85	32.73	0.16	4.83	40.55	3.13	0.01	0.81
	pvalue (adj)	1.000	0.010	1.000	0.000	1.000	0.701	0.000	1.000	1.000	1.000
Rated after 5pm	No	83%	9%	17%	65%	76%	63%	53%	33%	61%	18%
	Yes	89%	8%	15%	65%	71%	64%	48%	22%	63%	18%
	χ^2	5.56	0.32	0.68	0.00	2.79	0.27	2.50	12.16	0.75	0.00
	pvalue (adj)	0.459	1.000	1.000	1.000	1.000	1.000	1.000	0.012	1.000	1.000
Rated on Weekend	No	84%	10%	16%	64%	76%	61%	53%	31%	57%	19%
	Yes	85%	9%	16%	74%	79%	71%	49%	36%	64%	16%
	χ^2	0.90	1.05	0.04	30.04	3.68	33.54	6.19	5.97	13.76	3.92
	pvalue (adj)	1.000	1.000	1.000	0.000	1.000	0.000	0.321	0.364	0.005	1.000

Note: Table displays Pearson's χ^2 for the hypothesis that the rows and columns in each two-way table are independent.

Pvalue(adj) displays Bonferroni-adjusted significance level for k=25 simultaneous tests.

Tab. 6.31: County level Estimates of Disorder on Two Different Occasions. Unweighted estimates.

Neighborhood Items	Full Sample		Occasion #1		Occasion #2		[$Perc_1$-$Perc_2$=0]	
	Perc	n	$Perc_1$	n_1	$Perc_2$	n_2	t-test	p-value
Physical Disorder Items								
cars	9.8%	3,998	9.5%	1,999	10.1%	1,999	0.48	0.4890
trash	52.7%	3,998	54.5%	1,999	50.8%	1,999	5.50	0.0191
litter	73.6%	3,998	73.7%	1,999	73.6%	1,999	0.01	0.9428
drugs	3.4%	3,998	3.3%	1,999	3.5%	1,999	0.07	0.7929
bottles	21.0%	3,998	22.7%	1,999	19.3%	1,999	6.78	0.0093
cigars	59.7%	3,998	61.2%	1,999	58.2%	1,999	3.87	0.0492
graffiti	53.5%	3,998	54.3%	1,999	52.6%	1,999	1.09	0.2955
pograff	36.3%	3,998	37.0%	1,999	35.6%	1,999	0.79	0.3746
Social Disorder Items								
gang	1.2%	3,962	1.3%	1,981	1.2%	1,981	0.02	0.8913
loitering	8.6%	3,982	8.0%	1,991	9.2%	1,991	1.86	0.1728
prostit	0.3%	3,986	0.4%	1,993	0.3%	1,993	0.70	0.4038
homeless	2.0%	3,996	2.2%	1,998	1.9%	1,998	0.47	0.4953
selling	0.5%	3,996	0.6%	1,998	0.4%	1,998	0.89	0.3448
drinking	2.3%	3,996	3.2%	1,998	1.5%	1,998	12.90	0.0003
intox	1.3%	3,996	1.3%	1,998	1.3%	1,998	0.02	0.8908
Residential Decay Items								
bldgs	84.6%	3,627	84.9%	1,814	84.2%	1,814	0.27	0.6008
boarded	10.0%	3,627	10.8%	1,814	9.3%	1,814	2.24	0.1343
vacant	16.0%	3,627	15.6%	1,814	16.5%	1,814	0.49	0.4849
walls	66.7%	3,627	69.6%	1,814	63.7%	1,814	14.25	0.0002
yards	77.1%	3,627	78.5%	1,814	75.6%	1,814	4.23	0.0398
Residential Security Items								
barswin	63.8%	3,627	59.5%	1,814	68.2%	1,814	29.94	0.0000
secsign	52.2%	3,627	53.0%	1,814	51.4%	1,814	0.95	0.3302
dogsign	32.4%	3,627	32.1%	1,814	32.7%	1,814	0.15	0.6986
gates	59.2%	3,627	55.0%	1,814	63.4%	1,814	27.02	0.0000
ngwatch	17.7%	3,618	16.7%	1,809	18.8%	1,809	2.60	0.1068

Note: Table displays t-test for the hypothesis that the difference between the two ratings is zero.

Tab. 6.32: Multilevel Model of Probability of Perceiving Trash.

	Model (1)		Model (2)		Model (3)		Model (4)	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Fixed Effects								
Race White			0.10	(0.12)	0.13	(0.13)	0.13	(0.12)
Age 55+ yrs			-0.13	(0.16)	-0.16	(0.17)	-0.14	(0.16)
Has Kids			0.15	(0.11)	0.18	(0.11)	0.15	(0.11)
Comm. Involvement			-0.10	(0.08)	-0.08	(0.08)	-0.08	(0.08)
Ever Married			-0.13	(0.13)	-0.19	(0.13)	-0.16	(0.13)
Prior Experience					0.00	(0.03)	-0.01	(0.03)
Neighborhood close					0.03	(0.04)	0.04	(0.04)
Rated after 5pm					0.00	(0.03)	0.01	(0.03)
Rated on Weekend					0.02	(0.02)	0.00	(0.02)
Immigrant							0.07***	(0.02)
Disadvantage							0.10***	(0.02)
Affluence							-0.13***	(0.02)
Pop. Density							-0.01	(0.02)
Intercept	0.58***	(0.04)	0.60***	(0.07)	0.61***	(0.08)	0.58***	(0.07)
Random Effects[†]								
Interviewers	-1.73***	(0.15)	-1.71***	(0.16)	-1.68***	(0.16)	-1.72***	(0.16)
Tracts	-1.56***	(0.10)	-1.56***	(0.10)	-1.60***	(0.11)	-2.44***	(0.19)
Blocks	-1.98***	(0.07)	-2.03***	(0.08)	-1.96***	(0.09)	-1.96***	(0.09)
Residual	-0.95***	(0.01)	-0.96***	(0.01)	-0.98***	(0.01)	-0.98***	(0.01)
Summary Statistics								
n	3998		3621		2892		2892	
LL	-2133.90		-1900.80		-1495.15		-1460.05	
df	5		10		14		18	
AIC	4277.8		3821.6		3018.3		2956.1	
BIC	4309.3		3883.5		3101.9		3063.6	

$(*) = p < 0.05; (**) = p < 0.01; (***) = p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

Tab. 6.33: Multilevel Model of Probability of Perceiving Graffiti.

	Model (1)		Model (2)		Model (3)		Model (4)	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Fixed Effects								
Race White			0.10	(0.07)	0.13	(0.08)	0.13	(0.07)
Age 55+ yrs			-0.13	(0.10)	-0.17	(0.10)	-0.16	(0.09)
Has Kids			0.03	(0.07)	0.06	(0.07)	0.05	(0.07)
Comm. Involvement			0.03	(0.05)	0.03	(0.05)	0.04	(0.05)
Ever Married			-0.04	(0.08)	-0.05	(0.08)	-0.04	(0.08)
Prior Experience					-0.02	(0.02)	-0.03	(0.02)
Neighborhood close					-0.03	(0.03)	-0.00	(0.03)
Rated after 5pm					0.01	(0.03)	0.03	(0.03)
Rated on Weekend					0.01	(0.02)	-0.01	(0.02)
Immigrant							0.14***	(0.02)
Disadvantage							0.14***	(0.02)
Affluence							-0.23***	(0.02)
Pop. Density							0.02	(0.02)
Intercept	0.57***	(0.05)	0.54***	(0.06)	0.54***	(0.06)	0.49***	(0.05)
Random Effects[†]								
Interviewers	-2.21***	(0.16)	-2.26***	(0.18)	-2.30***	(0.21)	-2.32***	(0.21)
Tracts	-1.08***	(0.09)	-1.09***	(0.09)	-1.10***	(0.10)	-2.51***	(0.18)
Blocks	-1.99***	(0.06)	-1.98***	(0.07)	-1.93***	(0.07)	-1.93***	(0.07)
Residual	-1.18***	(0.01)	-1.16***	(0.01)	-1.17***	(0.01)	-1.17***	(0.01)
Summary Statistics								
n	3998		3621		2892		2892	
LL	-1318.08		-1244.08		-1005.02		-937.42	
df	5		10		14		18	
AIC	2646.2		2508.2		2038.0		1910.8	
BIC	2677.6		2570.1		2121.6		2018.3	

$(*) = p < 0.05; (**) = p < 0.01; (***) = p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

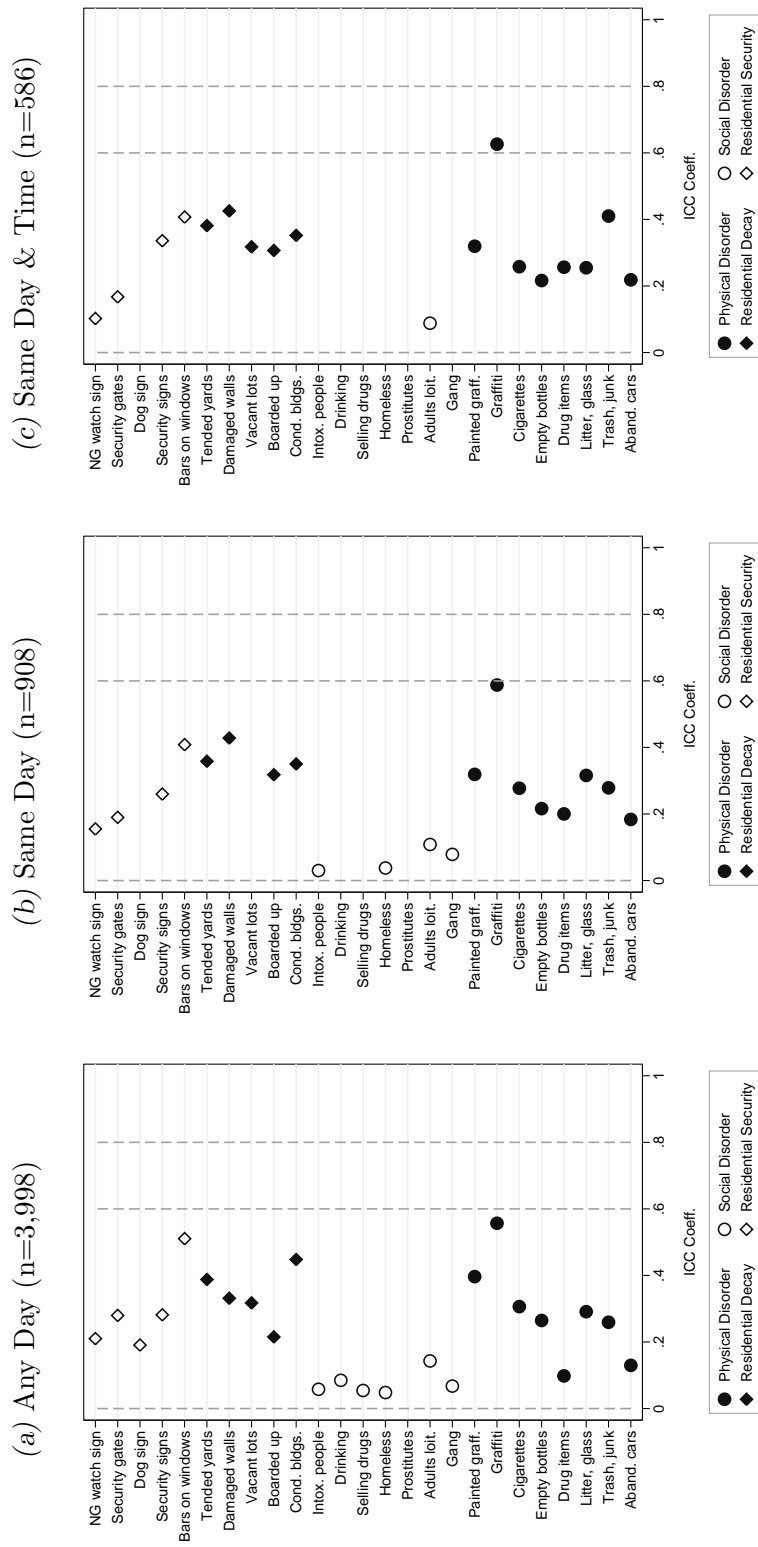*Tab. 6.34:* Multilevel Model of Probability of Perceiving Physical Disorder Items.

