The Neighbourhood Location of Source Plasma Clinics in the United States: 1980-1995.

A dissertation prepared in partial fulfilment
of the doctoral degree
from the
Faculty of Graduate Studies
University of Manitoba

Robert C. James
Winnipeg, Manitoba
Canada

© Copyright by Robert Charles James, 2003

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-79854-2

Canada

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION PAGE

The Neighbourhood Location of Source Plasma Clinics in the United States: 1980-1995

BY

Robert Charles James

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

ROBERT CHARLES JAMES ©2003

Permission has been granted to the Library of The University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilm Inc. to publish an abstract of this thesis/practicum.

The author reserves other publication rights, and neither this thesis/practicum nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

Abstract

This dissertation offers an analysis of the neighbourhood placement of commercial plasma clinics within the United States before, during, and after the recognition of HIV contaminating fractionated coagulation products. The dissertation tests an assertion by Chitwood, and parallel statements offered by the US government, that US source plasma clinics were located in "high-risk" areas. This unevaluated assertion has implications for US and international blood policy, and our understanding of the history of HIV contamination of plasma fractions.

The thesis is organized as a set of six independent chapters - only the last of which tests this assertion. Chapter 1 reviews the importance of "place" within social epidemiology; Chapter 2 offers an organizational structure and review of extant social epidemiology literature related to place. The third chapter offers a novel bivariate neighbourhood classification based on social disorganization theory - SECTOR, based on a factor analysis of 14 census-tract measures for all 1980 and 1990 census tracts. The predictive properties of three neighbourhood classification schemes to identify foci of drug use, sales, and availability are provided in Chapter #4. Consistent with social disorganization theory, an analysis of the 1993 National Household Survey on Drug Abuse showed neighbourhood context to be associated with extreme rates of drug selling, "very easy" access to several illicit drugs. No association between social environment and personal drug use was observed.

Chapter #5 advances a novel theoretical model of blood safety which

identifies eight distinct domains which determine blood safety - including the notion of "place". "Place" thus links the social epidemiology material of Chapters #1-4 with the study of plasma clinics found in Chapter #6.

Chapter 6 presents an analysis of US plasma clinic locations over the period 1980-1995 with respect to their neighbourhood social environment, using the SECTOR designation as well as two extant measures - extreme poverty and underclass area designations. Source plasma clinics are found to be 5-8 times more likely to be located in high-risk areas - characterized by active drug markets and drug availability - than would be expected by chance; source plasma clinics continued to be located in these high-risk areas through 1995. Policy implications are discussed.

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

A Thesis submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfilment of the requirement of the degree

of

Doctor of Philosophy

(c) 2003

Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

Acknowledgments

Acknowledgments are an important but necessarily incomplete, reflection of the debts that one incurs in writing a dissertation. I shall not attempt to hide the incompleteness of these thanks by offering what would ultimately prove to be a failed attempt at an exhaustive recitation of these contributions. So, to all who have helped, my sincerest thanks.

General Motors, and The Canadian Institute for Advanced Research are sincerely thanked for a doctoral completion award (1998) and the US National Institute for Drug Abuse is thanked for its financial and logistical support of the linkage study reported in Chapter#4. The now-defunct National Health Research and Development Program provided stipendiary support during the first years of my dissertation.

Four specific debts of gratitude merit specific mention. To James Mosley, my thanks for identifying an interesting question, and for support along the way.

To my committee - Raymond Currie, Margaret Fast, and Cameron Mustard (advisor) - my thanks for your indulgence and guidance. I have learned much during this time from you, not all of which will be found in the contents of this dissertation.

To my children, Hannah, Nicholas, and Molly, and Aidan, your dad can and will come outside to take you to the pool, get the toboggan out and fix your bicycles. Thanks for helping me with my "book". I hope some day you will read

this and find in it some modest compensation for your patience and indulgence.

To Manya, keeper of the family flame, my love burns brighter still.

Dedication

At its core, this dissertation speaks to events which have cost the lives of thousands of hemophiliacs, their partners, their children, and wreaked havoc on those who loved them. This was an epidemic, perhaps like most epidemics, that finds its origins in the interaction of biology and society. For those touched by these events it is my hope that this exploration of some of the social factors will bring about changes in the historiography of these events, and leads to improvements in the policy and technical environments which shape your future.

This dissertation honours the memory of Anna Sadouski who was, for too terribly brief a period, my sister-in-law. The dissertation is dedicated with great thankfulness to my parents, Roy and Marjorie James, who have always encouraged me to read, and read, and read some more. Isn't it amazing where *Curious George* will take you if you wait a while?

Duality Of Interest

On two occasions during the mid 1990s I prepared reports on the location of specific source plasma clinics which were requested by parties involved in civil litigation.

Glossary

Abbreviation	Expansion
CI	Confidence Interval
FDA	Food and Drug Administration, US Government
GAO	General Accounting Office, US Congress
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
NHSDA	National Household Survey on Drug Abuse
NORC	National Opinion Research Center
OMHS	Ontario Mental Health Survey

Table of Contents

Table of Figures

Table of Tables

Introduction

This dissertation applies techniques and insights from social epidemiology to characterize some aspects of the collection of human plasma for therapeutic purposes in the United States over the period 1980-1995. The intention of our work is to inform the historical understanding of the outbreaks of human immunodeficiency virus (HIV), and hepatitis C virus (HCV) amongst recipients of fractionated plasma products. Our particular focus is on the social geography of US source plasma clinics – from which plasma was acquired that was used to make therapeutic products that would be provided to persons with hemophilia in the United States and elsewhere.

The dissertation is focussed on testing the hypothesis, initially proposed by others, that these clinics were located in areas of "high risk". We have broken this assertion down into two temporal periods and explicitly examine two related hypotheses: 1) that US source plasma clinics were located in high risk areas during the early and mid 1980s, and 2) that these clinics remain in these high risk areas into the middle of the 1990s. By "high risk" areas, we are chiefly interested in areas within the United States which are characterized by high rates of illicit drug activity. This focus reflects the established transmissibility of both HIV and HCV through both injection and non-injection related drug practices, and the high prevalence of both HCV and HIV amongst certain types of illicit drug users.[1, 2]

Both direct and indirect links between various forms of drug use and infectious disease transmission are well established, and the chief problem to be

overcome in this dissertation was that there is no list of "high risk areas" within the United States – nor a national reference source where local rates of illicit drug activity, drug sales, or the prevalence of transfusion-transmissible infections, were recorded. Therefore, extant data resources did not permit the efficient testing of the hypothesis that plasma clinics were represented – or over-represented – in these "high risk" areas.

It is here, in developing and validating a national measure of "high risk" areas, that social epidemiology comes to play a key role in the dissertation. The hypotheses that we test here are ultimately social and ecological in nature and the tools needed to test the hypotheses arise from social epidemiology, and more specifically the techniques of classifying social environments. We devise and implement a social environment variable designed to find such areas, and together with two extant measures, test the predictive validity of these three measures to find "high risk" areas using a nationally representative survey of drug use. Only having accomplished this, do we then return to our initial questions regarding the placement of source plasma clinics.

The chapters of this dissertation are organized as follows: Chapter #1 sets out the argument that place is an important notion in epidemiology and highlights social disorganization theory as a promising approach to thinking about place – particularly the "high risk" areas we are interested in. Chapter #2 reviews how social environments have been measured in recent US public health papers and contrasts these methods with the theories of social environments presented in

Chapter #1. Finding a gulf between the theories of social environments, and the actual classification of these environments in the public health literature, the third chapter proposes and implements a novel classification of US census tracts - which we call SECTOR - which is based on social disorganization theory. Chapter #4 offers an analysis of a nationally representative survey of drug misuse in the United States which employs SECTOR and two extant measures of neighbourhood "place".

Chapter #5 represents an important departure from the social epidemiologic focus provided in Chapters 1-4, and offers a conceptual model of the determinants of blood safety. Intended as a comprehensive model of "blood system safety", it identifies eight domains which independently act as determinants of blood system safety; one of those domains is "place" and it is here, in the concept of place as a determinant of blood system safety that the two apparently disparate threads of the dissertation represented by chapters 1 through 4, and chapter 5 come together, allowing us to test in Chapter #6 the hypotheses regarding the preferential distribution of source plasma clinics in high risk areas.

We recognize that the organization of this dissertation covers a range of conceptual and methodological issues, and that the "thread" of the dissertation may be difficult to follow across the chapters. Therefore, each chapter includes a brief "Foreword" which attempts to establish the expectation for the forthcoming chapter, and to place what is to come within the larger structure of the dissertation. Figures and tables are found at the end of each chapter.

References

1. Hagan H, Des Jarlais DC. HIV and HCV infection among injecting drug users. *Mt Sinai J Med.* 2000;67:423-8.

2. Wilson TE, Minkoff H, DeHovitz J, Feldman J, Landesman S. The relationship of cocaine use and human immunodeficiency virus serostatus to incident sexually transmitted diseases among women. *Sex Transm Dis.* 1998;25:70-5.

Chapter #1: A Place for "Place" in Public Health

Foreword

The purpose of this chapter is to articulate why measures of context, and more specifically characterizations of neighbourhoods, are valid and important concepts for epidemiology. After a general review, the chapter focuses attention on the concept of social disorganization as a valuable, but largely unused, lens through which to assess the influence of neighbourhoods on health.

Introduction

Epidemiology is that shared branch of the medical and social sciences which is concerned with the population-based patterns of human health and illness, and the underlying causes of those patterns. Social epidemiology is specifically concerned with the organization of human health and illness with respect to social conditions - where social conditions range from the structure of nation states to the character of neighbourhoods and the qualities of inter-personal relationships. The underlying, but still elusive, goal of this enterprise is to determine if and how social conditions modify human health.

While social epidemiology might find historical roots in Virchow(1) and other 19th century writers, it has had a long period of relative quiescence during much of the twentieth century; however, social epidemiology is enjoying a resurgence of interest, supported by a developing body of quantitative evidence

which suggests that the organization of health and disease cannot be explained adequately by reference only to the distribution of individual-level factors.(2) A report by Haan and colleagues from the Alameda study is often taken as the first indication of a revival of interest in this area(3). In that paper, using nine years mortality data from the Alameda County study, Haan and colleagues observed a relative risk of 1.7 for mortality amongst those with residence in a poverty area - a risk that could not be explained away by multivariate modeling which incorporated characteristics of the individual including "baseline health status, race, income, employment status, access to medical care, health insurance coverage, smoking, alcohol consumption, physical activity, body mass index, sleep patterns, social isolation, marital status, depression, and personal uncertainty." These results, and subsequent studies, inform the recent interest in social environments and health, and prompt the inclusion of social environments within the dominant "health determinants" model informing Canadian health policy.(4, 5)

This dissertation builds on this line of investigation and argumentation- focusing on the relationship of health and illness to small residential social environments - neighbourhoods - one of the most frequently investigated types of social environments explored in social epidemiology research(6). However, this dissertation does not take up the question of social causation directly - the substantive questions under study in this dissertation do not require assessments of causation, nor does the cross-sectional nature of the neighbourhood-encoded

databases used in this dissertation support causal inferences; our interest is in characterizing and categorizing social environments, and testing our ability to organize specific types of health data with respect to social environments. Perhaps the classification strategy and metrics proposed in the dissertation will support future causal studies.

This chapter is organized into three subsequent sections. The next section is intended to argue for the salience of the concept of social environments - or more tersely "place" - within epidemiologic study and to focus attention on residential social environments as one example of the general idea of social environments. Our second goal is to review options for classifying neighbourhood social environments; our final goal is to highlight social disorganization theory as one viable lens through which to consider contemporary population health questions.

Composition and Context

Many commentators have identified and critiqued the individualistic focus of much of contemporary epidemiology(2, 7). Concurrent with these criticisms is a growing body of literature calling for a renewed consideration of social factors in the determination of population health. (3, 8, 9) The observation that social environments may modify the distribution and determinants of health in populations has focused attention on understanding and quantifying the nature of social environments, estimating their putative health impacts, and sorting out complex and confounded potential causal pathways.(9-16)

This analytic tradition is rooted in a distinction between the role of "composition" and context in determining individual level outcomes. Individual compositional variables might include age, sex, and blood pressure, whilst social context might range from estimates of neighbourhood social capital or national estimates of income inequality. Controlling for composition - typically age and sex- are standard tools in the epidemiologic armamentarium, but elucidating the contribution of social context - be that the nature of social and economic structures, the allocation of social value and resources within a society, or related social characteristics - to health has been less thoroughly considered in the epidemiologic literature.

This underlying distinction between context and composition has a long history in sociologic thought - most authors find its origins in Durkheim(17). However, the classic modern exploration of putative contextual effects comes from research in education by Alexander et al. Writing for an audience of educational researchers, Alexander et al propose that "Contextual effects are regarded as contributing to an ecological or organizational context, and not simply as the cumulation or summarization of individual SES levels."(18)(p. 225) Diprete and Forristal reflect this general approach, and have argued that the assertion of a potential for causal effects from social context represents "a defining claim of the sociological discipline" ((19)p. 331). Several definitions of "contextual effects" have been offered, but Blalock's definition enjoys some degree of acceptance and endurance; Blalock defines contextual effects to

represent "an allowance for macro processes that are presumed to have an impact on the individual actor over and above the effects of any individual-level variables that may be operating." (20)(p. 354).

An example of such a "contextual effect" has been identified by Carter, in a helpful paper on teaching contextual effects in sociology courses(21). Using data from the US National Longitudinal Study of the High School Class of 1972, he notes the apparent contextual effect associated with whether the student attended public or catholic secondary schools. Carter shows that when the survey results are aggregated, social class and cognitive achievement are highly correlated - lower social class origins are linked to lower cognitive assessments. Indeed the cognitive classification presented by Carter shows distributions that are almost exactly reversed across the 'class origin' measure. (Table 1.1) This table shows that 72.9% of those with poor social class origins in this survey had low assessed cognitive ability, while 79.9% of those with middle class origins had high assessed cognitive ability.

However, and here is where the contextual argument emerges, the data are also reported stratified by the school context. (Table 1.2) A clear change in the cognitive ability scores is evident when the contextual variable (public or catholic school) is included. While in public schools, the majority (75%) of poor students are reported to continue to have marked lower cognitive abilities, a similar majority of middle class students are reported to have high cognitive abilities. However, this distribution is not seen in catholic schools. In this context, some

75% of students from poor origins, and some 88% of students from middle class origins are rated as having high cognitive abilities.

The exact reasons why "context" counts so much in this analysis is beyond the scope of this precis, as are alternative interpretations of these data; however, in this example, we find a case for searching for contextual effects. Clearly, if a contextual reading of these data is justified, going to a Catholic school appears to confer a remarkable advantage on poor students, and less significant, but still positive, advantages on those children from middle class origins. Such an analysis appears to represent both a point of commendation for Catholic schools, and a critique of public schools; the observations also raise the intriguing possibility that the performances observed in the public schools do not necessarily represent innate, and inexorable, deficiencies in those with lower social class origins.

This simple example shows why contextual effects have received considerable attention by policy analysts and researchers in a variety of fields. Because social factors are (at least presumptively) amenable to social and policy change, this line of argument has been particularly attractive to scholars as they appear to promise a way of identifying modifiable public policy choices and social factors which impact on the life-chances of individuals.

The contextual effects of catholic school enrollment discussed above are consistent not only with Blalock's definitions, but also support Burbank's approach, proposed in a helpful discussion, entitled "How do contextual effects

work?"(20) Reflecting a substantive interest in behaviour, and particularly in voting behavior, Burbank contends that "contextual effects are systematic variations in the behavior of individual associated with variation across geographic settings" Moreover, "the logic of demonstrating a contextual effect seems clear: if the political behavior of individuals with similar attributes varies systematically with changes in the social context, this variation can be taken as evidence that context affects individual behavior. ... A statistically significant contextual coefficient [from a contexual analysis regression] is taken as evidence that properties of the social context are systematically related to individual vote choice (controlling for the effects of whatever other variables have been specified)." (22)(pp. 166-167) Carter's school performance data fit this description.

Testing the potential etiologic effect of context on individual outcomes requires both good measures of individuals and coherent measures of social context(6). It is here - in the classification and coherent measurement of social context - that our interest in social environments overlaps with interests in testing social causation hypotheses. Moreover, because our interest is not focused on assessing causation, we set aside a discussion of the requisite individual level variables which one might include in a causal model. (See Leventhal and Brooks-Gunn for a careful discussion on this point(23).) We focus on the question of how social context - residential social environments in particular - might be conceptualized, and quantified. Hierarchical models are now the statistical

technique of choice for testing causal hypotheses involving contextual data.(24) These statistical models require meaningful measurement properties for the hierarchical or social environment variables - and thus may benefit from the more foundational work presented here.

Clearly, theories of how social environments - "place" in the parlance of social geography - are organized and how they might modify health status should inform how social environments are classified. Having sketched a general concept of social context and contextual effects, and placed the notion of contextual effects in contrast to composition characteristics, we can turn now to consider conceptualizations and classifying residential social environments.

Conceptualizing and Classifying Place

The desire to quantitatively describe and categorize urban environments according to some measure of social context can be traced at least into the 19th century, and to the work of English scholar/reformers Mayhew and Rowntree(25, 26) and more directly to the work of the English reformer Charles Booth(27-30). Davies has identified Booth's creation and use of a multivariate index to differentiate London neighbourhoods by social characteristics as the first application of quantitative strategies to group together similar residential areas. In that index, Booth combined six neighbourhood rates – including poverty, domestic crowding, early marriages, "surplus unmarried", the birth and death rates – to build an overall measure of "social condition" for London neighbourhoods at the end of the 19th century. In these ways, Booth's index has been identified as a

precursor of 20th century neighbourhood classification strategies.

Although measures of neighbourhood social environments can be traced at least back to Booth, well-developed theories of urban environments are not evident before the work of Park, Burgess and the "Chicago school"(31); that school is characterized by its adoption of ecological models adapted from the biological sciences to describe human settlements, and by the use of those models to describe neighbourhoods in particular.(32, 33) This theoretical premise - rooted in evolutionary biology - led to models of urban differentiation which drew concentric circular zones about city centres, and the suggestion that different behaviours were adapted to each of these social "environments". Controversial even at the time(34), these theoretical models would subsequently be replaced by ever more complex theories of neighbourhoods over subsequent decades(2, 35).

An early alternative to the Chicago school approach was a non-concentric model of urban classification proposed by Shevky in the immediate post-World War II period(36, 37). Identifying three sources of variation of importance to urban differentiation – social rank, urbanization and ethnicity - Shevky's work is regarded as fundamental in establishing both a multidimensional conceptualization of urban differentiation, and proposing a multidimensional classification strategy linked to that theoretical perspective. Tyron, a psychiatrist, working concurrently but independent of Shevky, is credited with the introduction of modern multivariate methods into the study of urban differentiation - accomplished in an analysis of Oakland neighbourhoods and using some 33

census-tract level variables(38). Tyron's method involved the construction of relatively crude "indices of similarity" amongst subsets of the selected variables, and the application of cluster analysis against those indices to aggregate individual neighbourhoods into meaningful groups. White's study of 1980-era US neighbourhoods revisits and updates Shevky's residential differentiation model.(39)

More recent discussions of neighbourhood social environments and strategies for classifying neighbourhoods can be divided into two general categories - those which work primarily from census or similar data to craft quantitatively-defined neighbourhood classification schemes, and those which are focused on measuring underlying social process which are held to craft local social environments.

Examples of the former approach - the use of census and census-like data - to build neighbourhood classifications schemes are now typically associated with market segmentation strategies, such as those advanced by commercial entities including Claritas, and CACI. While the exact method used by these commercial firms is a trade secret, CACI's web-site clearly states that their methods of building neighbourhood typologies involve the application of both principal components analysis and clustering[‡]. These commercial agencies appear to employ ever more advanced revisions to the basic methods of factorial ecology,

[‡]http://demographics.caci.com/products/life_seg.htm

well described by Davies(40). In the UK, the ACORN system has been used in several public health publications(41-43) and alternative but non-proprietary systems have been proposed(44). Many of the commercial systems represent the integration of public datasets - notably the census - with transaction-based datasets arising from credit card transactions, and other measures of economic activity.

Given the input data, these typologies offer sets of classifications which are often based around the purchasing patterns and economic potential of specific populations, and these are used extensively in site-selection for locating businesses. However, the utility of these measures for public health is less well established, perhaps a function of the costs of leasing these segmentation systems, their proprietary methods, and the uncertain linkage of these classification systems to the neighbourhood determinants of public health.

The alternative to these - largely economically driven - classification schemes are schemes which attempt to classify areas on the basis of underlying social, rather than economic, patterns. It is here where scholarly attention has been focused. The alternatives are largely oriented around hypothesized latent constructs held to be conceptually distinct from compositional characteristics. Consistent with this approach, Lynch and Kaplan introduce their discussion of neighbourhood classification noting that "extraindividual socioeconomic factors closely related to the physical and social infrastructure of communities are thought to affect health above and beyond individual compositional aspects." This provide the rationale for characterizations of neighbourhoods; they present a non-

exhaustive precis of nine such "area-based" measures from the international literature(45).

Kreiger, Williams and Moss, in their review of measures of social class, suggest that "neighborhood-based measures of social class and other aspects of socioeconomic position merit greater use in public health research and surveillance."(46) This review also offer additional observations on neighbourhood measures of social class. In particular they note the benefits (finer spatial resolution) and limitations (suppression of socioeconomic data for these spatial aggregates because of limited population) of block-group measures, and the inadequacy of zip codes for social epidemiologic purposes. This review does not offer dominant theoretical perspectives or quantitative strategies for characterizing neighbourhoods, but implicitly contrasts several US indicators of neighbourhood status - which are primarily univariate in their construction, with three leading UK indicators which are uniformly multivariate.

Additional measures have also been published. Notable amongst these latent constructs are Skogan's concept of *neighbourhood disorder*(47), Massey and Denton's emphasis on *racial segregation*(48-57), Wilson's emphasis on *concentrated disadvantage*(58, 59), and the highly influential work of Coleman and Puttman with respect to *social capital*(60, 61).

Skogan has advanced the concept of neighborhood disorder, defined as "direct, behavioural evidence of disorganization"(47)(p.21). Ross and Mirowsky, have provided a useful review of some 12 operationalizations of this concept used

in various US locations.(62) Variables included in these operationizations necessarily vary but typically include measures intended to characterize the physical decay of neighbourhoods, and social disorder such as drug use and crimes against persons, noise and graffiti.

It is important, however, to observe that the emphasis in Skogan's work on the "evidence of disorganization" poses a measurement challenge - a limitation evidence in the published implementations reviewed by Ross and Mirowsky. Each of these measures requires the use of data which cannot be obtained from the census or other key small-area data sources, and therefore requires direct observation. As direct observation at a national level would likely only be possible within a census-like exercise, none of the extant measures of "neighbourhood disorder" present a mechanism to support national analyses. This observation does not diminish the effectiveness of measures of neighbourhood disorder as useful descriptors of this phenomenon for specific locations, but without a ready means to operationalize these measures for multiple areas, replication is difficult and the generalizability of findings is difficult

Collective efficacy is a relatively-recently defined concept, proposed by Sampson, *et al* in a 1997 *Science* paper(63). Collective efficacy was defined to reflect the ability of neighborhoods to act together to achieve collective goals. There is clear resonance between this concept and the work of Shaw and McKay. Higher levels of collective efficacy were held to "inhibit the occurrence of personal violence." In the inaugural studies, collective efficacy was constructed

from responses to several questions obtained by survey interviews amongst residents in various Chicago neighbourhoods. Having identified an association between collective efficacy and victimization rates, Sampson *et al* highlight the fact that, both conceptually and empirically, collective efficacy is independent of classification systems which arise from the analysis of census data.

While collective efficacy has strong parallels to the notion of social disorganization found in the work of Shaw and McKay - it attempts to *directly* measure the intention to generate informal social activity, and not to assume that this intention is uniform with particular strata of low economic status, ethnic heterogeneity and residential mobility - the traditional domains arising from Shaw and McKay's social disorganization model.

"Social capital" is an increasingly common construct used to capture notions of civic trust and reciprocal social relationships. Linked to work by Putnam(61), Coleman(60) and others, it was initially conceptualized to be an attribute of large social agencies such as nation states. But interest in this concept has now broadened considerably and authors have proposed measuring social capital for smaller social areas, including neighbourhoods(64).

A final example of social environment constructs is "social disorganization" - a term associated with the scholarship of Chicago School sociologists Shaw and McKay. (However, as Shoemaker has observed, the term "social disorganization" is not specifically used by these authors, nor are the now-canonized central domains of social disorganization specifically articulated by

them.(65)) Arising from Park's notion of concentric development, Shaw and McKay conducted studies related to the organization of juvenile delinquency rates across social environments.(31) Key work by these authors, notably *Delinquency Areas*(66), and their classic *Juvenile Delinquency and Urban Areas*(67), were viewed by contemporaries as supportive of Park's models of urban ecology; however, the lasting impact of their work was a sociological explanation for juvenile delinquency. Shaw and McKay's notion that neighbourhood context might alter human behaviour and outcomes represented a natural extension of Durkheim's assertion of the influence of social facts, and sits within the classic sociologic tradition that attributes of collectives modify the behaviours and outcomes of individuals.(17)

Their work is now widely associated with the "social disorganization" school, and is now taken to imply an emphasis on the social consequences of low economic resources, ethnic heterogeneity, and increased mobility. This choice of emphasis arises indirectly from their writings. For example, they argue:

> "Those among the foreign born and amongst the recent immigrants who from 1927 to 1933 lived in physically adequate residential areas of higher economic status displayed low rates of delinquents, while, conversely, those among the native born and among the older immigrants who in that period occupied physically deteriorated areas of low economic status displayed high rates of delinquents. Negroes living in the most deteriorated and

> disorganized portions of the Negro community possessed the highest Negro rates of delinquents, just as whites living in comparable white areas showed the highest white rates." (pp. 160-161)

Lower economic status, higher rates of ethnic heterogeneity and higher rates of residential turnover were all correlated with higher rates of juvenile delinquency by bivariate analysis. For Shaw and McKay, these three neighbourhood factors were taken to erode the ability of neighbourhoods to achieve *informal social control*, largely over unsupervised youth gangs - which were seen as the primary agents of juvenile delinquency.

Shaw and McKay's interest in social disorganization spawned an extensive period of neighbourhood research, and was an influence to many of the neighbourhood studies of the 1950-1970s. (See Shoemaker for review(65).) However, interest in this approach waned over the years, perhaps in part because of the lack of a clear theoretical specification in the initial work; Sampson and Groves (68) have argued that the neglect of this theoretical approach is in part a function of the lack of measures of neighbourhood context available to sociologically-inclined researchers in the key research datasets of this period.

The publication of a paper by Sampson and Grove heralded a resurgence of interest in social disorganization theory(68). This paper was specifically introduced as a formal test of the Shaw and McKay model - using the British Crime Survey. The authors held this to be the first formal test of the Shaw and

McKay hypothesis. Sampson and Grove's specifically remark on why it was over forty years after the work of Shaw and McKay before a published formal test had been undertaken. They write: "the crux of the problem is that previous macro-level research in crime and delinquency has relied primarily on census data that rarely provides measures for the variables hypothesized to mediate the relationship between community structure and crime." (p. 775) Their paper thus reflects not just a recovery of a particular theory about social areas, but also a revived interest in the causal contribution of neighbourhood context to criminality. The paper also suggested strategies for handling this sort of contextual inquiry in the quantitative analysis of social survey data.

In the paper, the authors draw out the distinction between the structural determinants of social disorganization, and "intervening dimensions of social disorganization". The structural determinants - low economic status, racial heterogeneity and residential mobility - are taken to erode the "intervening dimension" which include the ability of communities to supervise and control teenage peer groups, local friendship groups, and rates of participation in formal and voluntary organization. Low economic status, higher racial heterogeneity and increased residential mobility are seen to provide structural impediments which modify the intervening dimension of social disorganization. Sampson and colleagues more recent work in collective efficacy can be seen as an attempt to speak directly to those structural determinants.(63, 64) (It is important to note that the later definition of collective efficacy by Sampson *et al*(63), can be understood

as an attempt to define and measure that intervening dimension - distinct from its structural determinants.)

The field of cognitive psychology has offered important insights into what Sampson has called the "intervening" dimensions of neighborhoods. This body of scholarship has explored personal understandings of neighborhoods and other sorts of communities, and the inter-relationships of these contexts and life chances. A broad literature, much is found in the *Journal of Community Psychology,* or the *American Journal of Clinical Psychology.* At its core, the literature seeks to elucidate the concepts and constructs necessary to characterize the meaning of social and spatial context, and tests measures of these concepts for validity and statistical association with outcomes.

Writing in the *Journal of Community Psychology,* O'Brien-Caughy, *et al* have provided a review of neighborhoods, families and children, from a policy point of view(69). Particular attention is drawn to the work of O'Campo et al, in Baltimore in this regard. In an important *American Journal of Public Health* paper, O'Campo and colleagues used a combination of census (or census like) data, and administrative data to build neighborhood classification, and to study the relationships to low birthweight(70). Five contextual measures (four as interaction terms only) modified the outcome, together with four person-oriented measures. Interested readers are also referred to a recent commentary by O'Campo, and to a review offered by Coulton(71).

While promising, most, if not all of the approaches used or reviewed in these papers employ administrative data, surveys and/or interview methods to elucidate qualities of neighborhoods that have salience for residence. These survey, interview or administrative databases are not available on a national basis. For example, this approach is found in the work of Coulton and colleagues at Case-Western Reserve University in Cleveland. Building on a diverse dataset of administrative and survey data focused on the greater Cleveland area – Coulton has identified five domains important to neighbourhood measurement: Neighborhood Identity and Pride, Neighborhood Economic Opportunity, Safety and Security, Access to Institutions and Services, Family and Youth Development. These concepts, together with associated constructs and operationalized measures speak to the diversity of influences involved in the psychological relationship of persons to their residential environments,[‡] but most if not al require on-the-ground measurement.

While contributions from cognitive psychology are crucial to elucidating the psychological components that inform the relationship of context to health it is regrettably not currently possible to operationalize these measures at a national level. Sadly these contributions cannot be integrated into the pragmatic needs of this thesis.

The extant literature with respect to social environments presents a series of theoretical and practical options to the analyst. Which metrics should be chosen,

[‡] http://povertycenter.cwru.edu/research/docs_ct/Five_Aspects_Indicators_and_Measures.pdf

and why, will, of course vary from study to study, however, we argue here that a measure of social disorganization merits exploration and quantitative examination as a candidate means for describing social environments. In our view, social disorganization represents a promising theoretical and empirical candidate for analysis. First, with over half a century of sociological findings that low economic position, race/ethnic heterogeneity, and residential turnover are key factors which organize social outcomes, there is a long quantitative history to support the strategy. Second, most potential constructions of such a measure would vary across neighbourhoods, thereby allowing it to serve as a proxy for neighbourhood social environments. Finally, several authors have used census data to operationalize models of social disorganization, suggesting that a national database of social disorganization scores might be possible.

Racial and Ethnic Heterogeneity in Social Disorganization Theory

Before proceeding it is important to reflect briefly on the inclusion of ethnicity or "race" as a salient structural determinant of social disorganization. The inclusion of some measure of race or ethnicity in models of social disorganization rises directly from Shaw and McKay's analyses, where delinquency rates were stratified into white and black rates. Diversity in the underlying local composition of race and ethnic groups, argues Sampson and Grove, was "thought [by Shaw and McKay] to frustrate the ability of residents to achieve consensus" and thus inhibit collective social action(p. 781).

While such a logical explanation of the corrosive effect of ethnic

heterogeneity has an intuitive appeal - different languages, different social customs seem likely to lead to reduced communication amongst neighbourhoods - the introduction of race/ethnic heterogeneity raises critical questions of interpretation and measurement. Yet, Massey and Denton have shown that patterns of social and racial segregation in Shaw and McKay's era were less pronounced than the residential segregation patterns which emerged after the Second World War, and certainly the patterns of dense racial segregation in US communities in the 1960s and more recently. Massey and Denton have argued that it is racial homogeneity (and the segregation of economic and social resources) within the general umbrella of racial segregation, that represents the greatest threat to social order and opportunity.

The contrasting opinions of Shaw & McKay and Massey & Denton leave the analyst of contextual data with a confusion choice - racial heterogeneity is classically held to erode community social integrity, but so too is racial homogeneity. Moreover, both phenomena appear to be measured by similar sorts of metrics.(50) This would make it difficult to model a relationship between racial heterogeneity, racial homogeneity, and subsequent outcomes and, interpret regression coefficients arising from these race/ethnic heterogeneity variables. Therefore, we have chosen to leave aside considerations of race/ethnicity within our interpretation of social disorganization.