	trash		litter		cigars		graffiti	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Fixed Effects								
Race White	0.13	(0.12)	0.00	(0.10)	0.01	(0.15)	0.13	(0.07)
Age 55+ yrs	-0.14	(0.16)	-0.08	(0.13)	0.09	(0.20)	-0.16	(0.09)
Has Kids	0.15	(0.11)	-0.03	(0.09)	0.04	(0.13)	0.05	(0.07)
Comm. Involvement	-0.08	(0.08)	0.09	(0.06)	0.05	(0.10)	0.04	(0.05)
Ever Married	-0.16	(0.13)	0.12	(0.10)	0.10	(0.16)	-0.04	(0.08)
Prior Experience	-0.01	(0.03)	0.03	(0.02)	-0.01	(0.02)	-0.03	(0.02)
Neighborhood close	0.04	(0.04)	-0.00	(0.03)	-0.04	(0.03)	-0.00	(0.03)
Rated after 5pm	0.01	(0.03)	0.01	(0.03)	0.05	(0.03)	0.03	(0.03)
Rated on Weekend	0.00	(0.02)	0.03	(0.02)	0.01	(0.02)	-0.01	(0.02)
Immigrant	0.07***	(0.02)	0.09***	(0.02)	0.09***	(0.02)	0.14***	(0.02)
Disadvantage	0.10***	(0.02)	0.06***	(0.01)	0.11***	(0.02)	0.14***	(0.02)
Affluence	-0.13***	(0.02)	-0.16***	(0.02)	-0.13***	(0.02)	-0.23***	(0.02)
Pop. Density	-0.01	(0.02)	0.00	(0.02)	0.02	(0.02)	0.02	(0.02)
Intercept	0.58***	(0.07)	0.60***	(0.06)	0.47***	(0.08)	0.49***	(0.05)
Random Effects[†]								
Interviewers	-1.72***	(0.16)	-1.97***	(0.16)	-1.49***	(0.16)	-2.32***	(0.21)
Tracts	-2.44***	(0.19)	-2.54***	(0.20)	-2.61***	(0.24)	-2.51***	(0.18)
Blocks	-1.96***	(0.09)	-2.09***	(0.09)	-1.90***	(0.08)	-1.93***	(0.07)
Residual	-0.98***	(0.01)	-1.09***	(0.01)	-1.04***	(0.01)	-1.17***	(0.01)
Summary Statistics								
n	2892		2892		2892		2892	
LL	-1460.05		-1140.88		-1299.52		-937.42	

$(*) = p < 0.05; (**) = p < 0.01; (***) = p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

Tab. 6.35: Multilevel Model of Probability of Perceiving Residential Decay Items.

	blgds		walls		yards	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Fixed Effects						
Race White	0.06	(0.06)	0.04	(0.05)	0.03	(0.07)
Age 55+ yrs	-0.01	(0.08)	-0.25***	(0.06)	-0.12	(0.08)
Has Kids	0.09	(0.05)	0.11*	(0.04)	0.09	(0.06)
Comm. Involvement	0.05	(0.04)	0.04	(0.03)	0.01	(0.04)
Ever Married	-0.05	(0.06)	0.05	(0.05)	-0.06	(0.07)
Prior Experience	0.02	(0.02)	-0.02	(0.02)	0.05*	(0.02)
Neighborhood close	0.01	(0.03)	-0.01	(0.03)	0.12***	(0.03)
Rated after 5pm	0.07**	(0.02)	0.02	(0.03)	-0.03	(0.03)
Rated on Weekend	-0.04**	(0.02)	0.03	(0.02)	0.01	(0.02)
Immigrant	0.03	(0.02)	0.07**	(0.02)	0.10***	(0.02)
Disadvantage	0.07***	(0.02)	0.09***	(0.02)	0.09***	(0.02)
Affluence	-0.15***	(0.02)	-0.18***	(0.02)	-0.16***	(0.02)
Pop. Density	-0.01	(0.02)	-0.00	(0.02)	-0.02	(0.02)
Intercept	0.82***	(0.04)	0.56***	(0.04)	0.72***	(0.04)
Random Effects[†]						
Interviewers	-2.53***	(0.17)	-3.02***	(0.32)	-2.44***	(0.18)
Tracts	-2.09***	(0.12)	-2.10***	(0.14)	-2.22***	(0.15)
Blocks	-2.16***	(0.07)	-2.12***	(0.10)	-1.92***	(0.07)
Residual	-1.44***	(0.01)	-1.05***	(0.01)	-1.19***	(0.01)
Summary Statistics						
n	2631		2631		2631	
LL	-192.00		-1114.95		-816.78	

$(*) = p < 0.05; (**) = p < 0.01; (***) = p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

225

Tab. 6.36: Multilevel Model of Probability of Perceiving Residential Security Items.

	secsign		gates		ngwatch	
	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)	$\hat{\beta}$	SE($\hat{\beta}$)
Fixed Effects						
Race White	-0.17*	(0.08)	0.05	(0.12)	-0.02	(0.06)
Age 55+ yrs	-0.05	(0.10)	0.23	(0.15)	-0.07	(0.08)
Has Kids	-0.01	(0.07)	-0.14	(0.10)	0.06	(0.06)
Comm. Involvement	-0.00	(0.05)	0.16*	(0.08)	0.02	(0.04)
Ever Married	0.16	(0.09)	-0.01	(0.12)	-0.01	(0.07)
Prior Experience	0.01	(0.03)	-0.03	(0.03)	-0.03	(0.02)
Neighborhood close	0.05	(0.04)	-0.02	(0.04)	0.01	(0.03)
Rated after 5pm	-0.03	(0.03)	-0.00	(0.03)	-0.01	(0.03)
Rated on Weekend	0.01	(0.03)	0.01	(0.03)	-0.01	(0.02)
Immigrant	-0.07**	(0.02)	0.08**	(0.03)	-0.04*	(0.02)
Disadvantage	-0.07***	(0.02)	0.07**	(0.02)	-0.03*	(0.02)
Affluence	0.18***	(0.02)	-0.11***	(0.02)	0.01	(0.02)
Pop. Density	-0.00	(0.02)	0.01	(0.03)	-0.00	(0.02)
Intercept	0.47***	(0.05)	0.60***	(0.07)	0.21***	(0.04)
Random Effects[†]						
Interviewers	-2.25***	(0.22)	-1.78***	(0.17)	-2.52***	(0.27)
Tracts	-2.23***	(0.16)	-2.05***	(0.15)	-2.51***	(0.19)
Blocks	-1.93***	(0.09)	-1.68***	(0.07)	-1.98***	(0.08)
Block Faces (residual)	-0.90***	(0.01)	-1.04***	(0.02)	-1.06***	(0.01)
Summary Statistics						
n	2631		2631		2623	
LL	-1516.90		-1245.89		-1110.94	

$(*) = p < 0.05; (**) = p < 0.01; (***) = p < 0.001$

(†): Estimates of the random effects are provided in the logarithmic scale.

6.3   Appendix to Paper 3

Tab. 6.37: Estimates of Z-P Pearson Correlations

Variables	Roster Response			
	Corr	Min95	Max95	n
Neighborhood Observations				
Physical Disorder	0.095	0.064	0.125	4,010
Resid. Decay	0.116	0.086	0.147	4,010
Security Gates	0.064	0.033	0.095	4,010
Bars on Windows	0.056	0.025	0.087	4,010
Damaged Yards	0.107	0.076	0.137	4,010
Damaged Walls	0.096	0.066	0.127	4,010
Cond. Bldgs.	0.095	0.065	0.126	4,010
Graffiti	0.069	0.039	0.100	4,010
Litter, Glass	0.091	0.060	0.122	4,010
Census Records				
Immigrant Conc.	0.037	0.006	0.068	4,010
Conc. Disadv.	0.052	0.021	0.083	4,010
Conc. Affluence	-0.071	-0.102	-0.040	4,010
Perc. Pop< 18	0.086	0.055	0.117	4,010
Perc. Owner	-0.073	-0.103	-0.042	4,010
Perc. Non-White	0.058	0.027	0.088	4,010
Per. Multi-Unit	0.033	0.002	0.064	4,010
Pop. Density	0.029	-0.002	0.060	4,010

Tab. 6.38: Estimates of Z-Y Pearson Correlations (Part 1/4)

Variables	Reciprocated Exchange				Intergenerational Closure			
	Corr	Min95	Max95	n	Corr	Min95	Max95	n
Neighborhood Observations								
Physical Disorder	-0.142	-0.180	-0.104	2,582	-0.301	-0.335	-0.265	2,585
Resid. Decay	-0.079	-0.118	-0.041	2,582	-0.225	-0.262	-0.188	2,585
Security Gates	-0.146	-0.183	-0.108	2,582	-0.216	-0.253	-0.179	2,585
Bars on Windows	-0.119	-0.157	-0.081	2,582	-0.257	-0.293	-0.221	2,585
Damaged Yards	-0.075	-0.113	-0.036	2,582	-0.240	-0.276	-0.204	2,585
Damaged Walls	-0.077	-0.115	-0.038	2,582	-0.222	-0.258	-0.185	2,585
Cond. Bldgs.	-0.080	-0.118	-0.041	2,582	-0.207	-0.244	-0.170	2,585
Graffiti	-0.147	-0.184	-0.109	2,582	-0.304	-0.338	-0.268	2,585
Litter, Glass	-0.132	-0.170	-0.094	2,582	-0.290	-0.325	-0.255	2,585
Census Records								
Immigrant Conc.	-0.091	-0.129	-0.052	2,582	-0.182	-0.220	-0.145	2,585
Conc. Disadv.	-0.063	-0.101	-0.025	2,582	-0.147	-0.185	-0.109	2,585
Conc. Affluence	0.096	0.057	0.134	2,582	0.210	0.173	0.247	2,585
Perc. Pop< 18	-0.069	-0.107	-0.031	2,582	-0.213	-0.250	-0.176	2,585
Perc. Owner	0.155	0.117	0.193	2,582	0.278	0.242	0.313	2,585
Perc. Non-White	-0.128	-0.166	-0.090	2,582	-0.275	-0.311	-0.239	2,585
Per. Multi-Unit	-0.135	-0.172	-0.097	2,582	-0.188	-0.225	-0.150	2,585
Pop. Density	-0.139	-0.177	-0.101	2,582	-0.240	-0.276	-0.203	2,585

Tab. 6.39: Estimates of Z-Y Pearson Correlations (Part 2/4)

Variables	Church Member				Social Groups			
	Corr	Min95	Max95	n	Corr	Min95	Max95	n
Neighborhood Observations								
Physical Disorder	-0.138	-0.175	-0.100	2,572	-0.299	-0.334	-0.264	2,581
Resid. Decay	-0.142	-0.180	-0.104	2,572	-0.250	-0.286	-0.213	2,581
Security Gates	-0.121	-0.159	-0.083	2,572	-0.194	-0.231	-0.156	2,581
Bars on Windows	-0.089	-0.127	-0.050	2,572	-0.292	-0.327	-0.257	2,581
Damaged Yards	-0.146	-0.183	-0.107	2,572	-0.273	-0.308	-0.237	2,581
Damaged Walls	-0.113	-0.151	-0.075	2,572	-0.260	-0.296	-0.224	2,581
Cond. Bldgs.	-0.120	-0.158	-0.082	2,572	-0.242	-0.278	-0.206	2,581
Graffiti	-0.128	-0.166	-0.090	2,572	-0.304	-0.339	-0.269	2,581
Litter, Glass	-0.142	-0.179	-0.104	2,572	-0.291	-0.326	-0.255	2,581
Census Records								
Immigrant Conc.	-0.093	-0.131	-0.055	2,572	-0.161	-0.199	-0.124	2,581
Conc. Disadv.	-0.084	-0.122	-0.045	2,572	-0.107	-0.145	-0.068	2,581
Conc. Affluence	0.070	0.031	0.108	2,572	0.269	0.233	0.304	2,581
Perc. Pop< 18	-0.069	-0.107	-0.030	2,572	-0.247	-0.283	-0.210	2,581
Perc. Owner	0.134	0.096	0.172	2,572	0.243	0.206	0.279	2,581
Perc. Non-White	-0.077	-0.115	-0.038	2,572	-0.313	-0.347	-0.277	2,581
Per. Multi-Unit	-0.105	-0.143	-0.067	2,572	-0.127	-0.164	-0.089	2,581
Pop. Density	-0.105	-0.143	-0.067	2,572	-0.180	-0.217	-0.142	2,581

Tab. 6.40: Estimates of Z-Y Pearson Correlations (Part 3/4)

Variables	Poor Health				Unsafe Neighborhood			
	Corr	Min95	Max95	n	Corr	Min95	Max95	n
Neighborhood Observations								
Physical Disorder	0.221	0.184	0.258	2,532	0.444	0.413	0.475	2,568
Resid. Decay	0.199	0.161	0.236	2,532	0.379	0.346	0.412	2,568
Security Gates	0.132	0.094	0.170	2,532	0.289	0.253	0.324	2,568
Bars on Windows	0.189	0.152	0.227	2,532	0.376	0.342	0.409	2,568
Damaged Yards	0.192	0.154	0.229	2,532	0.361	0.327	0.394	2,568
Damaged Walls	0.195	0.157	0.232	2,532	0.362	0.328	0.395	2,568
Cond. Bldgs.	0.161	0.123	0.198	2,532	0.291	0.255	0.326	2,568
Graffiti	0.214	0.176	0.251	2,532	0.428	0.396	0.459	2,568
Litter, Glass	0.204	0.167	0.241	2,532	0.378	0.345	0.411	2,568
Census Records								
Immigrant Conc.	0.120	0.081	0.158	2,532	0.214	0.177	0.251	2,568
Conc. Disadv.	0.115	0.077	0.153	2,532	0.314	0.279	0.349	2,568
Conc. Affluence	-0.164	-0.202	-0.126	2,532	-0.271	-0.306	-0.234	2,568
Perc. Pop< 18	0.190	0.152	0.227	2,532	0.346	0.311	0.380	2,568
Perc. Owner	-0.160	-0.197	-0.121	2,532	-0.340	-0.374	-0.305	2,568
Perc. Non-White	0.210	0.172	0.247	2,532	0.370	0.336	0.403	2,568
Per. Multi-Unit	0.087	0.048	0.126	2,532	0.191	0.154	0.228	2,568
Pop. Density	0.119	0.081	0.157	2,532	0.230	0.193	0.266	2,568

Tab. 6.41: Estimates of Z-Y Pearson Correlations (Part 4/4)

	Uses Contraception Method				Disapproves Babies out Marriage			
Variables	Corr	Min95	Max95	n	Corr	Min95	Max95	n
Neighborhood Observations								
Physical Disorder	-0.007	-0.049	0.035	2,184	-0.133	-0.171	-0.096	2,609
Resid. Decay	0.017	-0.025	0.058	2,184	-0.124	-0.162	-0.086	2,609
Security Gates	-0.025	-0.067	0.017	2,184	-0.091	-0.129	-0.053	2,609
Bars on Windows	-0.034	-0.076	0.008	2,184	-0.099	-0.137	-0.061	2,609
Damaged Yards	0.018	-0.024	0.060	2,184	-0.117	-0.155	-0.079	2,609
Damaged Walls	0.000	-0.042	0.042	2,184	-0.120	-0.158	-0.082	2,609
Cond. Bldgs.	0.039	-0.003	0.081	2,184	-0.091	-0.129	-0.053	2,609
Graffiti	-0.010	-0.052	0.032	2,184	-0.119	-0.157	-0.081	2,609
Litter, Glass	-0.007	-0.049	0.034	2,184	-0.108	-0.146	-0.070	2,609
Census Records								
Immigrant Conc.	-0.021	-0.063	0.021	2,184	-0.084	-0.122	-0.046	2,609
Conc. Disadv.	0.016	-0.026	0.058	2,184	-0.089	-0.127	-0.051	2,609
Conc. Affluence	0.025	-0.017	0.067	2,184	0.059	0.021	0.098	2,609
Perc. Pop< 18	-0.007	-0.049	0.035	2,184	-0.060	-0.098	-0.021	2,609
Perc. Owner	-0.008	-0.050	0.034	2,184	0.123	0.085	0.160	2,609
Perc. Non-White	-0.020	-0.061	0.022	2,184	-0.083	-0.121	-0.044	2,609
Per. Multi-Unit	0.017	-0.025	0.059	2,184	-0.085	-0.123	-0.047	2,609
Pop. Density	0.006	-0.036	0.048	2,184	-0.091	-0.129	-0.053	2,609

Tab. 6.42: Descriptive Statistics for Predicted Response Propensities ($\hat{p}_z$)

Variable	Mean	Std. Dev.	Min.	Max.	N
Pr(response) Model 1	0.78	0.07	0.55	0.87	3083
Pr(response) Model 2	0.78	0.07	0.54	0.89	3083
Pr(response) Model 3	0.78	0.08	0.49	0.91	3083
Pr(response) Model 4	0.78	0.08	0.49	0.91	3083
Pr(response) Model 5	0.78	0.08	0.45	0.93	3083
Pr(response) Model 6	0.78	0.08	0.49	0.91	3083

Tab. 6.43: Descriptive Statistics for (Not-Trimmed) Household level weights

Variable	Mean	Std. Dev.	Min.	Max.	N
HH weight Model 1	1064.31	1441.52	19.83	8438.38	3083
HH weight Model 2	1073.75	1464.71	19.4	8524.68	3083
HH weight Model 3	1062.94	1438.44	19.19	8656.29	3083
HH weight Model 4	1057.47	1422.98	19.16	8464.99	3083
HH weight Model 5	1048.24	1399.83	19.84	8657.06	3083
HH weight Model 6	1057.24	1422.95	19.11	8488.81	3083

Tab. 6.44: Descriptive Statistics for (Not-Trimmed) RSA weights

Variable	Mean	Std. Dev.	Min.	Max.	N
RSA weight Model 1	1059.74	1486.35	5.35	14309.6	2609
RSA weight Model 2	1068.7	1511.99	5.21	14595.51	2609
RSA weight Model 3	1059.93	1491.25	5.12	14412.7	2609
RSA weight Model 4	1054.12	1476.06	5.10	14161.98	2609
RSA weight Model 5	1047.02	1463.11	5.4	14259.8	2609
RSA weight Model 6	1053.75	1474.11	5.07	14217.19	2609

Tab. 6.45: Descriptive Statistics for (Trimmed) RSA weights

Variable	Mean	Std. Dev.	Min.	Max.	N
RSA weight Model 1	957.36	1093.91	5.35	4009.21	2609
RSA weight Model 2	964.14	1108.37	5.21	4063.57	2609
RSA weight Model 3	954.48	1085.8	5.12	3970.21	2609
RSA weight Model 4	950.36	1076.69	5.10	3924.2	2609
RSA weight Model 5	945	1066.09	5.4	3897.05	2609
RSA weight Model 6	950.93	1078.3	5.07	3937.8	2609
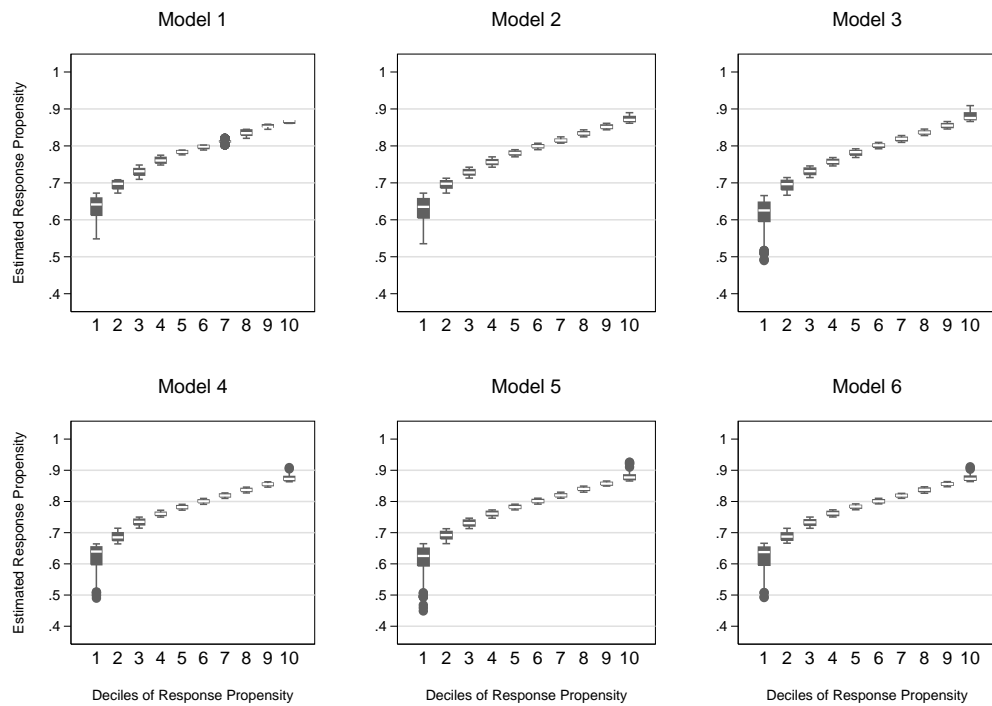
Fig. 6.6: Distribution of predicted propensities within propensity strata.
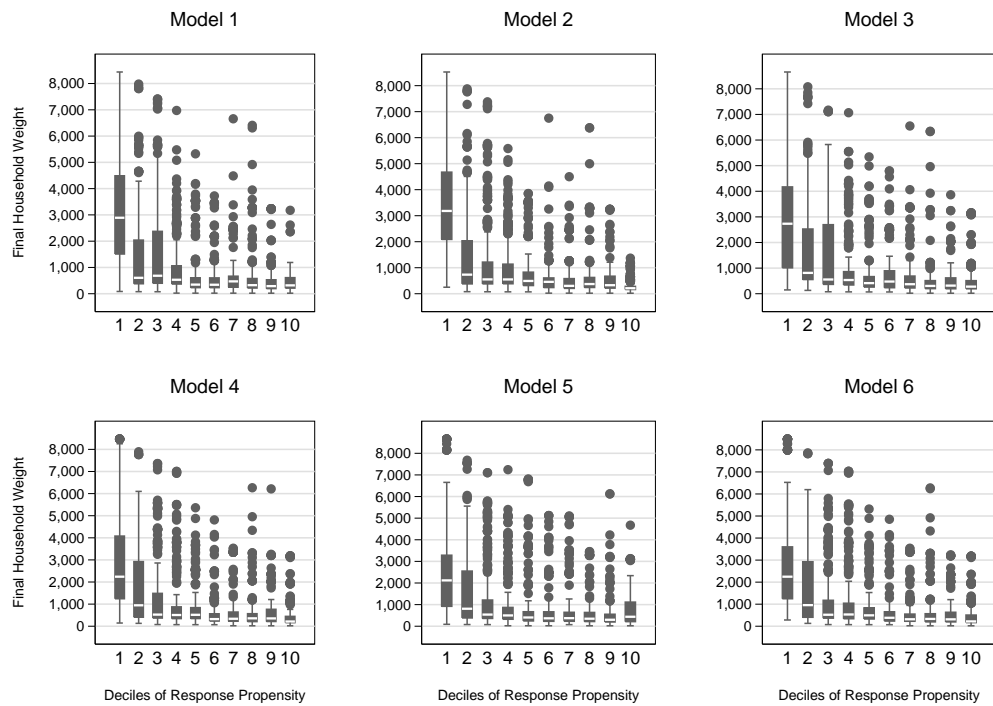
Fig. 6.7: Distribution of final household weights within propensity strata.
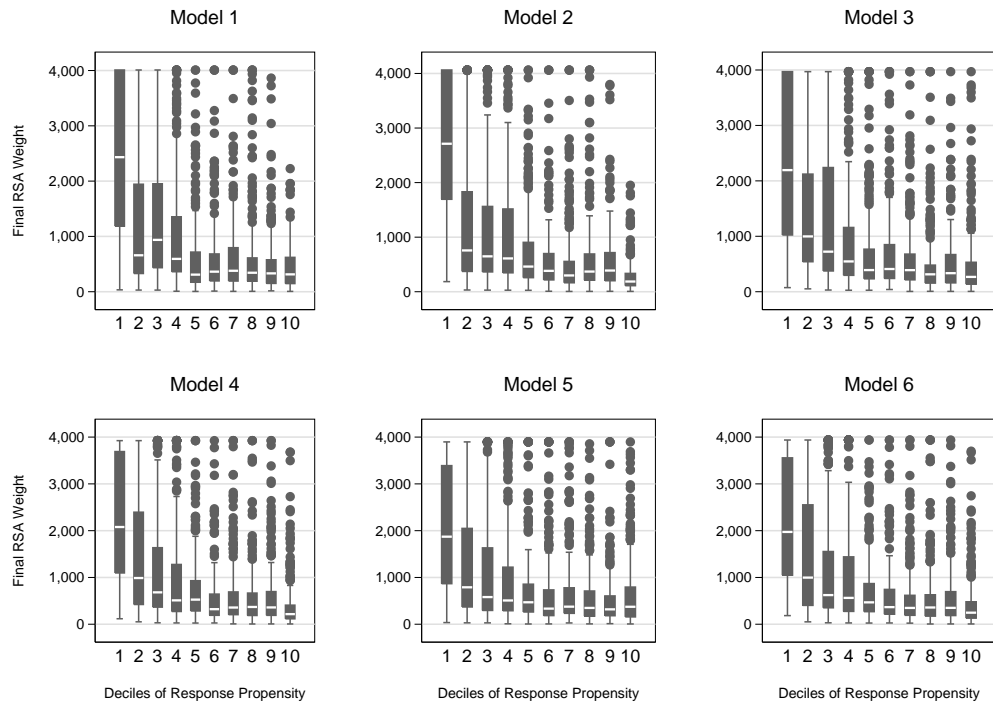
Fig. 6.8: Distribution of final RSA weights within propensity strata.