Conclusions

Interest in the social causation of individual outcomes can be traced at least

to Durkheim, but questions of if and how social causation exerts its influences remain open for debate and discussion. The revivial of context into debates on the etiology of health states raises both theoretical questions - "What are social environments?" and classification questions - "How can we measure and categorize social environments into meaningful groups?" With respect to residential neighbourhoods, the tradition over a century of study is that neighbourhoods are best conceptualized as determined by multiple characteristics, including urbanization, economic and social resources, racial and ethnic homogeneity, and life-stage. Attempts to elucidate the relationship of residential social environments to health status should remain attentive to the multivariate conceptualization of these contexts which can be found in the urban ecology literature.

Tables

Table 1.1 Cognitive Ability, stratified by Social Class Origins, as reported by Carter (1972)

Social Class Origins	**Cognitive Ability**	
	Low	High
Poor	72.9%	27.1%
Middle Class	20.1%	79.2%

Table 1.2 Contextual Effects: Cognitive Ability, stratified by social class origins and school setting, as reported by Carter (1972).

Social Class Origins and Educational Setting	Cognitive Ability	
	Low	High
Catholic Schools		
• Poor	24.5%	75.5%
• Middle Class	11.5%	88.5%
Public Schools		
• Poor	75.1%	24.9%
• Middle Class	20.9%	79.1%

References

1. Virchow RLK, Rather LJ. Collected essays on public health and epidemiology. Canton, MA: Science History Publications; 1985.

2. Yen IH, Syme SL. The social environment and health: a discussion of the epidemiologic literature. Annual Rev Public Health 1999;20:287-308.

3. Haan M, Kaplan GA, Camacho T. Poverty and health. Prospective evidence from the Alameda County Study. Am J Epidemiol 1987;125(6):989-98.

4. Evans RG, Stoddart GL. Producing health, consuming health care. Soc Sci Med 1990;31(12):1347-63.

5. Evans RG, Barer ML, Marmor TR. Why are some people healthy and others not? : the determinants of health of populations. New York: A. de Gruyter; 1994.

6. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. J Epidemiol Community Health 2001;55(2):111-22.

7. Krieger N. Epidemiology and the Web of Causation: Has anyone seen the spider? Soc Sci & Med 1994;39(7):887-903.

8. Coughlin SS, Beauchamp TL. Ethics, scientific validity, and the design of epidemiologic studies. Epidemiology 1992;3(4):343-7.

9. Jones K, Duncan D. Individuals and their ecologies: analysing the geography of chronic illness within a multilevel modelling framework. Health and

Place 1995;1(1):27-40.

10. Marmot MG. Improvement of social environment to improve health. Lancet 1998;351(9095):57-60.

11. Diez-Roux AV, Nieto FJ, Tyroler HA, Crum LD, Szklo M. Social inequalities and atherosclerosis. The atherosclerosis risk in communities study. Am J Epidemiol 1995;141(10):960-72.

12. Diez-Roux AV, Nieto FJ, Muntaner C, Tyroler HA, Comstock GW, Shahar E, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. Am J Epidemiol 1997;146(1):48-63.

13. Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. Am J Public Health 1998;88(2):216-22.

14. Duncan SC, Duncan TE, Biglan A, Ary D. Contributions of the social context to the development of adolescent substance use: a multivariate latent growth modeling approach. Drug Alcohol Depend 1998;50(1):57-71.

15. Brooks-Gunn J, Duncan GJ, Aber JL. Neighborhood poverty. New York: Russell Sage Foundation; 1997.

16. Duncan C, Jones K, Moon G. Psychiatric morbidity: a multilevel approach to regional variations in the UK. J Epidemiol Community Health 1995;49(3):290-5.

17. Durkheim E. Suicide : a study in sociology. 1st Free Press paperback -- ed. Glencoe, Ill.: Free Press; 1966.

18. Alexander KL, Fennessey J, McDill EL, D'Amico R. School SES Influences-Composition or Context? Sociology of Education 1979;52(4):222-237.

19. DiPrete T, Forristal J. Multilevel Models: Methods and Substance. Annal Review of Sociology 1994;20:331-357.

20. Blalock H. Contextual Effects Models: Theoretical and Methodological issues. In: Annual Review of Sociology. Palo Alto: Annual Reviews; 1984. p. 353-374.

21. Carter G. Teaching Students Contextual Effects. Teaching Sociology 1991;19:526-531.

22. Burbank M. How do contextual effects work? Developing a theoretical model. In: Eagles M, editor. Spatial and Contextual Models in Political Research. London: Taylor and Francis; 1995. p. 165-178.

23. Leventhal T, Brooks-Gunn J. The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. Psychol Bull 2000;126(2):309-37.

24. Goldstein H. Multilevel Models in Educational and Social Research. London: Charls Griffin & Co; 1987.

25. Rowntree BS. Poverty: a study of town life; 1901.

26. Mayhew H. London Labour and the London Poor; 1862.

27. Dorling D, Mitchell R, Shaw M, Orford S, Smith GD. The Ghost of Christmas Past: health effects of poverty in London in 1896 and 1991. British Medical Journal 2000;321(7276):1547-1551.

28. Davies W. Charles Booth and the Measurement of Urban Social Structure. Area 1978:290-296.

29. Booth C. Life and labour of the people in London. London: Macmillan; 1902.

30. Booth C. Presidential Address: Life and labour of the people of London. J. Royal Statistical Society 1893;55(4):557-91.

31. Park R, Burgess E. The City. Chicago: The University of Chicago Press; 1925.

32. Dotter DL. Beyond the Chicago School: Deviance and Social Conflict. 1978.

33. Snodgrass JON. Clifford R. Shaw and Henry D. Mckay: Chicago Criminologists. British Journal of Criminology 1976;16(1):1-19.

34. Quinn JA. The Burgess Zonal Hypothesis and its Critics. American Sociological Review 1940;5(2):210-218.

35. Davies W. Urban Social Structure: A Multivariate Structural Analysis of Cardiff and its Region. Cardiff: University of Wales Press; 1982.

36. Shevky E, Williams M, John Randolph Haynes and Dora Haynes Foundation. The social areas of Los Angeles, analysis and typology. Berkeley,: Pub. for the John Randolph Haynes and Dora Haynes Foundation by the Univ. of California Press; 1949.

37. Shevky E, Bell W. Social Area Analysis: Theory, Illustrative Application and Computational Procedures. Westport, Conn: Greenwood Press;

1955.

38. Tyron R. Identification of Social Area By Cluster Analysis: A General Methods with an Application to the San Francisco Bay Area. In: Brunwick E, Rosenzweig M, Tryon R, Tuddenham R, editors. University of California Publications in Psychology. Berkeley, CA: University of California Press; 1955.

39. White MJ, National Committee for Research on the 1980 Census. American neighborhoods and residential differentiation. New York: Russell Sage Foundation for the National Committee for Research on the 1980 Census; 1987.

40. Davies W. Factorial Ecology. Aldershot: Gower; 1984.

41. Elley KM, Langford JW. The use of a classification of residential neighbourhoods (ACORN) to demonstrate differences in dental health of children resident within the south Birmingham health district and of different socio-economic backgrounds. Community Dent Health 1993;10(2):131-8.

42. Whittle JG, Davies KW. A classification of residential neighbourhoods (ACORN) in relation to dental health and dental health behaviours. Community Dent Health 1992;9(3):217-24.

43. Bedi R, Quarrell I, Kippen A. The dental health of 10-year-old children attending multi-racial schools in Greater Glasgow. Br Dent J 1991;170(5):182-5.

44. Charlton M, Openshaw S, Wymer C. Some new classifications of census enumeration districts in Britain: a poor man's ACORN. Journal of

Economic and Social Measurement 1985;13:69-96.

45. Lynch J, G K. Socioeconomic Position. In: Berkman LF, Kawachi I, editors. Social Epidemiology. Oxford ; New York: Oxford University Press; 2000. p. 13-35.

46. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. Annu Rev Public Health 1997;18:341-78.

47. Skogan WG. Disorder and decline: crime and the spiral of decay in American neighborhoods. Berkeley: University of California Press; 1992.

48. Massey D, Cortese C, Falk R, Cohen J. On the Measurement of Segregation as a Random Variable. Demography 1978;16:553-563.

49. Massey D-S. Social Class and Ethnic Segregation: A Reconsideration of Methods and Conclusions. 1981.

50. Massey D, Denton N. The Dimensions of Residential Segregation. Social Forces 1988;67(2):281-313.

51. Massey D-S. American Apartheid: Segregation and the Making of the Underclass. 1990.

52. Massey D-S, Fong E. Segregation and Neighborhood Quality: Blacks, Hispanics, and Asians in the San Francisco Metropolitan Area. 1990.

53. Massey D-S, Eggers M-L. The Ecology of Inequality: Minorities and the Concentration of Poverty, 1970-1980. 1990.

54. Massey DS. American apartheid: segregation and the making of the

underclass. American Journal of Sociology 1990;96:329-57.

55. Massey D-S, Gross A-B, Eggers M-L. Segregation, the Concentration of Poverty, and the Life Chances of Individuals. 1991.

56. Massey D-S, Eggers M-L, Denton N-A. Disentangling the Causes of Concentrated Urban Poverty. 1994.

57. Massey D-S, Gross A-B, Shibuya K. Migration, Segregation, and the Geographic Concentration of Poverty. 1994.

58. Wilson WJ. Studying Inner-City Social Dislocations: The Challenge of Public Agenda Research: 1990 Presidential Address. American Sociological Review 1991;56(1):1-14.

59. Wilson WJ. The truly disadvantaged: The Inner City, The Underclass, and Public Policy. Chicago: University of Chicago Press; 1987.

60. Coleman J. Social Capital in the Creation of Human Capital. "American Journal of Sociology 1988;94:S95-S120.

61. Putnam RD, Leonardi R, Nanetti R. Making democracy work: civic traditions in modern Italy. Princeton, N.J.: Princeton University Press; 1993.

62. Ross R, Mirowsky J. Disorder and Decay: the Concept and Measurement of Perceived Neighbourhood Disorder. Urban Affairs Review 1999;34(3):4120432.

63. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 1997;277(5328):918-24.

64. Sampson R-J, Morenoff J-D, Earls F. Beyond Social Capital:

Spatial Dynamics of Collective Efficacy for Children. American Sociological Review 1999;64(5):633-660.

65. Shoemaker DJ. Theories of delinquency: an examination of explanations of delinquent behavior. 3rd ed. New York: Oxford University Press; 1996.

66. Shaw CR, Zorbaugh FM, McKay HD, Cottrell LS. Delinquency areas. Chicago, Ill.,: The University of Chicago press; 1929.

67. Shaw, McKay. Juvenile Delinquency and Urban Areas. Chicago: University of Chicago Press; 1942.

68. Sampson RJ, Groves WB. Community Structure and Crime: Testing Social-Disorganization Theory. American Journal of Sociology 1989;94(4):774-802.

Chapter #2: Classifying "Place"

Introduction

In the previous chapter we outlined a rationale for including social environments within epidemiologic study, and reviewed a century-long scholarly tradition of neighbourhood classification, including strategies focussed on developing classification models for business purposes and more recent attempts to characterize underlying processes which modify neighbourhood social environments - such as social capital and social disorganization. In this chapter we identify six distinct quantitative strategies for characterizing neighbourhood social environments using census data, and highlight four strategies which have appeared in recent US public health journals.

Our primary purpose is to organize extant quantitative strategies for classifying neighbourhoods, and to contrast these strategies with the theoretical traditions of urban sociology which we have briefly reviewed above. The thesis advanced in this chapter, simply put, is that there is a critical gap between the conceptualizations of neighbourhoods as offered in Chapter #1 - which are multivariate, and multidimensional - and the ways in which neighbourhoods are actually classified in quantitative studies of public health phenomenon. It is worth repeating that our focus is on the development of measures of small residential

social environments - neighbourhoods - that can be built from national census data, and which can be used to provide measures of social context for quantitative studies. We recognize that there may be additional value in obtaining qualitative impressions from respondents regarding their neighbourhood, but this is beyond the scope of this chapter.

Before proceeding, it is also important to describe what this chapter offers, and what it does not offer. This chapter identifies six general strategies for organizing social environment variables, and highlights recent papers which are good examples of the applications of those strategies. However, this chapter does not represent a systematic and quantitative assessment of social environment variables as found in the US public health literature. While we believe that our model is comprehensive classification of extant techniques, what follows is not an exhaustive list of papers which have used these techniques.

A Model for Organizing Neighbourhood Classifications Strategies

Neighbourhood characterization are intended to stratify the range of possible neighbourhood types into meaningful aggregates. The strategies used to undertake this classificaiton can be functionally classified with respect to their inputs and their outputs. A functional description highlights "how" neighbourhood classifications are conducted, and focuses attention on the sorts of inputs used, and how stratification and aggregation are accomplished. However, this strategy may obscure important differences in the underlying assumptions and conceptual

models which might share a common technique.

Figure 2.1 identifies six distinct strategies for classifying social environments - and presents those options in a simple tabular form; the strategies are labelled as "A" through "F" Three strategies found in the first row of the table, (Strategies A-C) employ only a single variable as the input to the classification strategy, while the second row (Strategies D-F) represent strategies which use multiple input variables. The first column (Strategies A, D) identifies strategies which do not classify or group neighbourhoods together in any meaningful way, but rather let one (Strategy A) or more input variables (Strategy D) - act without summarization to describe the neighbourhood. The second column (Strategies B & E) include strategies that use their input variables to generate a binary classification of neighbourhoods. The third column (Strategies C & F) reflect strategies which generate polytomous classification schemes; these schemes involve some degree of grouping or clustering. It is important to highlight that this simple classification scheme is a "functional" classification - how neighbourhoods are characterized - it does not focus on differences in the selection of variables across measures but within a strategy.

We identify key papers which use these approaches as found in the dominant US public health journals, including the *American Journal of Public Health*, and *The American Journal of Epidemiology, Science* and *Annals of Epidemiology* and discuss the advantages and disadvantage associated with each technique.

Classification Strategies A&C: Binary outputs

Strategies A & C both employ only a single variable to characterize social environments - strategy A offers no summarization or grouping of neighbourhoods around this variable, whilst strategy C offers a polytomous grouping. An example of strategy A might involve the exclusive use of one continuous variable - such as census-tract-specific household poverty rates - to characterize neighbourhoods; the application of strategy C would involve stratifying a single variable with reference to either an empirical distribution or to an a priori distribution - such as quartile bands of increasing poverty rate.‡ We have not yet been able to find a recent application of Strategy A - the use of a single continuous variable to describe neighbourhoods - in the US public health literature. Wilson employs Strategy C - highlighting three traditional poverty thresholds: 20%, 30% and 40% - and demonstrates a variety of graduated relationships of these poverty levels to social outcomes.(1) Both of these strategies(A & C) and strategy B share the attribute that they rely on a single output variable to capture neighbourhood social environments - an attribute that we discuss and critique more extensively in our introduction to Strategy B.

Classification Strategy B: A Singular input yielding binary classifications

Strategy B represents an approach to incorporating neighbourhood social environments into epidemiologic studies. While in principle the approach could

---

‡ It should be noted that the notion of "extreme poverty areas" as advanced by the Census Bureau, is an example of strategy C, as these areas are a subset of the larger set of all "poverty areas" neighbourhoods. However, we are unaware of extensive use of this graduated approach to neighbourhood classification.

use any variable the strategy appears to be typically associated with the use of poverty measures, and more specifically is limited to a single poverty construct fashioned by the US Census Bureau:"poverty areas". This particular construct involves a classification of neighbourhoods according to the concept of "poverty" and the determination of "poverty-areas" and/or "extreme poverty areas." (See Littman for a discussion of the history of the "poverty area concept"(2))

The "poverty areas" approach begins with the assessment of household income for each household as part of the census, and after reference to pre-established poverty-thresholds, yields a binary determination of household poverty. This variable is then aggregated across households within census tracts[‡] to estimate the proportion of households living in poverty within each census tract. Census tracts where the household poverty rate is higher than 20% are defined as "poverty areas."(3) The US census bureau also defines"extreme poverty areas" using the same variable (household income) and the same unit of aggregation (census tract) but with a 40% cutoff. Thus, extreme poverty areas are a subset of poverty-areas.[‡‡]

A recent application of this approach can be found in Cunradi *et al*, reviewing intimate partner violence data in the United States. Using the 20% poverty threshold cut-off, Cunradi sought to redress a perceived absence on

---

‡ Census tracts as defined by the US census are intended to be relatively small, socially homogeneous, and geographically stable residential areas of mean population 4,000.

‡‡ While household level income estimates are acquired for the census, these estimates are not made available in the usual public use files, because of confidentiality concerns.

contextual variables within the partner-abuse literature, proposing that this poverty-related binary variable would provide insight into "the role of socioenvironmental or ecological factors."(4)(p. 297.)

The officially sanctioned 20% and 40% cut-offs for poverty areas measurement have been used by academics both in public health, and those interested in other aspects of urban social problems.(5)‡ The general strategy of using a single input variable to characterize neighbourhoods and then dichotomizing the resulting distribution, as seen in the "poverty areas" approach, has a singular advantage and related problems. A singular advantage is its apparent clarity of definition. However, this clarity hides important ambiguity about the determination of household income, and the determination of associated poverty-cutoffs(8). Moreover, the poverty-area measure focuses on only a single aspect of the concept of economic resources -- income -- and does not measure other aspects of economic security, such as assets, unusual household expenditures, or extra-household resources that might be available. Any classification scheme which is rooted in a single variable is extremely vulnerable to limitations in the validity, reliability, and measurement of that variable.

While recognizing that poverty and concentrations of poverty represent a

---

‡ It is important to note that there is an alternative poverty-area construct which appears occasionally in the US public health literature, most notably in the reports from the longitudinal Alameda County Study(6), and the US National Health and Nutrition Examination Survey (1971-1974)(7). The neighbourhood "poverty area" definitions used for these studies are also binary classifications of neighbourhoods, but are based on multiple criteria, which include social, economic and geographic considerations. The continued use of this alternative definition of poverty areas, in contrast to the 20%, or 40% poverty-oriented measure canonized by the US census, reflects the initial adoption of these strategies into the initial sampling strategies for these studies, and thus necessitates their retention and continued use in all future analyses. Authors using these measures do not assert any theoretical or empirical advantage of their poverty area measure over the Census Bureau measure.

potent challenge to individuals, households and neighbourhoods, it does not follow that a binary measure based on poverty provides insight into a meaningful diversity of social environments. Clearly, any bifurcation of social environments hides from examination variations which occur *within* each group. The dominant use of a single income-based measure also implies that poverty "single-handedly" either drives, or at least describes, the range of neighbourhood social differentiation. In general, we suggest that the use of a dichotomized variable arising from a single measure will have limited promise as a general measure of neighbourhood social environments, because the leading theoretical perspectives on neighbourhoods neither theorize, nor find evidence of, a bifurcated social universe.

Classification Strategy D: Multiple Inputs yielding a continuous outcome

Neighbourhood classification strategy D involves the use of more than one continuous variable estimated at the neighbourhood level to describe neighbourhood social environments. With this strategy several continuous variables are acquired - typically from the census - and are used simultaneously to provide what might be characterized as a "gestalt" description of neighbourhood social environments. The key attribute of this strategy is that more than one neighbourhood measure is provided as input to some analytic process, for example linear regression, but that no classification or partitioning of neighbourhoods into groups or clusters of greater internal homogeneity is

undertaken.

For example, Roberts employed this approach in a study of low-birth weight in Chicago, after identifying a set of measures of neighbourhood social environment which "were suggested by the diversity of social science literature." (9) Roberts included nine distinct measures of community characteristics as his measures of neighbourhood social environments: a composite score which he denotes as socioeconomic status[‡], economic hardship, the percentage of young residents, the proportion of old residents, residential stability, the proportion of African-American residents, the proportion of Hispanic residents, median rent, and crowded housing rate,. He then applied these measures of community characteristics as independent variables in a least-squares regression analyses of low birth-weight rates in Chicago neighbourhoods.

Similarly, Wasserman *et al* have recently reported on the relationship of neighbourhood social conditions and its relationship to neural tube defects, based on reports of spatial clusterings of neural tube defects that are not fully explained by individual-level socioeconomic factors.(10) The neighbourhood measures included in this study included the proportion of adults without high school graduation, the proportion aged 16 years or older without employment, the proportion aged 16 years or older with labourer occupations, poverty rate, overcrowding, and the proportion of households which were rented. Wasserman does not offer a rationale for why these specific variables might be informative to

‡ A composite variable using the combined Z score for the average proportion of white-collar workers, median family income and median educational level.

the study of the epidemiology of neural tube defects. Again these variables are submitted as independent variables into regression analyses.

Diez-Rous, *et al,* working with data from the Athlerosclerosis Risk in Communities study provide yet another application of this approach – using a mix of census variables to explore a social-causation hypothesis – with an outcome of heart disease.(11) In their paper, Diez-Rous *et al* report that they selected four census variables "a priori" as indicators of neighbourhood context: percentage of adults over age 25 with incomplete high school, median household income, and the percent of persons in selected professional and technical categories, and median household value. (p. 49) They use these variables in logistic regression analyses of heart disease mortality among women.

Like other papers using this technique, Diez-Rous *et al* frustrate the interpretation of their results by failing to provide a rationale for the selection of the contextual variables. At one point they tantalizingly suggest that the "four variables were used as alternative indicators of the same underlying construct (social inequalities across neighbourhoods)." Yet at another point they suggest that these variables measured a more general, and potentially multidimensional concept of "neighborhood socioeconomic characteristics." Unless social inequalities in neighbourhoods are taken to be synonymous with the potentially broader concept of neighbourhood socioeconomic characteristics, the authors do not otherwise reconcile these two views of the meaning of their variables.

A common problem identified across these three papers is that the authors

have elected not to describe directly the construct(s) of neighbourhood social environments that they hoped to operationalize with the selected variables. This frustrates the assessment of construct validity, and critical methodological questions cannot be answered directly:"Are these variables adequate measures of the salient domains under discussion?" "How do these measures give insight into the etiology of the outcomes of interest?" Silence on these points leaves researchers and policy audiences adrift when it comes to the critical question: "What components of social environments *might cause* variation in health?" or "Where are the potential mechanisms to alter these relationships."

The lack of theoretical clarity is not a necessary defect of this method – there is nothing *conceptually* invalid about using multiple variables to operationalize a construct of neighbourhood social environments - clearly it has advantages over the use of a single variable, and potentially allows linkages to multidimensional social theories about neighbourhood social environments. However, the absence of an articulated theoretical framework to inform the selection of particular variables can be read to reflect a conceptual uncertainty about what *exactly* is a neighbourhood social environment, and what makes it important for health.

A recent paper by LeClerc *et al,* provides an important limit to this general critique. In this paper the authors explore the relationship of neighbourhood social context to racial differences in heart disease mortality in US women, and articulates strong theoretical links between the package of measures selected to describe neighbourhoods and the outcome - demonstrating that causally coherent

models are possible.(12) LeClerc's premise is that "residential segregation and economic isolation are the structural manifestation of racial discrimination" After relating segregation and economic isolation to a concentration of female-headed households, the authors theorize that "residential concentration of poor single parents or multi-generational households implies a network of stressful obligatory social ties that may hasten the death of black women from heart disease."

In Le Clerc's paper, the measured neighbourhood social environment variables included the proportion of female headed families, the proportion of black race/ethnicity, median family income, receipt of public assistance, unemployment, and deep poverty. All the variables can be easily integrated into the proposed causal pathway - "stressful obligatory social ties" and have strong face-validity for describing a demanding social context. The selection of variables and the interpretation of the results in this paper can be reconciled with the conceptual model of neighbourhood social environments and the etiologic assertion of the authors. Therefore, we would observe that while other authors have neglected to articulate a strong rationale for the choice of contextual variables when multiple variables are selected, this is not a structural weakness of the method itself.

We have, thus far, largely stressed issues of theoretical coherence in our critique of this method. However, it is also important to recognize an important statistical problem which arises under this approach: multiple colinearity. This approach requires the regression analysis of multiple variables - variables which

are highly correlated - particularly at the census tract level. Indeed, multiple colinearity is almost inescapable in neighbourhood social and economic data. This creates problems for the estimation of effect sizes for specific measures, and even for the ability to achieve convergence in some forms of multivariate analysis. LeClerc *et al,* for example, identify colinearity as a significant challenge to their analysis.

In summary, the "strategy D" approach to characterizing neighbourhood social environments is distinguished by the concurrent use of several continuously measured variables but does not include any steps which group or cluster neighbourhoods together. A subset of those papers using strategy D take the initial set of neighbourhood-specific measures, and then employ either factor analysis or some other scale-building strategy to provide a more restricted series of variables for subsequent analyses.

Appealing though the strategy is for its potential to represent a multidimensional conceptualization of social environments, papers using this method have often, although not exclusively, glossed lightly over the need to describe the theorized neighbourhood social environment they are seeking to model with the selected variables. The selection of variables thus can appear *ad-hoc*, frustrating the assessment of the construct validity of the resulting characterization of neighbourhood social environments. Where attention is paid to articulating a rationale for the variables used in the analysis, these criticisms do not apply. When factor analysis or other data-reduction strategies are not

undertaken, the approach is also susceptible to the problem of multiple colinearity.

Classification Strategy E: Multiple inputs yielding binary classifications

Still another strategy to describe neighbourhood social environments can be seen in the use of multiple input variables to yield a binary determination of neighbourhood social environment. This approach was adopted by Ricketts and Sawhill in their definition of "underclass areas".(13) In fairness, it should be noted clearly that the authors did not set out to provide a generalized neighbourhood social environment area, but rather to find concentrations of "deviance". Their interest in concentrations of deviance represented an explicit attempt to move debate about concentrations of social problems away from "poverty" as an explanatory variable, and point toward other salient concepts.

The "underclass areas" strategy sought to identify one particular form of neighbourhood social environments - areas with perceived acute departures from dominant behavioural norms - arguing that not all such areas were necessarily areas of extreme poverty, nor were all poverty-areas foci of "concentrated social deviance". However, no evidence of an association between areas designated "underclass" and any independent measure of social deviance was offered by the authors.

The Underclass areas construct was operationalized at the census tract level, using four variables available on the 1970, the 1980, and 1990 census. Four

proportions were identified and linked to the concept of social deviance: the proportion of households with female-heads, the proportion of adult males with with long-term detachment from the labour force, the proportion of young adults who were high school drop-out rates, and proportion of households on welfare. The associated "social deviances" were, respectively, out-of-wedlock childbearing, adult male participation in the workforce, the completion of at-least secondary education, and self-sufficiency. Those census tracts that were one standard deviation (as defined from the 1980 census) above the national mean (1980 census) with respect to all four of the above measures were denoted underclass areas.

Several authors have offered technical critiques of the Ricketts and Sawhill underclass areas proposal, arguing that the use of a national measure of dispersion (the 1980 standard deviation of each measure) is insensitive to state-specific variation, or that the fixed reference to the 1980 distribution – that is, the 1980 estimates of mean and variance - frustrates identifying evolving forms of deprivation (14-16) However, these criticisms largely accept the conceptual model and many of the operational decisions proposed by Ricketts and Sawhill.

(A more substantive, and more politicized, engagement of the Ricketts and Sawhill approach would note that concentrations of elevated adult male unemployment, receipt of welfare, female-headed households and high school drop-out rates, may not reflect so-called social deviance, but rather the working-out of personal circumstances and the impact of public policy choices. The

argument that concentrations of social problems represent the fallout of public policy choices, not personal deviance, has been forcefully articulated by Wallace and Wallace. (17, 18))

The Ricketts and Sawhill "underclass areas" measure, while clearly articulating its theoretical premises, has generally not attracted ongoing interest from social scholars, although a dissertation by Qi is an exception(19). Finally, its binary outcome measure — similar to the poverty-areas measures offered by the US Census Bureau — can be criticized for obscuring variation within areas designated underclass or non-underclass.

More generally, Strategy E shares with Strategy B the limitation of bifurcating the landscape of social environments, a quantitative strategy that does focus attention on a particularly problematic social environment, but which does not provide a nuanced depiction of a broad range of social environments. While Ricketts and Sawhill's underclass areas measure avoids this limitation by specifically disclaiming any attempt to offer a general assessment of social environments, the general strategy of dichotomizing social environments seems to offer limited flexibility for inquiries into social environments and health.

Classification Strategy F: Factorial Ecology

The sixth and final strategy identified in our framework - Strategy F - involves the selection of multiple input variables to characterize neighbourhood social environments; however, unlike strategy D, strategy F groups neighbourhoods into similar groups before yielding up some polytomous

classification of neighbourhood types. The most popular of these strategies is known as factorial ecology, and typically invokes factor analysis to reduce a set of continuous census-based measures of neighbourhoods to a reduced set of latent variables, and then cluster analysis to group similar neighbourhoods (based on the factor structure) in ways that maximize within-group homogeneity and between group heterogeneity. This approach to classifying neighbourhoods arises from the disciplines of sociology and urban geography, and can be traced back for at least a half-century(20-26). Curiously, this approach has received only passing attention in the public health literature.

In our view, the techniques of factorial ecology, carefully reviewed by Davies (27, 28), represent an important strategy for neighbourhood classification of social environments. The end-result of this approach is a small set of neighbourhood "types" based on minimizing the heterogeneity within clusters (of the measured domains), and maximizing the heterogeneity between and across clusters, and some sense of neighbourhoods that are more or less alike. Having a limited number of neighbourhood types, enhances the likelihood that representative sampling across large areas would yield sufficient observations in most neighbourhood types to support statistical analysis of differences across neighbourhood types. Secondly, the use of factor analysis means that different neighbourhoods can be characterized by degrees of similarity thereby permitting explorations of whether increasing levels of exposure to a social characteristic result in consistent changes in health status - suggesting a dose-response

relationship.

Lecair and Innes provide a useful example of a factorial-ecology approach in an effort to construct a model of neighbourhood social ecology and its association with child and adolescent mental health; they employ principal component analysis and then cluster methods for small-area data in Windsor, Ontario.(29) They develop their model of neighbourhood social environments using 19 variables, measured at the enumeration-area level, reflecting employment status, income, mobility, family-organization, immigrant density, high-school leaving, and residential mobility. The authors then report a six factor solution, which explains some 73.1% of the variance amongst the 19 variables. The first component, representing 28.5% of the variance, is interpreted as "social status/mobility"; the second component was labelled "employment level/employment sector". The residual components (each representing 10% or less of the variance) were labelled housing quality, income source, post-secondary attainment/immigration status, and finally non-professional attainment. Each enumeration area could then be scored with respect to each of these six factors, and these six inputs were then submitted to cluster analysis - from which the authors identify 16 clusters of neighbourhoods.

The authors then report correlations between rates of *mood-conduct disorders* and these 16 types of social environments, but not between rates of *neurophysiological disorders* and social environment. The authors take this mixed evidence to argue for social environment causation of mood-conduct disorders,

and argue that the negative finding for neuro-physiologic disorders argues against "reverse causation" as a valid explanation for the mood-conduct finding. That is, the lack of evidence of neurophysiologic disorders suggests that the mentally ill are not moving into specific social areas, thereby challenging the potential influence of "downward drift" in the production of these social inequalities in the. distribution of mood conduct disorders.

The approach taken by Leclair and Innes represents a strong example of the factorial ecology approach; with its move from a plethora of measured variables to a set of latent constructs, or factors, and from there to the association of "like" neighbourhoods with respect to those latent factors.

Harries has presented several papers using factorial ecology methods to fashion a model of neighbourhood social environments for Baltimore – building a measure he calls an Index of Social Stress - to describe the relationship of social environments to violence.(30-32) Using 24 variables measured at the census-block-group level, the stated aim was to operationalize a concept of "social stress." Harries suggests that 15 of his variables "may be regarded as direct, indirect, or inverse indicators of poverty in some form," (p. 1256) reflecting his assertion of the centrality of poverty as the "most quintessential" of stressors. (p. 1254) Other variables included measures of housing quality, and neighbourhood stability, crowding, and vacancy rates. Harries also includes two variables which measures immigration from the Southern US and information on the frequency of violence which was obtained from non-census sources.

Harries goes on to present a rotated, three factor solution, arising from his 24 initial variables, with factors he labels "poverty-violence", unemployment, and "stable neighbourhood". The factor solution explained a total of 52.5% of the variance in the input matrix. Harries then submitted the factor scores for cluster analysis, identifying five clusters, composed of between 3 and 571 census blocks-groups to each cluster. Harries documents differences in the three measured domains across the clusters. Not surprisingly given the methods employed, differences in the poverty-violence, unemployment and stability domains can be seen between the clusters.