Tab. 6.46: Estimates of the Nonresponse Adjusted Mean $(\hat{\bar{y}}_{wxz})$ - Not-Trimmed RSA weights.

Nonreponse Adjustment Model	rexch	intclos	church	social	health	unsafe	ctruse	babies2
Estimates of the Mean								
Original	2.6352	3.6133	38.09	38.02	16.59	20.89	52.95	57.16
HH only	2.6469	3.6154	37.80	38.48	17.11	21.17	51.99	57.51
Census scales	2.6467	3.6165	37.88	38.55	17.03	20.97	51.99	57.55
Census Items	2.6473	3.6167	37.92	38.47	16.96	20.99	51.82	57.55
Ngobs Scales	2.6452	3.6168	37.95	38.46	16.92	20.98	51.92	57.57
Ngobs Items	2.6439	3.6175	37.97	38.49	16.92	20.99	51.82	57.67
Census + Ngobs scales	2.6455	3.6171	37.96	38.49	16.93	20.98	51.96	57.58
Estimates of the Standard Error								
Original	0.0401	0.0362	1.87	2.52	1.14	2.11	2.23	1.72
HH only	0.0404	0.0368	1.86	2.63	1.22	2.15	2.48	1.64
Census scales	0.0405	0.0369	1.87	2.65	1.22	2.15	2.49	1.65
Census Items	0.0410	0.0372	1.88	2.65	1.20	2.15	2.47	1.64
Ngobs Scales	0.0399	0.0370	1.90	2.67	1.19	2.15	2.45	1.65
Ngobs Items	0.0394	0.0372	1.88	2.66	1.18	2.12	2.45	1.64
Census + Ngobs scales	0.0398	0.0370	1.90	2.67	1.19	2.15	2.46	1.65

6.4   L.A. FANS Questionnaires

Exhibit 5.1 Screening Form - page 1

**L.A.FANS
SCREENING/CONTACT BOOKLET**

HOUSEHOLD INFORMATION BOX

Tract # _____   Block # _____

CaseID: _____

Address/Notes: _____

_____

City: _____

HH w/o kids: Yes or No

**INTERVIEWER OBSERVATION:
SCREENER RESPONDENT
WHO COMPLETED S1 AND S2**

Gender: ☐ Male ☐ Female

Age: ☐ 18-24

☐ 25-34

☐ 35-54

☐ 55-69

☐ 70+

**INTERVIEWER OBSERVATION:
CODE WHAT RACE OR RACES YOU WOULD SAY
THE SCREENER RESPONDENT IS IF YOU DID NOT
KNOW ANYTHING ABOUT HIM/HER**

CIRCLE ALL THAT APPLY

1. Latino
2. White
3. African-American, Black
4. Asian
5. Pacific Islander
6. Native American/American Indian

**A. INTERVIEWER OBSERVATIONS:
DWELLING CHARACTERISTICS**

(Record after <u>first</u> visit to HH)

A1. WHAT TYPE OF HOUSING IS THIS?

1. Apartment
2. Single family home
3. Mobile home or trailer
4. Unit in a rooming house
5. Other, specify: _____
6. Duplex

**A2. WHAT IS YOUR BEST ESTIMATE OF
HOW MUCH THIS PLACE WOULD
COST TO RENT PER MONTH? (YOUR
BEST GUESS.)**

1. Less than $500 per month
2. $500 to $999 per month
3. $1000 to $1999 per month
4. $2000 to $2999 per month
5. $3000 PER MONTH OR MORE

**B. CASE INFORMATION (RECORD
<u>AFTER</u> SCREENING IS COMPLETE)**

**May I have your full name and telephone
number so that one of my supervisors can
contact you to verify the quality of my work?**
¿Podría decirme su nombre completo y su número de
teléfono para que mi supervisor pueda comunicarse
con usted para verificar la calidad de mi trabajo?

B1. _____

First Name   MI   Last Name

B2. Phone: (   ) _____ - _____

☐ SR would not provide name

☐ SR would not provide phone number

B3. SCREENER LANGUAGE

1. English
2. Spanish
3. No interview

B4. HOUSEHOLD TYPE

1. Household <u>without</u> children
2. Household <u>with</u> child
3. Don't know/UNABLE TO DETERMINE

Fig. 6.9: L.A. FANS Screening Form (part 1/2).

Exhibit 5.2 Screening Form - page 2

PART 1. INTRODUCTION: YOU SHOULD ALWAYS:

- Give your name and show your RTI ID card.

- Explain that you are conducting a neighborhood survey for RTI, a not-for-profit research organization in Research Triangle Park, NC.

- Ask to speak to an adult household member (or household head) who is at least 18 years old.

- If no adult (or household head is home), find out the best day/time to come back. Probe for day of week and time of day (mornings, afternoon, or evenings).

- Explain that we mailed the household a letter about the survey. Be prepared to give the resident another copy of the letter and brochure in case they didn't receive the initial packet or they want more information.

- B̲riefly describe the purpose of your visit**:**
  Today, I'd like to ask you two questions to find out if you're eligible to take part in our neighborhood survey.
  Ahora, me gustaría hacerle dos preguntas para saber si su familia tiene las características para tomar parte en la encuesta del vecindario.

- If the respondent asks for more information, use suggested question and answer guide to respond to his/her concerns.

- Ask the two screener questions (S1 and S2). See questions in box below:

PART 2. ASK TWO SCREENER QUESTIONS: RECORD ANSWERS IN SPACE BELOW

To find out if your household is eligible for the survey, I just need to ask you two questions.

S1. **Including yourself, how many adults age 18 and older usually live or stay in this household?**
  Incluyéndolo/a a usted, ¿cuántos adultos de 18 años o mayores normalmente viven o se quedan en este hogar?
  Number of Adults (18 and older): _____ or [ ] None   [ ] Refused

S2. **And how many children age 17 and younger usually live or stay in this household?**
  ¿Y cuántos niños de 17 años o menores normalmente viven o se quedan en este hogar?
  Number of Children (17 and younger): _____ or [ ] None   [ ] Refused

PART 3. DETERMINE IF HOUSEHOLD IS ELIGIBLE :

[   ] HOUSEHOLD HAS C̲HILDREN AGE 17 OR YOUNGER → HH IS ELIGIBLE! **CODE 291; GO TO PART 4**

[   ] HOUSEHOLD DOES N̲OT HAVE CHILDREN → Check eligibility in Household Information Box on page 1.

  ☞ ____ **If HH is eligible to participate in survey)** → **CODE 292; GO TO PART 4**

  ☞ ____ **If HH is n̲ot eligible** →   **CODE 290**

  Your household was not selected for the study, but we may contact you later in the year to see if you would like to take part. Thank you for your time.
  Su hogar no fue seleccionado para participar en este estudio, pero es posible que no comuniquemos con usted más adelante para saber si le gustaría participar. Gracias por su tiempo.

- Make sure that you have collected verification information in B1 & B2 on page 1.

PART 4. FOR ELIGIBLE HOUSEHOLDS:

- FOLLOW THE RECRUITMENT PROTOCOL TO EXPLAIN STUDY AND PERSUADE HH TO PARTICIPATE IN THE SURVEY.

Fig. 6.10: L.A. FANS Screening Form (Part 2/2).

4/10/00
NEIGHBORHOOD OBSERVATION
COVER PAGE

Interviewer ID No.: _____

Interviewer Name: _____

Tract Number: _____

Block Number: _____

No. of Block Faces: ☐☐ → **See attached Master List of Block Faces**

No. of Alleys: ☐ **or None.....0**

Record of Block Visit:	**Your Experiences on Block**
a. Date: _____ / _____ / _____ Mo Day Yr	1. What experiences have you had on this block? *(CIRCLE ALL THAT APPLY)* None .... 0 Screened or interviewed households .... 1 Validated household screening or interviews .... 2 Listed .... 3 Validated listings .... 4 Familiar with area .... 5 Have friends/relatives in area .... 6 Live in area or near by 7
b. Day of Week: *(CIRCLE ONE)* Monday ....1 Tuesday ....2 Wednesday ....3 Thursday ....4 Friday ....5 Saturday ....6 Sunday 7	2. How many [hours/minutes] did you spend on this block face **this visit**? └──┴──┘ hours └──┴──┘ minutes 3. How many times have you visited this block face? └──┴──┘ visits 4. Have you visited this block during any of the following times? *(CIRCLE ALL THAT APPLY)* Weekday evening/night .... 1 Weekend days .... 2 Weekend evening/night 3
c. Start Time: _____ : _____ ₁ AM ₂ PM d. End Time: _____ : _____ ₁ AM ₂ PM	5. **Notes**: Were there any special situations or circumstances that may have affected how you filled out this form (e.g., unusual weather, trash day, etc.)? *(CIRCLE ONE)* Yes ....1 → Describe on back of this form. No 5

-1-

Fig. 6.11: L.A. FANS Neighborhood Observations Cover Page.

242

BLOCK FACE OBSERVATION FORM

Street Name/Description:_____

Between Street/Landmark 1:_____

And Street/Landmark 2: _____

1. How many **lanes for traffic** are there on this street?

   Number of lanes .... |___|

2. What is the **traffic flow** on this street?
   - Very light 1
   - Light.. 2
   - Moderate 3
   - Heavy 4
   - Very heavy 5

3. Are there **speed bumps** on this street?
   - Yes 1
   - No 5

4. How would you rate the condition of the **street surface** (for driving)?
   - Very poor 1
   - Fair.. 2
   - Moderately good 3
   - Very good 4
   - Under construction................................... 5

5. How would you rate the condition of the **sidewalks** (for walking)?
   - Very poor 1
   - Fair.. 2
   - Moderately good 3
   - Very good 4
   - Under construction................................... 5
   - No sidewalks.. 6

6. Are there permit-only **parking restrictions** on this street?
   - Yes 1
   - No 5

7. Is there **public transportation** (e.g., a bus stop) on this block?
   - Yes 1
   - No 5

8. Is this street **barricaded** to prevent through-traffic?
   - Yes 1
   - No 5

9. Are there **trees** lining the street of the block face?
   - None.. 1
   - A few... 2
   - Some 3
   - Many 4

10. Are there **abandoned cars** on the street or in alleys or lots?
    - None 1
    - A few 2
    - Some...3
    - Many...4

11. Is there **trash or junk** on the street or sidewalks, in yards/lots?
    - None 1
    - A little...2
    - Some...3
    - A lot..4

12. Is there **garbage, litter, or broken glass** on the street or sidewalk, in yards, or vacant lots?
    - None 1
    - A little...2
    - Some...3
    - A lot..4

13. Are there **needles, syringes, condoms, or drug-related paraphernalia** on the street or sidewalk, in yards/lots?
    - None 1
    - A few 2
    - Some...3
    - Many...4

14. Are there **empty beer containers or liquor bottles** on the street or sidewalks, in yards, or vacant lots?
    - None 1
    - A few 2
    - Some...3
    - Many...4

15. Are there **cigarettes or cigar butts or discarded cigarette packages** on the street or sidewalks, in yards/lots or gutters?
    - None 1
    - A few 2
    - Some...3
    - A lot..4

16. Is there **graffiti** on buildings, sidewalks, walls, or signs?
    - None 1
    - A little...2
    - Some...3
    - A lot..4

Fig. 6.12: L.A. FANS Block Face Observation Form (Part 1/4).

17. Is there **painted-over graffiti** on buildings, sidewalks, walls, or signs?
    None..1
    A little 2
    Some 3
    A lot 4

18. Are there obvious **strong odors** anywhere in the block face (urine stench, rotting garbage, etc.)?
    Yes 1
    No 5

19. How would you characterize the **land use** on this block face? *(CIRCLE ONE)*
    Primarily residential (houses and apartments) 1
    Primarily commercial (stores and businesses) 2
    Primarily industrial (warehouses and factories)..............3
    Primarily vacant lots or undeveloped open space 4
    Mixed residential and commercial..................................5
    Mixed residential and industrial....................................6
    Mixed residential and vacant lots..................................7
    Other, specify:_____ 8

20. What are the **main types of housing** along this block face? *(CIRCLE ALL THAT APPLY)*
    No residential units (GO TO 33)....................................1
    Stand-alone houses 2
    Duplexes (two-household structures)..............................3
    Multiple household occupancy (3-6 units).......................4
    Housing units over commercial store fronts. 5
    Low rise apartment or condominium buildings (7 or more units, one to three floors) 6
    Mid-rise apartment or condominium buildings (four to six floors) 7
    High-rise apartment or condominium buildings (more than 6 floors) 8

21. What is the overall **condition** of the residential buildings?
    Very poor 1
    Poor...2
    Fair...3
    Very good 4
    Excellent 5

22. How many houses/apartments are **burned out, boarded up, or abandoned**?
    None..1
    Very few 2
    Some 3
    Many 4
    All 5

23. How many **vacant lots** are there on this block?
    None..1
    Very few 2
    Some 3
    Many 4
    All 5

24. How many houses/apartments have **peeling paint or damaged exterior walls**?
    None 1
    Very few 2
    Some 3
    Many...4
    All...5

25. How many houses/apartments have **well-tended yards or gardens**?
    None 1
    Very few 2
    Some 3
    Many...4
    All...5

26. How many houses/apartments have **window bars or gratings** on doors or windows?
    None 1
    Very few 2
    Some 3
    Many...4
    All...5

27. How many houses/apartments have signs indicating they are protected by **private security services**?
    None 1
    Very few 2
    Some 3
    Many...4
    All...5

28. How many houses/apartments have signs indicating they are **protected by dogs**?
    None 1
    Very few 2
    Some 3
    Many...4
    All...5

29. How many houses/apartments have **security gates or security fences**?
    None 1
    Very few 2
    Some 3
    Many...4
    All...5

30. Are there signs indicating there is a **neighborhood watch** on this block?
    Yes...1
    No 5

Fig. 6.13: L.A. FANS Block Face Observation Form (Part 2/4).