Whilst a classic example of the application of the techniques of factorial ecology - Harries' index of social stress is clear in its central theoretical premises, and the construction of the index of social stress follows the factorial ecology method very closely, but its application and interpretation are ultimately problematic. The inclusion of variables to measure the hypothesized violent nature of "southern" emigres frustrates the generalizability of this "social stress" scale to other locations — notably the southern states. More problematic, however, is the decision to include violence rates as an input to the factor analysis. Violence rates were acquired from non-census sources, and their inclusion as a covariate in the factor analysis makes it impossible to show an unconfounded predictive relationship between "social stress" (which is measured by an aggregate of violence rates) and physical trauma rates. Additionally, having one of the five clusters populated with only 3 census tracts restricts - by simple small numbers -

the assessment of whether there are distinctive patterns of traumatic injury in this cluster. Finally, within the paper it is difficult to discern the quantitative association of the factors and the outcome, as the results are largely presented in map form, with points showing locations of traumatic injury overlaid on maps shaped to represent the five neighbourhood social environments. This point-overlayed choropleth mapping strategy makes it difficult to determine if any apparent associations are statistically significant.

While Harries has presented an ultimately unsatisfying and confounded presentation of social stress, and one which is of not suited to national application, its factorial ecology foundations are clearly evident - the reduction of correlated census variables into a restricted set of latent factors, and the clustering of neighbourhoods based on these latent factors.

Before concluding, an important variant of Strategy F needs to be identified, and represents a combination of Strategy F - in its use of multiple inputs and data reduction - and Strategy D in the persistence of continuous variables into the analysis phase. This variation is distinguished by an intermediate step - the use of multiple continuous variables to derive one or more formally constructed indices, and the subsequent use of these continuous indices to quantify neighbourhood social environments. For example, Sampson *et al* select 10 census variables and develop a 3 factor solution which highlights the traditional influences of social disorganization theory: low economic status, racial heterogeneity and mobility. Sampson then goes on to use these three factor scores (which are invariant within

neighbourhoods) in regression analyses of violence outcomes in Chicago neighbourhoods.

Cohen *et al* have taken a similar approach, offering two neighbourhood indicies: one a poverty index composed of income, unemployment, and low education; the other index was dubbed the"broken windows" index and measured housing quality, abandoned cars, graffiti, trash, and other signs of neighbourhood deterioration. While the poverty index could be obtained from census data, the "broken windows" index required local data ascertainment(33).

This strategy - creating indices from a set of variables, and then using those indices rather than the underlying variables necessarily raises questions about index construction - but also apprears to represent a promising strategy. The application of factor analysis such as is reported by Sampson, or other data reduction strategies (such as is used by Cohen *et al)* are aimed to reduce correlation across the indices - if successful, this should reduce the vulnerability of the resulting indices to colinearity considerations.

Discussion

We have identified six approaches to characterizing neighbourhoods, and have provided an amplified discussion of four of these as used in the public health literature. Two general observations arise: First, poverty, and particularly the concentration of poverty within neighbourhoods, is systematically acknowledged as "the most quintisential" (Harries) aspect of neighbourhood social environments - an effect which Sampson describes as "centrifugal" for its impact on residents of

high-poverty areas. Measures of neighbourhood social structure should be attentive to the effects of poverty, and particularly of concentrated poverty.

A second observation is that while a rich theoretical tradition exists in urban ecology - dating from the 1930s and forward - and has emphasized the multidimensional nature of neighbourhood social environments, little of the theoretical insights or empirical methods from urban ecology has evidently filtered into the public health literature. Dichotomized outcomes - typically poverty/non-poverty measures- are regularly used in the public health literature to capture the nuanced concept of "social environments". Such measures not only fail to speak to the theoretical models of social environments, but they necessarily obscure variation within both poor and non-poor areas.

Polytomous classifications arising from analyses of multiple ecologically descriptive variables are potentially more satisfying because of their ability (not always actualized) to harmonize theorized and measured domains of social environments. However, this approach presents technical challenges associated with multiple colinearity. The introduction of data reduction techniques such as factor analysis, and principal components analysis can mitigate these problems - so too the construction of intermediate indices - these methods have not been extensively used, particularly in the US public health literature.

Strategy F approaches - including factor analysis and cluster analysis, or at least data reduction and the creation of indices - offers three comparative advantages as a technique for the construction of classifications of neighbourhood

social environments: the strategy allows a large number of initial variables to be reduced to a more manageable set of "factors"; it encourages a reconciliation of the observed factor structure with pre-specified theoretical models; and finally, (at least when cluster analysis is used) it creates a limited plurality of neighbourhood types, thereby avoiding both the dichotomous approaches of the poverty or underclass "areas" approaches, and the conceptual ambiguity and technical problems associated by the amalgamation of several collinear and continuous ecological variables.

Public health applications which seek to use neighbourhood social environments, either to organize health data, or to test etiologic hypotheses about neighbourhoods should reacquaint themselves with the techniques of factorial ecology to aid the construction of meaningful metrics of neighbourhood social environments.

Figures

Number of inputs	Number of Distinct Output Values		
	None	Two	More than Two
One	A	B	C
More than One	D	E	F

Figure 2.1 Classification Scheme for Neighbourhood typologies.

Note: Cell contents are labels ('A' -'F') denoting six potential strategies for describing neighbourhood social environments.

References

1. Wilson WJ. The truly disadvantaged: The Inner City, The Underclass, and Public Policy. Chicago: University of Chicago Press; 1987.

2. Littman MS. Poverty Areas and the 'Underclass': Untangling the Web. Monthly Labor Review 1991;114(3):19-32.

3. Geographic Areas Reference Manual. Washington: US Bureau of the Census; 1994.

4. Cunradi CB, Caetano R, Clark C, Schafer J. Neighborhood poverty as a predictor of intimate partner violence among White, Black, and Hispanic couples in the United States: a multilevel analysis. Ann Epidemiol 2000;10(5):297-308.

5. Jargowsky P, Bane M. Ghetto Poverty in the United States, 1970-1980. In: Jencks C, Peterson P, editors. The Urban Underclass. Washington, D.C: The Brookings Institution; 1991. p. 235-273.

6. Haan M, Kaplan GA, Camacho T. Poverty and health. Prospective evidence from the Alameda County Study. Am J Epidemiol 1987;125(6):989-98.

7. Waitzman NJ, Smith KR. Phantom of the area: poverty-area residence and mortality in the United States [published erratum appears in Am J Public Health 1998 Jul;88(7):1122]. Am J Public Health 1998;88(6):973-6.

8. Citro CF, Michael RT, Panel on Poverty and Family Assistance (United States). Measuring poverty : a new approach. Washington, D.C.: National

Academy Press; 1995.

9. Roberts EM. Neighborhood social environments and the distribution of low birthweight in Chicago. Am J Public Health 1997;87(4):597-603.

10. Wasserman CR, Shaw GM, Selvin S, Gould JB, Syme SL. Socioeconomic status, neighborhood social conditions, and neural tube defects. Am J Public Health 1998;88(11):1674-80.

11. Diez-Roux AV, Nieto FJ, Muntaner C, Tyroler HA, Comstock GW, Shahar E, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. Am J Epidemiol 1997;146(1):48-63.

12. LeClere FB, Rogers RG, Peters K. Neighborhood social context and racial differences in women's heart disease mortality. J Health Soc Behav 1998;39(2):91-107.

13. Ricketts E, Sawhill I. Defining and Measuring the Underclass. Journal of Policy Analysis and Management 1988;7(2):316-325.

14. Hughes MA. Concentrated Deviance and the "Underclass" Hypothesis. Journal of Policy Analysis and Management 1989;8(2):274-81.

15. Ricketts E, Mincy R. Growth of the Underclass: 1970-1980. J Hum Res 1990;25:137-145.

16. Sawhill IV. Comment on "Concentrated deviance and the 'underclass' hypothesis". Journal of Policy Analysis and Management 1989;8:282-3.

17. Wallace R. Social Disintegration and the Spread of AIDS: Thresholds for Propagation along 'sociographic' networks. Soc Sci Med 1991;33(10):1155-1162.

18. Wallace R. Social Disintegration and the Spread of AIDS-II: Meltdown of Sociogeographic Structure in Urban Minority Neighborhoods. Social Science and Medicine 1993;37(7):887-896.

19. Huang Q. Urban Ecology and the Underclass In New York City [PhD Dissertation]. New York City: City University of New York; 1994.

20. Shevky E, Williams M, John Randolph Haynes and Dora Haynes Foundation. The social areas of Los Angeles, analysis and typology. Berkeley,: Pub. for the John Randolph Haynes and Dora Haynes Foundation by the Univ. of California Press; 1949.

21. Shevky E, Bell W. Social Area Analysis: Theory, Illustrative Application and Computational Procedures. Westport, Conn: Greenwood Press; 1955.

22. Tyron R. Identification of Social Area By Cluster Analysis: A General Methods with an Application to the San Francisco Bay Area. In: Brunwick E, Rosenzweig M, Tryon R, Tuddenham R, editors. University of California Publications in Psychology. Berkeley, CA: University of California Press; 1955.

23. Openshaw PJM. When we sneeze, does the immune system catch a cold? Viral infections may have important effects on the body's response to

bacteria and other viruses. Br Med J 1991;303:935-936.

24. Reading R, Openshaw S, Jarvis S. Are multidimensional social classifications of areas useful in UK health service research? J Epidemiol Community Health 1994;48(2):192-200.

25. Openshaw S, Blake M. Geodemographic segmentation systems for screening health data. Journal of Epidemiology and Community Health 1995;49(Supplement 2):S34-S38.

26. Charlton M, Openshaw S, Wymer C. Some new classifications of census enumeration districts in Britain: a poor man's ACORN. Journal of Economic and Social Measurement 1985;13:69-96.

27. Davies W. Factorial Ecology. Aldershot: Gower; 1984.

28. Davies W. Urban Social Structure: A Multivariate Structural Analysis of Cardiff and its Region. Cardiff: University of Wales Press; 1982.

29. Leclair JA, Innes FC. Urban ecological structure and perceived child and adolescent psychological disorder. Soc Sci Med 1997;44(11):1649-59.

30. Harries K, Powell A. Juvenile Gun Crime and Social Stress: Baltimore, 1980-1990. Urban Geography 1994;15(1):45-63.

31. Harries K. The Ecology of Homicide and Assault: Baltimore City and County, 1989-1991. Studies in Crime and Crime prevention 1995;4:44-60.

32. Harries K. Social stress and trauma: synthesis and spatial analysis. Soc Sci Med 1997;45(8):1251-1264.

33. Cohen D, Spear S, Scribner R, Kissinger P, Mason K, Wildgen J.

"Broken windows" and the risk of gonorrhea. Am J Public Health 2000;90(2):230-6.

Chapter #3: SECTOR: A classification for US census tracts based on social disorganization theory

Foreword

The previous two chapters have argued for the importance of social environments as an important concept in epidemiology, and have shown that there is a long line of scholarship (although mostly outside of public health) which has thought about classifying neighbourhood social environments (Chapter #1). We have also shown (in Chapter #2) that existing measures of social environments are often conceptualized poorly or are operationalized in ways that either conflict with dominant theories of neighbourhood, or that present quantitative challenges. In this chapter we offer a new metric - SECTOR - which goes some way to meeting the key deficiencies of extant census-based characterizations of neighbourhoods; Chapter #4 will test the validity of SECTOR.

Introduction

In the two immediately previous chapters we have reviewed theories of neighbourhood social environments and contrasted those theories with the actual measurement of neighbourhoods in the recent US public health literature. We have identified a gap between the dominant theoretical perspectives about social environments - which are multivariate and multidimensional - and the empirical

"standards of practice" for coding social environment variables evident in the US public health literature.

In this chapter we seek to bridge the gulf between theory and practice by offering a classification of US neighbourhood based on social disorganization theory which uses the multivariate inputs and polytomous outputs. The goal is to provide a reasonable census-data-oriented, theory-based measurement scheme which will support both descriptive presentations of health data, and etiologic studies of the potential causal role played by neighbourhood social environments in patterns of population health.

Our approach has several important and distinguishing attributes that depart from the main body of published methods for the classification of US neighbourhood social environments in the US public health literature. First and foremost, it is grounded in a leading theory of social environments - social disorganization theory; it is also multidimensional (and hence multivariate) in its conceptualization and in its operationalization; it is national in scope, but measures census tracts, and is based exclusively on census data; finally it identifies a plurality of neighbourhood social environments, and organizes them in a manner which allows at least for a limited consideration of "dose-response". These features commend the resultant measure as a possible mechanism for describing neighbourhood social context, that is consistent with at least one leading theoretical approach.

The remainder of this chapter is organized into three sections. First, we

articulate a theoretical framework for thinking about US neighbourhood social environments; the subsequent section reports on the operationalization of that model; some summary remarks are offered at the end.

Social disorganization: a theoretical perspective on neighbourhoods and health

Any classification of neighbourhoods for public health purposes requires some attention be paid to the theorized aspects of neighbourhoods' social environments which are held to modify human behaviour and health. Social disorganization theory is one such theory, associated with the inter-war research of Shaw and McKay; this approach stressed the influence of neighbourhood characteristics in the production of a particular social phenomenon - local juvenile delinquency rates.(1) In particular, the theory stressed the importance of low economic status, racial heterogeneity, and residential mobility as independent influences on social outcomes.

After a long hiatus, interest in Shaw and McKay's work is being revitalized, most evidently in the criminology studies work of Robert Sampson(2-4); by Brookes-Gunn with respect to child outcomes(5, 6); and in the work of Wallace & Wallace(7-9), and Kawachi and Kennedy(10) with respect to other measures of social pathology and human health.

As discussed in chapter #1, there is necessarily some overlap between social disorganization theory and other posited social factors, including social capital, theories of social attachment(11), Sampson's recently articulated theory of

collective efficacy,(3) and broader social measures such as income inequality(12). While each approach has its distinctive empirical approach, cross-cutting themes of social inter-dependence, and reciprocity are common amongst these social theories. However, social disorganization theory is arguably unique amongst these domains for its emphasis on neighbourhood level phenomena, and in particular for its identification of the potential import of residential turnover in the production of social pathologies. (Although not surprising given his intellectual debts to Shaw and McKay, some evolution in this uniqueness of social disorganization's rootedness in the neighbourhoods can be seen in a recent paper by Sampson *et al* which argues that social capital should best be thought of as a neighbourhoood phenomenon.(13)).

This dissertation does not propose to set up a test of the relative efficacy of the various leading theories of social influences on a specific set of outcomes.[‡]

In this assumption, we join with Kawachi *et al*(10) in identifying social disorganization theory as a viable theoretical model for epidemiology. Acknowledging a debt to the late Sol Levine, they write:"[w]here social relations are strained, we also tend to observe increased rates of crime, as well as unhealthier and unhappier citizens." (14) As social disorganization theory has

---

[‡]The various theories underlying these concepts are generally proposed to operate at distinct scales of social collectives, ranging from neighbourhoods to nations - so head-to-head comparisons would not be possible. Rather, the dissertation takes as its point of departure the key assertion of social disorganization theory – that neighbourhood level influences matter to human health and behaviour, and more specifically that economic deprivation and residential mobility are orthogonal domains of influence on human health and behaviour.

been shown to inform the distribution of criminality across neighbourhoods, (1, 3, 4) we suggest it represents a viable theoretical lens for epidemiologic analysis.

Consistent with social disorganization theory, and the general thrust of traditional urban ecology, we further suggest that neighbourhood social environments are determined by multiple factors, that there are a plurality of such social environments, and that this plurality can be organized into intellectually coherent groups and clusters. Therefore, we reject the assertion that binary classifications, whether based on a single input variable, or a set of variables, are adequate measurement strategies for representing social envirnoments in quantitative studies.[‡]

Under social disorganization theory, low economic resources and high residential mobility are expected to confer disadvantage, indeed independent disadvantage, with respect to the ability of neighbourhoods to achieve specific social goals - most notably the control of youth crime. Neighbourhoods with the lowest social resources and the highest residential mobility are expected to be the most disadvantaged with respect to achieving these social goals.

This observation has important implications for the numerical representation of a classification scheme. Clearly it must have at least two separate dimensions, one for economic resources and another for mobility - and each dimension needs sufficient variation to show trends in outcomes within the dimension.

[‡] In suggesting that social environments are continuously distributed, we do not exclude the possibility of "threshold points" beyond which rates of events change in non-linear fashions.

However, a model of neighbourhood social environments which operationalized only these two domains - economic resources and mobility - might be characterized as under-determined. Certainly it would not incorporate domains which have traditionally been thought important to neighbourhood differentiation - such as urbanization and family-structure(4). This recognition parallels the strategy adopted by Sampson and Groves in their classic British Crime Survey paper(4) in that paper the authors augment measures of the classic three dimensions of social disorganization theory (economic deprivation, mobility, racial heterogeneity) with measures of family structure and urbanization. Therefore, it was important to operationalize and quantitatively test these other two domains which either have a tradition in the urban ecology literature, or which have also been associated with the social disorganization tradition.

A general reading of the public health literature also suggests that a set of phenomena which we will group under the label "social-press" collectively act to erode the ability of residential collective groups to achieve collective goals; examples of "social press" include overcrowding, lack of adequate plumbing, etc. Overcrowding has been used in a variety of public health studies (15). Roberts, for example, has used estimates of overcrowding in his characterization of Chicago neighbourhood social environments.(16)

We have identified five main domains (economic resources, mobility, urbanization, family disruption, and social press) which are either directly associated with social disorganization theory, or are established constructs in

public health or urban sociology; all these domains might well be operationalized in a neighbourhood classification scheme.

A sixth organizing dimension - racial heterogeneity - has been dropped from the classic social disorganization approach. Certainly it has been used as an organizing dimension of various studies, from Shaw and McKay to Sampson. However, including racial heterogeneity presents theoretical, empirical and interpretive challenges that we have attempted to avoid by its exclusion. The interpretive tradition following Shaw and McKay is that racial heterogeneity encourage the formation of informal social control, and thereby reduced a salient deterrent to crime; racial heterogeneity is held to reduce the formation of informal social control, and therefore is associated with increases in disorder.

However, racial *homo*geneity has also been recognized as a negative force in social life - the out-working of racial segregation policies(17-20). Confusion clearly exists as to which end of the race/ethnicity distribution is important; perhaps more accurately, there seems to be agreement that both racial heterogeneity and racial homogeneity have negative social consequences.

There are daunting (and thus-far unaddressed) technical challenges to operationalizing segregation measures for all *census tracts* in the United States. All segregation measures imply a comparison of a local area to a larger area with respect to the distribution of some outcome. While for many census tracts, the larger areal area would likely be a metropolitan statistical area (MSA), census tracts outside of MSA would have missing values for such a measure of

segregation.(17-19, 21) Finally, including race/ethnicity measures in our neighbourhood classification scheme would complicate the interpretation of any race-oriented observations that we, or others, might report using our proposed measure of neighbourhood social environments. Taken together, these observations argue against the inclusion of race or measures of racial homogeneity or heterogeneity in our model.

Before proceeding to operationalize a model of neighbourhood social environments, reference need be made to the work of Drs. Roderick and Deborah Wallace(7-9, 22-29). A central theme in their work is that public policy, and specifically the withdrawal of city services from urban cores, has created urban "dessertification." Using data from New York City they argue that several consequences of this withdrawal of civic services have included epidemics of fire, migrations of a scale rarely seen in peace-time, and a breakdown of social relationships and social fabric which they argue has facilitated the transmission of infectious agents such as tuberculosis and HIV. They further argue, referencing Shaw and McKay, that these migrations were associated with profound disruption of social and familial networks, with a consequent reduction of informal social influence and control. The Wallaces' approach with its emphasis on residential mobility is explicitly framed within a social disorganization perspective, but finds the root cause of social problems in parts of New York in patterns of massive migration related to capitalist economics and subservient public policy, not in cultural or social explanations of crime. In this emphasis on the negative effects

of mobility on civil order, the Wallaces' reflect the social disorganization tradition.

The central problem with developing a national neighbourhood classification scheme which specifically addresses the Wallace hypothesis is that the domains of their causal model: the destructive influences of public policy, inter-census tract migration rates, housing loss, withdrawal of city services, disruption of social networks either cannot be measured from public census data or are too difficult to operationalize accurately at a national level using extant census data. It is therefore important to highlight that the decision not to employ the Wallace models is strictly a pragmatic one, and does not discount the importance that these related theories have for public health. The lack of data to operationalize these hypotheses represents a critical gap in public data collection.

Methods

Data

To operationalize a concept of neighbourhood social environments we selected 15 measures which could be derived from both the 1980 and 1990 US census STF-3A files(30, 31). These measures reflected five domains related to social disorganization theory: economic resources, mobility, family structure, urbanization and social press, as defined above. Data were obtained for all populated US census tracts from the 1980 and 1990 census. Census tracts are small geographic areas, contained within US counties (or their equivalent); census tracts are defined to be:

> A small, relatively permanent statistical subdivision of a county in a metropolitan area (MA) or a selected nonmetropolitan county, delineated by a local committee of census data users (a CSAC) for the purpose of presenting decennial census data. Census tract boundaries normally follow visible features, but may follow governmental unit boundaries and other nonvisible features in some instances; they always nest within counties. Designed to be relatively homogeneous units with respect to population characteristics, economic status, and living conditions at the time the CSAC [Census Statistical Areas Committee] established them, census tracts usually contain between 2,500 and 8,000 inhabitants. They may be split by any subcounty geographic entity(32).

Census tracts have long been the most popular choice of small-area census geographic units for the concept of neighbourhood.(32) We restrict our analysis to the states of the union, and do not include Puerto Rico, the Virgin Islands, or Pacific Island territories. Census tracts were excluded if they had no associated landmass (such as ships) or had no population.

The social disorganization neighbourhood model used data acquired directly from the US STF-3A census file, and from two intermediate sources of STF-3A data – the Urban Institute's Underclass database and 1990 census tract data which was bundled with the geographic information system, Maptitude.(33). Definitions of variables are reported in Table 3.1; data were also obtained to operationalize the Underclass Areas (34), and the poverty-areas and extreme poverty areas measures.

In order to operationalize the urbanization domain, it was necessary to estimate population density; therefore, estimates of census tract area and population counts are required; however, for 1980 no measures of census tract area was available from the US Census.(32) To redress this, the 1980

census tract boundary files were extracted from a layer of the 1992 TIGER files (35). A digital national base-map for 1980 US census tracts was constructed using Maptitude, and census-tract estimates of the area were undertaken using its area-estimation algorithms.

Technical problems prevented reconstruction of census tract boundary files for parts of Los Angeles County and the states of Hawaii and Utah for the 1980 census. Thus, these areas could not be included in our analyses. For 1990, all census tracts had area estimates and complete boundary files; these area estimates were obtained from the digital census-tract map distributed with Maptitude; these boundaries are based on US Census Bureau TIGER files.

It is important to note that at no point were data from the two census years pooled. Variables which measured the same construct in the two different census periods (e.g. welfare rates) were stored in separate datasets, and given variable names which incorporated evidence of the census year (e.g. welfare8, welfare9).

Data Analysis

Following Davies (36, 37) we initially adopted a factorial-ecology strategy in our attempt to build a social environment model– we began with multiple variables measuring hypothesized axes of neighbourhood differentiation acting as the primary inputs; thereafter we employ factor analysis to reduce the initial data vector to a more limited set of latent factors.

The 1980 and 1990 census datasets were always manipulated and analyzed independently.

Factor analysis, with varimax rotation, was performed on the derived measures from the 1980 and 1990 census datasets, using SAS's PROC FACTOR. A parsimonious factor model was identified for each census dataset, with factor-selection based on the rule that retained factors need have eigenvalues greater than 1.

In order to support comparability across the two time-periods, variables were retained into the "parsimonious" model if the variable had a loading greater than 0.5 for either decade, for any of the retained factors. Factor scores were calculated for each census tract in each census period with non-missing values for the variables which acted as inputs to the factor analysis. No imputation was done for missing values; census tracts with missing values for any of the independent variables were dropped from the final analyses for that census year.

Because a common number of factors (2) was observed across the two census periods, and because there was initial qualitative evidence to suggest that each factor corresponded to a constant conceptual domain, we sought to determine if quantitative evidence of consistent loadings could be established. To that end, correlation coefficients were calculated for each factor across the two census periods. That is, variable-specific Factor 1 loadings for 1980 ($F80_1$) were correlated with the Factor 1 loadings for the same variable for

1990 ($F90_1$) similarly $F80_2$ loadings were correlated with $F90_2$ loadings. High correlation coefficients were taken to suggest a common construct across the two census periods. The unit of observation in both census periods was the census tract. We also report on the distribution of differences in the 1980 and 1990 factor loadings, obtained by subtracting the 1980 loadings from the 1990 loadings; although crude, this was intended to highlight variable and factor-specific loadings that varied dramatically across the two census periods.

It was initially intended that the resultant factor scores would be submitted to cluster (38) However, we observed that clusters labels and centroids arising from the independent analysis of the 1980 and 1990 data could not be compared across the two census periods. For example, cluster "A" in 1980 might have contradictory quantitative characteristics to the cluster "A" identified in the analysis of the 1990 census data - attempts to harmonize the data structures across the two census periods proved problematic.

Lacking a way to harmonize the designations across the two census periods prohibited us from making coherent statements about neighbourhood types (as represented by the cluster) across the two census periods, and this argued against proceeding with a cluster analysis.

Frustration with the ability to build consistent clusters across two distinct census datasets forced an unanticipated methodological innovation.

We noted that the correlation of the factor scores across the two periods was very high - suggesting that a coherent factor structure existed across the two time periods, and that the number of factors identified in each census-year-based factor analysis was constant. The interpretation of these two factors will be discussed in the results section. However, these observations encouraged us to develop an alternative summary strategy that would allow for cross-census comparisons.

In order to achieve classifications of neighbourhoods that were consistent across the two census periods, we divided the continuous two-dimensioned "space" which arose from the factor analysis into "sectors" where sectors were defined (arbitrarily) with cut-points at the values of -1 and +1 normalized units for each factor score. We call the resulting classification SECTOR. SECTOR has nine values, labelled in the manner presented in Figure 3.1. The letters allocated to each sector are arbitrary, but proceed across the columns and then down the rows. (Note that the letter "I" is not used in labelling the sectors.) All census tracts from the 1980 and 1990 census with no missing values for variables needed to estimate Factor 1 and Factor 2 could then be assigned to a specific value of SECTOR, on the basis of the census-specific factor loadings, and the census-specific values of the 14 measured variables (Table 3.1).

We report on the number of census tracts allocated to each SECTOR in each census period. We also determined the distribution of the 14 input

variables used in the factor analysis across the two census datasets, and estimated the median centile rank of each value of Sector, with respect to each of the 14 sociodemographic variables. These median centile ranks for each SECTOR value are reported. In order to look for significant "drift" in the social characteristics within a SECTOR across the two census periods, the 1980 centile median rank was subtracted from the median 1990 centile rank for each variable, within each value of SECTOR, and univariate measures were employed to quantify the resulting distribution of the shifts in median centile ranks over the period. No statistical testing was undertaken on the resulting difference.

Results

Descriptive

A total of 43,274 census tracts were identified from the 1980 census and 61,389 census tracts from the 1990 census. Some 2,613 census tracts in 1,980 (6.1% of total), and 2,523 tracts in 1990 (4.1% of total) had missing values for one or more variables used in the factor analyses and are thus excluded from this analysis. Table 3.2 presents univariate descriptions of the 14 variables measured at the census tract level for 1980 and 1990.

Factor Analysis

Factor analysis was undertaken independently on the 1980 and 1990 census tract data. Parsimonious factor models (Table 3.5) were developed and a two-factor structure was observed in both periods, and all variables were

retained into the parsimonious model, save the "no-plumbing" variable which was dropped because in neither factor and in neither census period did the loading exceed 0.5. The parsimonious rotated two-factor model accounted for 61.3% of the variance in both the 1980 and 1990 data. The variance explained by $F1_{80}$ was 42.1%, and for $F1_{90}$ the explained variance was 44%; the variance explained by $F2_{80}$ was 19.2% and for $F2_{90,}$ 17.3%.

When comparing the loadings of the 13 variables retained in the parsimonious model, the correlation coefficient between the 1980 and 1990 loadings (arising from census-tract measures) for factor 1 was 0.99; for factor 2 the correlation coefficient was 0.90.

The key variables associated with the Factor 1 loaded strongly on economic resources. Poverty rate had the highest positive loading value in both datasets, followed by welfare rates, female headed households, males with long-term detachment from the workforce, and finally overcrowding. Large negative loadings were observed between Factor 1 and increased household median income, the proportion of the adult population with professional or supervisory employment, and the proportion owner-occupied. We call the factor "economic deprivation" because of the dominance of economic variables which load on this term, and the fact that increasing values of this factor are associated with diminishing economic resources.

Factor 2 (F2), representing between 17-20% of the variation in the two census periods, shows a more complex pattern of associations, and more

marked evolution between the 1980 and 1990 results for Factor 1. There are two distinct sets of variables in this factor, which can be organized with respect to whether the correlations of the variables to the factor are positive or negative. Positive correlations are organized around three demographic factors in 1980 – education, proportion of students, and managerial/professional occupation; density enters the model in 1990. Two other variables (residential mobility and owner occupancy) load negatively in both periods.

Factor 2 admits to no ready and simplistic explanation. Increasing proportions of higher educated individuals, and their higher incomes, seems to erode, not sustain this measure of community cohesion and identification. Our hypothesis is that higher education, professional occupation and proportion of students are surrogates and stimulants to residential mobility; likewise increasing concentrations of those with higher education or managerial/professional occupation may lead directly to more regular residential migration, and reduced community ties. This would suggest that concentrations of professionals and those with university educations does NOT lead necessarily to enhanced social cooperation within neighbourhoods.

We propose the label "Residential Mobility" for this factor, but recognize that other labels might also be proposed, and that further exploration of this factor would be desirable, together with a review of alternative factor analysis strategies

It is also important to recognize that this two factor solution is intended to operationalize a version of Shaw and McKay's social disorganization theory, and not to provide a systematic model for residential differentiation in the United States. That larger task is beyond the scope of the dissertation; readers searching for such a model may wish to consult White(39). The factor solution presented here is but one of several that might arise from factor analysis of these variables. The selection of this model over other models reflects both quantitative considerations (explained variation) and qualitative assessments of the concordance of the model with the social disorganization model. Within this restricted reading, the model represents nothing more, nor less, than a quantitative attempt to test Shaw and McKay's social disorganization theory against national census data.

Sector Approach

Using the "sector" approach, Table 3.6 reports on the distribution of tracts into the nine sectors, for the two census periods. In both census periods, approximately 50% of the census tracts are found in Sector E, and there is general consistency in the proportions of tracts in each Sector across these two periods. Sector A contained only one census tract in 1980 and two tracts in 1990.

Table 3.5 and 3.6 report, respectively, on the sector-specific median centile for each of the 14 measured covariates across the nine sectors in the 1980 and 1990 census years.

Using these data, it is useful to review the key characteristics of each Sector before proceeding. (For clarity, we report here on the 1990 median centile ranks, and remark on the 1980 centile ranks only where they differ by more than 20 points. Sector A reflects neighbourhoods with low residential mobility and low rates of economic deprivation (as compared with national means) - however, there are only a handful of neighbourhoods in this classification, too few to draw any meaningful statistical inference.

Sector B neighbourhoods have near-median levels of welfare (57th centile), high school leaving (48th centile), median household income (37th centile) and other economic indicators, but this Sector has a low proportions of its adult population with post-secondary education (21st centile), managerial/professional employment (26th centile), and has high rates of residential stability (81st centile).This is the highest level of residential stability amongst the Sector types, and the lowest median log-population-density of the Sector types.

Sector C neighbourhoods are characterized by extremely high ranking for poverty (89th centile), welfare (85th centile),male unemployment (90th centile),the lowest household incomes (1st centile), together with equally extreme low values for post-secondary education (11th centile) and professional employment (8th centile). The highest rates of overcrowding occur in Sector C.

High levels of household income (94th centile), post-secondary

education (91st centile) , and managerial/professional employment (92nd centile) characterize Sector D, as well as extremely low rates of households lacking plumbing (1st centile); however, these neighbourhoods have median values for residential mobility (50th centile).

Sector E neighbourhoods, effectively representing the majority of US census tracts in both census periods are, not surprisingly, characterized by its similarity to the national distributions (and thus near median values) for each of the indicated variables.

Sector F neighbourhoods are characterized by high rates of unemployment (90th centile) and poverty (92nd centile), and both low household income (12th centile) and low proportions of managerial/professional employment (12th centile), and median levels of owner-occupancy (51st centile).

Sector G neighbourhoods are characterized by comparatively high rates of income (85th centile), the highest rates of both post-secondary education (96th centile), and managerial/professional employment. These neighbourhoods also have low rates of neighbourhood stability (12th centile)

Sector H neighbourhoods are characterized by the second-highest population-densities (83rd centile), high proportions of college-students (87th centile), low rates of owner-occupied housing (10th centile), and median rates for high school leaving (58th centile), welfare (51st centile) and median household incomes (48th centile).

Finally, Sector J neighbourhoods are associated with elevated and extreme values for female-headed households (96th centile), poverty rate (95th centile), and male unemployment (91st centile), as well as the highest rates for overcrowding (92nd centile). Concurrently, Sector J neighbourhoods have the lowest proportion of owner-occupancy (5th centile) and low median household income (9th centile) in both census periods. It is worth noting, however, that while residential stability is low in both census periods (33rd centile in 1980, 27th centile in 1990) in neither census period do Sector J neighbourhoods have the lowest rates of residential stability – this is linked to Sectors G and H in both census periods. What makes Sector J remarkable as compared with Sectors H, or F is the extreme economic deprivation (which we do not see in Sector H), and the low proportion of home ownership (which we do not see in Sector F), and the lower rates of residential stability.

Detailed study of tables 3.5 and 3.6 will reveal considerable concordance in the within-sector rankings across the two census periods. To illustrate this more clearly, table 3.7 reports on the differences between the 1980 and 1990 median centile ranks rank for the 14 socio-demographic variable and the eight populated values of SECTOR; the 1980 values have been deducted from the 1990 ranks. The difference in median rankings ranges from -33 (for overcrowding in Sector C) to +23 for the proportion overcrowding in Sectors G & H. However, the mean difference in ranks, was a modest -0.77, suggesting overall consistency in rankings across Sectors and

measures.

Discussion

Polytomous multivariate neighbourhood classification systems have been used extensively in the United Kingdom to place health data "in context."(40) However, such classification schemes have not been available to public health researchers in the United States. Rather, the US literature has been dominated by binary classifications of neighbourhood context, chiefly variations of the "poverty area" approach. Alternative approaches have been reported (e.g. Harries), but often suffer from conceptual ambiguity or the statistical problems associated with colinearity, or are limited geographically to one location.

That said, polytomous classification systems, based on the techniques of factorial ecology, have much to commend them, including their potential consonance with theoretical models of neighbourhood social differentiation, and the removal of notions of a bifurcated world of neighbourhood social environments such as "poor or not poor" "underclass or not underclass". etc. We adopted a theoretical approach to neighbourhoods rooted in social disorganization theory that stresses the primary importance of economic resources and residential mobility – an approach which is consistent with the work of Shaw and McKay, and more recently of Sampson *et al.* and Kawachi *et al.*

Fourteen derived measures from the 1980 and 1990 census periods were

submitted to standard factor analysis, and two factors were dominant in each census period. Constistent with factor analysis resported by Harries for Baltimore, Sampson for Chicago, and Land for US cities (footnote!), our Factor 1 - which we call economic deprivation - loaded negatively on measures of economic resources, while our Factor 2, called "residential mobility", was inversely associated with increasing rates of residential mobility and household ownership, but positively associated with professional occupation and higher education. There is consistency in these findings with Sampson *et al* who found that "residential stability" emerged as a third factor in their analysis of Chicago data(3), and the work of Harries, which identified "Stable Neighbourhoods" as their third factor(41, 42).

The factor loadings across the 13 retained variables showed high correlation coefficient when comparing loadings for 1980 and 1990: 0.99 for Factor 1 and .90 for Factor 2. When the centile ranks for the various sociodemographic variables were compared across the two census periods, typically modest differences (mean -0.77) in the centile ranks were noted, and the three largest movements in median centile rank were associated with one variable- over-crowding. Taken together this evidence argues for the persistence of two underlying constructs - economic deprivation and residential instability - across the two census periods.

It is important to note that the factor structure for 1980 and 1990 for the residential stability factor differs in important ways. In interpreting these

differences, it is useful to recall that the census tract approach to classifying geography was not fully in place in 1980, and large parts of the US land-mass were not included in census-tract level counts in that census. By 1990, the whole population was allocated to census tracts. Thus, the 1990 factor structure is based on a more complete data structure. The emergence of managerial/professional classes and adult post-secondary education rates in the 1980 data may reflect a complex relationship between the mobility of a trained workforce and the more evidently urban, mobility captured in the 1990s.

Consistent with our observations on the nature of the underlying data, it is also important to observe that one variable, DENSITY, was more salient in the 1990 analysis, than in 1980. This may reflect the previously mentioned incomplete introduction of census tract methods in 1980, and its subsequent and complete use in the 1990 census.

Our initial intention to submit the census-tract-specific factor scores to cluster analysis was abandoned when it was recognized that the cluster analysis would provide neither consistent centroid coordinates nor naming strategies across the two census periods - thus frustrating any attempt to associate social characteristics associated with a specific centroid location across the two census periods. We adopted the SECTOR measure created from overlaying a simple 3 by 3 grid over the two-factor space; census tracts were assigned membership in a specific sector for each census period, based

on their factor score - the product of the variable loadings and the census data for that census tract in that census period.

The SECTOR classification strategy we have advanced has several important features. First, it is national in scope, but describes neighbourhoods using multiple measures , and is based completely on census data. Thus, it can be revisited with the provision of future census data. The classification strategy is also based on a well-respected model of neighbourhood social environments, and its empirical qualities accord with the theorized structure of neighbourhood social environments. Interestingly, we added three domains of measurement which were conceptually distinct from either economic deprivation or community attachment - these included family structure, urbanization, social press. These factors were either not important to either of the salient factors which emerged or were subsumed within the dominant domains emerging from the factor analysis.

That said, this chapter does not provide evidence that these classification schemes actually predict rates relevant to social disorganization. This is addressed in the next chapter.

There are important limitations to our work. First, the reliance on census data is a pragmatic virtue. Theories such as those advanced by Wallace and Wallace cannot be operationalized at this time for the whole United States using extant census data. Their work, in particular, raises important public policy and scientific questions, which are consonant with the

social disorganization approach, but are not directly addressed by our work. Secondly, the work by Sampson *et al* suggests that direct measurement of collective qualities such as collective efficacy are likely to complement and enhance the predictive ability of models which might otherwise include only measures of social environment constructed from census or census-like datasets.(3, 13)

A central challenge in this dissertation was developing a nosology that could be used across the 1980 and 1990 census periods. The approach initially envisioned a ready reconciliation of a set of empirically defined clusters arising from the separate analyses of the 1980 and 1990 census datasets. This reconciliation assumed, incorrectly, that the definition of clusters (that is, the estimation of centroids and ranges) would be consistent across the two datasets. Under this assumption, it would have been theoretically possible to speak of the characteristics of a cluster "X", across the two census periods, and to look for associations with outcomes in each cluster.

However, the separate analyses of the 1980 and 1990 census datasets did not yield a common number of clusters, nor were the quantitative characteristics of those centroids similar enough to support the proposed analysis. Three alternatives were explored. First, we considered the prospect of linking the 1980 and 1990 census tracts together and identifying constant spatial units across the two census periods. However, no ready reconciliation

of the 1980 and 1990 tracts was possible (Tatian). We learned that attempting to specify a longitudinal approach to census tracts was complex; new tracts arise, some tracts disappear, and the spatial boundaries of many tracts are revised. Very recent work (2002) by the Urban Institute and a private sector firm Geolytics have resulted in progress in this area, but these efforts were not available to this study.

The second option considered was to bring the 1980 and 1990 analyses together and conduct our cluster analysis using the combined data. This approach would have required complex manipulations of each census dataset to eliminate temporal changes in metrics over the decade (for example in average wages earned) Although normalization is part of factor analysis, cross-census normalization would have left unaddressed the question of how well normalization deals with the highly skewed distributions of many census-based measures of socioeconomic statutes. We could find no model in the literature which used this approach, and it too was abandoned.

The third approach – which we have adopted – was a necessary compromise. In this approach, we first confirmed that the factor structures across 1980 and 1990 for the two census periods shared commonality both in terms of qualitative measures and in terms of face-validity. Then we imposed a crude stratification of results (at -1 and +1 SD for each of Factors 1 and 2). This stratification allowed us to speak of variation in each axis, and had a degree of objectivity about it. The failure to find a significant number of

tracts in Sector A was a disappointment, but could not have been predicted when the method was devised. Further methodological work in providing consistent cross-census classification schemes for neighborhood typologies would be useful to social epidemiology.

In that vein, it is important to observe that Sampson's "collective efficacy" concept was both independent of, and empirically a more powerful explanatory variable than either the concentrated disadvantage, immigrant concentration or residential stability with respect to violence, victimization, or homicide rates.(3) While their work has yet to be replicated at a national level, it may be prudent for census agencies to consider the collection of collective efficacy and related measures in future census studies. Pending the availability of that sort of data, we suggest that the above approach continues to have merit.

With the SECTOR schema we have described, tests of the influence of social disorganization theory can now be advanced from census data on national geocoded survey data. In particular, social disorganization theory suggests that if increasing economic deprivation, and lack of community attachment are associated with poorer outcomes, then it seems likely that outcomes should be most desirable in sector A, and then proceed step-wise to the least desirable outcomes in sector J. Social disorganization theory does not tell us if thresholds exist across the "factor space" but quantitative tests of this hypothesis could be constructed.

Figures

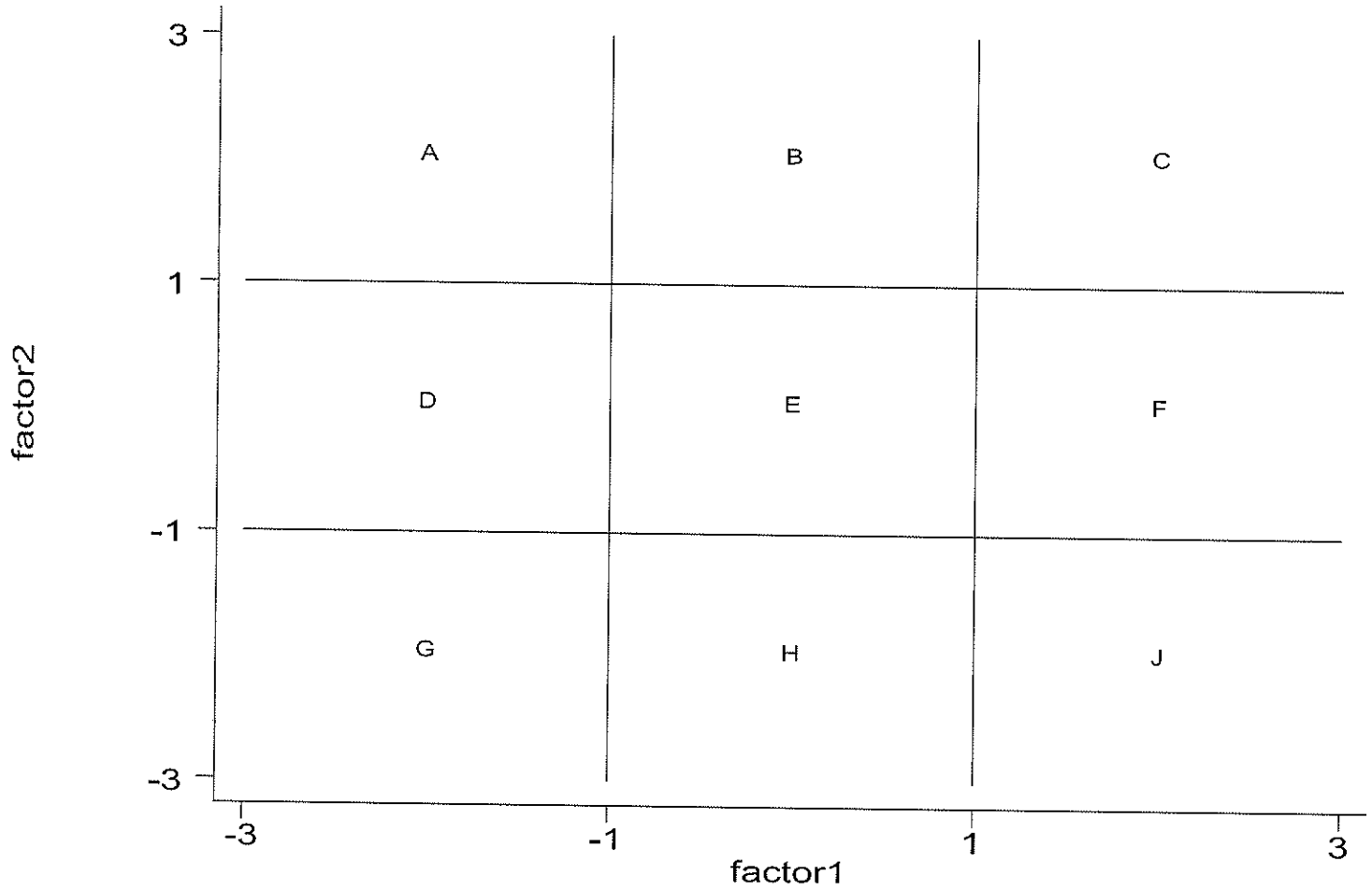

Figure 3.1 Schematic representation of SECTOR Model.

Table 3.1 Inputs to the Factor Analysis, organized by initial proposed domains of a social environment model.

Domain	Measures
Economic Deprivation	Poverty Rate Welfare Rate Long-term(Male) Detachment from the Work Force Household Median Income High School Drop Out Rate Proportion of the population with professional or managerial positions
Residential Mobility	Same house as five years prior to census Owner-Occupancy
Family Disruption	Female Headed Households
Urbanization	Population Density [Measured as log(Pop/sq mi)]
Social/Environmental Press	Proportion College students Overcrowding No plumbing Proportion over aged 25 with post-secondary education

Table 3.2 Univariate measures (weighted means and standard deviations) for census-tract inputs to the factor analysis, as derived from the 1980 and 1990 STF-3A census files.

Variable	1980			1990			Definition of Variable
	Number of Non-Missing Observations	Mean	Standard Deviation	Number of Non-Missing Observations	Mean	Standard Deviation	
Density	41233	7.57	1.85	59653	6.79	2.37	Log of Persons per Square Mile
FFH8	42351	0.19	0.14	59170	0.23	0.17	P. of households with female heads
HH_MEDIA	42813	18806	7142.49	61255	32123	14396.00	Median household income (in tract)
HSDROP	42593	0.13	0.11	59227	0.12	0.11	High school drop-out rate.
MNOPART	42865	0.31	0.12	59851	0.34	0.13	P. of males with no part-time work
MP	42868	0.16	0.09	59890	0.17	0.10	P. of adult males with professional occupation
NOPLUMB	42544	0.02	0.03	59525	0.01	0.02	P. of households with no private plumbing
OO	42600	0.64	0.24	59525	0.64	0.23	P. of households which are owner occupied
OC	42435	0.04	0.06	59519	0.05	0.07	P. of households with more than one person per room.
P25PS	42858	0.34	0.18	59845	0.43	0.19	P. of population age>= 25 years who have completed more than Secondary School
POVRAT	42726	0.12	0.11	59676	0.15	0.13	Poverty rate in tract
SMHSE	42846	0.53	0.15	59890	0.54	0.15	P. of population (age 5 years +) living same house as 5 years previous
STUDENTS	42868	0.06	0.08	59890	0.07	0.08	P. of population reporting current attendance at college
WELFARE	42544	0.08	0.08	59525	0.08	0.09	P. population reporting receipt of welfare

Table 3.3 Factor loadings from varimax factor analysis of 1980 and 1990 STF-3A census tract Data, Parsimonious Model.

Year	1980		1990	
Variable	**Factor 1 Economic Deprivation**	**Factor 2 Residential Mobility**	**Factor 1 Economic Deprivation**	**Factor 2 Residential Mobility**
Log population density	32	32	10	58
Female Headed Households	80	20	72	33
Household Median Income	-82	9	-78	11
HS Drop out	51	-16	47	3
Male Long-Term Unemployment	74	1	73	-4
Managerial/Professional	-67	58	-77	44
No Plumbing	Dropped			
Owner Occupied	-70	-46	-53	-70
Proportion in overcrowding	58	-8	48	19
Proportion 25y+ with Post/Sec	-60	78	-75	57
Poverty Rate	91	9	88	15
Proportion in same house as 5 year before census.	-3	-43	-1	-60
Proportion college Students	1	55	-9	50
Welfare Rate	84	-6	83	7
Percentage of variance explained by factor	41.2%	19.2%	44.0%	17.3%
EIGENVALUES	5.75	2.22	5.05	2.13
TOTAL VARIANCE EXPLAINED	61.3 %		61.3%	

N.B. All variables except overcrowding met criteria for inclusion in factor score calculation.

Table 3.4 Allocation of 1980 and 1990 Census tracts, by SECTOR variable.

Cluster	1980		1990	
	N	% of all census tracts	N	% of all census tracts
A	1	0	2	0%
B	4,871	12%	7,035	11.5%
C	232	0.6%	879	1.4%
D	3,827	9.4%	6,506	10.6%
E	21,615	53.2%	30,491	49.7%
F	4,304	10.6%	4,864	7.9%
G	885	2.2%	1,205	2
H	3,709	9.1%	5,488	8.9%
J	1,217	3.0%	2,396	3.9%

Table 3.5 Centile rank by sectors, of 14 socio-demographic characteristics measured at census tract level, by sector for 1980.

	SECTOR								
Census-Measure	A	B	C	D	E	F	G	H	J
Log population density		22	37	39	48	78	51	74	88
Proportion of population who are college students		20	14	71	49	29	87	89	50
Proportion of Female Headed Households		28	67	16	49	92	45	75	97
Male Long-Term Unemployment rate		51	88	16	47	90	18	54	93
Poverty rate		46	91	10	47	92	25	65	95
Welfare rate		52	90	12	49	93	10	43	95
High school drop out rate		51	80	17	51	81	17	50	83
Household Median Income		54	14	95	54	10	86	34	5
Proportion in same house as 5 year before census		74	85	52	52	56	25	14	33
Proportion of population with managerial/professional occupation		29	8	92	53	11	97	77	17
Proportion of population aged 25y+ with post-secondary education		22	4	91	52	12	97	82	26
No plumbing		57	90	1	48	78	35	59	88
Proportion of population in overcrowding		56	90	15	48	88	15	50	88
Proportion of households which are owner occupied-occupied		82	61	88	52	20	35	12	4

Table 3.6 Centile rank by sectors, of 14 socio-demographic characteristics measured at census tract level, by sector for 1990.

	SECTOR								
Census-Measure	A	B	C	D	E	F	G	H	J
Log population density		14	18	52	48	68	65	83	92
Proportion of population who are college students		20	12	73	49	29	89	87	51
Proportion of Female Headed Households		28	55	17	51	91	40	75	96
Male Long-Term Unemployment rate		61	90	15	47	90	8	43	91
Poverty rate		59	89	9	45	92	22	62	95
Welfare rate		57	85	12	48	93	10	51	96
High school drop out rate		48	64	23	51	80	22	58	87
Household Median Income		37	1	94	57	12	85	48	9
Proportion in same house as 5 year before census		81	66	50	51	51	12	12	27
Proportion of population with managerial/professional occupation		26	8	92	54	12	97	72	13
Proportion of population aged 25y+ with post-secondary education		21	11	91	53	13	96	76	18
No plumbing		79	88	1	47	72	1	46	78
Proportion of population in overcrowding		45	57	19	47	84	38	73	92
Proportion of households which are owner occupied-occupied		82	67	86	51	24	21	10	3

Table 3.7 Difference (1990-1980) in median centile ranks for 14 socio-demographic characteristics measured in 1980 and 1990 by SECTOR, and univariate description of differences.

	SECTOR								
Census-Measure	A	B	C	D	E	F	G	H	J
Log population density		-8	-19	13	0	-10	14	9	4
Proportion of population who are college students		0	-2	2	0	0	2	-2	1
Proportion of Female Headed Households		0	-12	1	2	-1	-5	0	-1
Male Long-Term Unemployment rate		10	2	-1	0	0	-10	-11	-2
Poverty rate		13	-2	-1	-2	0	-3	-3	0
Welfare rate		5	-5	0	-1	0	0	8	1
High school drop out rate		-3	-16	6	0	-1	5	8	4
Household Median Income		-17	-13	-1	3	2	-1	14	4
Proportion in same house as 5 year before census		7	-19	-2	-1	-5	-13	-2	-6
Proportion of population with managerial/professional occupation		-3	0	0	1	1	0	-5	-4
Proportion of population aged 25y+ with post-secondary education		-1	7	0	1	1	-1	-6	-8
No plumbing									
Proportion of population in overcrowding		-11	-33	4	-1	-4	23	23	4
Proportion of households which are owner occupied-occupied		0	6	-2	-1	4	-14	-2	1

Univariate analysis: Mean difference= -0.78; range -33 (decrease in over-crowding in Sector C) to +23 (rise in overcrowding in Sectors G and H)

References

1. Shaw, McKay. Juvenile Delinquency and Urban Areas. Chicago: University of Chicago Press; 1942.

2. Sampson R. What 'Community' Supplies. Washington: Brookings Institution; 1997 February, 1997.

3. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 1997;277(5328):918-24.

4. Sampson RJ, Groves WB. Community Structure and Crime: Testing Social-Disorganization Theory. American Journal of Sociology 1989;94(4):774-802.

5. Brooks-Gunn J, Duncan GJ, Aber JL. Neighborhood poverty. New York: Russell Sage Foundation; 1997.

6. Leventhal T, Brooks-Gunn J. The neighborhoods they live in: the effects of neighborhood residence on child and adolescent outcomes. Psychol Bull 2000;126(2):309-37.

7. Wallace R. Social Disintegration and the Spread of AIDS-II: Meltdown of Sociogeographic Structure in Urban Minority Neighborhoods. Social Science and Medicine 1993;37(7):887-896.

8. Wallace R. Social Disintegration and the Spread of AIDS: Thresholds for Propagation along 'sociographic' networks. Soc Sci Med 1991;33(10):1155-1162.

9. Wallace R. Traveling waves of HIV infection on a low dimensional 'socio-geographic' network. Soc Sci Med 1991;32(7):847-52.

10. Kawachi I, Kennedy B, Wilkinson R. Crime: social disorganization and relative deprivation. Soc Sci Med 1999;48(6):719-231.

11. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000;51(6):843-57.

12. Wilkinson RG. Unhealthy societies : the afflictions of inequality. London ; New York: Routledge; 1996.

13. Sampson R-J, Morenoff J-D, Earls F. Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children. American Sociological Review 1999;64(5):633-660.

14. Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality, and mortality [see comments]. Am J Public Health 1997;87(9):1491-8.

15. Wong YK, Dawkins KD, Ward ME. The association between deaths from myocardial infarction and household size in England and Wales. J Cardiovasc Risk 2001;8(3):159-63.

16. Roberts EM. Neighborhood social environments and the distribution of low birthweight in Chicago. Am J Public Health 1997;87(4):597-603.

17. Massey D, Cortese C, Falk R, Cohen J. On the Measurement of Segregation as a Random Variable. Demography 1978;16:553-563.

18. Massey DS. American Apartheid: Segregation and the Making of the Underclass. 1990.

19. Massey D, Denton N. The Dimensions of Residential Segregation. Social Forces 1988;67(2):281-313.

20. Massey DS, Eggers ML. The Ecology of Inequality: Minorities and the Concentration of Poverty, 1970-1980. 1990.

21. Massey DS. Social Class and Ethnic Segregation: A Reconsideration of Methods and Conclusions. 1981.

22. Wallace R. Contagion and incubation in New York City structural fires: 1964-1976. Human Ecology 1978;9:423-433.

23. Wallace R. The New York City fire epidemic as a toxic phenomena. International Archives of Environmental Health 1982;50:33-51.

24. Wallace R, Wallace D. Urban fire as an unstabilized parasite: the 1976-1978 outbreak in Bushwick, Brooklyn. Envir. Planning A 1983.;15:207-226.

25. Wallace R. A synergism of plagues: "planned shrinkage," contagious housing destruction, and AIDS in the Bronx. Environ Res 1988;47(1):1-33.

26. Wallace R, Wallace D. Origins of public health collapse in New York City: the dynamics of planned shrinkage, contagious urban decay and social disintegration. Bull N Y Acad Med 1990;66(5):391-434.

27. Wallace R, Fullilove M. AIDS deaths in the Bronx 1983-1988: spatiotemporal analysis from a sociographic perspective. Environment and Planning A 1991;23:1701-1723.

28. Wallace R, Wallace D. AIDS, Tuberculosis, Violent Crime and Low Birthweight in Eight US Metropolitan Areas: Public policy and the regional diffusion of inner city markers. In:; 1995.

29. Wallace D, Wallace R. Scales of geography, time, and population: the study of violence as a public health problem. Am J Public Health 1998;88(12):1853-8.

30. Census of Population and Housing, 1980 [United States]: Summary Tape File 3A: United States Department of Commerce, Bureau of the Census; 1980.

31. Census of Population and Housing, 1990 [United States]: Summary Tape File 3A: United States Department of Commerce. Bureau of the Census; 1990.

32. Geographic Areas Reference Manual. Washington: US Bureau of the Census; 1994.

33. Caliper Maptitude Users's Guide: Version 4.0 for use with Microsoft Windows. In:. Newton, MA: Caliper Corporation; 1996.

34. Ricketts E, Sawhill I. Defining and Measuring the Underclass. Journal of Policy Analysis and Management 1988;7(2):316-325.

35. Topologically Integrated Geographic Encoding and Referencing system. http://www.census.gov/geo/www/tiger/. Washington, DC: US Department of Commerce, Bureau of the Census; 1992.

36. Davies W. Factorial Ecology. Aldershot: Gower; 1984.

37. Davies W. Urban Social Structure: A Multivariate Structural Analysis of Cardiff and its Region. Cardiff: University of Wales Press; 1982.

38. SAS Institute. SAS/STAT user's guide, version 6. 4th -- ed. Cary, NC: SAS Institute; 1990.

39. White MJ, National Committee for Research on the 1980 Census. American neighborhoods and residential differentiation. New York: Russell Sage Foundation for the National Committee for Research on the 1980 Census; 1987.

40. Carstairs V. Deprivation indices: their interpretation and use in relation to health. J Epidemiol Community Health 1995;49 Suppl 2:S3-8.

41. Harries K. The Ecology of Homicide and Assault: Baltimore City and County, 1989-1991. Studies in Crime and Crime prevention 1995;4:44-60.

42. Harries K. Social stress and trauma: synthesis and spatial analysis. Soc Sci Med 1997;45(8):1251-1264.

Chapter #4: Drug Use and Drug Activity Organized by Social Environment

Foreword

The purpose of Chapter #3 was to build a polytomous variable which categorized neighbourhood social environments. The purpose of Chapter #4 is to test if illicit-drug-related measures are elevated in specific neighbourhood social environments. This is arguably the critical chapter of the dissertation because the definition and quantitative development of the SECTOR measure of neighbourhood social environments, and the extant extreme poverty and underclass areas definitions, have not been linked to any outcome data. Thus far, we have not presented any evidence that any phenomenon, and serious illicit drug use more specifically, is focused in specific social environments.

With respect to the SECTOR variable, the chapter therefore represents an important test of the social-disorganization hypothesis - that rates will vary both with respect to economic deprivation and community attachment.

This chapter marks the last chapter of the dissertation with a formal focus on social epidemiology. Subsequent chapters return to our initial concerns about the epidemic of hepatitis and HIV infection in recipients of some plasma fractions; the measure of neighbourhoods developed in previous chapters and

validated in this chapter will allow us to approach that substantive interest with a well-characterized measure of social environments.

Introduction

"Social disorganization" is a theoretical perspective on social problems initially articulated by Shaw and McKay to explain the persistence of juvenile delinquency in US neighbourhoods observed in the 1920 and 1930s(1). The persistence of delinquency rates was theoretically problematic to contemporary scholars because rates did not change with change in the assumed etiologic variables including neighbourhood ethnic composition. Social disorganization theory represented an alternative perspective that discounted the importance of the composition of a neighbourhood and in its place stressed the concurrent and independent importance of economic deprivation and residential mobility as distinct contextual determinants of juvenile delinquency.

After a period of relative neglect, "social disorganization" is currently enjoying a revival as an informative theoretical lens for considering the social organization of divergent phenomena, including social problems, child development, and health.(2-4) In this chapter we apply a measure of social disorganization - operationalized at the census tract - to a specific contemporary social problem -- illicit drug use -- using drug-related data obtained from a representative sample of US adults residing in households.

Social Disorganization

Shaw and McKay's social disorganization theory placed emphasis on characteristics of neighbourhoods as determinants of juvenile delinquency, as opposed to the characteristics of individual offenders. Of particular interest to Shaw and McKay was the influence of increased residential mobility and economic deprivation in eroding informal social control; eroded informal social control was regarded as the key determinant of high juvenile delinquency rates.(1) Sampson has been the leading contemporary voice advocating a recovery of social disorganization theory as a lens for viewing social pathologies. He adopts the view that the role of economic deprivation and mobility have been consistent empirical findings over the past half-century, with higher rates of economic deprivation and lower rates of community attachment as corrosive to social life.(5)

The Social Context of Drug Use

Previous attempts have been made to contextualize drug-use patterns, both in the United States, and internationally. Within the United States there has been a particular interest in contextual variables which shed light on the officially-reported over-representation of blacks on some drug-use measures. Of particular note, Lille-Blanton *et al* conducted a "post-stratification" analysis of the 1988 US National Household Survey on Drug Abuse (NHSDA), focusing on the reported positive association of black race with crack-cocaine use.(6) After restricting the analysis to those primary sampling units (taken as

neighbourhoods) where crack-cocaine was used, Lille-Blanton *et al* reported that an increased risk of crack-use amongst blacks was not observed.

The authors argued that this evidence suggested that blacks were no more likely to use crack cocaine, given that there was crack in their neighbourhood; rather crack availability was localized in predominantly black neighbourhoods, and this explained the reported over-representation of blacks among crack-users observed in the national data. Unfortunately, the 1988 NHSDA did not provide data to allow the authors to describe the neighbourhoods in which crack cocaine was found. Thus, the intriguing suggestion that "neighbourhood matters" in the organization of some drug-behaviours was not followed-up with information on what types of neighbourhoods might promote this specific form of drug use.

Similar observations have been reported in the UK - noting the association of serious illicit drug abuse with neighbourhood conditions. There are several UK-based studies linking extreme deprivation with drug use; the most systematic of which is reported by Squires, who demonstrates that"serious" drug use was concentrated in the most extremely deprived social circumstances in Liverpool.(7)

Hypotheses

The central contention of this chapter is that a measure of social disorganization for US census tracts will provide explanatory insights into both personal illicit drug behaviours and ecological aspects of illicit drug use -

insights that are not available when the standard person-level variables are used. Three testable hypotheses emerge:

1. That neighbourhood social context explains variation with respect to recent personal use of specific illicit drugs or related behaviours.
2. That neighbourhood social disorganization explains variation with respect to ecological aspects of drug use, including perceived access to specific illicit drugs, and the frequency with which drug sales and intoxicated individuals are observed in neighbourhoods.
3. That illicit drug use will be organized in a particular way within a social disorganization framework. We expect rates of drug use and drug availability will be lower where economic deprivation is lower, and that the same measures will be lower when community attachment is enhanced- particularly as expressed through low mobility rates. We expect the highest rates of social problems to occur where economic deprivation is most pressing and concurrently community attachment most limited.

Methods

Data

Overview

Data relevant to illicit drug use were acquired from the 1993 National Household Survey on Drug Abuse (NHSDA) which is a periodic representative survey of American adults (12 years and older) residing in households and group-quarters in the continental United States (n=26,489)(8). The survey is conducted by the Substance Abuse and Mental Health Services Administration (SAMHSA) an agency within the National Institute for Drug Abuse. Based on person-to-person interviews, and confidential written questionnaires administered within the interview, the survey reports measures in a variety of domains related to substance abuse, including illicit drugs used, methods of abuse, access to illicit drugs, and behavioural, demographic and social characteristics of respondents.

Extant Social Environment Variables in the 1993 NHSDA

Consistent with most social surveys, the NHSDA concentrates its measurement of "independent" variables on demographic characteristics of respondents, and to a lesser degree on characteristics of their households. Limited information is provided on the social context in which respondents live. The 1993 NHSDA public-use file contains only one variable which might meaningfully be described as a measure of *neighbourhood* social environment – the underclass area measure, as defined by Ricketts and Sawhill(9); this measure highlights census tracts with concurrently high rates of unemployment in adult males, high-school leaving, female-headed households and receipt of welfare.

The NHSDA does not provide the census tract of the respondent in its public use files, but this variable is used in the sampling and is thus available in the confidential master survey files retained by the SAMHSA, within the US National Institutes of Health.

Ricketts and Sawhill suggest that high rates on these four measures reflect an underlying social deviance. It is important to note that while the underclass areas variable is on the public use file, the measure has not been validated to predict drug-related outcomes. However, given the measure's theoretical foundation in "social deviance" it seems reasonable to test the underclass areas measure for its ability to identify foci of illicit drug-related behaviours.