31. How many houses/apartments have "**for sale**" or "**for rent**" **signs**?

None... 1
Very few 2
Some 3
Many 4
All 5

32. Are there **old, beat-up cars** on the street or in driveways or yards?

None... 1
Very few 2
Some 3
Many 4

33. What is the overall **condition** of the commercial/industrial buildings?

No Commercial/Industrial Building (GO TO 38)..... 1
Excellent 2
Very good 3
Fair.. 4
Poor... 5
Very poor 6

34. How many of the commercial/industrial buildings are **abandoned, burned out, or boarded up**?

None... 1
Very few 2
Some 3
Many 4
All 5

35. How many of the commercial/industrial buildings have **windows that are barred or boarded** against entry?

None... 1
Very few 2
Some 3
Many 4
All 5

36. How many of the commercial/industrial properties have **security fences**?

None... 1
Very few 2
Some 3
Many 4
All 5

37. How many commercial / industrial buildings have "**for sale**" or "**for rent**" **signs**?

None... 1
Very few 2
Some 3
Many 4
All 5

38. Are there any **recreational facilities** in the block face (see list in next question)?

Yes.. 1
No (GO TO 40) 5

39. What kinds of **recreational facilities** are in the block face? (*CIRCLE ALL THAT APPLY*)

Park... 1
Playground..2
Sports/playing fields/courts/swimming pool 3
Community gardens.......................................4

40. Is there a **public telephone** easily visible on the block face?

Yes.. 1
No 5

41. Please look at the list below and indicate the presence of **commercial establishments** that you observed in this face block. Are there any…(WRITE 1 FOR YES, 0 FOR NO)

01. Street vendors (on sidewalk or in vehicles) ... |___|
02. Banks .... |___|
03. Check cashing services .... |___|
04. Pawn shops .... |___|
05. Second hand stores / thrift shops .... |___|
06. Massage parlor... |___|
07. Sex stores/porno shops/peep shows............... |___|
08. Video store .... |___|
09. Video games/pool halls .... |___|
10. Liquor stores.. |___|
11. Bars.. |___|
12. Restaurants (sit-down).................................... |___|
13. Fast food/take out places .... |___|
14. Hotels/motels .... |___|
15. Cinema/theatre.. |___|
16. Parking lot (commercial) .... |___|
17. Barber shops and beauty salons...................... |___|
18. Dry cleaners/tailors.. |___|
19. Laundromats.. |___|
20. Clothing store .... |___|
21. Discount stores (e.g., Target, WalMart) .... |___|
22. Convenience stores/7-11s............................. |___|
23. Food stores (e.g., bakery, butcher) .... |___|
24. Grocery stores: large chain .... |___|
25. Grocery stores: independent .... |___|
26. Drug store/pharmacy .... |___|
27. Specialty stores (books, software) .... |___|
28. Variety stores.. |___|
29. Electronics stores... |___|
30. Appliance sales/rental/repair/etc................... |___|

Fig. 6.14: L.A. FANS Block Face Observation Form (Part 3/4).

31. How many houses/apartments have "**for sale**" or "**for rent**" **signs**?

    None 1
    Very few 2
    Some 3
    Many 4
    All 5

32. Are there **old, beat-up cars** on the street or in driveways or yards?

    None 1
    Very few 2
    Some 3
    Many 4

33. What is the overall **condition** of the commercial/industrial buildings?

    No Commercial/Industrial Building (GO TO 38) 1
    Excellent 2
    Very good 3
    Fair 4
    Poor 5
    Very poor 6

34. How many of the commercial/industrial buildings are **abandoned, burned out, or boarded up**?

    None 1
    Very few 2
    Some 3
    Many 4
    All 5

35. How many of the commercial/industrial buildings have **windows that are barred or boarded** against entry?

    None 1
    Very few 2
    Some 3
    Many 4
    All 5

36. How many of the commercial/industrial properties have **security fences**?

    None 1
    Very few 2
    Some 3
    Many 4
    All 5

37. How many commercial / industrial buildings have "**for sale**" or "**for rent**" **signs**?

    None 1
    Very few 2
    Some 3
    Many 4
    All 5

38. Are there any **recreational facilities** in the block face (see list in next question)?

    Yes 1
    No (GO TO 40) 5

39. What kinds of **recreational facilities** are in the block face? (*CIRCLE ALL THAT APPLY*)

    Park 1
    Playground 2
    Sports/playing fields/courts/swimming pool 3
    Community gardens 4

40. Is there a **public telephone** easily visible on the block face?

    Yes 1
    No 5

41. Please look at the list below and indicate the presence of **commercial establishments** that you observed in this face block. Are there any…(WRITE 1 FOR YES, 0 FOR NO)

    01. Street vendors (on sidewalk or in vehicles) ... |___|
    02. Banks .... |___|
    03. Check cashing services .... |___|
    04. Pawn shops .... |___|
    05. Second hand stores / thrift shops .... |___|
    06. Massage parlor .... |___|
    07. Sex stores/porno shops/peep shows .... |___|
    08. Video store .... |___|
    09. Video games/pool halls .... |___|
    10. Liquor stores .... |___|
    11. Bars .... |___|
    12. Restaurants (sit-down) .... |___|
    13. Fast food/take out places .... |___|
    14. Hotels/motels .... |___|
    15. Cinema/theatre .... |___|
    16. Parking lot (commercial) .... |___|
    17. Barber shops and beauty salons .... |___|
    18. Dry cleaners/tailors .... |___|
    19. Laundromats .... |___|
    20. Clothing store .... |___|
    21. Discount stores (e.g., Target, WalMart) .... |___|
    22. Convenience stores/7-11s .... |___|
    23. Food stores (e.g., bakery, butcher) .... |___|
    24. Grocery stores: large chain .... |___|
    25. Grocery stores: independent .... |___|
    26. Drug store/pharmacy .... |___|
    27. Specialty stores (books, software) .... |___|
    28. Variety stores .... |___|
    29. Electronics stores .... |___|
    30. Appliance sales/rental/repair/etc .... |___|

Fig. 6.15: L.A. FANS Block Face Observation Form (Part 4/4).

SOCIAL OBSERVATION FORM

Street Name/Description:_____

Between Street/Landmark 1:_____

And Street/Landmark 2: _____

1. Did you see a **police officer** on the block face?
   (CIRCLE ALL THAT APPLY)
   In a vehicle.. 1
   On a bicycle/horseback 2
   On foot 3
   Did not see a police officer 9

2. Did you see any **private security guards** on the block face?
   Yes 1
   No 5

3. Did you see any **children** on the block face?
   (CIRCLE ALL THAT APPLY)
   Playing in the front private yards.............................. 1
   Playing on the sidewalk or in the street..................... 2
   Under adult supervision / accompanied by an adult.. 3
   Arguing, fighting, acting hostile or threatening 4
   Saw children but not in above activities.................... 5
   Did not see any children.. 9

4. Did you see any **teenagers** on the block face?
   Yes 1
   No (GO TO 9).. 5

5. Did you see any **teenagers in groups** of three or more?
   Yes 1
   No (GO TO 9).. 5

6. Were teenagers in the groups you saw **male, female, or mixed**?
   All male... 1
   All female 2
   Mixed male/female 3
   Did not see teenagers in peer groups......................... 9

7. Did you see teenagers in the group who were…?
   (CIRCLE ALL THAT APPLY)
   Wearing the same style clothes? 1
   Wearing the same color(s)?.. 2
   Wearing sports insignias? 3
   Wearing the same hats, jewelry, or shoes?................ 4
   Saw teenagers in groups but none of the above 5
   Did not see any teenagers in groups.......................... 9

8. Did any of the groups of teens you saw appear to be a **gang**?
   Yes 1
   No 5
   Did not see teenagers in groups 9

9. Did someone tell you that there was a **gang** or gang activity on the block face?
   Yes...1
   No 5
   No one spoke about gangs on the block face.............9

10. Did you see any **adults** on the block face?
    Yes...1
    No (GO TO 12) 5

11. Did you see any adults on the block face **loitering, congregating or hanging out**?
    Yes...1
    No 5

12. Did you see any **prostitutes** on the block face?
    Yes...1
    No 5

13. Did someone tell you that **prostitutes** work on the block face?
    Yes 1
    No 5
    No one spoke about prostitutes on the block face......9

14. Did you see any **homeless** people or people **begging** on the block face?
    Yes...1
    No 5

15. Did someone tell you that there are **homeless** people or people **begging** on the block face?
    Yes 1
    No 5
    No one spoke about homeless people or people
        begging on the block face...................................9

16. Did you see people who were **selling illegal drugs** on the block face?
    Yes...1
    No 5

17. Did someone tell you that people **sell illegal drugs** on the block face?
    Yes 1
    No 5
    No one spoke about illegal drug sales on the block
        face 9

18. Did you see any people **drinking alcohol** openly on the block face?
    Yes...1

Fig. 6.16: L.A. FANS Social Observation Form (Part 1/2).

No 5

19. Did you see any **drunken** or otherwise **intoxicated people** on the block face?

    Yes 1

    No 5

20. Did someone tell you that **drunk** or **intoxicated people** loiter on the block face?

    Yes 1

    No 5

    No one spoke about drunk or intoxicated people

        loitering on the block face....................................... 9

21. Did you hear **loud music** playing from boom boxes or any of the buildings on the block face?

    Yes 1

    No 5

22. Did you hear or see another **language** other than English on the block face? *(CIRCLE ALL THAT APPLY)*

    Heard or saw other language(s) but don't know

        which one.. 01

    Spanish... 02

    Armenian 03

    Khmer 04

    Vietnamese.. 05

    Korean... 06

    Filipino... 07

    Chinese 08

    Japanese 09

    Other, Specify _____ 10

    Other, Specify _____ 11

    No people around or did not hear or see any non-

        English language(s)... 99

23. How did people on the block face regard you? *(CIRCLE ALL THAT APPLY)*

    Paid little or no attention by those around.................... 1

    Treated with suspicion 2

    Friendly responses, greetings, helpful.......................... 3

    Polite responses to your queries................................... 4

    Queried about what you were doing in

    neighborhood 5

    No people around 9

Fig. 6.17: L.A. FANS Social Observation Form (Part 2/2).

ALLEY OBSERVATION FORM

Alley Name/Description: _____

Between Street/Landmark 1: _____

And Street/Landmark 2: _____

1. Are there any **houses** with main entrances on the alley?
   - Yes 1
   - No 5

2. Are there any **businesses** with main entrances on the alley?
   - Yes 1
   - No 5

3. Is this alley **locked-down**?
   - Yes 1
   - No 5

4. How would you rate the condition of the **street surface** in the alley?
   - Very poor 1
   - Fair 2
   - Moderately good 3
   - Very good 4
   - Under construction 5

5. Are there **abandoned cars** in the alley?
   - None 1
   - A few 2
   - Some 3
   - Many 4

6. Is there **trash or junk** in the alley?
   - None 1
   - A little 2
   - Some 3
   - A lot 4

7. Is there **garbage, litter, or broken glass** in the alley?
   - None 1
   - A little 2
   - Some 3
   - A lot 4

8. Are there **needles, syringes, condoms, or drug-related paraphernalia** in the alley?
   - None 1
   - A few 2
   - Some 3
   - A lot 4

9. Are there **empty beer containers or liquor bottles** in the alley?
   - None 1
   - A few 2
   - Some 3
   - Many 4

10. Are there **cigarettes or cigar butts or discarded cigarette packages** in the alley?
    - None 1
    - Very little 2
    - Some 3
    - A lot 4

11. Is there **graffiti** on buildings, walls, or signs?
    - None 1
    - Very little 2
    - Some 3
    - A lot 4

12. Is there **painted-over graffiti** on buildings, walls, or signs?
    - None 1
    - A little 2
    - Some 3
    - A lot 4

13. Are there obvious **strong odors** anywhere in the alley (urine stench, rotting garbage, etc.)?
    - Yes 1
    - No 5

Fig. 6.18: L.A. FANS Alley Form.