Supplemental Social Environment Variables Added to the 1993 NHSDA

In addition to the Underclass variable, we were interested in applying other social context variables measured at the census tract level to our analysis of drug patterns. Two additional variables were identified: the first was a binary variable, identifying respondents who lived in neighbourhoods meeting the US Census Bureau's definition of an "Extreme Poverty Area"– that is census tracts where 40% or more of the population live below the US household-based poverty line.(10)

The second supplemental social context measure is based on social disorganization theory. Its theoretical base and empirical development are

discussed in Chapter #3; briefly, a factor analysis of 14 measures obtained at the census tract level for all 1990 US census tracts resulted in a two-factor solution where Factor 1 (F_1) increased with increasing economic deprivation; Factor 2 (F_2) was related to residential mobility, and this was influenced most predominantly by owner-occupied housing rates and by residential mobility rates.[‡] For each census tract where estimates of economic deprivation and residential mobility could be completed, the census tract was allocated into one of nine "sectors" based on a simple categorization of the bivariate distribution of the factor scores. Simply described, all non-missing census tracts were plotted using a simple bivariate graph; "lines" were established at -1 and +1 for both Factors 1 and 2, thereby effectively creating a "3-by-3" grid. Census tracts were then allocated to one of these nine sectors, with sectors labeled from left to right and top to bottom, using the letters A through J, and omitting the letter I. We denote the variable itself as SECTOR, and specific neighbourhood types are described as Sector A, or Sector J, etc. (Figure 4.1) This strategy had several advantages, amongst which it allows for the ready tabular presentation of data. Most of our tabulations adopt this approach.

A total of 58862 of a total 61387 (95.9%) 1990 US census tracts were classified using this approach; those tracts which were not classified typically

---

[‡] The factor solution corresponds very closely to the dominant factors published at sub-national levels; see Harries' (11-13) For factor-analysis of Baltimore-area census blocks, and Sampson *et al*'s factor analysis of Chicago neighbourhoods(5)

had small or no population; the small population triggers the selective suppression of data by the US Census Bureau to support respondent confidentiality, and this meant that one or both of the factor scores necessary to place a census tract within a sector could not be directly calculated. Finally, as only two out of the 61387 census tracts from 1990 were located in Sector A, any respondents from that sector were excluded from the analysis.

Linkage of New Social Environment Variable to the 1993 NHSDA

In order to attach additional social context variables for the public use version of the NHSDA, arrangements were made between the investigator and the National Institute for Drug Abuse to undertake a linkage at the National Opinion Research Center (NORC) using the NHSDA master file (which includes the census tract of respondents), and a file containing the two social context measures (Extreme Poverty Area, and SECTOR) for each 1990 US census tract. A privacy impact assessment was performed by NORC to ensure that adding these social environment variables would not reduce the privacy guarantees provided to respondents. The file which was returned to the investigator included only the unique identifier from the public use file, and the social context measures for the respondents. At no time was the census tract of the respondent released to the investigator.

Drug-Related Measures

A wide variety of measured domains are available in the NHSDA. We restrict our analyses to a selected subset of these domains: recent personal drug use, needle use, and drug-selling; reported ease of access to illicit drugs, and perceptions of neighbourhood drug activity. Three basic types of measures were identified, to operationalize the hypotheses: 1) measures of personal drug use and 2) measures of personal drug-related behaviours; 3)measures of neighbourhood drug activity.

For personal drug use, we measure self-reported use of cocaine, heroin, PCP(Phencyclidine) and crack cocaine both in the respondent's lifetime and in the last year. These are variables with true/false values. For drug-related behaviours, we report on self-reported selling of drugs in the last year, and lifetime needle use; again these are represented as binary variables in the public use file.

For neighbourhood level drug activity, we focus on self-reported assessments of access to four drugs: cocaine, heroin, LSD (Lysergic acid diethylamide) and PCP; these answers were recoded from the 5-point Likert scale found in the public use file, so that only the answer indicating "Very Easy" access to drugs was regarded as the non-baseline condition. We use a similar strategy for the question regarding the frequency of seeing "drunk or high" persons in the respondents' neighbourhood, categorizing "Very often" as the non-baseline condition, and setting all other non-missing answers to be the

baseline. For each outcome we present in cross-tabulated form the estimated proportion of the population reporting the outcome and the associated 95% Confidence Interval (CI)[‡].

Statistical Analyses

The NHSDA (1993) employs a complex, multi-stage sampling design; therefore, adjustments both for weighting and to modify the usual estimation of standard errors are required.(14) These adjustments were undertaken on all estimates reported below. All analyses were performed using STATA (Versions 6 & 7, Stata Corporation, College Station Texas). All analyses use the weighted individual respondent as the unit of observation.

We present four panels of (survey-adjusted) logistic regression models to test the formal hypotheses articulated above. Each panel of models includes six related regression estimates for a total of 24 distinct regression analyses. Each panel has six models; within that panel of six models, two models use the underclass area measure as a description of social context, two use extreme poverty areas, and two use the "Sector J" designation. Within each pair of models dedicated to a particular social context measure, one does not include a race/ethnicity variable, and the other does. A standard numbering strategy is used to distinguish amongst these 24 models. Each model is given a three-part

---

[‡]In some cases, the outcome is not a true prevalence – the existence of an attribute of an individual - but a self-report of some phenomenon, for example very easy access to cocaine.

label, of the form *a.b.c* where *a* is the panel number (1-4), *b* reflects which social environment variable that was used (1=underclass areas; 2= extreme poverty area; 3=Sector J), and *c* denotes if a race/ethnicity variable was excluded (c=1) or included (c=2).

The first panel of models (numbered 1.1.1-1.3.2) test associations of individual demographic variables and neighbourhood social environment with respect to self-reported use in the last year of one or more of crack cocaine, heroin or cocaine; this represents a test of hypothesis #1. The second panel of models (2.1.1-2.3.2) model respondents' reporting that they "Very often" observe drug selling in their neighbourhoods." The third panel of models model the likelihood that respondents will report "Very easy" access to selected illicit drugs. The second and third panel of models are provided to concurrently test the second hypothesis that neighbourhood drug activity varied with neighbourhood social context.

(We have also provided cross-tabulated evidence concerning several outcomes that we do not also submit to logistic regression. With respect to drug-related behaviours we have not reported models for lifetime needle use, nor for selling drugs, largely in the interest of parsimony; similarly we have focussed on perceived drug sales and "very easy" access to selected illicit drugs, and have not included "Very often" seeing drunk individuals because the wording of the

question uses the term "drunk" which might suggest a focus on alcohol-related impairment and not on all forms of drug-related intoxication.

Within each of these first three panels, we offer six models – half of the models include race/ethnicity variables, and the other half do not. Two of the models (one with race/ethnicity variables and one without) use the underclass areas variable, two use the extreme poverty areas variable, and finally the last two models use the SECTOR measure. The inclusion of six models within each panel is intended to allow the reader to assess the impact of race/ethnicity variables within each of the three social environment variables, and to allow users to compare and contrast the parameter estimates for Extreme Poverty Areas, Underclass Areas, and the plurality of estimates associated with the SECTOR variable.

The third hypothesis - that is, that those with the lowest measures of economic resources and the highest levels of residential mobility- those individuals in Sector J - are at particularly high risk for specific social deviance outcomes - requires the introduction of an interaction term. In essence this provides us with a means to look for a non-linear increase in risk associated with increased economic deprivation and reduced community attachment. We explore this hypothesis using the SECTOR variables as dummy variables with the addition of a single term which is the product of the Factor 1 (economic resource) and Factor 2 (community attachment) scores for the census tract.

Baseline conditions for models were: male sex, non-black and non-Hispanic race, age-range 12-17 years, and (depending on the model within the panel) residence in (one of) non-underclass areas, non-extreme-poverty areas, or "Sector E" areas[‡]. (Age, sex, and race were previously used by Lille-Blanton *et al* in their paper on contextual influences on crack cocaine use.)

Results

The SECTOR Variable

Table 4.1 provides a brief overview of the distribution of SECTOR, amongst all 1990 US census tracts, and amongst respondents to the 1993 NHSDA. Because Sector A only represents 2 of the approximately 63,000 US census tracts in 1990, and only approximately 300 persons in the entire 1990 US population reside in these tracts, any observations for this neighbourhood type are not considered in the rest of these analyses.

The number of census tracts in a sector varies widely, from 449 tracts (Sector C) to 29,199 tracts (Sector E), but each SECTOR value reflects the residential census tract of at least 1 million persons (not shown). Table 4.1 also reports on the comparative distribution of the sampled population to the populations in the census tracts, and specifically compares the proportion of the US population which is found in each sector (third row) with the proportion of

---

[‡]Sector E neighbourhoods are the most common type of neighbourhoods under the SECTOR approach.

the 1993 NHSDA respondent population (after weighting) found in that sector (fifth row). A correlation coefficient was calculated comparing the proportion of respondents from any sector and the proportion of the 1990 US population in that same sector; its value was 0.966 (not shown). Despite the high overall correlation between the distribution of (weighted) respondents to the survey and the census-derived population distribution, the table does highlight a modest apparent over-sampling in Sector E (56.4% of respondents .vs. 51.3% of population), and an apparent under-sampling in Sector G (2.3% of respondents .vs. 8.2% of population).

The number of unweighted respondents to the 1993 NHSDA varies by sector, from 244 respondents in Sector C to 12090 in Sector E. Sector J has a total of 2406 individual respondents. There were less than 1 million persons living in the 2523 census tracts that were not allocated to the SECTOR variable, representing 0.3% of the US population, and a mean population in these tracts of 330 individuals, which was approximately 1/12th the mean population per tract for 1990 census tracts. With respect to the number of respondents by Sector, Sector C has the fewest unweighted respondents, at 244, while the remainder have nearly 900 or more respondents.

Lifetime Drug Use Rates

Table 4.2 reports lifetime self-reported use of three illicit drugs: crack, cocaine, and heroin for the US population, and the rates in Underclass areas and

Extreme Poverty areas; table 4.3 reports the same measure, but is organized by the SECTOR variable. The national estimate of lifetime "crack" use was 1.8% (95%CI: 1.5%, 2.1%); the estimate for underclass areas was 3.5% (95%CI: 1.6%,7.2%), and for extreme poverty tracts 2.0% (95%CI: 1.2%, 3.1%). When organized by SECTOR, we see a 46-fold difference in lifetime crack use, with the lowest rate in Sector C (0.1%, 95%CI: 0, 0.5%) and the highest observed rate was in Sector J at 3.5% (95%CI: 2.5%, 4.9%).

Lifetime cocaine use had an estimated national prevalence of 11.3% (95%CI: 10.3%, 12.5%), with underclass areas having a rate of 6.9% (95%CI: 4.1%, 11.5%), and extreme poverty tracts having a rate of 6.2% (95% CI: 4.8%, 8.0%). The variation observed in lifetime cocaine use across the SECTOR variable was slightly less than six-fold. Rates of lifetime cocaine use ranged from 3% in Sector C, to 17.8% in Sector H. Respondents in Sector J reported 10.2% (95% CI: 7.6%, 13.4%) lifetime exposure to cocaine.

The national lifetime use of heroin was estimated at 1.1% (95% CI: 0.8%, 1.5%); the rate for underclass areas was 1.8% (95% CI 1.1%, 2.9%), and the rate in extreme poverty tracts 1.1%, (95% CI: 0.6%, 2.1%). There was a 33-fold variation in lifetime heroin use across the eight populated sectors with the highest rate of lifetime heroin use observed in Sector J with a rate of 2.5% (95% CI: 1.5%, 4.0%) .

For both crack and heroin the highest rates of lifetime drug use were reported amongst respondents living in Sector J; the 95% confidence intervals for both lifetime heroin and lifetime cocaine use amongst residents of that neighbourhood type either completely exclude or only narrowly include the estimated rates in the other neighbourhood types. The distribution of lifetime use of cocaine is substantively different from exposure to crack and heroin use, both with respect to its distribution - the highest rates are in Sector H, not J, and the overall prevalence of cocaine use is 6-7 times greater than that reported for crack and heroin.

Drug use in the Last Year

Table 4.4 reports on the rates of recent (last-year) self-reported use of three illicit drugs for the United States, and the rates in extreme poverty and underclass areas; table 4.5 reports the same rates, but stratified by the SECTOR variable. The national estimate for crack use in the last year was 0.5% (95%CI 0.4%, 0.6%); the rate in underclass areas was 1.0% (0.4%, 2.6%), while the rate in extreme poverty-tracts was 0.6% (95%CI: 0.3%, 1.0%.) When organized by SECTOR, the rates varied from 0% in Sectors C and G[‡] to 1.6% in Sector J (95%CI: 0.9%, 2.7%). A range of rates across the sectors cannot meaningfully be calculated because Sectors C and G had no respondents reporting this outcome.

[‡] When the observed rate is truly zero, confidence intervals for proportions are not reported by STATA.

For cocaine use in the last year, the national estimate was 2.2% (95%CI: 1.9%, 2.5%); the rate in underclass areas was 3.4% (95% CI 1.9%, 5.9%), and the rate in extreme poverty tracts was 1.9% (95%CI: 1.3%, 2.8%). When organized by SECTOR, the highest estimated use of cocaine in the last year was in Sector H, with a prevalence of 3.9% (95%CI, 2.8%, 5.2%) Sector J had an estimated annual rate of 3.4% (95%CI: 2.5%, 4.6%).

Minimal heroin use in the year before interview was disclosed by respondents (only 85 unweighted respondents reported heroin use); the national rate was 0.1%, with 0.3% in extreme poverty areas; underclass areas had a heroin use rate of 0.5%. When organized by SECTOR, the highest rate was observed in Sector J neighbourhoods at 0.4%.

Table 4.6. reports on the results of Model 1 – modelling self-reported use of one or more of heroin, crack cocaine, or cocaine in the last year. As model 1.1.1 shows, the dominant influences on recent reported drug use are clearly demographic, with odds ratios of 6.7 for those aged 18-25, and 5.8 for those aged 25-34, and odds ratios of 2.8 for those aged 35-49 (the baseline age range is 12-17 years). Age >50 years is highly protective of recent drug use – with an odds ratio of 0.13. Female sex is protective of recent drug use with an odds ratio of 0.4 in all six models.

With respect to social environment, underclass areas are significantly associated with increased risk, but only when race variables are excluded; this

association does not stand when race variables are added. Somewhat incongruously, extreme poverty areas are protective, with an odds-ratio of 0.74 (n.s) in model 1.2.1 and significant and with a point estimate of 0.66 when race variables are added (model 1.2.2). SECTOR does not show statistically significant associations with recent self-reported drug use, save for Sector H which is associated with an odds ratio of 1.44 ($p< 0.05$) in the no-race models (1.3.1) and is marginally significant in the race-included model(1.3.2). When SECTOR is used, race is no longer a significant determinant of these types of recent illicit drug use.

Needle use and Drug Selling

Table 4.7 reports on the rates of lifetime needle use and recent drug-selling for the US, and for underclass and extreme poverty areas are found in table; rates stratified by the Sector variable are found in table 4.8.

Lifetime needle use was an infrequently reported behaviour, reported by only 1.4% of the national population, 1.8% of respondents from underclass areas, and 1.3% amongst residents of extreme poverty areas. When organized by SECTOR, needle use are highest in Sector J at 1.9%, and lowest in Sectors B, C, and G, all with rates less than or equal to 0.7%.

At a national level, 0.8% (95%CI: 0.6%, 1%) of respondents reported selling drugs in the last year; the rate in underclass areas was 1.8%, and in extreme poverty tracts 1.5%. When organized by SECTOR, the rates vary

essentially with respect to economic deprivation - the horizontal axis, or "factor 1" of the factor analysis reported in Chapter #3. Sectors D and G (the left-most column) had rates less than 0.5%, Sectors B, E and H (the middle column) had rates greater than 0.5% and less than 1%, while Sectors C, F and J (the right-most column) had rates in excess of 1% (C: 1.7%; F: 1.6%, and J: 1.6%) A logistic-regression model for this outcome is not reported here.

Neighbourhood Drug Activity

Neighbourhood drug-selling and the presence of intoxicated individuals are reported for the US population, and for poverty-areas and for underclass areas in table 4.9; results stratified by SECTOR are found in table 4.10. The national proportion of respondents reporting that drug-selling was observed in their neighbourhoods "very often" was 5.6% (95%CI: 4.8%, 6.4%); the rate in underclass areas was 46.7%, and in extreme poverty tracts, the rate was 24.9% (95%CI: 18%, 33.4%). When organized by SECTOR, rates varying from slightly less than 1% in SECTOR D to 32% in Sector J.

When examining the rates across sectors, a semi-quantitative assessment of the proportions suggests that the prevalence of very-frequent sightings of such individuals increases as one moves down each column, and across each row. For example, in the right-most column the upper bound of the distribution for Sector C (0.26) is just slightly less than the lower bound of the estimated distribution in Sector F (0.27), and the upper bound for Sector F (0.356) is less than the lower

bound for Sector J (0.375); similar, but slightly less consistent and pronounced trends observations can be made in the middle column, and for the contrast between Sectors D and G. These observations support the impression of a trend from left to right and top to bottom in the measurements as stratified by Sector

The organization of respondent's reports of "very often" observing drug-selling in their neighbourhood is, perhaps not surprisingly, very similar to reports of frequently observing intoxicated individuals in the respondent's neighbourhood. These rates are reported in tables 4.12 (for nation, underclass areas and extreme poverty areas) and 4.13 (by Sector). Nationally 11.5% (95%CI: 10.5%, 12.6%) of respondents report that intoxicated individuals are observed "very often" in their neighbourhoods. The rate of these observations in underclass areas was 49.3%, and the rate in extreme poverty tracts was 41.1% (95%CI: 34.8%, 47.8%) When stratified by SECTOR, the lowest rates were observed in Sector D (3.5%) and the highest rate (43%; 95% CI: 37.4%, 49.3%) was observed in Sector J. There is clear evidence of a trend in these rates in both axes of the SECTOR variable; in all five of the "horizontal contrasts" (eg. Sector D .vs. E, or G. .vs. H, etc.) the point estimate for the right adjacent cell is equal to, or exceeds, the upper 95% CI of the left adjacent cell. With respect to the vertical axis, in four of the five "vertical" contrasts (e.g. D .vs. G, or B .vs. E) the point estimate is greater in the lower adjacent cell than the upper bound of the distribution from the higher adjacent cell.

Model 2 (table 4.11) presents the results of logistic regression modelling of the "Very Often" response to a question about observing drug selling in the respondent's neighbourhoods. Again, the panel is composed of six models, with three sets of two model, one with a race/ethnicity variable, and one without.

These models contrast with the models of personal drug use presented above; they are characterized first and foremost by large odds ratios for each of the three social environment variables, odds ratios that remain large even when race variables are included; race/ethnicity variables are also associated with large, statistically significant estimates. Increasing age is associated with decreasing risk of observing drug sales. Being female is associated with an increased risk of this outcome.

The race/ethnicity of the respondent plays an important and independent role in the perception of "very often" observing drug selling in neighbourhoods, across all three social environment variables, and regardless of whether or not race/ethnicity variables are included. Black race/ethnicity is associated with odds ratios of 6-9 across the three models (2.1.2, 2.2.2, 2.3.2), while Hispanic race is associated with an odds ratios ranging from 2.5 to 3.9, again depending on the social environment variable under study.

With respect to the role of social environments in organizing increased the likelihood of "very often" observing drug sales, underclass areas are associated with an odds ratio of 15.9 when race is not included in the model, and retains a

7-fold increase when race variables are included. Extreme poverty-areas show 5.8-fold differences from non-extreme-poverty areas when race is not included, and 2.7 when race is included.

With respect to the SECTOR variable, Sector D is clearly protective of drug-selling, with odds ratios between 0.2 and 0.3, and Sector F is clearly at heightened risk, with odds-ratios of 5.5 (no race variables) and 2.7 (with race variables). Sector H is associated with an odds ratio of 1.5 when race variables are included, and Sector J residents are at the highest risk (amongst the sectors) for this outcome, with odds-ratios of 10.4 when race is not included and 5.0 when race variables are added.

Drug Availability

Table 4.12 reports on the proportion of respondents who reported that they had "Very Easy" availability to four distinct illicit drugs: cocaine, heroin, LSD, and PCP, by Sector; the corresponding rates for the country, for underclass and extreme poverty areas are reported in table 4.13.

Nationally, 20% of respondents reported "very easy availability" for cocaine or crack, while approximately 10% reported very easy access to heroin, LSD, or PCP. In extreme poverty areas, the rate was approximately 50% higher, with 30% of respondents reporting ready availability for cocaine, and mid-teen rates for the other drugs. Underclass areas were associated with a doubling to tripling of the national rates. When the SECTOR variable is applied, both

apparently protective and at-risk areas emerge. Sector D, generally speaking, appears to have lower rates than the Sector E, and there appears to be a general trend toward higher rates as one moves down columns, or to the right across rows. Sector F and J have markedly higher rates of respondent reported ready availability for all drug-specific uses; this is particularly true respect to Sector J where the lower bound of the confidence interval (0.347) typically excludes the drug-specific upper bound of the confidence intervals for Sector E (0.221).

Results from model 3 are presented in Table 4.14 The models show that younger age (<35 years) is almost uniformly associated with higher perceived ease of access to these drugs, except for the oldest age group. This group reports a reduced likelihood of easy access, as compared with the baseline age. Across all models, female sex is associated, at least at the $p < 0.1$ level, with a slight reduction (0.85) of perceived availability of these drugs. With respect to the social environment variables, model 3.1.1 and 3.1.2 suggest that underclass areas are linked to an increased odds ratio, (1.8-3) while extreme poverty conveyed an increased risk only when race variables were not included (OR: 1.58). The SECTOR variables show a consistent and significant protective effect (at least at the $p<0.1$ level) for Sectors C and D, and increased risk for Sector F (1.5-2) and Sector J (1.7-2.5).

Non-linear effects in Sector J

Exploratory evidence concerning the third hypothesis - that the combination of economic deprivation(factor 1) and low community attachment (factor 2) would be associated with higher adverse outcomes rates - is reported in Table 4.15. Two models are presented; both models report parameter estimates for the interaction term that are greater than 1; for "very often" observing local drug selling (model 4.1) the parameter estimate for this interaction term is 1.36 (95%CI 1.19, 1.54), while the parameter estimate for model 4.2 is 1.10 (95%CI 0.98, 1.23). This parameter estimate is multiplied by the product of factors 1 and 2 for census tracts found in Sector J; the product of those two factors has a mean of 3.16 (not shown) amongst Sector J neighbourhoods, suggesting a marked interactive effect once the beta estimate are multiplied by the mean value of the interaction term.

Discussion

In this paper we have explored the utility of adding measures of social environment to a nationally representative survey of illicit drug use. Three measures of social environment were considered: the extreme poverty areas measure, the underclass areas measure, and the "SECTOR" variable linked to social disorganization theory. It is important to highlight that all three measures were developed before linkage to the NHSDA data, and that this exercise represents a test of the validity of measures as organizing devices for understanding some drug-related phenomena.

We review our results in five general areas: personal drug use, ecological aspects of drug use, drug access, the apparent contrast between personal drug use and ecological aspects of drug economies, and finally offer some specific comments with respect to the SECTOR variable and its links to social disorganization theory.

Personal Drug Use

With respect to personal drug use - lifetime or within the last year - we considered both self-reported lifetime and self-report recent drug use of three or four serious illicit drugs, and the role that social environments might play in organizing those variables. With respect to lifetime use of selected illicit drugs, the cross-tabulatory evidence was, at best, inconsistent and where associations appeared to exist, they were quantitatively weak. Lifetime use of crack was elevated in Sector J, and in Underclass neighbourhoods, but lifetime heroin rates were at most doubled in underclass and Sector J areas. For cocaine, residence in an extreme-poverty areas and underclass areas was apparently protective. In all cases the estimates are based on very small numbers - both because of the relative infrequency with which individuals report these behaviours and the geographic stratification arising from our focus on social groupings strata which represent very small proportions of the population.

The results of the Model #1 logistic regression suggest the following characterization of recent drug use of these three classes of drugs: late

adolescence and early adulthood predominates in terms of odds ratios, female sex is protective, and black race implies a modestly increased risk (OR=1.2-1.4); the black-race variable varies in its statistical significance, depending on the social environment variable used. While Underclass areas and Sector H are at statistically higher risk for these outcomes, extreme poverty areas are estimated to be protective in both race-inclusive and race-exclusive models.

We suggest that the evidence is only modestly supportive (for Underclass and Sector J neighbourhoods) of hypothesis #1 that the addition of social environment variables will provide new information on the organization of personal drug use. This conclusion must be qualified by the caveat that there remains considerable potential for unspecified error and bias - particularly in sampling and measurement- to be confounding the observed results. We will discuss some potential confounding influences in a later section of this discussion.

Ecological Aspects of Drug Use

When the analysis turns to reports on neighbourhood drug activity, the modest and equivocal differences seen in the cross-tabulations of social environments and personal drug-use are replaced by much clearer and larger differences With respect to observing drug-selling or intoxicated individuals, the cross-tabulated evidence shows that underclass and extreme poverty neighbourhoods are associated with odds ratios of 4-8, as compared with the national average. When organized by SECTOR, the

anticipated gradients are clear – rates increase for both outcomes when going from left to right and from top to bottom, with the highest rates observed in Sector J, and the next-highest rates in Sectors F and H. This finding is particularly notable because the baseline condition incorporates all but the most extreme response from various 5-point Likert scales. The findings are also consistent with social disorganization theory and its argument that both economic deprivation and decreased community attachment will result in higher rates of social deviance.

When examined using logistic regression, and in contrast to reports of personal drug-use, younger age had only a modest impact on perceptions of the frequency of drug selling and intoxicated individuals in the respondent's neighbourhood. Age greater than 25 years was associated with an ever-reducing perception of drug selling; black and hispanic race/ethnicity markedly increase the perception of very frequent drug-selling. With respect to social environments, Underclass Areas, and Sector J areas were associated with odds ratios >10, before race was entered into the model, and odds ratios > 5 thereafter. The association with Extreme Poverty areas was weaker – estimated at 5.7 in the non-race models and 2.7 in the race-included model. Sector D is statistically significant as protective of this outcome in both models, while Sectors F, H and J, (as would be predicted from social disorganization theory) are at elevated risk for the outcome, again in both race-excluded and race-included models.[‡]

[‡](It is interesting to speculate why race should be so markedly associated with the *perception* of local drug-sales - an outcome that one might expect to be relatively race-independent. Our analyses do not answer this question, but suggest the intriguing hypothesis that local drug economies may have race-specific components – perhaps

Clearly social ecology variables could be employed to further explore these relationships. An attractive candidate variable would be segregation - however it is important to note that extant measures of segregation were not suited to this investigation because it would require the production of national segregation indices. Segregation indicies are, however, typically constructed with respect to municipal or regional population distributions. We are not aware of any national data resources providing estimates of segregation for all US census tracts at this time. That said, regional studies could well make use of this variable to better understand the social organization of illicit drug use.

Drug Access

The cross-tabulations with respect to drug-access tell a similar story to the analysis of local drug-sales. From these tables, we can see that residents of underclass and extreme poverty areas are anywhere from 1.5-3 times as likely to report very easy access to each of these four illicit drugs. Residents of Sector J neighbourhoods are also roughly twice as likely to report "very easy" access to these drugs, as compared with residents of Sector E.

The logistic regression model (#3) combines easy access to each specific drug into a single variable indicating easy access to at least one of these drugs. The results continue to point to an important effect of race (with odds ratios of 2.7-3 for black race and approximately 1.5 for Hispanics) associated with this outcome. Residents of

---

blacks and hispanics are approached more frequently than others to buy drugs, or perhaps blacks and hispanics are more aware of the signs of drug sales activity.)

underclass areas had an odds ratio of 3, when race was not included, and odds ratios of 1.8 when race was included. The modest (1.5) and statistically significant effect associated with extreme-poverty areas disappears when race is added. However, the protective effect of Sector C, and the increased risk associated with Sectors F (2) and J (2.5) remain, but are diminished, after the inclusion of race-related variables.

With respect to Hypothesis #2, the analysis of self-reported "very easy" access to drugs, and "very frequent" observation of drug sales in underclass areas, and Sector J, and the concurrent evidence of a protective effect in Sector B, C, or D neighbourhoods provide evidence that is strongly supportive of the hypothesis that social environments modify ecological aspects of drug activity, and more specifically that SECTOR and its underlying bivariate theoretical framework are important to understanding these social behaviours. Evidence to support this hypothesis for Extreme Poverty areas is less consistent and quantitatively more modest. We conclude that household poverty alone is an inadequate measure of social environment to explain these outcomes.

Comparing Personal and Ecological Reports of Drug Activity.

The results present above yield a consistent impression – that specified social environments (notably, Underclass areas and Sector J neighbourhoods) are associated with higher rates of neighbourhood drug activity. The results also present a puzzling contrast: residents of Sector J (and to a lesser extent Sectors F and H), Underclass and Extreme Poverty areas all report rates of drug sales, and access to illicit drugs which

are at least four times the national average. However, respondents from these neighbourhood types report only very modest differences in recent personal drug use from the national samples. This raises the puzzling question of why the respondents appear to have only modestly increased rates of drug use, but report the areas they live in to clearly have markedly increased illicit drug activity. Simply put: who is buying and using these drugs, if not the residents of the neighbourhoods in which the drug sales are most common?

We are not alone in our inability to find foci of active illicit drug use in representative survey data. Friedman, *et al* have also had difficulty in finding endemic areas for current drug use or foci of lifetime drug users. (15) Using data from a household probability sample in Bushwick, Brooklyn, this group found modest rates of *lifetime* cocaine (9%) and heroin use (3%), results that are very similar to our observations for Sector J neighbourhoods of 10.2% for *cocaine* use, and 2% for *lifetime* heroin use. Fullilove *et al* in the Harlem Health Study, have documented a 14% lifetime crack-cocaine use and crack-use in the last month by 4.6% of the adult population; we report a rate of 3% for current crack use amongst respondents in our Sector J neighbourhoods.(16) Ensminger *et al* have reported "few differences between the lifetime prevalence of drug use [of an at-risk sample] and a nationally representative sample of adults of the same age range.(17) Our results also conform with the recent observations of Sake *et al*, using a different year of the NHSDA, that local drug visibility does not predict higher rates of drug use amongst residents.(18)

Several possibilities arise with respect to these data. First, it is possible that the results are correct, and that residents of these areas are, in fact, modest users of illicit drugs, despite the active drug economies around them; this would suggest that users come from outside the neighbourhood to purchase and use these drugs. Another possibility is that respondents are not forthright in their reports of recent personal drug use, under-reporting the true rates of their personal participation in these socially proscribed activities. A third possibility is that respondents are not representative of the drug-user patterns of the populations which live in these areas. A fourth and final possibility is that the survey design – focusing as it does on household-living individuals – necessarily excludes those most impaired by drug use and consequently least likely to reside in households.

Gfroerer(19), proposes four general areas of error in household surveys: coverage, sampling, non-response and measurement. It is not difficult to imagine that these sources of bias and error are all acting in Sector J, and Underclass neighbourhoods. Kandel offers additional options when she writes:

> "it must be emphasized, however, that surveys based on households ...generally exclude the individuals most likely to be involved in nonconforming activities, including drug use: those without regular addresses, the homeless, .. Those living in institutions... these deviant individuals constitute presumably a relatively small proportion of the general population, however, their exclusion does not significantly bias the overall epidemiologic

> estimate reported..., although the less frequently used drugs and heaviest patterns of use may be under-represented. Another limitation of these data is that self-reports are sensitive behaviours, such as drug use, may be subject to reporting bias, which is not randomly distributed throughout the population. In particular blacks appear to be more likely than other ethnic groups to underreport their infrequent use of illicit drugs." ((20), p. 371)

Kandal's comments on the small quantitative influence that might arise from excluding these populations is generally correct, but the quantitative impact is likely to be larger and the bias more pronounced when results are stratified by variables designed to highlight areas of social deviance and social disorganization.

It is also important to note that Gfroerer's list of sources of error in surveys might be augmented by Kandal's observation that the sampling frame itself – in this case households – erodes the capacity of the survey to be representative of the complete population residing in an area, notably excluding those who are not in households, including those in shelters, abandoned housing, and the homeless. Similar caveats have been raised by Boboshev and Anthony(21) in their analysis of the NHSDA. While time of year and southern geography may both play a role in homelessness, the NHSDA survey was conducted across the whole year; the limited number of non-baseline responses to many of our outcomes measures would not support disaggregation by season or climate.

The use of households in the sampling frame represents a technical constraint of multi-stage designs – this design cannot be implemented without some sort of residential unit which can be randomly selected. The potential of a survey to exclude important populations because of the focus on households does not yield a recognized error associated with survey methods – simply surveys are designed to provide information on the universe of interest and non-household dwelling individuals are out of the scope of household-based surveys. However, not being representative of a population larger than the stated universe of interest is not a source of error for a survey from a methodological point of view, it does represent a significant impediment from a public policy perspective. Additional research is needed to determine whether these traditional and novel forms of bias, error and non-response are organized by social environments.

William Kornblum makes an important historical and political observation that is relevant to the interpretation of our data. He argues that the localization of deviance - which we have documented here for drug economics and drug access - represents the working out of a specific public policy strategy - ghetto containment of deviance. "Although never an avowed feature of public policy, the containment of market transactions for drugs and other deviant goods and services rested on a pattern of collusion between corrupt authorities and illegal merchants. ((22), p. 423) Elsewhere Kornblum writes: "This century's history of ghetto containment of deviance - a specific feature of institutional racism - has compounded the extreme

social and psychological difficulties of minority persons, specially the African-American male. It has stimulated the creation of thriving illicit markets and illicit economic institutions in the nation's racial ghettos and has gone far to produce a population debilitated by addictive drugs, alcohol and fratricidal violence." (p. 431) Our results - and specifically the concentration of drug markets into underclass and Sector J neighbourhoods - certainly provides quantitative evidence for the localization of illicit drug markets into a highly limited set of neighbourhoods; potential methodological limitations, and biases, may inform the apparent discordance between personal drug use findings and neighbourhood drug-activity. (This work might usefully be replicated using more recent NHSDA surveys which use computer-assisted interviewing techniques to reduce respondent bias and to increase accurate reporting of proscribed behaviours.[‡])

Social Disorganization

Overall, we believe that our "SECTOR" approach has performed satisfactorily – in providing a useful heuristic for classifying social environments, and in demonstrating the utility of social environment variables in organizing the distribution of these social phenomena. The focus of drug activity is clearly in Sector J, and Sectors B, D and sometimes C were often protective for outcomes for which Sectors J, F and H represent increased risk. This is entirely consistent with social disorganization theory which suggests that both community attachment and economic

[‡] Development of Computer-Assisted Interviewing Procedures for the National House [sic] Survey on Drug Abuse (BKD397). Copyright information to follow.

deprivation matter for these outcomes. Of particular note are the rates in Sector C, which for almost all outcomes were either comparable with the national rates or lower than the national rates – this even though these neighbourhoods are equivalent with respect to economic deprivation, as measured by F_1. These findings challenge the assertion that neighbourhood poverty is the sole determinant of drug-use rates or drug-activity rates; the results from Sector C suggest that some extremely poor neighbourhoods are able to control patterns of illicit drug use and drug activity quite effectively.

The importance of residential mobility – a key component of our F_2 measure – to the initiation of drug activity has recently been supported by DeWit in their analysis of data from the Ontario Mental Health survey, for adolescents - the chief risk group of initiants to illicit drug-activity.(23) Taken together these data suggests that neighbourhood level mobility has a strong association with drug use, and in focusing on adolescents, DeWit has begun to make an argument for an etiologic role for residential mobility in drug initiation; however, both the NHSDA and the Ontario Mental Health Survey data are cross-sectional, restricting the potential for etiologic assertions.

Finally, the identification of a significant interaction term for the product of Factor 1 and Factor 2 for two outcomes: self-reported very easy access to heroin, LSD and PCP, and "very often" seeing neighbourhood drug sales represents intriguing evidence that Sector J neighbourhoods may be at particular risk for social deviance -

at least more than would be expected from a linear interpolation across the SECTOR matrix. Important additional work will be required to determine if this effect can be replicated with other specifications of the interaction term, and to determine if the interaction term is simply capturing variance which is lost in the 3*3 classification matrix which informs the SECTOR variable.

General Comments

Finally, we would note that the objective of this chapter was not so much to characterize the factors which encourage drug taking, or even the neighbourhood ecology of drug use, but rather to propose classifications of neighbourhoods and test whether specific neighbourhoods are areas of frequent drug activity. In this regard, the Underclass measure has generally been more successful than the Extreme Poverty measure in finding such areas. Sector J neighbourhoods, overall, appear to have similar characteristics to Underclass areas in terms of odds ratios and rates at which various drug-related behaviours are reported. However, being polytomous, the Sector variable also allows researchers to find other areas of elevated risk (Sectors F and H) and areas of reduced risk, notably Sectors B, C, D. The logistic regression models for neighbourhood drug activity are supportive of the argument that the observed levels of neighbourhood drug activity are not simply contextual – but contextual or not, with the cross-tabulated evidence there is now limited room to challenge the existence of focussed deviance markets with respect to these illicit drug activities. The evidence presented here establishes the value of both the Underclass or Sector approaches in

identifying areas at risk for very active neighbourhood drug activity, with Sector showing important gradients and interaction terms. Moreover, Sector identifies areas (Sector C in particular) which appear to be at least somewhat protected from these outcomes, despite overwhelming economic deprivation

Conclusions

The evidence presented above suggests that a variety of patterns of drug use vary systematically with measures of social environment, and particularly so with respect to SECTOR, a variable derived from social disorganization theory. A striking pattern of high prevalence of drug-related activities was specifically observed in Sector J neighbourhoods – neighbourhood characterized by economic deprivation and low levels of community attachment. Alternatively, sectors "advantaged" by either economic affluence or high levels of community attachment showed lower levels of drug-related activity. These results argue for the independent effect of economic deprivation and community attachment on these outcomes.

While no causal role from the influence of Sector J residence on drug use can be inferred from cross-sectional data, it seems reasonable to assert that Sector J neighbourhoods are areas with very active drug economies, including drug sales of hard street drugs such as crack, PCP, and heroin, and are foci of active drug use. Evidence that residents of Sector J are the users of the drugs being sold in these neighbourhoods is less clear, the interpretation clouded by the significant practical problems of undertaking surveys in these neighbourhoods, the methodologic issues

raised by the survey methods, and the challenges of ensuring survey quality for socially and legislatively proscribed activities in marginalized populations.

Finally, our approach has shown that existing representative surveys can be augmented with meaningful contextual variables in a way which preserves the confidentiality of respondents and the methodological integrity of the survey, while substantially enhancing analytic capacity.

Figures

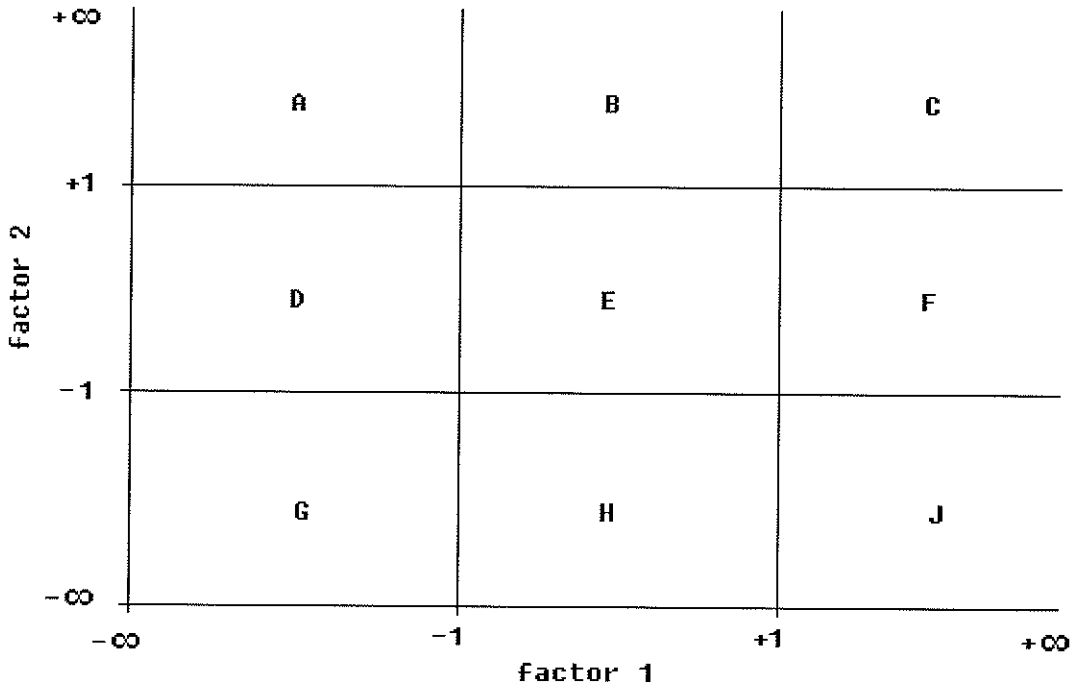

Figure 4.2 Schematic representation of SECTOR Model.

Tables

Table 4.1 Comparative distribution of 1990 US population (from the 1990 US census) and respondents to the 1993 National Household Survey on Drug Abuse, by three social environment variables: Sector, Extreme Poverty and Underclass Area.

Social Environment Variable	Sector										Underclass	Extreme Poverty
	A	B	C	D	E	F	G	H	J	Missing		
Number of Census Tracts in 1990 by Social Environment Variable		8330	449	6566	29199	5710	1678	4882	2049	2523	928	3413
Percentage of all 1990 US Census Tracts by Social Environment Variable		13.6	0.7	10.7	47.6	9.3	2.7	8	3.3	4.1	1.5	5.7
Percentage of total 1990 US population residing in specified social-environment		12.3	0.6	12.5	51.3	8.2	8.3	8.9	2.8	0.3	1.1	4.2
Number of Respondents to 1993 NHSDA		899	244	3149	12090	2447	894	4271	2406	0	536	1979
Percentage of NHSDA Sample (Weighted)		8.6	1.4	13.7	56.4	5.3	2.3	9.9	2.6	0	0.8	3.4

N.B.: No values are reported for Sector A, because of the small number (2) of census tracts which were identified in this Sector in 1990.

Table 4.2 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting lifetime use of crack, cocaine and heroin, national results, and for Underclass Areas and Extreme Poverty areas.

Social Environment Variable / Drug	National	Underclass	Extreme Poverty
Crack	.018 (0.015,0.021)	.035 (.016, 0.072)	.02 (0.012, .031)
Cocaine	.113 (.103, .123)	.069 (.041, .115)	.062 (048, .08)
Heroin	.011 (.008, .015)	.018 (011, .029)	.011 (.006, .021)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.

Table 4.3 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting lifetime use of crack, cocaine and heroin, by SECTOR.

Sector	A	B	C
Crack	.	.009 (.004, .02)	.001 (.0001, .005)
Cocaine	.	.071 (.049, .102)	.030 (.014, .065)
Heroin	.	.004 (.001, .014)	.001 (.894, .006)
Sector	**D**	**E**	**F**
Crack	.014 (.01, .022)	.018 (.014, .022)	.025 (.014, .043)
Cocaine	.121 (.099, .146)	.109 (.096, .123)	.088 (.056, .136)
Heroin	.01 (.005, .020)	.011 (.007, .017)	.011 (.005, .024)
Sector	**G**	**H**	**J**
Crack	.012 (.005, .029)	.026 (.017, .04)	.035 (.025, .049)
Cocaine	.165 (.117, .229)	.178 (.137, .228)	.102 (.076, .134)
Heroin	.018 (.004, .087)	.016 (.011, .023)	.025 (.015, .04)

Notes:

1. Cell contents are estimated proportions. 95% confidence intervals are reported in parenthesis after the proportion.

Table 4.4 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting recent use of crack, cocaine and heroin: national results and results for Underclass Areas and Extreme Poverty areas.

Social Context / Drug	National	Underclass	Extreme Poverty
Crack	0.005 (0.004, 0.006)	0.01 (.004, 0.026)	0.006 (0.003, 0.01)
Cocaine	0.022 (0.019,0.025)	0.034 (0.019,0.059)	0.019 (0.013, 0.028)
Heroin	0.001 (0.001,0.002)	0.005 (0.002, 0.015)	0.003 (0.001, 0.007)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.
2. Recent use is defined as within one year of the survey.

Table 4.5 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting recent use within of crack, cocaine and heroin, by SECTOR.

Sector	A	B	C
Crack	.	0.002 (0.0, 0.012)	0
Cocaine	.	0.016 (0.01, 0.027)	0.015 (0.004, 0.052)
Heroin	.	0.00	0
Sector	**D**	**E**	**F**
Crack	0.002 (0.001, 0.004)	0.005 (0.004, 0.008)	0.008 (0.004, 0.017)
Cocaine	0.015 (0.009, 0.023)	0.021 (0.017, 0.026)	0.023 (0.016, 0.032)
Heroin	0	0.001 (0.001, 0.002)	0.003, (0.001, 0.007)
Sector	**G**	**H**	**J**
Crack	0	0.007 (0.003, 0.012)	0.016 (0.009, 0.027)
Cocaine	0.016 (0.006, 0.038)	0.039 (0.028, 0.052)	0.034 (0.025, 0.046)
Heroin	0.001 (0,0.003)	0.004 (0.001, 0.009)	0.001 (0.001, 0.002)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.
2. Recent is defined as use within one year of administering the survey.

Table 4.6 Odds ratios for self-reported recent use of either crack or cocaine, or heroin arising from six survey-adjusted logistic regression models using respondents to the 1993 National Household Survey on Drug Abuse. (Models vary with respect to inclusion of race/ethnicity variables and social environment variable).

	Demographic Variables							Social Environment Variables								
	Age				Sex	Race/Ethnicity		Under-	Extreme	Sector						
Model	18-25y	26-34y	35-49y	50+y		Black	Hispanic	class	Poverty	B	C	D	F	G	H	J
1.1.1	**6.67**	**5.81**	**2.84**	**0.13**	**0.41**			1.63								
1.2.1	**6.71**	**5.89**	**2.9**	**0.13**	**0.41**	**1.39**	1.21	1.39								
1.2.1	**6.72**	**5.79**	**2.81**	**0.13**	**0.42**				0.74							
1.2.2	**6.78**	**5.87**	**2.88**	**0.13**	**0.41**	**1.45**	1.25		**0.66**							
1.3.1	**6.22**	**5.66**	**2.82**	**0.12**	**0.41**					0.86	0.70	0.68	1.12	0.61	**1.44**	1.38
1.3.2	**6.29**	**5.72**	**2.86**	**0.13**	**0.41**	**1.27**	1.11			0.88	0.69	0.69	1.03	0.62	1.41	1.24

The outcome variable modelled is self-reported use of either crack, or cocaine or heroin within one year of the 1993 National Household Survey on Drug Abuse. Cell contents are odds ratios arising from logistic regression modelling. The six models differ with respect to the independent variables included. Models 1.1.1, 1.2.1 and 1.3.1 do not include race/ethnicity terms, whilst models 1.1.2, 1.2.2, 1.3.2 include these covariates. Models 1.1.1 and 1.1.2 use the Underclass Areas measure to represent social environments; model 1.2.1 and 1.2.2 employ the Extreme Poverty Areas measure; models 1.3.1 and 1.3.2 use the Sector strategy for encoding social environment variables.

Statistical significance: **Bold** $p<0.05$; Underline $p<0.1$

Table 4.7 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting lifetime needle use, or drug-selling in 12 months prior to survey; national results and results for Underclass Areas and Extreme Poverty areas.

Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.

Social Environment Context Drug -Related Activity	National	Underclass	Extreme Poverty
Lifetime Needle Use	0.014 (0.012, 0.017)	0.018 (0.009, 0.038)	0.013 (0.007, 0.025)
Sold Drugs in last year	0.008 (0.006, 0.01)	0.011 (0.006, 0.021)	0.015 (0.008, 0.026)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.

Table 4.8 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting lifetime needle use, or drug-selling in the 12 months prior to the survey, by SECTOR.

Sector	A	B	C
Lifetime Needle Use Sold Drugs in last year		0.007 (0.003,0.017) 0.006 (0.003,0.014)	0.006 (0.001,0.033) 0.017 (0.005,0.055)
Sector	**D**	**E**	**F**
Lifetime Needle Use Sold Drugs in last year	0.009 (0.986,0.995) 0.004 (0.002,0.006)	0.016 (0.013,0.021) 0.007 (0.005,0.010)	0.013 (0.006,0.029) 0.016 (0.010,0.028)
Sector	**G**	**H**	**J**
Lifetime Needle Use Sold Drugs in last year	0.005 (0.002,0.012) 0.001 (0.001,0.003)	0.017 (0.011,0.028) 0.010 (0.006,0.017)	0.019 (0.012,0.031) 0.016 (0.006,0.029)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.

Table 4.9 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting "Very Often" seeing drug selling and drug or alcohol-related intoxication in their neighbourhood: national results and results for Underclass Areas and Extreme Poverty areas.

Social Context / Drug	National	Underclass	Extreme Poverty
Frequency of seeing drunk/high people in neighbourhood	0.115 (0.105, 0.126)	0.493 (0.377, 0.610)	0.411 (0.347, 0.478)
Frequency of seeing people selling drugs in neighbourhood	0.056 (0.048, 0.056)	0.467 (0.353, 0.585)	0.249 (0.188, 0.344)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.

Table 4.10 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting "Very Often" seeing drug selling and drug or alcohol-related intoxication in their neighbourhood, by SECTOR.

Sector	A	B	C
Drunk		.09 (.059, .135)	.15 (.082, .261)
Selling Drugs		.029 (.018, .05)	.033 (.016, .067)
Sector	**D**	**E**	**F**
Drunk	.035, (.026, .048)	.098 (.084, .113)	.306 (.260, .356)
Selling Drugs	.0097 (.005, .018)	.041, (.031, .056)	.197 (.155, .247)
	G	**H**	**J**
Drunk	.08 (.037, .170)	.17 (.139, .199)	.43 (.375, .495)
Selling Drugs	.047 (.021, .101)	.085 (.067, .107)	.32 (.273, .382)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.

Table 4.11 Odds ratios for self-reported "Very Often" seeing drug selling and drug in respondent's neighbourhoods arising from six survey-adjusted logistic regression models using respondents to the 1993 National Household Survey on Drug Abuse. (Models vary with respect to inclusion of race/ethnicity variables and social environment variable).

	Demographic Variables							Social Environment Variables								
	Age				Sex	Race/Ethnicity		Under-	Extreme	Sector						
Model	18-25y	26-34y	35-49y	50+y		Black	Hispanic	class	Poverty	B	C	D	F	G	H	J
2.1.1	1.07	**0.74**	**0.66**	**0.35**	**1.23**			**15.90**								
2.1.2	1.12	**0.81**	0.77	**0.45**	1.19	**8.98**	**3.86**	**6.92**								
2.2.1	0.93	**0.78**	**0.68**	**0.37**	**1.22**				**5.77**							
2.2.2	1.08	0.83	0.78	**0.46**	1.19	**8.62**	**3.67**		**2.70**							
2.3.1	0.91	**0.74**	**0.71**	**0.37**	**1.25**					0.71	0.78	**0.22**	**5.47**	1.14	**2.00**	**10.36**
2.3.2	1.02	0.80	0.78	**0.43**	1.18	**6.09**	**2.52**			0.94	0.67	**0.27**	**2.70**	1.23	**1.58**	**4.97**

The outcome variable modeled is self-reported "Very Frequently " observing drug sales in their own neighbourhoods from the 1993 National Household Survey on Drug Abuse. Cell contents are odds ratios arising from logistic regression modeling. The six models differ with respect to the independent variables included. Models 2.1.1-2.1.3 do not include race/ethnicity terms, whilst models 2.2.1-2.2.3 include these covariates. Models 2.1.1 and 2.2.1 use the Underclass Areas measure, and models 2.1.2 and 2.2.2 use the SECTOR variable; models 2.1.3 and 2.2.3 use the Extreme Poverty social environment variable.

Statistical significance: **Bold** $p<0.05$; Underline $p<0.1$

Table 4.12 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting "Very Easy" access to PCP, LSD, Heroin, or Cocaine: national results and results for Underclass Areas and Extreme Poverty areas.

Context / Drug	National	Underclass	Extreme Poverty
Cocaine or Crack	0.2 (0.186, 0.215)	0.419 (0.289, 0.561)	0.292 (0.212,0.386)
Heroin	0.119 (0.107, 0.132)	0.301 (0.207, 0.415)	0.174 (0.125,0.237)
LSD	0.114 (0.104, 0.124)	0.244 (0.167, 0.342)	0.139 (0.106,0.180)
PCP	0.099 (0.09, 0.109)	0.243 (0.160, 0.351)	0.128 (0.090,0.179)

Notes:

1. Cell contents are proportions; 95% confidence intervals are reported in parentheses after the proportion.

Table 4.13 Weighted proportion of respondents to 1993 National Household Survey on Drug Abuse self-reporting "Very easy" access to PCP, LSD, Heroin, or Cocaine, by SECTOR.

Sector	A	B	C
Cocaine or Crack		0.189 (0.134, 0.259)	0.179 (0.134, 0.237)
Heroin		0.121 (0.086, 0.169)	0.088 (0.056, 0.136)
PCP		0.119 (.084, .167)	0.083 (0.047, 0.135)
LSD		0.139 (0.094, 0.199)	0.083 (0.044, 0.152)
Sector	**D**	**E**	**F**
Cocaine or Crack	0.163 (0.138, 0.191)	0.201 (0.183, 0.221)	0.358 (0.296, 0.426)
Heroin	0.097 (0.071, 0.13)	0.119 (0.103, 0.137)	0.207 (0.15, 0.28)
PCP	0.078 (0.06, 0.099)	0.099 (0.085, 0.114)	0.143 (0.093, 0.213)
LSD	0.099 (0.08, 0.121)	0.113 (0.1, 0.129)	0.15 (0.101, 0.219)
Sector	**G**	**H**	**I**
Cocaine or Crack	0.198 (0.141, 0.270)	0.206 (0.175, 0.242)	0.411 (0.347, 0.479)
Heroin	0.094 (0.046, 0.184)	0.121 (0.098, 0.149)	0.289 (0.232, 0.355)
PCP	0.092 (0.045, 0.178)	0.104 (0.081, 0.132)	0.256 (0.205, 0.315)
LSD	0.102 (0.051, 0.195)	0.119 (0.096, 0.147)	0.255 (.208, 0.309)

Notes:

1. Cell contents are proportions of the respondent population reporting very easy access to the indicated drug; 95 % confidence intervals are reported in parentheses after the proportion.

Table 4.14 Odds ratios for respondents' self-reporting "Very ease" access to one or more of cocaine, PCP, LSD or heroin, arising from six survey-adjusted logistic regression models using respondents to the 1993 National Household Survey on Drug Abuse. (Models vary with respect to inclusion of race/ethnicity variables and social environment variables.).

	Demographic Variables							Social Environment Variables							
	Age				Sex	Race/Ethnicity		Under-	Extreme	Sector					
Model	18-25y	26-34y	35-49y	50+y		Black	Hispanic	class	Poverty	B	C	D	F	G	H	J
3.1.1	**1.46**	**1.42**	1.14	**0.67**	0.87			**2.97**								
3.1.2	**1.51**	**1.5**	**1.22**	**0.74**	**0.85**	**3.02**	**1.41**	**1.83**								
3.2.1	**1.43**	**1.43**	1.15	**0.67**	0.87				**1.58**							
3.2.2	**1.5**	**1.50**	**1.22**	**0.74**	**0.85**	**3.06**	**1.42**		1.05							
3.3.1	1.44	**1.46**	**1.18**	**0.68**	0.87					0.89	0.67	**0.79**	**2.03**	0.86	0.95	**2.54**
3.3.2	1.51	**1.52**	**1.23**	**0.73**	**0.85**	**2.73**	**1.28**			0.98	**0.64**	0.85	**1.42**	0.88	0.86	**1.72**

The outcome variable modeled is self-reported "Very easy access" to one or more of Cocaine, heroin, LSD, or PCP from the 1993 National Household Survey on Drug Abuse. Cell contents are odds ratios arising from logistic regression modeling. The six models differ with respect to the independent variables included. Models 3.1.1-3.2.1, and 3.3.1 do not include race/ethnicity terms, whilst models 3.1.2, 3.2.2, 3.3.2, include these covariates. Models 3.1.1 and 3.1.2 use the Underclass Areas measure; models 3.2.1 and 3.2.2 use the Extreme Poverty areas coding, while models 3.3.1 and 3.3.2 use the SECTOR approach. Baseline covariates are age 12-17, male sex, non-hispanic, non-black race/ethnicity. For social environment variables, baseline states are (model 3.1.2) residence in non-underclass areas, (3.2.2) non-extreme poverty areas, and Sector E (3.3.3).

Statistical significance: **Bold** $p<0.05$; Underline $p<0.1$

Table 4.15 Exploratory Models for Interaction Effect of Economic Resources and Community Attachment - Drug Selling and self-reported "Very Easy Access" to PCP, LSD and Heroin.

		Demographic Variables							SOCIAL ENVIRONMENT VARIABLES							
Model		Age				Sex	Race/Ethnicity		Sector							Interaction Term
		18-25y	26-34y	35-49y	50+y		Black	Hispanic	B	C	D	F	G	H	J	
4.1	Est	**1.02**	**0.80**	**0.78**	**0.44**	**5.82**	**2.52**	**1.18**	**0.98**	**1.31**	**0.29**	**2.50**	**2.25**	**1.56**	**1.93**	1.36
	S.E.	0.14	0.09	0.11	0.12	0.94	0.35	0.13	0.29	0.35	0.10	0.51	0.99	0.29	0.55	0.09
	p value	0.87	0.04	0.09	0.00	0.00	0.00	0.13	0.95	0.32	0.00	0.00	0.07	0.02	0.02	0.00
	95% CI	0.79	0.64	0.59	0.26	4.23	1.91	0.95	0.54	0.77	0.15	1.67	0.94	1.07	1.09	1.19
		1.33	0.99	1.04	0.74	8.01	3.32	1.48	1.76	2.23	0.58	3.73	5.37	2.27	3.41	1.54
4.2	Est	**1.12**	**1.11**	**1.12**	**0.95**	**1.87**	**1.35**	**0.94**	**1.15**	**0.66**	**0.92**	**1.32**	**0.92**	**1.02**	**1.53**	1.10
	S.E.	0.10	0.10	0.10	0.15	0.20	0.14	0.09	0.24	0.20	0.14	0.26	0.35	0.13	0.37	0.06
	p value	0.21	0.25	0.22	0.77	0.00	0.00	0.53	0.50	0.17	0.57	0.16	0.83	0.87	0.08	0.08
	95% CI	0.93	0.93	0.93	0.70	1.52	1.10	0.77	0.76	0.37	0.68	0.90	0.44	0.79	0.94	0.99
		1.35	1.33	1.35	1.30	2.31	1.66	1.14	1.74	1.19	1.23	1.94	1.95	1.33	2.49	1.23

Notes: The table reports regression products arising from survey-adjusted logistic regression. The outcome variable vary with the model: Model 4.1 models respondent's self-reported of "very often" observing drug selling in the respondent's neighbourhood. Model 4.2 models self-reported "very easy" access to PCP, LSD or heroin amongst respondents. The interaction term is the product of the two scores arising from a factor analysis of selected census variables. S.E. denotes Standard error; CI denotes Confidence Interval.

References

1. Shaw, McKay. Juvenile Delinquency and Urban Areas. Chicago: University of Chicago Press; 1942.

2. Wallace R. Social Disintegration and the Spread of AIDS-II: Meltdown of Sociogeographic Structure in Urban Minority Neighborhoods. Social Science and Medicine 1993;37(7):887-896.

3. Wallace R. Social Disintegration and the Spread of AIDS: Thresholds for Propagation along 'sociographic' networks. Soc Sci Med 1991;33(10):1155-1162.

4. Kawachi I, Kennedy B, Wilkinson R. Crime: social disorganization and relative deprivation. Soc Sci Med 1999;48(6):719-231.

5. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 1997;277(5328):918-24.

6. Lillie-Blanton M, Anthony J, Schuster C. Probing the Meaning of Racial/Ethnic Group Comparisons in Crack Cocaine Smoking. JAMA 1993;269(8):993-997.

7. Squires NF, Beeching NJ, Schlecht BJ, Ruben SM. An estimate of the prevalence of drug misuse in Liverpool and a spatial analysis of known addiction. J Public Health Med 1995;17(1):103-9.

8. National Household Survey on Drug Abuse; Public Release Codebook, 1993 & CDROM. Rockville, MD: Substance Abuse and Mental Health Services Administration, US Department of Health and Human Services; 1993.

9. Ricketts E, Sawhill I. Defining and Measuring the Underclass. Journal of Policy Analysis and Management 1988;7(2):316-325.

10. Census Bot. Characteristics of Families Residing in Poverty Areas: March 1966. Current Population Reports Series P-23 1966;19.

11. Harries K, Powell A. Juvenile Gun Crime and Social Stress: Baltimore, 1980-1990. Urban Geography 1994;15(1):45-63.

12. Harries K. The Ecology of Homicide and Assault: Baltimore City and County, 1989-1991. Studies in Crime and Crime prevention 1995;4:44-60.

13. Harries K. Social stress and trauma: synthesis and spatial analysis. Soc Sci Med 1997;45(8):1251-1264.

14. Deaton A. The analysis of household surveys : a microeconometric approach to development policy. Baltimore, MD: Published for the World Bank [by] Johns Hopkins University Press; 1997.

15. Friedman SR, Curtis R, Jose B, Neaigus A, Zenilman J, Culpepper-Morgan J, et al. Sex, drugs, and infections among youth. Parenterally and sexually transmitted diseases in a high-risk neighborhood. Sex Transm Dis 1997;24(6):322-6.

16. Fullilove RE, Fullilove MT, Northridge ME, Ganz ML, Bassett MT, McLean DE, et al. Risk factors for excess mortality in Harlem. Findings from the Harlem Household Survey. Am J Prev Med 1999;16(3 Suppl):22-8.

17. Ensminger ME, Anthony JC, McCord J. The inner city and drug use: initial findings from an epidemiological study. Drug Alcohol Depend 1997;48(3):175-84.

18. Saxe L, Kadushin C, Beveridge A, Livert D, Tighe E, Rindskopf D, et al. The visibility of illicit drugs: implications for community-based drug control strategies. Am J Public Health 2001;91(12):1987-94.

19. Gfroerer J, Lessler J, Parsley T. Studies of nonresponse and measurement error in the national household survey on drug abuse. NIDA Res Monogr 1997;167:273-95.

20. Kandel DB. The social demography of drug use. Milbank Q 1991;69(3):365-414.

21. Bobashev GV, Anthony JC. Clusters of marijuana use in the United States. Am J Epidemiol 1998;148(12):1168-74.

22. Kornblum W. Drug legalization and the minority poor. Milbank Q 1991;69(3):415-35.

23. DeWit DJ. Frequent childhood geographic relocation: its impact on drug use initiation and the development of alcohol and other

drug-related problems among adolescents and young adults. Addict Behav 1998;23(5):623-34.

Chapter #5: A Conceptual Model of Blood System Safety: The 8P Model

Foreword

This chapter marks an important departure in the dissertation. The previous chapters have concerned themselves with social epidemiology and neighbourhoods. We have shown in Chapter #4 that two classification schemes - the Underclass Areas approach proposed by Ricketts and Sawhill - and SECTOR J neighbourhoods both identify a small subset of US neighbourhoods which have very active drug economies, and where residents report dramatically high rates of both drug selling and very-easy access to particular drugs of abuse, including PCP, LSD and heroin.

In this chapter we turn away from social epidemiology to propose a framework for considering "blood safety" - arguing that extant proposals appear to miss important determinants of the safety of the blood system. We identify eight salient domains which determine the safety of the blood system, and "place" is one of these eight dimensions. The notions of place articulated in Chapters 1-4, and the notion of place articulated here in Chapter #5 will be brought together in a quantitative analysis in Chapter #6.

Introduction

"Is the blood supply safe?" This simple question has been a "hot button" in health policy in Canada, the US, the UK and elsewhere over much of the past two

decades. Despite marked advances in the measured safety of blood and blood products over this period in many developed countries, [1] the prominence afforded blood safety in the media, and the public's continuing interest, suggest that it will continue to attract both public and policy concern into the future.

While there is a voluminous scientific literature which speaks to blood safety, the literature has largely focussed on the many specific questions associated with blood safety. An abundance of papers address questions such as the sensitivity of laboratory tests, or the benefits of various methods for the reduction of viral contamination of plasma; this abundance of materials can be contrasted with the relative dearth of discussions of the "macro" level determinants of blood safety. In our view, this imbalance reflects the lack of an overarching conceptual framework which sets out to integrate the various particular factors which modify blood safety

The lack of an articulated "road-map" is problematic for all aspects of blood banking, including basic research, epidemiologic studies, outbreak management, information strategies, policy analysis and risk communication [2]. The problem, simply put, is that lacking an accepted and comprehensive model of blood safety, discussions of the topic are prone to be directed by the particular "crisis of the day" and not by a sustained strategy for blood safety. Moreover, lacking a comprehensive conceptual model, important terms in the safety "equation" may be omitted, and other variables may be given inappropriate prominence; this suggests the possibility of mis-specification (and particularly under-specification) of epidemiologic and risk models;

the lack of a comprehensive conceptual model may also mean that crucial etiologic and control factors are not considered in outbreaks of infectious disease in blood products and among blood product recipients.

In the policy realm, the lack of a comprehensive conceptual model suggests that the full range of policy levers available to enhance blood safety may not be clearly articulated. Finally, a coherent conceptual model might provide a framework for discussion and education on blood safety issues with audiences as varied as government Ministers, juries in civil litigation, donors, recipients, and various scientific and public-interest groups.