LA FANS Interviewer Survey

We would like to know something about our interviewers' backgrounds and ask you to answer the questions below.

Your participation in this survey is completely *voluntary* and does *not* affect your employment as an LA FANS interviewer in any way. Your answers also will have *no* effect on your employment. If you do not wish to answer a particular question, just skip over it.

 The information you provide will be used *only* by project personnel to get to know our interviewers and as part of the research to be conducted by the LA FANS project. No information which could identify you will be released publicly.

Please write in your name _____

A. *FIRST A FEW QUESTIONS ABOUT YOUR NEIGHBORHOOD*

1. What city or town do you live in?

_____

2. Suppose you were talking to someone who lives here in the *same* city or town that you do and you were telling them where you live. What name would you use for your neighborhood?

 _____

3. . When you are talking to someone about your neighborhood, what do you mean? Is it….

(CIRCLE ONE ANSWER)

1   The block or street you live on?
2   Several blocks or streets in each direction?
3   The area within a 15-minute walk from your house?
4   An area larger than a 15-minute walk from your house?

Fig. 6.19: L.A. FANS Interviewer Survey (Part 1/7).

250

4. All things considered, would you say you are very satisfied, satisfied, dissatisfied or very dissatisfied with your neighborhood as a place to live?

CIRCLE *ONE* RESPONSE

1. Very Satisfied
2. Satisfied
3. Neutral – not satisfied or dissatisfied
4. Dissatisfied
5. Very Dissatisfied

5. How long have you lived in your current neighborhood?

1. Less than one year
2. _____ Years (FILL IN ABOUT HOW MANY YEARS)
3. Lived in neighborhood my whole life **(SKIP TO Q7)**

6. How long have you lived in Los Angeles County?

1. Less than one year
2. _____ Years (FILL IN ABOUT HOW MANY YEARS)
3. Lived in Los Angeles County my whole life

7. About how many of your relatives or in-laws live in your neighborhood, but do not live with you?

(CIRCLE ONE ANSWER)

1. None
2. Few
3. Many
4. Most or all

8. About how many of your friends live in your neighborhood, but do not live with you?

(CIRCLE ONE ANSWER)

1. None
2. Few
3. Many
4. Most or all

Fig. 6.20: L.A. FANS Interviewer Survey (Part 2/7).

B. NEXT SOME QUESTIONS ABOUT YOUR BACKGROUND

9. Please look at this list and tell me what *group or groups* describe your race or ethnic origin.

(CIRCLE ALL THAT APPLY)

1. BLACK/AFRICAN-AMERICAN
2. WHITE
3. LATINO/ HISPANIC/ LATIN AMERICAN
4. ASIAN INDIAN/SOUTH ASIAN
5. CHINESE
6. FILIPINO
7. JAPANESE
8. KOREAN
9. VIETNAMESE
10. OTHER ASIAN
11. NATIVE AMERICAN/ AMERICAN INDIAN
12. INUIT/ESKIMO/ALEUT
13. HAWAIIAN
14. PACIFIC ISLANDER
15. OTHER, SPECIFY _____ (WRITE IN YOUR ANSWER)

Fig. 6.21: L.A. FANS Interviewer Survey (Part 3/7).

**10. ***Answer this question *only* if you circled more than one answer in 9.
Otherwise skip to question 11 below. ********
If you had to choose *one single* group which *best* describes your race or national origin,
which one would you choose?

(CIRCLE ONLY <u>ONE</u>)

1. BLACK/AFRICAN-AMERICAN
2. WHITE
3. LATINO/ HISPANIC/ LATIN AMERICAN
4. ASIAN INDIAN/SOUTH ASIAN
5. CHINESE
6. FILIPINO
7. JAPANESE
8. KOREAN
9. VIETNAMESE
10. OTHER ASIAN
11. NATIVE AMERICAN/ AMERICAN INDIAN
12. INUIT/ESKIMO/ALEUT
13. HAWAIIAN
14. PACIFIC ISLANDER
15. OTHER, SPECIFY

11. How old are you?

_____ Years Old

12. Where were you born? What city, state, and country?

_____ CITY
_____ STATE/PROVINCE/TERRITORY
_____ COUNTRY

**13. ***Answer this question only if you were born outside of the United States.
Otherwise skip to question 14.*******
How old were you when you <u>first</u> came to the United States to live or work? Please do
not include short trips for shopping, vacation or family visits.

1. Less than one year old
2. _____ Years old (WRITE IN HOW OLD YOU WERE)

Fig. 6.22: L.A. FANS Interviewer Survey (Part 4/7).

14. How much school have you completed?

(*CIRCLE ONE ANSWER*)

1. LESS THAN HIGH SCHOOL
2. HIGH SCHOOL GRADUATE OR COMPLETED GED
3. SOME VOCATIONAL SCHOOL
4. COMPLETED VOCATIONAL SCHOOL
5. SOME COLLEGE
6. ASSOCIATES' DEGREE (AA)
7. BACHELORS' DEGREE (BA, BS)
8. SOME GRADUATE OR PROFESSIONAL SCHOOL (AFTER COMPLETING COLLEGE)
9. COMPLETED GRADUATE/PROFESSIONAL DEGREE

15. Think about the highest grade of regular school or highest degree that you completed. In what year did you complete this grade or degree?

_____Year

16. Are you currently in school? If so, what level are your enrolled in?

1. Yes, in high school
2. Yes, in college
3. Yes, in graduate or professional school
4. Yes, in vocational, training or apprenticeship program
5. Yes, in some other type of program
6. No, not currently in school

17. Have you received any other degree or a certificate through a vocational school, a training school, or an apprenticeship program? Please do not include ESL, citizenship classes or Job Club.

1. YES, I've received _____
5. NO

18. What is your current marital status?

(*CIRCLE ONE*)

1. Currently married or living with a partner
2. Separated from a marriage
3. Widowed
4. Divorced
5. Never (legally) married

Fig. 6.23: L.A. FANS Interviewer Survey (Part 5/7).

19. Do you have any children?

1. Yes
2. No **(SKIP TO Q22)**

20. How many children do you have?

_____ Children

21. How old is your oldest child?

1. Less than one year
2. _____years old

22. Do you currently have another job (other than with LA FANS)?

1. Yes
2. No (SKIP TO Q25)

23. What kind of work are you doing at this other job? (For example: electrical engineer, stock clerk, typist, farmer)

NOTE: IF YOU HAVE MORE THAN ONE OTHER JOB, ANSWER FOR THE MAIN JOB

_____

_____

24. What are your most important activities or duties at this other job? (For example: typing, keeping account books, filing, selling cars, operating printing press, finishing concrete)

NOTE: IF YOU HAVE MORE THAN ONE OTHER JOB, ANSWER FOR THE MAIN JOB

_____

_____

Fig. 6.24: L.A. FANS Interviewer Survey (Part 6/7).

25. Do you know anyone personally who is currently receiving welfare (such as CalWORKS, AFDC, General Relief)?

1. Yes (SKIP TO Q27)
2. No

26. Have you ever known anyone personally who was receiving welfare (such as CalWORKS, AFDC, General Relief)?

1. Yes
2. No

27. How active are you in community activities or organizations or in volunteer activities? Would you say….

1. Very active
2. Somewhat active
3. Not active **(SKIP TO END)**

28. Which community organizations or activities or volunteer organizations are your involved in?

_____

Thanks for your help!!

Fig. 6.25: L.A. FANS Interviewer Survey (Part 7/7).

BIBLIOGRAPHY

Abraham, K., S. Helms, and S. Presser (2009). How social processes distort measurement: The impact of survey nonresponse on estimates of volunteer work in the United States. *American Journal of Sociology 114*, 1129–1165.

Abraham, K., A. Maitland, and S. Bianchi (2006). Nonresponse in the American Time Use Survey. Who is missing from the data and how much does it matter? *Public Opinion Quarterly 70*, 676–703.

Aellio, J. R. and A. Baum (1979). *Residential Crowding and Design*. New York: Plenum Press.

Alwin, D. (2008). *Margins of Error*. New York: John Wiley and Sons.

Andresen, E., T. Malmstrom, D. Miller, and F. Wolinsky (2006). Reliability and validity of observer ratings of neighborhoods. *Journal of Aging and Health 18*(1), 28–36.

Aneshensel, C. S. (2002). *Theory-Based Data Analysis for the Social Sciences*. Thousand Oaks, CA: Pine Forge Press, Sage Publications, Inc.

Araya, R., F. Dunstan, R. Playle, H. Thomas, S. Palmer, and G. Lewis (2006). Perceptions of social capital and the built environment and mental health. *Social Science and Medicine 12*, 3072–3083.

Atrostic, B., N. Bates, and A. Silberstein (2001). Nonresponse in U.S. government household surveys: Consistent measures, recent trends, and new insights. *Journal of Official Statistics 17*, 209–226.

Bailey, L., T. F. Moore, and B. A. Bailar (1978). An interviewer variance study for the eight impact cities of the national crime survey cities sample. *Journal of the American Statistical Association 73*, 16–23.

Bates, N., J. Dahlhamer, and E. Singer (2008). Privacy concerns, too busy, or just not interested: Using doorstep concerns to predict survey nonresponse. *Journal of Official Statistics 24*, 591–612.

Bethlehem, J. (1988). Reduction of nonresponse bias through regression estimation. *Journal of Official Statistics 4*(3), 251–260.

Biemer, P. and L. Lyberg (2003). *Introduction to Survey Quality*. Wiley Series in Survey Methodology. Hoboken, NJ: John Wiley and Sons, Inc.

Blom, A., P. Lynn, and A. Jackle (2008). Understanding cross-national differences in unit non-response: the role of contact data. Technical report, Institute for Social and Economic Research.

Brehm, J. (1993). *The phantom respondents: Opinion surveys and political representation.* Ann Arbor: University of Michigan Press.

Brick, J. M. and G. Kalton (1996). Handling missing data in survey research. *Statistical Methods in Medical Research 5*, 215–238.

Brooks-Gunn, J., G. Duncan, and J. Aber (1997). *Neighborhood Poverty: Vol. I: Context and Consequences for Children.* New York: Russell Sage Found.

Brown, B., D. Perkins, and G. Brown (2004). Incivilities, place attachment and crime: Block and individual effects. *Journal of Environmental Policy 24*(3), 359–371.

Browning, C. (2002). The span of collective efficacy: Extending social disorganization theory to partner violence. *Journal of Marriage and the Family 64*(4), 833–850.

Browning, C., T. Leventhal, and J. Brooks-Gunn (2005). Sexual initiation in early adolescence: The nexus of parental and community control. *American Sociological Review 70*(5), 758–778.

BUREAU, U. C. (2009). Design and methodology, american community survey. Technical report, U.S. Government Printing Office, Washington, DC.

Burgess, E. W. (1930). The value of sociological community studies for the work of social agencies. *Social Forces 8*(4), 481–491.

Bursik, R. J. (1986). Ecological stability and the dynamics of delinquency. In A. J. Reiss and M. Tonry (Eds.), *Communities and Crime.* Chicago, IL: University of Chicago Press.

Bursik, R. J. (1989). Political decisionmaking and ecological models of delinquency: Conflict and consensus. In S. F. Messner, M. D. Krohn, and A. E. Liska (Eds.), *Theoretical Integration in the Study of Deviance and Crime.* Albany: State University of New York Press.

Campanelli, P., P. Sturgis, and S. Purdon (1997). *Can You Hear me Knocking: An Investigation into the Impact of Interviewers on Survey Response Rates.* London: National Centre for Social Research.

Carroll-Scott, A. (2008). *Neighborhood Structural Characteristics, Social Processes, and Child Wellbeing: A Multilevel Study in Two Urban Contexts.* Ph. D. thesis, University of California, Los Angeles.

Carton, A. (2008). Are our paradata always of good quality? In *19th International Workshop on Household Survey Nonresponse*, Ljubljana, Slovenia.

Casas-Cordero, C. (2010). Assessing the quality of interviewer observations of neighborhood characteristics. In *Paper presented at the 2010 International Total Survey Error Workshop*, Stowe, Vermont.

Caughy, M., P. O'Campo, and J. Patterson (2001). A brief observational measure for urban neighborhoods. *Health and Place 7*(3), 225–236.

Caughy, M. O., P. J. O'Campo, and C. Muntaner (2003). When being alone might be better: Neighborhood poverty, social capital, and child mental health. *Social Science and Medicine 57*, 227–237.

Clemente, F. and M. B. Kleiman (1977). Fear of crime in the United States: A multivariate analysis. *Social Forces 56*, 519–31.

Cohen, D., B. Finch, A. Bower, and N. Sastry (2006). Collective efficacy and obesity: The potential influence of social factors on health. *Social Science and Medicine 62*, 769–778.