Unfortunately, neither the recent Commission of Inquiry on the Blood System in Canada[3,4] nor a review of the US blood-system's handling of HIV by the US Institute of Medicine [5] has offered such a model. Ludlam has argued that the safety of blood products is premised exclusively on three domains – donor selection, (laboratory) screening and decontamination procedures. Aubuchon and Kruskall advance a similar argument to explain improvement in the safety of the blood supply in the United Sates since the early 1980s; they highlight the importance of "reducing viral risks through expansion of donor history screening, improvement in infectious disease marker testing, and [the] development of virus inactivation"[6] Elsewhere AuBuchon, Brookemeyer and Busch have included the use of allogenic transfusion in their list of key innovations to improve blood-product safety.[1]

As we shall show below, while both Ludlam and the two AuBuchon papers identify important domains of blood product safety, the specified domains are not, in our view, an *exhaustive* description of the factors which determine blood-product safety. Further, while the identified domains of influence on blood-product safety are widely accepted, and of undoubted relevance, the domains are not part of a pre-specified description of blood safety.

The alternative approach which we adopt, seeks to supplement considerations of "blood-product" safety of the sort offered by Ludlam and AuBuchon, with additional factors which define aspects of safety and risk which are distinct from the product itself. For example, the effectiveness of regulatory and operational agencies to identify and quickly recall questionable blood products is clearly of importance to blood safety, but is one of several factors that is not captured by an exclusive focus on "blood-product" safety. Likewise, the competence of collection agencies to avoid risks if possible, to detect risks when they occur, and to respond appropriately to identified risks is equally salient as an aspect of blood safety, but again not explicitly captured within a strictly blood-product paradigm. We suggest that a broader perspective is needed, and that this perspective should include both key issues in blood-product safety; we call this more inclusive strategy the "blood-system" safety approach.

The remainder of this paper is organized into three sections. The next section presents the proposed model of blood safety, outlines its key domains and articulates

the underlying logic; thereafter, we identify salient dimensions for each of these domains. The final section offers some observations on the limitations and applications of this model, and suggests a program of research.

Blood-system Safety: A Provisional Model

We understand blood-system safety to include all aspects of blood banking, from the safeguarding of potential injuries to donors, through to the provision of safe products to recipients, the effective oversight of blood collection agencies by regulators, and to include the remediation of such harms as might occur in recipients. While we have distinguished blood-product and blood-system safety, it is important to state clearly that a model of blood-system safety requires and is predicated on incorporating "blood product" safety considerations.

We propose that the safety of the blood-system can be understood by reference to eight distinct domains of influence; we entitle these domains: Persons (Donors and Recipients), Place, Pathogens & Pathology, Processes, Products & Production, Provider Practice, Policy, and Period. The model, which we informally refer to as the "8P" model, is depicted graphically in Figure 5.1.

Before proceeding to define these eight domains, some important features of the model, as depicted in the associated figure, should be noted. The figure highlights a typical chronological nature of the production of blood products beginning with donors, proceeding through the various steps in the acquisition, testing, and manufacture of blood products, and ending with recipients. Although not all risks

associated with blood products follows this sequential order, nor do all risk arise with donors, the depicted order represents the dominant etiologic pathway for the majority of transfusion-associated infections. The "Persons" domain appears twice in this figure, once to denote donors and the risks which are transferred from donors into the blood system; the second use of the term "Person" domain is associated with recipients and those who might be placed at a secondary risk of infection because of the infection of a recipient with a specific pathogen. The "Policy" domain likewise appears twice, indicated by two nested circles. The first circle is intended to capture the "corporate" policies of blood banks - both for-profit and not-for-profit; here we mean to include both the formal policies and the corporate cultures of blood collection and manufacturing agencies. The second use of "Policy" denotes the public policy environment in which blood donor recruitment, blood banking, regulation and blood product use is organized; thus the second circle is intended to capture such policy activities as the promotion of appropriate blood donating behaviour, the regulatory environment, the funding of blood services, and the organization of liability and compensation for transfusion-associated injuries. By placing these corporate policies *within* the larger public policy context, we are making a normative assertion that blood agency policies should serve larger public policy objectives. Our model does not preclude the potential influence that corporate policies and blood agencies may have on modifying public policy, nor does it assume that a single policy approach exists within either the corporate or public realms.

With these caveats in place, the general logic of the graphic representation can now be articulated. We turn first to the definition of the eight domains. The "Persons" domain is intended to capture both donors and recipients, and to recognize that the biological attributes of both donor and recipient play some role in determining the nature and consequences of transfusion-associated disease transmission. Within the concept of recipients we include both individuals who have received blood product (direct recipients) and other individuals who may be placed at some risk by the infection of a direct recipient; typically (but not exclusively) these individuals would include the sexual partners, household members, medical staff attending to patients, and future progeny of recipients.

The "Place" domain is intended to capture both the epidemiologic and social aspects of the location of blood collection activities and blood donors, such as the local incidence and prevalence of key infectious agents, as well as the social circumstances in which donation occurs - this would include the proximity of clinic locations to high-risk activities such as drug abuse.

The "Pathogens and Pathology" domain includes both the array of known transmission-transmissible agents, as well as syndromes which appear infectious but for which specific agent(s) are not yet identified. (This domain can also be expanded to include allergens, impurities, toxins and other noxious substances which might be found in blood products – not all of these risks arise from donors, but these impurities may have implications for blood-product and blood-system safety.[7]

The "Process" dimension includes all those practices which are specifically designed to reduce infectious risks from blood products. Thus defined, "Processes" are found throughout the collection, testing, manufacture, and distribution of blood products, and range from "soft" interventions such as donor screening strategies, to "hard" interventions such as laboratory testing for infectious diseases or the viral inactivation of blood products.

The "Products and Production" dimension is designed to distinguish both the various blood products, and the various methods used to produce any particular product. Thus, the model encourages explicit distinctions between cellular and acellular products, between fractionated and unfractionated versions of plasma products, and differences in methods of production - within both cellular and acellular products.

"Provider practice" is defined simply as the practices of providers of blood products (chiefly physicians) which modify the exposure of potential recipients to blood products. This would include the determination of when blood products are required, which blood products are needed for a specific therapeutic goal, the quantities of these products, and the safe administration of those products.

As discussed above, the "Policy" domain recognizes two distinct types of policy: a corporate policy realm associated with the agencies involved in the acquisition, testing, manufacture, and distribution of blood products, and a general public policy realm which defines or at least informs how blood services operate, including the

statutory specification and regulatory requirements for blood services, as well as the funding, the nature of blood product liability and litigation, and the public promotion of appropriate blood donation in the population.

A third policy realm (not shown in the figure) can also be identified which would capture the international dimensions of blood banking -- and in particular the commercial trade in blood products across international boundaries. Little information is available about the international trade in blood products, other than the well-known fact that many plasma fractions have international markets, and that both red-cells and plasma are bought and sold in international markets. It is difficult to offer more substantial comment on this policy realm other than the recognition that the commercial trade in blood and blood products has an international dimension, which appears to be avoid some, if not most, national regulatory schemes.

Finally, "Period" is included to provide a mechanism for change in the seven other domains through time; in some cases, change in a domains is sudden and dramatic (such as the introduction of a new laboratory test) and occurs over a period of days or hours. In other cases, change is very gradual - occurring over a period of years - for example the widespread introduction of a new pathogen into the blood donor population, or the adoption of new standards for blood product use by physicians.

The model presented here differs in three important respects from the informal models advanced by Ludlam and by Aubuchon *et al*, discussed earlier in this paper.

First, this model represents a formal specification of the domains of blood product safety, and makes a claim to providing an exhaustive categorization of the key influences on both blood-product and blood-system safety. Second, our model provides a more exhaustive definition of blood-system safety, one that incorporates and subsumes model of blood-product safety. Third, our model of blood-system safety allows consideration of facets above and beyond the safety of blood products. While these additional although outside of the traditional ken of transfusion medicine, they can have important implications for the safety of blood systems. We turn now to identify key dimension for each of the domains.

Key Dimensions of the Domains

For each of the eight domains identified above, there are likely to be one or more dimensions (or aspects). In this section we identify and explore some of the key dimensions for each of the eight domains. While it is argued that the *domains* are constant influences on blood safety over time, the possibility exists that the specified dimensions for each domain might evolve over time – through the advent of new dimensions, the disappearance of others, or changes in the salience of the dimensions in light of temporal changes, new pathogens, new tests, etc.

Persons - Donors

With respect to donors and blood-system safety, there appear to be five central dimensions:

1. the likelihood that the donor is infected with transfusion-transmissible infections
2. the likelihood that the donor is at risk for transfusion-transmissible infections on the basis of current or prior events, behaviour or environments
3. the motivation of the donor to donate
4. the comprehension, veracity and comprehensiveness of the donor's disclosures through questionnaires and interviews, particularly with regard to lifetime behavioural and other risk factors for infectious disease
5. the pattern of donation behaviour over time

The first four dimensions identify distinct elements of the donor's suitability to donate, and the donor's ability and willingness to provide a health history that is complete and accurate. Of particular concern here are the comprehensibility of any interviews or questionnaires[8], and the potential for "conflicts of interest" to arise within the donor when successfully donating provides benefits (be they social benefits or monetary benfits) which are of higher value to the donor than the virtue of disclosing important and relevant aspects of their risk history. Such conflicts can arise in situations where there is no payment for blood donation, such as when deferral as a blood donor might bring about social stigma. There is also the long-standing concern that where cash payment is offered for donations, that the financial

inducement may be more valued than the obligation to report a complete health history.

The fifth and final dimension of the "Persons-Donors" domain is blood donation behaviour; here we would particularly stress the spacing of donations over time. While the relationship between blood donor behaviour and blood safety is poorly conceptualized, it is clear that individuals who have relatively shorter inter-donation intervals can undertake more donations in any interval between infection and detection; this suggests that closely spaced donations may present dangers.[9] These dangers may be offset by donors with patterns of frequent donation being more exposed to prevention messages, or more socially committed to blood donation. Quantitative work in this area is lacking.

Place

Two distinct dimensions are relevant to the assessment of the place of donation. Clearly the epidemiologic profile of the location of donation is of importance, including the prevalence and incidence of important pathogens, and the nature of risks acquired from the environment (notably the presence of vector-borne diseases such as malaria and Chagas Disease). Local social patterns such as extreme poverty, or patterns of behavioural risk (such as injection drug use) may independently impart risks into the blood-system, especially if these behavioural risks interact with other factors, for example cash payment for donations. Place-related factors should be regarded as being "located" outside of the individual themselves; as such the domain

complements and enhances, but is not identical with, individual-level risk factors. Place-based analyses of transfusion-associated risk have been undertaken at the trans-national level for Central America [10] and the small-area level, with particular reference to the Bronx [11]. The distinctive importance of place is also recognized in existing regulations which prohibit blood donations from individuals who have lived in certain HIV and now CJD and vCJD-endemic areas. In Canada, Justice Krever has also cautioned against the collection of blood in areas at high risk for HIV, even though cash inducements are virtually non-existent in that country.

Pathogens

Several dimensions of pathogens need be considered. These include:

1. the exact nature of the pathogen (bacteria, virus, etc.)
2. the epidemiologic correlates of the pathogen, and particularly the degree to which the epidemiologic pattern of the pathogen coincides with donor sub-populations
3. the degree to which the pathogen is (or might be) detected directly or indirectly by laboratory measures
4. the potential viability of the pathogen through the various manufacturing steps associated with the production of various blood products
5. the potential to resolve a recognized infection of a pathogen in a recipient

6. the potential pathogenicity of the agent in infected recipients of specific blood products

While these characteristics are typically specific to a single pathogen, it is worth noting that different strains of the same pathogen may demonstrate different properties in one or more of these dimensions. For example, different laboratory tests have varying sensitivity to different genotypes of HCV [12], and different genotypes of HCV have distinct clinical consequences [13].

Process

Two basic types of processes can be distinguished. There are processes which are primarily decision-oriented, and are typically designed to support binary decisions about the acceptance or rejection of a specific donor or their donation; there are also general processes (primarily downstream of the donor and most commonly used in plasma fraction production) which act on all donations, and are typically designed to reduce the pathogen load amongst donations which have already passed the decision-type processes described above.

Donor screening and laboratory tests are the most obvious of the decision-type types of processes. Key dimensions of such testing processes are their sensitivity and specificity, cost and cost-effectiveness.

When considering these decision-type processes, it is also important to consider the "targets" of these decision-oriented processes, as the construct actually measured by the test may be only a partial marker for the underlying risk. For example, antibody

tests may be highly sensitive and specific for antibodies, but as antibodies follow *after* infection, even exquisite sensitivity in an antibody test may not provide the desired protection. Similarly, tests for liver dysfunction represent only an intermediate measure of the potential for infectious hepatitis to be present in donated blood or plasma.

It is important to observe that repeated infectious disease testing in low-prevalence populations (which describe most non-remunerated blood donor populations), can give rise to the loss of safe donors through the inevitable problems of false-positive screening tests, and generally through the necessary loss of positive predictive value associated with screening in low-prevalence populations, even if screening tests are highly-specific. The associated loss of donors and donations has important operational implications for blood banks,[14] although the influence on blood-system safety occurs only indirectly - if and when unnecessary loss of "safe" inventory result in the selection of less optimal raw or finished materials.

The second class of processes are designed to reduce the amount of viable pathogen in blood product. Typically these processes are applied universally to selected types of inputs and outputs of plasma fractionation. The two most common examples of this sort of process are the application of "solvent and detergent" to wash plasma pools, and the heat-treatment of plasma during the production of plasma fractions. Salient dimensions for these sorts of purification processes include their effectiveness in reducing pathogen load, the range of pathogens that are susceptible to

the particular decontamination process, and the consistency and quality of the application of the intervention. [15]

Product and Production

The two key dimensions of this domain are captured in the title. It is critical to distinguish amongst the over thirty common blood products that exist, because each product may have a distinct profile for infectious disease transmission, a profile that may well be contingent on the particular pathogen(s) involved. It is important to emphasize, particularly to lay audiences, that the introduction of a pathogen into the blood system by a donor does not necessarily result in contamination occurring in products, nor infection occurring in recipients, nor that all blood products are at similar risk for outbreaks. For example, while it is a certainty that HIV contaminated some of the plasma pools from which anti-D immunoglobulins[‡] were prepared, no case-reports of HIV transmission from this product have been published. While HIV was efficiently transmitted in coagulation plasma fractions, transmission has not been linked to immunoglobulins; however, anti-D immunoglobulins have been shown to transmit hepatitis C virus [16, 17].

With respect to production, there are often several methods by which any specific product might be made; thus, while the general production methods necessary to render blood or plasma into a product are constrained, there is often considerable

[‡]Anti-D immunoglobulin is used for the prevention of Rh-hemolytic disease of the newborn, a condition which arises in multigravida females where the mother is immunized to anti-D in previous pregancies, and develops a immune response to an Rh positive fetuses, often with catastrophic consequences for the fetus.

choice and variation in the details. Modification of production steps may play a role in occasional outbreaks of infectious diseases in specific blood products, for example, outbreaks of HCV in immunoglobulins.[18]

The other critical production-related issue is the use of pooling in the production of plasma products. As many as 500,000 different plasma donations (and thus 500,000 potentially distinct individuals) may be involved in the production of some plasma fractions. While the potential benefits of reducing plasma pool size are controversial[19, 20] it clearly represents a source of risk for recipients.

Provider Practice

Provider practice is an important aspect of blood-system safety, but it is distinct from the other domains in that physicians (typically) cannot modify the infectivity of the blood products available to them. With two exceptions, the provider's role is largely limited to four distinct dimensions:

- the identification of appropriate recipients of blood products
- the selection of blood products appropriate to that patient, including the use of autologous blood donations, with or without the use of medications to increase red-blood cell production, and/or the use of techniques such as intra-operative cell-salvage to reduce the need for exogenous blood transfusions (Goodnough-LT, 1996)
- the selection of specific name-brand of blood product where multiple brand-names exist

• the amounts and mixtures of blood products used

Important contributions during the last decade have begun to define both the underlying physiological mechanisms of anemia and other precipitating causes of transfusion, to describe criteria for the identification of patients, and the appropriate product mix in recipients.

There are at least two steps which providers can take to further reduce the potential of infection; irradiation of some cellular blood products can be ordered by physicians when a recipient is immuno-compromised and at particular risk of infection. Similarly, leukocyte reduction is a technique to remove white blood cells, and can occur either during the collection and manufacturing of blood products, or it can be applied at the bedside. The use of leucocyte depletion can also be ordered by providers, and has recently been demonstrated as effective for the removal of CMV from products being given to immuno-compromised individuals,[21] and is being adopted in the UK as a method to reduce the as-yet hypothetical potential risk of CJD and nv-CJD transmission in blood products.(reference?)

Persons-Recipients

Although recipients have often been depicted as having no role in the production of blood safety, recent evidence suggests that this may merit reconsideration. Key dimensions of the recipient domain include the prevalence of both natural and vaccine-induced immunity in recipient population, the likelihood and consequences of secondary transmission, and the potential of an infection to do harm in the recipient

population. Finally there is recent evidence that some forms of genetically-liked resistence may exist, at least for HIV.[22] The role of genetic characteristics in resisting HIV and other infections is still poorly understood.

Policy

The policy domain of our conceptual model has been subdivided into two basic dimensions: the corporate policies of blood collection agencies, and the public policies which inform the operation of blood services. We have earlier highlighted our normative assertion that corporate policy should typically be accountable to, and in accord with, public policy. This does not suggest however, that the corporate policies of blood agencies do not influence public policy, nor does it suggest that there is a single "public policy" in which blood services are seen. Blood services are matters not just of health policy but also include unexpected policy realms such as economic development, national defence, industry and trade, in addition to the obvious public health, pharmaceutic regulation and medical care concerns; blood services also raise political issues, quite apart from their policy considerations. Our model does not assume that the various policy realms would agree with respect to a specific blood system safety goal, nor does it assume that different blood collection agencies would agree amongst themselves. (The key dimensions of both corporate and public policy related to blood-system safety have not been widely addressed in the published literature. Thus, our identification of key dimensions for this important domain should be regarded as preliminary.)

Clearly the most important dimension of corporate policy is the ability to apply effective managerial and scientific oversight to the activities of agencies with respect to the six previously identified domains. Corporate culture can also be seen as an important part of corporate policy, and include an agencies' encouragement of error-reporting and quality assurance.

Key dimensions of public policy include any statutory context in which blood agencies operate, as well as the (often distinct) statutory context in which the regulation of blood collection is organized. The degree of interdependence and independence between operational agencies and regulatory agencies is also salient, as are national blood-related policies such as the legality of paid donations, and the pursuit of national and regional self-sufficiency with respect to plasma products [23-26]. In many countries, national or regional governments also fund the operation of blood services; in countries with no public funding of blood agencies, there is often support for transfusion-related research or at least the promotion of appropriate blood donation systems in the population. Public policy can modify blood system safety in a variety of ways, from how blood agencies are financed, and at what levels, to the regulation of commercial blood donation, or the promotion of self-sufficiency in blood products.

Period

Period is the last domain of our model, and the only domain that does not have distinct dimensions. The domain is included to capture changes over time in any of

the previous seven domains. Change in some domains is likely to be quite slow, having time scales typically measured in years – for example the adoption of new standards for transfusion triggers may take years to define, assess and adopt, while other changes will occur in the matter of days or weeks - such as new statutory environments, or the introduction of new laboratory tests.

Discussion

We have advanced a model which identifies eight distinct determinants of blood-system safety. The determinants, or domains, are the primary means by which blood-system safety might be maintained or altered. The model is premised on a distinction between "blood-product" and "blood-system" safety. We have argued that the first concept is too narrowly defined and focuses largely on the potential infectivity of blood products; the second concept, "blood-system" safety, is intended to be broader and attempts to subsume blood-product safety within a larger model which also includes recipients, provider practice, and policy considerations. For each of the eight distinct domains, we have identified key dimensions which both justify the inclusion of the domain and suggest parameters for the measurement and management of both blood product and blood-system safety.

There are several important implications to this model that deserve mention. In our view, the foremost implication is the recognition that blood-systems safety is multi-factorial and involves more than the safety of blood products. Moreover, the model suggests that the potential exists to modify blood-system safety even when

there are no immediate improvements possible to blood-product safety. The model effectively refutes the possibility of a simple and single answer to the question of "How safe is the blood supply?". The model suggests that evaluations of safety are highly contingent, and demand concurrent consideration of these eight dimensions. Although this result may be unsatisfying from a communications point of view, it more fairly reflects the phenomena under consideration.

While our model is multi-factorial, our reading of the extant literature in transfusion medicine suggests that it has historically focussed on three of the eight domains identified here: pathogens, processes, and product & production. That is, the literature has concentrated on the study of the microbiologic characteristics of specific pathogens and the related issue of the clinical impact of infection, the creation and refinement of methods to test for specific pathogens, and approaches to manufacturing which might mitigate infection while still ensuring the potency and therapeutic benefit of a blood product. The last decade has seen new publications about blood donors, provider practice, and (to a very modest degree) public policy, but newer research still represents an exception to the continuing hegemony of the three dominant domains, rather than a realignment and re-weighting of the literature. Almost no attention has been paid to the "place" domain, and the roles and influences of the policy domain are only dimly sketched.

Finally, it is also important to note that we conceive of these eight domains as exhaustive of the determinants of blood-system safety, and that each domain is

conceptually independent. That said, it seems likely that future risks to, and improvement in, the blood supply will arise both from single-domain changes, (such as the introduction of new laboratory tests, or the introduction of new pathogens), and from the interaction of two or more domains. Public policy changes, for example, can modify the mixture of paid and unpaid donations in a country, or can leverage changes in the corporate culture of blood agencies through regulatory actions. Thus, while our eight domains are distinct, the potential for powerful interactions (for good and ill) is very strong. Future outbreak investigations may wish to use this model to organize epidemiologic investigations.

It is important to recognize that the model has not made separate accommodation for the economic implications of blood safety, and the provision of blood products is, simply put, an expensive business. We see economic issues acting across all eight domains, from donors who either are, or are not, paid, to the potential costs for recipients for blood products, and the financing of operational and regulatory agencies.

That said, concern has been raised about the cost-effectiveness of further improvements to the blood system. Work is urgently needed in defining the economic impact of various interventions and to determine their relative cost-effectiveness both by comparison to other transfusion and non-transfusion related initiatives, and by comparing adjustments in each of these domains to ensure that benefits are maximized, and risks minimized.

Finally, we have not proposed a single quantitative measure by which blood-system safety could be assessed. The model implicitly rejects the simplistic notion that infectivity rates alone should be the sole quantifier of a blood-system safety measure. Selecting a measure or measures to reflect blood system safety represents an area for future research, and will demand collaboration amongst ethicists, economists, epidemiologists and transfusionists. It is hoped that this model is useful in describing, understanding and managing risks and benefits in the blood-system safety.

Figures

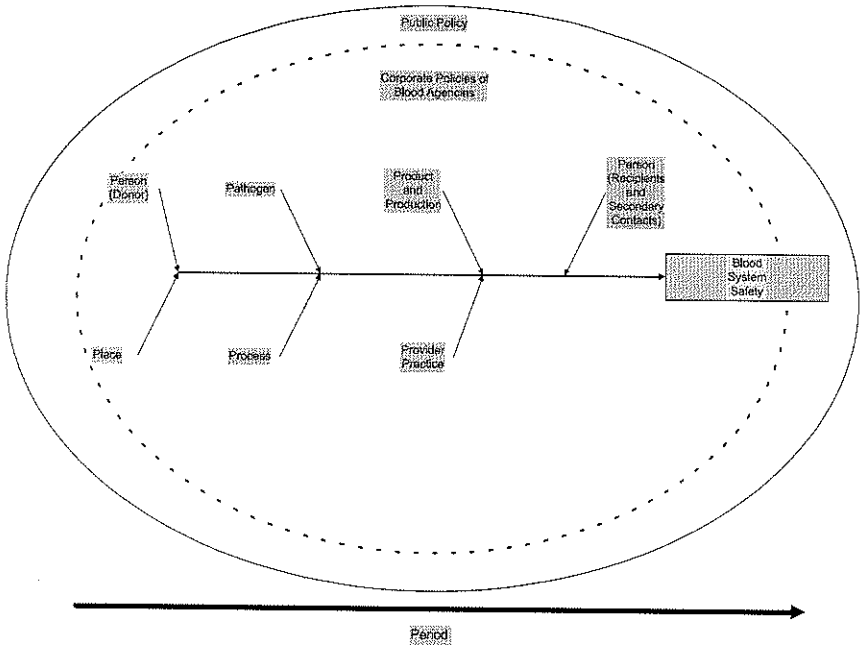

Figure 5.1 Schematic representation of Blood System Safety Model.

The figure describes a sequential and logical ordering of the influence of eight key factors on the production of blood system safety. The factors are: persons, place, processes, pathogens, products, providers, policies, and period. Both Persons and Policy domains appear twice in the figure. For persons, this reflects the distinct contribution of donor and recipient factors in the production of blood-system safety, while the duplication of the policy domain reflects the distinct influence of corporate policies of blood-related agencies, and the public policy and regulatory processes which often are designed to oversee those corporate policies.

References

1. AuBuchon JP, Birkmeyer JD, Busch MP. Safety of the blood supply in the United States: opportunities and controversies. *Ann Intern Med.* 1997;127:904-9.

2. Renn O. The role of risk communication and public dialogue for improving risk mamagement. *Rick Decision and Policy.* 1998;3:5-30.

3. Interim Report. . Toronto: Commission of Inquiry on the Blood System in Canada; 1995.

4. Report of the Commission of Inquiry on the Blood System in Canada. . Ottawa: Queen's Printer; 1997.

5. Leveton LB, Sox HC, Stoto MA, Institute of Medicine (U.S.) Committee to Study HIV Transmission Through Blood and Blood Products. HIV and the blood supply : an analysis of crisis decisionmaking. . Washington, D.C.: National Academy Press; 1995:xi, 334.

6. AuBuchon JP, Kruskall MS. Transfusion safety: realigning efforts with risks [see comments]. *Transfusion.* 1997;37:1211-6.

7. Anonymous. Adverse Ocular Reactions Following Transfusions — United States, 1997–1998. *MMWR.* 1998;47:49-50.

8. Mayo DJ, Rose AM, Matchett SE, Hoppe PA, Solomon JM, McCurdy KK. Screening potential blood donors at risk for human immunodeficiency virus. *Transfusion.* 1991;31:466-474.

9. Ward JW, Deppe DA, Samson SM, et al. Risk of human immunodeficiency virus infection from blood donors who later developed the acquired immunodeficiency syndrome. *Ann Intern Med.* 1987;106:61-62.

10. Schmunis G, Zicker F, Pinheiro F, Brandling-Bennett D. Risk for Transfusion-Transmitted Infectious Diseases in Central and South America. *EID.* 1998;4.

11. Irwin KL, Olivo N, Schable C, Weber JT, Janssen R, Ernst J. Absence of human immunodeficiency virus type 2 infection among patients in a hospital serving a New York community at high risk for infection. Centers for Disease Control and Prevention/Bronx-Lebanon Hospital Center HIV Serosurvey Team. *Transfusion.* 1996;36:731-3.

12. Gretch DR. Diagnostic tests for hepatitis C. *Hepatology.* 1997;26:43s-47s.

13. Preston FE, Jarvis LM, Makris M, et al. Heterogeneity of hepatitis C virus genotypes in hemophilia: relationship with chronic liver disease. *Blood.* 1995;85:1259-1262.

14. Ownby HE, Korelitz JJ, Busch MP, et al. Loss of volunteer blood donors because of unconfirmed enzyme immunoassay screening results. Retrovirus Epidemiology Donor Study. *Transfusion.* 1997;37:199-205.

15. Remis RS, MV OS, Tsoukas C, et al. HIV transmission to patients with hemophilia by heat-treated, donor-screened factor concentrate. *Can Med Assoc J.* 1990;142:1247-1254.

16. Yap PL. Viral transmission by blood products: a perspective of events covered by the recent tribunal of enquiry into the Irish Blood Transfusion Board [editorial]. *Ir Med J.* 1997;90:84, 86, 88.

17. Power JP, Lawlor E, Davidson F, Holmes EC, Yap PL, Simmonds P. Molecular epidemiology of an outbreak of infection with hepatitis C virus in recipients of anti-D immunoglobulin [see comments]. *Lancet.* 1995;345:1211-3.

18. Yap PL. Intravenous immunoglobulin and hepatitis C virus: an overview of transmission episodes with emphasis on manufacturing data. *Clin Ther.* 1996;18 Suppl B:43-58.

19. Barker LF. Plasma pool size and the safety of plasma derivatives [editorial; comment]. *Transfusion.* 1996;36:768-9.

20. Lynch TJ, Weinstein MJ, Tankersley DL, Fratantoni JC, Finlayson JS. Considerations of pool size in the manufacture of plasma derivatives [see comments]. *Transfusion.* 1996;36:770-5.

21. Bowden RA, Slichter SJ, Sayers M, et al. A comparison of filtered leukocyte-reduced and cytomegalovirus (CMV) seronegative blood products for the prevention of transfusion-associated CMV infection after marrow transplant [see comments]. *Blood*;86:3598-603.

22. Huang Y, Paxton WA, Wolinsky SM, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression [see comments]. *Nat Med.* 1996;2:1240-3.

23. Adey J. Is self-sufficiency financially viable and ethically justifiable?--the role of the National Blood Authority. *Blood Coagul Fibrinolysis.* 1994;5 Suppl 4:S35-6.

24. Fischer J. Is the unpaid/paid donation debate for better or for worse?--a commercial viewpoint. *Blood Coagul Fibrinolysis.* 1994;5 Suppl 4:S29-31.

25. Schadlich PK. Self-sufficiency--20 years on. *Blood Coagul Fibrinolysis.* 1994;5 Suppl 4:S9-14.

26. Keown J. The gift of blood in Europe: an ethical defence of EC directive 89/381 [see comments]. *J Med Ethics.* 1997;23:96-100.

Chapter #6: Placing US Source Plasma Clinics: 1980-1995

Foreword

This final chapter is intended to bring together the two streams of work developed in the previous chapters. In Chapter #5, we proposed a model for blood systems safety that included the concept of "place" amongst its eight domains, and we identified neighbourhoods and the social milieu as an important consideration in the promotion of blood system safety. More specifically, we highlighted the importance of avoiding high risk areas when looking for donors. High risk areas can be taken to include areas where there is active drug activity, and drug availability.

Previous chapters of this dissertation have proposed, developed, and validated the use of three measures - extreme poverty areas, underclass areas and Sector J neighbourhoods as being areas where there is ready drug access and active drug economies. The association of neighbourhood and drug activity is most pronounced for underclass areas and Sector J neighbourhoods.

In this final chapter, we apply these neighbourhood classification schemes to the placement of US source plasma clinics for the years 1980-1995. Our primary purpose is to test the hypothesis that these source plasma clinics are disproportionately located in these high risk areas - a finding that would seem to erode, not enhance, the safety of the blood system.

Introduction

An epidemic of HIV infection swept through the hemophila community of the United States in the early years of the 1980s; concurrently, yet unrecognized at the time, was the exposure of hemophiliacs to hepatitis C virus. While HIV prevalence estimates necessarily decline with time since exposure (because of HIV-related mortality) it is important to characterize the extent of the epidemic at the time that initial control measures were introduced. In 1985 Goedert noted that in a prevalence survey conducted between September 1982 and April 1984, initial serum samples from 74% of hemophiliacs who had ever been treated with commercial factor VIII concentrate, 90% of those frequently treated with factor VIII concentrate... had antibodies reactive against antigens of HTLV III(1)

The consequences of that epidemic are well known. Sourie *et al* have reported that approximately 65% of recent (year 2000) mortality amongst hemophiliacs in the United States is related to HIV, (2); using alternative methods Chorba has reported similar results(3); Walker and Julian have also reported comparable results for Canada.(4) Curiously elusive in the medical literature, however, are estimates of the actual number of infected individuals, either for single nations or across nations.

The prevalence of Hepatitis C Virus - associated with hepatocellular carcinoma - nears saturation in hemophiliacs. For example, Yee *et al* have reported a 98% antibody prevalence amongst a cohort of British hemophiliacs(5), and Troisi *et al*

have likewise reported near saturation prevalence of HCV antibody in a large America hemophilia cohort.(6)

The underlying causes and context of this iatrogenic epidemic are complex, and the subject of continuing controversy. Responses to these outbreak - and parallel outbreaks amongst recipients of other blood products - franctionated and unfractionated - have ranged from public inquiries (in Canada, and the US) to criminal convictions in France, and to civil litigation in many countries. There has also been widespread reorganization of many transfusion services over the last two decades, including the termination of the Canadian Red Cross Society's role in the provision of transfusion activities in that country.