Cohen, D., K. Mason, A. Bedimo, R. Scribner, V. Basolo, and T. Farley (2003a). Neighborhood physical conditions and health. *American Journal of Public Health 93*(3), 467–471.

Cohen, D., S. Spear, R. Scribner, P. Kissinger, K. Mason, and J. Wildgen (2000). 'Broken Windows' and the risk of gonorrhea. *American Journal of Public Health 90*(2), 230–236.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement 20*, 37–46.

Cook, T. D., S. C. Shagle, and S. M. Degirmencioglu (1997). Capturing social process for testing mediational models of neighborhood effects. In J. Brooks-Gunn, G. Duncan, and J. Aber (Eds.), *Neighborhood Poverty: Vol. I: Context and Consequences for Children*, pp. 94–119. Russell Sage Foundation.

Coulton, C., J. Korbin, T. Chan, and M. Su (2001). Mapping residents perceptions of neighborhood boundaries: A methodological note. *American Journal of Community Psychology 29*(2), 371–383.

Couper, M. P. and R. M. Groves (1996). Social environmental impacts on survey cooperation. *Quality and Quantity 30*, 173–188.

Couper, M. P., E. Singer, and R. A. Kulka (1998). Participation in the 1990 Decennial Census - Politics, privacy, pressures. *American Politics Quarterly 26*(1), 59–80.

Craig, C., R. Brownson, S. Cragg, and A. Dunn (2002). Exploring the effect of the environment on physical activity - A study examining walking to work. *American Journal of Preventive Medicine 23*(2), 36–43.

Crawford, T. (2002). Spatial modeling of village functional territories to support population-environmental linkages. In S. Walsh and K. Crews-Meyer (Eds.), *Linking People, Place and Policy: a GIScience Approach*, pp. 91–112. Norwell, Mass.: Kluwer Academic Publishers.

Curtin, R., S. Presser, and E. Singer (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly 69*(1), 87–98.

Davis, J. A. and T. W. Smith (2005). General Social Surveys, 1972-2006. Technical report, NORC ed. Chicago: National Opinion Research Center, producer, 2005; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut, distributor.

De Leeuw, E. and W. DeHeer (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. Groves, D. Dillman, J. Eltinge, and R. Little (Eds.), *Survey Nonresponse*, New York, pp. 41–54. Wiley.

DeMaio, T. J. (1980). Refusals: Who, Where and Why? *Public Opinion Quarterly 44*, 223–233.

Deming, W. E. (1953). On a probability mechanism to attain an economic balance between the resultant error of response and the bias of nonresponse. *Journal of the American Statistical Association 48*, 743–772.

Dever, J., F. Kreuter, and R. Valliant (2011). *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer, Pre Print.

Deville, J. C. and C. E. Särndal (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association 87*, 376–382.

Diez-Roux, A. V. (2001). Investigating neighborhood and area effects on health. *American Journal of Public Health 91*(11), 1783–1789.

Dunstan, F., N. Weaver, R. Araya, T. Bell, S. Lannon, G. Lewis, J. Patterson, H. Thomas, P. Jones, and S. Palmer (2005). An observation tool to assist with the assessment of urban residential environments. *Journal of Environmental Psychology 25*(3), 293–305.

Durrant, G. B., R. M. Groves, L. Staetsky, and S. Fiona (2010). Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opinion Quarterly 74*(1), 1–36.

Durrant, G. B. and F. Steele (2009). Multilevel modelling of refusal and non-contact in household surveys: evidence from six uk government surveys. *Journal Of The Royal Statistical Society Series A 172*(2), 361–381.

Earls, F. J., J. Brooks-Gunn, S. W. Raudenbush, and R. J. Sampson (1997). Project on human development in chicago neighborhoods: Community survey, 1994-1995. Technical report, Inter-university Consortium for Political and Social Research, Ann Arbor, MI.

Fowler, F. J. (1991). Reducing interviewer-related error through interviewer training, supervision, and other means. In P. Biemer, R. M. Groves, L. Lyberg, N. A. Mathiowetz, and S. Sudman (Eds.), *Measurement Errors in Surveys*. New York: Wiley.

Franck, K. A. (1980). Friends and strangers: The social experience of living in urban and non-urban settings. *Journal of Social Issues 36*(3), 52–71.

Franzini, L., M. O. Caughy, S. M. Nettles, and P. O'Campo (2008). Perceptions of disorder: Contributions of neighborhood characteristics to subjective perceptions of disorder. *Journal of Environmental Psychology 28*(1), 83–93.

Furr-Holden, M. J., J. L. Smart, N. S. I. Pokorni, P. J. Leaf, H. D. Holder, and J. C. Anthony (2008). The nifety method for environmental assessment of neighborhood-level indicators of violence, alcohol, and other drug exposure. *Prevention Science 9*, 245–255.

Furstenberg, F. F. J., T. Cook, J. Eccles, G. Elder, and A. Sameroff (1999). *Managing to Make It: Urban Families and Adolescent Success*. Chicago: Univ. Chicago Press.

Garofalo, J. and J. Laub (1978). The fear of crime: Broadening our perspective. *Victimology 3*, 242–253.

Gauvin, L., L. Richard, C. L. Craig, M. Spivock, M. Riva, M. Forster, S. Laforest, S. Laberge, M.-C. Fournel, H. Gagnon, S. Gagné, and L. Potvin (2005). From walkability to active living potential: An "ecometric" validation study. *American Journal of Preventive Medecine 28*, 126–133.

Gfroerer, J., J. Lessler, and T. Parsley (1997). Studies of nonresponse and measurement error in the national household survey on drug abuse. In L. Harrison and A. Hughes (Eds.), *The Validity of Self-Reported Drug Use: Improving the Accuracy of Survey Estimates*, pp. 273–295. Rockville, MD: National Institute on Drug Abuse.

Goldstein, H. (1987). Multilevel covariance components models. *Biometrika 74*, 430–431.

Goldstein, H. (1995). *Multilevel Statistical Models*. London: Arnold.

Goyder, J., J. Lock, and T. McNair (1992). Urbanization effects on nonresponse: a test within and between cities. *Quality and Quantity 26*(1), 39–49.

Groves, R. (1989). *Survey errors and survey costs*. New York: Wiley.

Groves, R. and M. Couper (1998). *Nonresponse in Household Interview Surveys.* New York: Wiley.

Groves, R. and K. McGonagle (2001). A theory-guided interviewer training protocol regarding survey participation. *Journal of Official Statistics 17*, 249–266.

Groves, R. and E. Peytcheva (2008). The impact of non-response rates on non-response bias: a meta-analysis. *Public Opinion Quarterly 72*, 167–189.

Groves, R., E. Singer, and A. Corning (2000). Leverage-salience theory of survey participation: Description and an illustration. *Public Opinion Quarterly 64*, 299–308.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly 70*(5), 646–675.

Groves, R. M. and S. G. Heeringa (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society, Series A 3*, 439–457.

Guest, A. and B. Lee (1984). How urbanites define their neighborhoods. *Population and Environment 71*(1), 32–56.

Hansen, M. H. and W. N. Hurwitz (1946). The problem of nonresponse in sample surveys. *Journal of the American Statistical Association 41*, 517–529.

Hansen, M. H., W. N. Hurwitz, and M. A. Bershad (1961). Measurement errors in censuses and surveys. *Bulletin of the International Statistical Institute 38*, 359–374.

Hintze, J. M. (2005). Psychometrics of direct observation. *School Psychology Review 34*(4), 507–519.

House, J. S. and S. Wolf (1978). Effects of urban residence on interpersonal trust and helping behavior. *Journal of Personality and Social Psychology 36*(9), 1029–1043.

Hox, J., E. De Leeuw, and I. Kreft (1991). The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (Eds.), *Measurement Errors in Surveys.* New York: John Wiley and Sons, Inc.

Hunter, A. (1985). Private, parochial and public social orders: The problem of crime and incivility in urban communities. In G. Suttles and Z. Mayer (Eds.), *The Challenge of Social Control: Citizenship and Institution Building in Modern Society*, pp. 230–242. Norwood, N.J.: Ablex Publishing.

Jencks, C. and S. Mayer (1990). The social consequences of growing up in a poor neighborhood. In L. Lynn and M. McGeary (Eds.), *Inner-City Poverty in the United States*, Washington, DC, pp. 111–186. National Academy Press.

Johnson, T., Y. Cho, R. Campbell, and A. Holbrook (2006). Using community-level correlates to evaluate nonresponse effects in a telephone survey. *Public Opinion Quarterly 70*(5), 704–719.

Johnson, T., D. O'Rourke, J. Burris, and L. Owens (2002). Culture and survey nonresponse. In R. Groves, D. Dillman, J. Eltinge, and R. Little (Eds.), *Survey Nonresponse*, New York, pp. 55–70. Wiley.

Kalton, G. (1983). Compensating for missing survey data. Technical report, Survey Research Center, University of Michigan, Ann Arbor, Michigan.

Kalton, G. and I. Flores-Cervantes (2003). Weighting methods. *Journal of Official Statistics 19*, 81–97.

Kalton, G. and D. Kasprzyk (1986). The treatment of missing survey data. *Survey Methodology 12*, 1–16.

Kane, E. and L. Macaulay (1993). Interviewer gender and gender attitudes. *Public Opinion Quarterly 57*, 1–28.

Kawachi, I. and L. Berkman (2003). *Neighborhoods and Health.* New York: Oxford University Press Inc.

Kennickell, A. B. (1999). Analysis of nonresponse effects in the 1995 survey of consumer finances. *Journal of Official Statistics 15*(2), 283–303.

Kennickell, A. B. (2003). Reordering the Darkness: Application of Effort and Unit Nonresponse in the Survey of Consumer Finances. In *Proceedings of the American Statistical Association, Section on Survey Research Methods.*

Kish, L. (1962). Studies of interviewer variance for attitudinal variables. *Journal of the American Statistical Association 57*(297), 92–115.

Kish, L. (1965). *Survey Sampling.* New York: Wiley.

Kish, L. (1992). Weighting for unequal pi. *Journal of Official Statistics 8*, 183–200.

Kornhauser, R. (1978). *Social Sources of Delinquency: An Appraisal of Analytical Models.* Chicago, IL: University of Chicago Press.

Kreuter, F., M. Lemay, and C. Casas-Cordero (2007). Using proxy measures of survey outcomes in post-survey adjustments: Examples from the European Social Survey ESS. In *Proceedings of the American Statistical Association, Survey Research Methods Section.*

Kreuter, F., K. Olson, J. Wagner, T. Yan, T. M. Ezzati-Rice, C. Casas-Cordero, M. Lemay, A. Peytchev, R. M. Groves, and T. E. Raghunathan (2010). Using proxy measures and other correlates of survey outcomes to adjust for nonresponse: examples from multiple surveys. *Journal Of The Royal Statistical Society Series A 173*(2), 389–407.

Laraia, B. A., L. Messer, J. S. Kaufman, N. Dole, M. Caughy, P. O'Campo, and D. A. Savitz (2006). Direct observation of neighborhood attributes in an urban area of the U.S. south: characterizing the social context of pregnancy. *International Journal of Health Geographics 5*, 11.

Lee, B. and K. Campbell (1997). Common ground? urban neighborhoods as survey respondents see them. *Social Science Quarterly 78*(4), 922–936.

Lee, B. A., K. E. Campbell, and O. Miller (1991). Racial differences in urban neighboring. *Sociological Forum 6*(3), 525–550.

Lee, G. (1983). Social integration and fear of crime among older persons. *Journal of Gerontology 6*, 745–50.

Lepkowski, J. M. and M. P. Couper (2001). Nonresponse in longitudinal household surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little (Eds.), *Survey Nonresponse*. New York: Wiley and Sons.

Lepkowski, J. M., W. D. Mosher, K. E. Davis, R. M. Groves, and J. Van Hoewyk (2010). The 20062010 National Survey of Family Growth: Sample design and analysis of a continuous survey. Vital Health Statistics Series 2, Number 150, National Center for Health Statistics.

Lessler, J. and W. D. Kalsbeek (1992). *Nonsampling Error in Surveys*. New York: John Wiley and Sons.

Little, R. (1986). Survey nonresponse adjustments for estimates of means. *International Statistical Review 54*, 139–157.

Little, R. and D. Rubin (2002). *Statistical Analysis with Missing Data*. Hoboken: Wiley.

Little, R. and S. Vartivarian (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodology 31*, 161–168.

Little, R. J. and S. Vartivarian (2003). On weighting the rates in non-response weights. *Statistics in Medicine 22*(9), 1589–1599.

Lochner, K., I. Kawachi, R. Brennan, and S. Buka (2003). Social capital and neighborhood mortality rates in Chicago. *Social Science and Medicine 56*, 1797–1805.

Logan, J. R. and O. A. Collver (1983). Residents' perceptions of suburban community differences. *American Sociological Review 48*(3), 428–433.

Lynn, P., P. Clarke, J. Martin, and P. Sturgis (2002). The effects of extended interviewer efforts on nonresponse bias. In R. Groves, D. Dillman, J. Eltinge, and R. Little (Eds.), *Survey Nonresponse*. New York: Wiley.

Maitland, A., C. Casas-Cordero, and F. Kreuter (2009). An evaluation of nonresponse bias using paradata from a health survey. In *Proceedings of the Section on Government Statistics, Joint Statistical Meetings*, Washington, D.C.

Mayer, S. E. and C. Jencks (1989). Growing up in poor neighborhoods: How much does it matter? *Science 243*(4897), 1441–1445.

McCulloch, S., F. Kreuter, and S. Calvano (2010). Interviewer observed vs. reported respondent gender: Implications on measurement error. In *Paper presented at the 2010 Annual Meeting of the American Association for Public Opinion Research*, Chicago, IL.

McGuire, J. (1997). The reliability and validity of a questionnaire describing neighborhood characteristics relevant to families andyoung children living in urban areas. *Journal of Community Psychology 25*, 551–566.

Morenoff, J. (2003). Neighborhood mechanisms and the spatial dynamics of birth weight. *American Journal of Sociology 108*(5), 976–1017.

Mujahid, M. S., A. V. D. Roux, J. D. Morenoff, and T. Raghunathan (2007). Assessing the measurement properties of neighborhood scales: From psychometrics to ecometrics. *American Journal of Epidemiology 165*(8), 858–867.

Newsome, T., W. Walcott, and P. Smith (1998). Urban activity spaces: Illustrations and application of a conceptual model for integrating the time and space dimensions. *Transportation 25*, 357–377.

Nicotera, N. (2007). Measuring neighborhood: A conundrum for human services researchers and practitioners. *American Journal of Community Psychology 40*(1-2), 26–51.

Oh, H. and F. Scheuren (1983). Weighting adjustment for unit nonresponse. In W. Madow, I. Olkin, and D. Rubin (Eds.), *Incomplete Data in Sample Surveys, Volume 2: Theory and Bibliography*, pp. 143–184. New York: Academic Press.

O'Muircheartaigh, C. and P. Campanelli (1998). The relative impact of interviewer effects and sample design effects on survey precision. *Journal of the Royal Statistical Society, Series A 161*(1), 63–77.

O'Muircheartaigh, C. and P. Campanelli (1999). A multilevel exploration of the role of interviewers in survey non-response. *Journal of the Royal Statistical Society, Series A 162*(Part 3), 437–446.

Pebley, A. R. and N. Sastry (2003, February). Neighborhoods, poverty and children's well-being: A review. Working Papers 03-04, RAND Corporation Publications Department.

Perkins, D. D., W. Meek, and R. B. Taylor (1992). The physical environment of street blocks and resident perceptions of crime and disorder: Implications for theory and measurement. *Journal of Environmental Psychology 12*, 21–34.

Peterson, C. E., A. R. Pebley, and N. Sastry (2007). The Los Angeles County Neighborhood Services and Characteristics Database. Technical Report WP-240/12-LAFANS, Labor and Population Program, RAND Corporation.

Peterson, C. E., N. Sastry, and A. R. Pebley (2007). The Los Angeles Family and Neighborhood Survey: Neighborhood Observations Codebook. Technical Report WR-240/13-LAFANS, Labor and Population Program, RAND Corporation.

Peytchev, A., S. Riley, J. Rosen, J. Murphy, and M. Lindblad (2010). Reduction of nonresponse bias through case prioritization. *Survey Research Methods 4*(1), 21–29.

Pfeffermann, D., C. Skinner, D. Holmes, H. Goldstein, and J. Rasbash (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society 60*(1), 23–40.

Pickering, K., R. Thomas, and P. Lynn (2003, July 2003). Testing the shadow sample approach for the english house condition survey. Technical report, Prepared for the Office of the Deputy Prime Minister by the National Centre for Social Research, London.

Pickery, J. and G. Loosveldt (2001). An exploration of question characteristics that mediate interviewer effects on item nonresponse. *Journal of Official Statistics 17*, 337–350.

Pickery, J. and G. Loosveldt (2002, NOV). A multilevel multinomial analysis of interviewer effects on various components of unit nonresponse. *Quality and Quantity 36*(4), 427–437.

Pickery, J. and G. Loosveldt (2004a). A simultaneous analysis of interviewer effects on various data quality indicators with identification of exceptional interviewers. *Journal of Official Statistics 20*(1), 77–89.

Pickery, J. and G. Loosveldt (2004b). A simultaneous analysis of interviewer effects on various data quality indicators with identification of exceptional interviewers. *Journal of Official Statistics 20*(1), 77–89.

Rabe-Hesketh, S. and A. Skrondal (2006). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society, Series A 169*(4), 805–827.

Rabe-Hesketh, S. and A. Skrondal (2008). *Multilevel and Longitudinal Modeling using Stata*. College Station, TX: Stata Press.

Rabe-Hesketh, S., A. Skrondal, and A. Pickles (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *Stata Journal 2*, 121.

266

Rabe-Hesketh, S., A. Skrondal, and A. Pickles (2004). GLLAMM manual. Technical Report 160, Division of Biostatistics, University of California, Berkeley.

Raghunathan, T. E., J. M. Lepkowski, J. Van Hoewyk, and P. Solenberger (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology 27*(1), 85–95.

Raudenbush, S. (2003). The quantitative assessment of neighborhood social environment. In I. Kawachi and L. Berkman (Eds.), *Neighborhoods and Health*, pp. 112–131. New York, NY: Oxford University Press.

Raudenbush, S. and A. Bryk (2002). *Hierarchical Linear Models*. Thousand Oaks: Sage Publications.

Raudenbush, S. and R. J. Sampson (1999). "Ecometrics": Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology 29*, 1–41.

Rindfuss, R., B. Entwisle, S. Walsh, P. Prasartkul, Y. Sawangdee, T. Crawford, and J. Reade (2002). Continuous and discrete: Where they have met in nang rong, thailand. In S. Walsh and K. Crews-Meyer (Eds.), *Linking People, Place and Policy: a GIScience Approach*, pp. 7–38. Norwell, Mass.: Kluwer Academic Publishers.

Rosenbaum, P. and D. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*, 41–55.

Rosner, B. (2000). *Fundamentals of Biostatistics*. Portland, OR: Cengage Learning.

Ross, C. (2000). Neighborhood disadvantage and adult depression. *Journal of Health and Social Behavior 41*(2), 177–187.

Ross, C., J. Reynolds, and K. Geis (2000). The contingent meaning of neighborhood stability for residents' psychological well-being. *American Sociological Review 65*, 581–597.

Sampson, R., J. Morenoff, and F. Earls (1999). Beyond social capital: Spatial dynamics of collective efficacy for children. *American Sociological Review 64*(5), 633–660.

Sampson, R., J. Morenoff, and T. Gannon-Rowley (2002). Assessing "neighborhood effects": Social processes and new directions in research. *Annual Review of Sociology 28*, 443–478.

Sampson, R. and S. Raudenbush (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology 105*(3), 603–651.

Sampson, R. and S. Raudenbush (2004). Seeing disorder: Neighborhood stigma and the social construction of "Broken windows". *Social Psychology Quarterly 67*(4), 319–342.

Sampson, R., S. Raudenbush, and F. Earls (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science 277*, 918–924.

Särndal, C.-E. and S. Lundström (2008). Assessing auxiliary vectors for control of nonresponse bias in the calibration estimator. *Journal of Official Statistics 24*(2), 167–191.

Sastry, N., B. Ghosh-Dastidar, J. Adams, and A. Pebley (2003). The design of a multilevel longitudinal study of children, families and communities: the Los Angeles Family and Neighborhood Study. Technical Report DRU-2400/1-LAFANS, Labor and Population Program, RAND Corporation.

Sastry, N. and A. Pebley (2003). Non-response in the Los Angeles Family and Neighborhood Study. Technical Report DRU-2400/7-LAFANS, Labor and Population Program, RAND Corporation.

Schnell, R. and F. Kreuter (2005). Separating interviewer and sampling-point effects. *Journal of Official Statistics 21*(3), 389–410.

Schouten, B., F. Cobben, and J. Bethlehem (2009). Indicators for the representativeness of survey response. *Survey Methodology 35*(1), 101–113.

Schrapler, J.-P., J. Schupp, and G. G. Wagner (2010). Individual and neighborhood determinants of survey nonresponse. Technical report, SOEP papers on Multidisciplinary Panel Data Research, Deutsches Institut for Wirtschaftsforschung (DIW Berlin).

Schuman, H. and J. Converse (1971). The effects of black and white interviewers on black responses in 1968. *Public Opinion Quarterly 35*, 44–68.

Shaw, C. R. and H. D. McKay (1942). *Juvenile Delinquency in Urban Areas*. Chicago, IL: University of Chicago Press.

Shaw, C. R. and H. D. McKay (1969). *Juvenile Delinquency in Urban Areas: A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities*. Chicago, IL: University of Chicago Press.

Shrout, P. and J. Fleiss (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin 86*(2), 420–428.

Sinibaldi, J. (2010). Measurement error in objective and subjective interviewer observations. In *Paper presented at the 2010 Annual Meeting of the American Association for Public Opinion Research*, Chicago, IL.

Skogan, W. (1990). *Disorder and Decline: Crime and the Spiral of Decay in American Cities.* New York: Free Press.

Smith, T. W. (1983). The hidden 25 percent: An analysis of nonresponse on the 1980 General Social Survey. *Public Opinion Quarterly 87*, 386–404.

Snijders, T. and R. Bosker (1999). *Multilevel Analysis. An introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA: Sage.

StataCorp (2007). *Stata: Release 10. Statistical Software.* College Station, TX: StataCorp LP.

Steeh, C. G. (1981). Trends in nonresponse rates, 1952-1979. *Public Opinion Quarterly 45*, 40–57.

Stoop, I. (2005). *The hunt for the last respondent.* The Hague: Social and Cultural Planning Office.

Suttles, G. D. (1972). *The social construction of communities.* Chicago: University of Chicago Press.

Taylor, R. B. (1997). Order and disorder of streetblocks and neighborhoods: Ecology, microecology and the systemic model of social organization. *Journal of Research in Crime and Delinquency 34*, 113–155.

Taylor, R. B. (2001). *Breaking Away from Broken Windows: Baltimore Evidence and Implications for the Nationwide Fight Against Crime, Grime, Fear and Decline.* New York: Westview Press.

Taylor, R. B., S. D. Gottfredson, and S. Brower (1984). Neighborhood naming as an index of attachment to place. *Population and Environment 7*, 101–111.

Taylor, R. B., B. A. Koons, E. M. Kurtz, J. R. Greene, and D. D. Perkins (1995). Streetblocks with more nonresidential land use have more physical deterioration: Evidence from baltimore and philadelphia. *Urban Affairs Review 31*, 120–136.

Taylor, R. B., S. A. Shumaker, and S. D. Gottfredson (1985). Neighborhood-level links between physical features and local sentiments: Deterioration, fear of crime, and confidence. *Journal of Architecture and Planning Research 2*, 261–275.

Van Goor, H., F. Jansma, and R. Veenstra (2005). Differences in undercoverage and nonresponse between city neighbourhoods in a telephone survey. *Psychological Reports 96*(3, Part 2), 867–878.

Vincenty, T. (1975). Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey Review 22*(176), 88–93.

Weich, S., E. Burton, M. Blanchard, M. Prince, K. Sproston, and B. Erens (2001). Measuring the built environment: Validity of a site survey instrument for use in urban settings. *Health and Place 7*, 283–292.

West, B. (2010). An examination of the quality and utility of interviewer estimates of household characteristics in the national survey of family growth. In *Paper presented at the 2010 Annual Meeting of the American Association for Public Opinion Research*, Chicago, IL.

Wilson, J. Q. and G. Kelling (1982). Broken Windows. *Atlantic Monthly 211*, 29–38.

Wilson, T. C. (1985). Settlement type and interpersonal estrangement: A test of theories of wirth and gans. *Social Forces 64*(139-150), 1.

Wyant, B. R. (2008). Multilevel impacts of perceived incivilities and perceptions of crime risk on fear of crime: Isolating endogenous impacts. *Journal of Research in Crime and Delinquency 45*, 39–64.

Xue, Y., T. Leventhal, J. Brooks-Gunn, and F. Earls (2005). Neighborhood residence and mental health problems of 5- to 11-year-olds. *Archives of General Psychiatry 62*(5), 554–563.

Zenk, S. N., A. J. Schulz, G. Mentz, J. S. House, C. C. Gravlee, P. Y. Miranda, P. Miller, and S. Kannan (2007). Inter-rater and test-retest reliability: Methods and results for the neighborhood observational checklist. *Health and Place 13*(2), 452–465.