The degree to which blood and plasma donors have, or are at risk for acquiring, transfusion-transmissible infectious agents is an important determinant of the overall safety of blood products (Chapter #4). Moreover, minimizing the risks associated with transfusion-transmissible agents is an important scientific and moral imperative of blood collection agencies and regulators. However, a long-standing critique of American source plasma collection is that risk is not minimized. Specifically it has been claimed that commercial source plasma clinics attract high risk donors (7), and that commercial source plasma are located near to areas with a high prevalence of illicit drug use(8). These two related claims map to two of the domains - place and persons - of a blood system safety model which we have articulated earlier in this dissertation.

Direct evidence to adjudicate this claim should be available through comparison of various infection-rates as observed amongst commercial plasma donors, and comparing these to test-results in voluntary donors. However, very limited data is available on test-positivity rates for commercial donors (7), this despite the routine collection of this information by commercial source plasma clinic operators. The near complete lack of information on seropositivity rates in commercial plasma donations contrasts sharply with a vast literature which reports on rates of seropositivity in various voluntary donor populations(9-13).) The lack of direct evidence regarding commercial donors is particularly puzzling because this information is necessarily and routinely collected by commercial source plasma agencies in fulfilment of regulatory requirements for testing of donations.[‡] What data does exist supports a general pattern of increased seropositivity amongst commercial plasma donors, as compared with volunteer donors. The US government's General Accounting Office has suggested that these increased rates arise because "...monetary incentives such as those offered by commercial plasma-collection centres may be tantalizing to some of those who are

---

[‡]While the GAO did not have comprehensive comparative test-positivity data, Steinhardt did report on two pieces of data which support continued concern about increased risk amongst commercial donors. With respect to HIV, Steinhardt was able to report on California data while the "HIV prevalence rates per 100,000 commercial plasma donations in California have decreased in recent years but [the rates] remain substantially higher than those same rates of volunteer whole blood donors." (p. 5). Steinhardt also testified that analyses of information submitted by plasma manufacturers in support of viral testing techniques showed "The test-positive rates for commercial plasma donors were substantially higher than those of volunteer whole blood donors, ranging from 2 to 20 times higher on the different tests."(pp. 6-7)

known to be at risk for infectious diseases, such as intravenous drug users and prostitutes." (p.7) However the GAO did not substantiate this assertion with reference to published studies, despite earlier published studies of paid blood/plasma donation in cohorts of Injection Drug users conducted separately by Nelson in Baltimore (14) and Chitwood in South Florida(8).

In addition to high risk donors, Chitwood *et al* have specifically raised the issue of the location of paid blood donation centres in high risk areas occurs, and suggested that this encourages donation by injection drug users:

> "The elimination of the sale of blood by [intravenous drug users] IVDUs is a difficult matter. As long as commercial centres exist, IVDUs will attempt to sell their blood. Two characteristics of commercial blood centres facilitate this selling behaviour: commercial blood centres which advertise that they pay cash for blood or plasma attract active drug users who do not have stable sources of income; commercial blood centres often are located in inner-city neighbourhoods where poverty and drug use are prevalent"(p. 632) (8)

However, despite this commentary on the location of source plasma clinics in the conclusion of their paper, no evidence is presented within the paper regarding the location of paid blood/plasma donor clinics. Moreover, the scope of their evidence was apparently limited to the south-Florida area. The assertion that paid "blood" donor clinics were located in areas with high rates of drug use and poverty was

unsubstantiated by the data presented in the paper. Other authors have also historically or recently commented on the location of these paid blood and plasma clinics, asserting that they are over-represented in specific social areas; none, however, present supporting data.(15-17)

While unproven, the Chitwood assertion is the most cogent claim, and has particular historical importance with respect to the transmission of HIV and HCV to plasma product recipients, most notably those with coagulation disorders. Moreover, if the claim was found to be true for very recent years, this would represent an important critique of self-regulation by the plasma industry and also of governmental regulation of source plasma collection in the United States; such a finding would also have international implications – as the US is the chief supplier of source plasma and plasma-derived pharmaceutical in the world market.

Of particular interest is the potential testability of the Chitwood assertion. To undertake such a quantitative test requires one to be able to find areas of endemic drug use, and to test the proximity of paid blood/plasma donor clinics to these foci of drug use. While there is no direct method to do this – that is there is no national database that list areas where drug use is prevalent – indirect methods are possible, provided there is validation. For example, if an operational definition of "inner-city neighbourhoods where poverty and drug use are prevalent" could be proposed and validated, then an empirical test of the Chitwood assertion could be undertaken.

The remainder of this paper offers a quantitative test of the Chitwood assertion using three validated typologies of neighbourhood. In the methods section we review the validation work which shows the associations between these specific neighbourhood types and a wide variety of personal and neighbourhood-level drug phenomena. We then use these neighbourhood typologies to test the Chitwood assertion.

Hypotheses

This paper tests two empirical questions which arise from the Chitwood assertion:

1. Is there evidence that source plasma clinics were disproportionately located in areas with high rates of risk behaviours associated with transfusion-transmissible infections within the United States during the years 1980-1990?
2. Is there evidence that source plasma clinics were disproportionately located in areas with high rates of risk behaviours associated with transfusion-transmissible infections within the United States in 1995?

Statistically, we test the null hypothesis that there is no disproportionate location in these high risk areas.

It is important to note that these questions focus on the disproportionate over-representation of clinics in specific neighbourhood types,– which is to say, the hypothesis that the proportion of clinics found in these areas are in excess of what

would have been expected if source plasma clinics were located randomly across the observed distribution neighbourhood types in the United States. By focusing on excess rates, however, we have moved the question from a simple question of fact – were any clinics located in these areas? - to a statistical question – were the rates at which source plasma clinics are found in specific neighbourhood types higher than would have been expected by chance? We think this is a reasonable interpretation of Chitwood use of the term "often" when describing the placement of these clinics in high risk areas. We have thus set a quantitative standard for evaluation and statistical analysis.

Methods:

Inclusion Criteria:

All addresses of FDA-regulated source plasma clinics operating in the period 1980-1990 and 1995 were eligible for inclusion in this analysis, provided that the address was located in the continental US, or Hawaii or Alaska. Addresses were obtained from public repositories holding the organ of the *American Blood Resources Association* – an industry association; this journal was published under at least three distinct names during the period 1980-1990 and 1995. Some years were not available from public domain sources at the time this analysis was begun. Table 3 provides bibliographic information on the sources of these addresses. At the time that the data were initially input for the dissertation, lists were not available for 1984 or 1985.

Approximately once per year the journal published a list of addresses of FDA-licensed source plasma clinics. Commentary included with these lists suggest that the information was based on Freedom of Information requests filed by the journal with the US Food and Drug Administration, the regulating body for US blood and plasma collection. Clinics licensed by the FDA on the effective data of the FDA filing the request for addresses would be included. Clinics which both opened and closed between two such reporting dates are not included in the data.

The determination of a "unique" address was done by manual comparison of addresses across years. An address was judged unique if it appeared in one or more listings, and was not likely to be a single-year variant caused by clerical error. Some judgement was necessary to determine if an address was consistent across years, and if variations in addresses across years represented typographical errors, or real changes in an address. Where-ever there was significant doubt, addresses were assumed to be unique. All decisions about what constituted a unique address versus a typographical variant of a previously identified address were made without reference to whether the address was within, or outside of, any of the neighbourhood classifications noted above. Where there were variant addresses which were accepted, we did not attempt to determine which address was "correct" as no other public domain information source was available.

Exclusion Criteria

Because our interest lay with paid plasma donation clinics and the environments in which these clinics were located, several types of plasma clinics were excluded from our analysis. Clinics operated by the American Red Cross were excluded because these did not offer payment for blood or plasma donations, and thus are not part of our population of interest. (Some community-based blood collection agencies which would not have offered payment for source plasma donations remain in the list, as these could not be objectively distinguished from the vast majority of clinics offering payment, based on the data available in the annual listings.) Secondly, plasma clinics operating within penal institutions are excluded, even though behaviours associated with HIV, viral hepatitis, and behaviours associated with parenteral disease transmission are common in penal institutions (footnote); our rationale for excluding these clinics is that we wished to use US census data to characterize neighbourhoods, but census data arising from penal institutions could not be used to support standard interpretations of these neighbourhood classification schemes.

We further excluded those plasma clinics where the reported address was not suitable for geocoding; this exclusion applied to addresses where the mailing address was given as a postal box, or where the address was a building name, functional descriptions of buildings (e.g. Bus Depot) or otherwise not a street address. The non-excluded addresses were submitted to geocoding. A subset of these addresses was successfully geocoded, and (in 1980 only) a subset of geocoded addresses were not linked to a census tract. Only those addresses which were successfully geocoded to a

census tract were included in the analysis dataset. There is no clear metric by which to compare the excluded clinics from the included clinics.

Geocoding

Once a definitive list of unique addresses was prepared, the addresses were geocoded, thereby associating the address with either (or both) the 1980 and 1990 US census tract geography for that address. The choice of census geography depended on the year in which a clinic operated at an address. The 1980-era addresses were geocoded by GDT, a commercial company and used the 1980 census geography. For those addresses with active clinics in the 1990s, the Maptititude geographic information system (GIS) package was used(18); this GIS uses US census bureau TIGER (Topologically Integrated Geographic Encoding and Referencing system) base-maps to determine the latitude and longitude of specific street addresses, and then determines the census tract which contains that point. This process also requires files which reflect the geography of census tracts; these are likewise available under Maptitude, as supplied by the US census bureau.

Once the 1980 and/or 1990 census tract for a given address was known, data from either the 1980 and/or 1990 STF-3A census files(19, 20), as appropriate, was attached to the address. The census tract identified, appropriate census variables for that census tract could be linked to the address, and a determination of the three neighbourhood typologies (see below) could be concluded. Some addresses were in census tracts which have missing values for one or more of the variables used in the

determination of the neighbourhood typologies; these neighbourhoods are given missing values for the typologies that could not be operationalized.

Neighbourhood Classification: Methods and Results

Before proceeding it is important to undertake a substantial digression and review the neighbourhood classification schemes which we will use in this chapter, and to establish their validity for finding areas with high drug use rates.

We have recently tested three neighbourhood classification schemes for their ability to find small areas (census tracts) with high levels of drug-related activity, using the standard reference for population-based studies of drug abuse in the United States: the National Household Survey on Drug Abuse (NHSDA). This survey is a representative sample of household-dwelling adults aged 12 years and older, and uses a standard multistage sampling design. Our analyses were restricted to the 1993 survey, with a sample size of 26,489, and focuses on neighbourhood drug activity, availability of"hard" drugs, and personal drug use. By special arrangement with the Substance Abuse and Mental Health Services Administration, three neighbourhood classification schemes were integrated with the public use file, and provided to the author. This augmentation of the 1993 NHSDA public use file did not compromise the anonymity of respondents.

Details of three neighbourhood classification schemes are provided in Table 6.1, and briefly reviewed here. Two of the three neighbourhood classification used in the validation were defined by outside agencies: the US Census Bureau's "extreme-

poverty-areas" designation (21), and the underclass areas designation, proposed by researchers at the Washington-based Urban Institute (22). The US census defines extreme poverty areas to be those census tracts where the poverty rate is greater than 40%. The underclass areas designation does not use poverty as part of the classifying algorithm, but rather measures high rates of four measures of so-called "deviance."

A third classification scheme was derived within this dissertation to identify areas with high rates of so-called "social disorganization." (This classification scheme was finalized prior to linkage with the NHSDA and before linkage was made to source plasma donor collection sites. Thus the classification scheme was not "tuned" to either the validation work or tested for its application in an analysis of source plasma clinics.) This classification strategy was informed by sociological theory which suggested that two separate social forces influence the rates of social problems: social resources and community attachment. Neighbourhoods with low social resources and low "community attachment" are proposed to be centres of social problems, including illicit drug use. We therefore constructed the "Sector" classification using a factor analysis of 14 measures from the US census, which was structured along these two dimensions. The resulting factor analysis was then stratified into nine sectors, and the subset with the lowest social resources and lowest levels of community attachment - which we designate as "Sector J neighbourhoods." We hypothesized that Sector J neighbourhoods, extreme poverty areas, and underclass areas, would show high rates of illicit drug use, and active local drug economies, and

undertook an analysis of a special geocoded version of the 1993 National Household Survey on Drug Abuse to test this hypothesis.

For measurement of personal drug use, we were interested in those practices and drug choices that have either been linked directly to infectious disease transmission (heroin, crack, PCP, needle use), or reflect social proscribed behaviours (selling drugs) suggestive of active local drug economies. We were also interested in individual reports of neighbourhood drug activity, and drug availability, apart from what the individual reported about their own illicit drug use and illicit drug-related activities.

Our analysis of the NHSDA data is reported in Table 6.2 which reviews unadjusted results from our analysis of the 1993 NHSDA. A cursory review of this table shows that residents of underclass areas, extreme poverty areas, and Sector J areas all report rates for a wide variety of measures of drug activity which are higher than the rates reported at the national level. The rate ratios (not shown) varied widely from 2 to 8 fold higher than those reported at a national level. With rare exception (cocaine use in the last year amongst respondents in underclass areas), the rates in these designated areas were always higher than the rates in the rest of the country. Statistical significance testing confirms this impression; detailed analyses are reported elsewhere (Chapter #4).

Evidence from the 1993 NHSDA suggests that these specific neighbourhood types are characterized by very active drug-selling, and very ready availability of a

broad range of "hard" street drugs. We therefore define US census tracts included under any of these three designations as "high risk areas" for transfusion-transmissible diseases.

Analysis Plan

The presentation of the analysis plans is divided into two key sections. In the first section we report basic information on data handling and univariate data about the clinics themselves. With respect to data handling, we document the number of unique addresses identified in each year, the number of clinics which were excluded for each of the exclusion criteria, the proportion of the non-excluded addresses which were geocoded, and finally the proportion of addresses which, though geocoded, could not be linked to the census.

In the second section, comparisons of the distribution of clinics in the three neighbourhoods types are presented, with the national distribution of these neighbourhood types from the appropriate census year. Addresses operating in the period 1980-1989 were compared to the 1980 census, while addresses operating in 1990 or 1995 were compared to the 1990 distribution. Statistical testing of the two rates was accomplished using the exact binomial distribution (23) – this test assumes that the decision to locate a clinics within any of these neighbourhood types can be seen as a "toss of a coin" with the likelihood of placement in these neighbourhoods defined from the underlying proportion of census tracts in this neighbourhood type.

Results:

Table 6.4 provides an overviews of the data handling. Addresses from the years 1980-1983, 1986-1990 and 1995 were abstracted from the indicated sources (Table 6.3). A total of 3962 addresses were reviewed, from which a total of 915 unique addresses were identified; of the which 712 unique addresses were associated with clinics operating in the years 1980-1989 and 586 unique addresses were associated with clinics operating in either 1990 or 1995. Amongst the 712 unique 1980-era addresses, 16 were penal institutions, 16 American Red Cross sites and 11 addresses were unsuitable for geocoding. This left a total of 669 potentially "geocodeable" operating in the 1980s. Geocoding was successful for 601 of the addresses, yielding a geocoding rate of 89.8%. Of these 601 geocoded addresses, 20 addresses failed to link to the 1980 census tract data; investigation showed that the geocoding linked the address to a block-numbering area and not to a census tract.[‡] A total of 581 unique addresses were available for analysis.

With respect to the 1990-era addresses, a total of 588 unique addresses were identified. Of these, 9 addresses were penal institutions, 36 addresses were operated by the American Red Cross, and 9 were inappropriate for geocoding. This left a total of 534 1990-era addresses, of which 47 did not geocode to a census tract, for a geocoding rate of 91.9%. All 491 geocoded addresses were linked to a 1990 census

---

[‡] Block numbering areas were geographic areas defined on an interim basis for the 1980 census. They tended to have larger areas and population sizes than census tracts, and are therefore excluded from our study because the associated measures are not for census tracts.

tract and associated census data. These geocoding rates are consistent with other studies.(24)

With respect to the location of these clinics, Table 6.5 reviews the distribution of source plasma clinics with respect to the three binary classifications of neighbourhoods – extreme poverty areas, underclass areas, and Sector J neighbourhoods. The degree to which source plasma clinics are disproportionately located in these areas clearly varies with time, and with the classification method, but the over-representation of clinics in these neighbourhood types is persistent across all years and all classification schemes, and typically represents at least a 5-fold increased representation than would have been expected had plasma clinics been allocated to these neighbourhood types at the same rate that they occurred within the country. For each of the 10 years available for study, source plasma clinics are concurrently over-represented in areas designated as underclass areas, extreme poverty tracts, and Sector J neighbourhoods. For example, extreme poverty-areas represented 4.36% of all census tracts in the 1980s and 5.6% of all 1990 census tracts, but represented the location of between 22.6% and 39.8% of all source plasma clinics in the years studied during the period 1980-1995. This represents a 5.2 and 7.1 times over-representation. The underclass areas definition tells a similar story; these areas represented approximately 2% of 1980 census tracts and 1.5% of 1990 census tracts, but between 9.3% and 11.3% of all plasma clinics were located in these areas. This measure also shows a 5-7.4-fold excess representation of clinics in these tracts. Finally the Sector J

measures show a similar pattern, of over-representation of source plasma clinics in census tracts identified as Sector J. Three percent of 1980 tracts and 3.9% of 1990 tracts could be designated as Sector J, but from 19.1% to 29% of all source plasma clinics are found in these neighbourhood types, representing a 6.4-8.2 fold excess over what would have been expected by chance alone. All differences between the expected proportion and the observed were tested against the binomial distribution, using STATA, with p values consistently < 0.00001.

Discussion

In their 1990 paper, Chitwood and associates suggested that source plasma clinics offering cash payment for plasma and blood donations were preferentially located in areas characterized by poverty and drug use. However, their assertion was not supported by evidence in their paper. This paper set out a quantitative standard by which to test this assertion.

Two hypotheses were advanced, one for the period 1980-1990, and one for the year 1995. For all three measures, in all years studied, source plasma clinics were disproportionately likely to be located in these neighbourhood types, and that association was statistically significant. Both hypotheses were strongly supported using each of the three conceptually distinct neighbourhood typologies; this is strong evidence to endorse the Chitwood assertion.

There are limitations to our study. Technological limitations did not allow us to geocode all clinics, and the risk profile of those clinics may not have been randomly

distributed between more and less risk areas. However, we can identify no reason why geocoding failures may not be randomly distributed across neigbourhood types As our high risk areas represent a minority of all census tracts (generally <5%) we think any bias is likely to be conservative. We have excluded clinics in penal institutions from the analysis, thereby reducing a clear and present source of risk. We have handled the inclusion and exclusion issues conservatively.

During the analysis it was noted that many clinics are located adjacent to census-tract boundaries. Our geocoding used the standard point-in-polygon approach where a given address is translated into a latitude and longitudine, and this point is then associated with a particular polygon (census tract). Errors in point-in-polygon imputations necessarily increase when the point is immediately adjacent to the edge of a polygon. To the extent that these clinics were more likely to be located adjacent to these edges than might be expected by chance alone, these methods may misallocate a clinic into a neighbouring census tract. However, because the types of tracts that we have designated as "high risk" represent a minority of all tracts, this error should result in a lessening of the proportion of clinics located in high risk tracts, not an increase.

These data clearly suggest that the location of these source plasma clinics is markedly non-representative of the spectrum of neighbourhood socioeconomic circumstances and social environments in the United States, at least over a 15 year period. The observation that US source plasma clinics were disproportionately located in high risk areas in the early 1980s is not surprising, and appears to reflect well

recognized historical strategies for locating these clinics(15-17). What is surprising and troubling is that such clinics continue to operate in these areas well after the epidemiology of HIV was defined, and in particular its association with injection drug use. We can articulate no rationale by which the placement of source plasma clinics in these neighbourhood increases the safety of the blood supply. Cogent and compelling arguments can, however, be made that this strategy of the location of source plasma clinics which provide cash compensation for plasma donation significantly compromises the blood supply safety. That these clinics remain in these areas as late as 1995 is inconsistent with epidemiologic and ethnographic insights into these areas and inexplicable from the point of view of promoting blood system safety.

In our view, the unavailability of test positivity rates for commercial source plasma donors, as evidenced in Steinhardt's testimony to the US House of Representatives, and our data, suggest that existing regulatory efforts lack fundamental information on which to guide US domestic blood policy. This information is clearly known to source plasma collectors - and is likely available to commercial fractionators; however, patients, providers and regulators appear to have limited access to these important data. Powerful information asymmetries such as these frustrate the process of initial and ongoing regulatory oversight, and erode the potential for informed consent by users of the resultant products. We would suggest that reporting of test results should be mandated on a clinic-by-clinic basis, and real and effective controls, including closure, should be imposed on those clinics which are not able to

show test positivity rates in keeping with those observed in non-paid donor cohorts. Comparable results from penal institutions should also be acquired by regulatory agencies, and publicly reported.

This thesis has not addressed the issue of the international context for plasma collection, trade, and fractionation. Important questions are beyond the scope of this thesis: Where does the plasma go that is collected in US source plasma clinics? What proportion is manufactured under US FDA oversight, and what proportion is sold to brokers or to manufacturers not operating within regulated environments. For national regulatory agencies in countries which important plasma fractions, is appears difficult to establish how these products fare with respect to the domains of our "8P" model. In our view, too little is known about the international market in blood and plasma to offer further scholarly comment; we therefore doubt that there is sufficient information on which to establish effective national regulatory control of these products.

The current geographical organization of these clinics places a significant minority of commercial plasma collection at the epicenters of illicit drug markets within the United States. These locations are focii for the sale of illicit drugs - including drugs linked directly and indirectly to transfusion-transmissible infectious diseases, notably cocaine, heroin, crack and PCP. It is difficult to imagine that these clinics could not but come to represent an attractive potential source of money for users of these drugs, and that the money available from successfully navigating the donor screening would not but represent a powerful duality of interests for drug

seekers in need of money. For these reasons, we would promote and encourage a profound reduction in the placement of clinics in these neighbourhoods. However, more is required.

Moving clinics to new areas would clearly reduce the geographical proximity of illicit drug markets and source plasma clinics. However, it seems clear that it is the availability of cash payment (or modern equivalents thereof) - that ties together in inappropriate ways, clinics, drug markets, and donors. Therefore we would encourage the regulated elimination of immediate cash-payment for source plasma donations in order to substantially reduce any short-term motivation for source plasma donation by those at risk for transfusion-transmissible diseases because of current or prior illict drug use. While we support the elimination of payment for source plasma donation, we recognize this to be a point of contention within US blood policy. Enforcing delayed payment for source plasma donation can be imposed without eliminating payment for donations, and whilst only a preliminary step in our view, this option may be more acceptable within the US policy context.

Our study documents a systematic and enduring pattern in the location of source plasma centres in non-representative, and high-risk, locations within the United States during the years 1980-1995. We observe a 5-8 fold increased likelihood of a clinic operating in each of three distinct typologies of high-risk areas, as compared against the mixture of neighbourhood types in the United States. We conclude that source plasma clinics are not avoiding these high risk areas, as both ethics and epidemiology

would suggest, but the evidence suggest strongly that those responsible for locating these clinics appear to have actively and preferentially located them in these areas. Regulatory, and potentially legislative, action may be indicated to redress both the lack of critical information needed for regulation, and to change the immediate-payment incentive and the high-risk location of source plasma donation, in order to enhance blood system safety.

Tables

Table 6.1 Three US Census-tract based neighbourhood classification schemes.

Measure	Reference	Characteristics
Extreme poverty areas	**(4)**	**Poverty rates > 40%.**
Underclass Areas	**(5)**	**Concurrent high rates for female-headed households, welfare recipiency, high-school leaving, and adult male non-participation in the workforce. High rates are defined (for both 1980 and 1990 census years) to be grater than the 1980 mean plus one 1980 standard deviation for the measure. All four metrics had to be "high" in order to meet definition**
Sector J neighbourhoods	**James, unpublished**	**A 3-by-3 classification of neighbourhoods, based on a 2 factor analysis of 14 common census variables, independently estimated for 1980 and 1990.**

Table 6.2 Rates of Selected Measures of Illicit Drug Abuse, by Neighbourhood Type, using an augmented version of the 1993 National Household Survey on Drug Abuse (USA).

Response	National	Underclass Areas	Extreme poverty	Sector J neighbourhoods
Last-year self-reported use of crack cocaine.	0.5%	1%	0.6%	1.6%
Last-year self-reported use of cocaine.	2.2%	3.4%	1.9%	3.4%
Last-year self-reported use of heroin.	0.1%	0.5%	0.3%	0.4%
"Very frequent" drug sales in neighbourhood.	5.6%	46.7%	24.9%	32%
"Very frequent" observations of intoxicated individuals in neighbourhood.	11.5%	49.3%	41.1%	43%
"Very easy" access to cocaine.	20%	41.9%	29.2%	40.8%
"Very easy" access to heroin.	11.9%	30.1%	17.4%	28.6%
"Very easy" access to LSD	11.4%	24.4%	13.9%	25.2%
"Very easy" access to PCP	9.9%	24.3%	12.8%	25.4%
Lifetime history of injection drug use.	1.4%	1.8%	1.3%	1.9%
Drug Selling in last year	0.8%	1.1%	1.5%	1.6%

Table 6.3 Bibliographic Information for Address lists.

Year	Journal Name	Bibliographic Details
1980	*Plasma Quarterly*	June, 1980; pp. 46-62.
1981	*Plasma Quarterly*	June 1981; pp. 42-59.
1982	*Plasma Quarterly*	Summer 1982; pp. 46-63.
1983	*Plasma Quarterly*	Spring 1983; pp. 48-64.
1986	*Plasma Quarterly*	[n.d.]; pp. 53-62.
1987	*Plasmapheresis*	[n.d.]; pp. 9-27.
1988	*PLASMApheresis* [sic]	June 1988; pp. 158-168.
1989	*PLASMApheresis* [sic]	October 1989; pp. 160-170.
1990	*PLASMApheresis* [sic]	Fall 1990; pp. 72-82.
1995	*The Journal of the American Blood Resources Association*	4(1):1995; pp. 21-31.

Table 6.4 Data Management of Licensed Source Plasma Clinic Addresses, 1980-1995, including geocoding and linkage rates.

YEAR	1980	1981	1982	1983	1986	1987	1988	1989	1980's	1990	1995	1990's
Number of Unique Addresses (A)	396	379	363	342	386	408	399	401	712	450	438	588
Addresses not suitable for geocoding (10)	4	4	3	2	5	6	6	6	11	5	6	9
Number of Penal Institutions (12)	10	9	10	10	10	12	10	10	16	9	1	9
American Red Cross Licenses (C) (11)	1	3	12	11	11	11	11	11	16	19	29	36
Number appropriate for geocodeing	381	363	338	339	360	379	372	374	669	417	402	534
Successfully geocoded	348	331	307	293	327	345	336	335	601	393	367	491
Geocoding Rate	91.3%	91.2%	90.8%	86.4%	90.8%	91.0%	90.3%	89.6%	89.8%	94.2%	91.3%	91.9%
Linked to 1980 Census (B)	342	322	300	289	319	337	326	325	581	393	367	491
Data Recovery Date (B/A)	98.3%	97.3%	97.7%	98.6%	97.6%	97.7%	97.0%	97.0%	96.7%	100.0%	100.0%	100.0%

Table 6.5 Location of source plasma clinics, for selected years and neighbourhood type.

Neighbourhood Type		Extreme Poverty Areas				Underclass Areas				Sector J Neighbourhoods			
Year	Number of Clinics	Rate of NT in Census (A)	Number of Clinics in NT	Proportion of Clinics in NT (B)	Ratio (B/A) **	Rate of NT in Census (C)	Number of Clinics in NT	Proportion of Clinics in NT (D)	Ratio (D/C) **	Rate of NT in Census (E)	Number of Clinics in NT	Proportion of Clinics in NT (F)	Ratio (F/E) **
1980	342	4.36%	77	22.5%	5.16	2.01%	37	10.82%	5.38	3.0%	84	24.56%	8.19
1981	321	4.36%	73	22.7%	5.22	2.01%	33	10.28%	5.11	3.0%	72	22.43%	7.48
1982	299	4.36%	68	22.7%	5.22	2.01%	30	10.03%	4.99	3.0%	64	21.40%	7.13
1983	288	4.36%	68	23.6%	5.42	2.01%	32	11.11%	5.53	3.0%	63	21.88%	7.29
1986	319	4.36%	75	23.5%	5.39	2.01%	33	10.34%	5.15	3.0%	65	20.38%	6.79
1987	335	4.36%	80	23.9%	5.48	2.01%	38	11.34%	5.64	3.0%	69	20.60%	6.87
1988	324	4.36%	77	23.8%	5.45	2.01%	35	10.80%	5.37	3.0%	65	20.06%	6.69
1989	324	4.36%	75	23.1%	5.31	2.01%	34	10.49%	5.22	3.0%	62	19.14%	6.38
1990	392	5.6%	156	39.8%	7.11	1.51%	44	11.22%	7.43	3.9%	114	29.08%	7.46
1995	367	5.6%	136	37.1%	6.62	1.51%	34	9.26%	6.14	3.9%	93	25.34%	6.50

*** N is total number of clinics observations in year. NT is neighbourhood type: Extreme Poverty Area (>40% poverty rate) Underclass Area(Y/N) Over this period, 5 and 10% of clinics could not be associated with a Sector value, because of missing values for one or more of the covariates which set the underlying factor analysis. All rates are estimated with total N as denominator, thereby assuming that none of the clinics not assigned a Sector value would have been in Sector J.**

**** All ratios of (Proportion of Clinics in NT/Proportion of NT amongst all census tracts) in all years are $p < 0.000001$ (exact binomial)**

References

1. Goedert JJ, Sarngadharan MG, Eyster ME, Weiss SH, Bodner AJ, Gallo RC, et al. Antibodies reactive with human T cell leukemia viruses in the serum of hemophiliacs receiving factor VIII concentrate. Blood 1985;65:492-495.

2. Soucie JM, Nuss R, Evatt B, Abdelhak A, Cowan L, Hill H, et al. Mortality among males with hemophilia: relations with source of medical care. The Hemophilia Surveillance System Project Investigators. Blood 2000;96(2):437-42.

3. Chorba TL, Holman RC, Clarke MJ, Evatt BL. Effects of HIV infection on age and cause of death for persons with hemophilia A in the United States. Am J Hematol 2001;66(4):229-40.

4. Walker IR, Julian JA. Causes of death in Canadians with haemophilia 1980-1995. Association of Hemophilia Clinic Directors of Canada. Haemophilia 1998;4(5):714-20.

5. Yee TT, Griffioen A, Sabin CA, Dusheiko G, Lee CA. The natural history of HCV in a cohort of haemophilic patients infected between 1961 and 1985. Gut 2000;47(6):845-51.

6. Troisi CL, Hollinger FB, Hoots WK, Contant C, Gill J, Ragni M, et al. A multicenter study of viral hepatitis in a United States hemophilic population. Blood 1993;81(2):412-8.

7. Blood Safety: Enhancing Safeguards Would Strengthen the Nation's Blood Supply. n.p.: United States General Accounting Office; 1997 June 5, 1997.

8. Chitwood DD, Page JB, Comerford M, Inciardi JA, McCoy CB, Trapido EJ, et al. The donation and sale of blood by intravenous drug users. Am J Public Health 1991;81:631-633.

9. Gill MJ, Towns D, Allaire S, Meyers G. Transmission of human immunodeficiency virus through blood transfusion: the use of lookback and traceback approaches to optimize recipient identification in a regional population. Transfusion 1997;37(5):513-6.

10. Gunson HH, Rawlinson VI. Screening of blood donations for HIV-1 antibody: 1985-1991. CDR (Lond Engl Rev) 1991;1(13):R144-6.

11. Glynn SA, Kleinman SH, Schreiber GB, Busch MP, Wright DJ, Smith JW, et al. Trends in incidence and prevalence of major transfusion-transmissible viral infections in US blood donors, 1991 to 1996. Retrovirus Epidemiology Donor Study (REDS) [see comments]. Jama 2000;284(2):229-35.

12. Chamberland ME. Surveillance for transfusion-transmitted viral infections in the United States. Biologicals 1998;26(2):85-8.

13. Williams AE, Thomson RA, Schreiber GB, Watanabe K, Bethel J, Lo A, et al. Estimates of infectious disease risk factors in US blood donors. Retrovirus Epidemiology Donor Study. Jama 1997;277(12):967-72.

14. Nelson KE, Vlahov D, Margolick J, Bernal M, Taylor E. Blood and plasma donations among a cohort of intravenous drug users. JAMA 1990;263(16):2194-7.

15. Titmuss R. The gift relationship: from human blood to social policy. London: Allen & Unwin; 1971.

16. Domen R. Paid-versus-Volunteer Blood Donation in the United States: A Historical Review. Transfusion Medicine Reviews 1995;IX(1):53-59.

17. Kretzmann MJ. Bad Blood The Moral Stigmatization of Paid Plasma Donors. Journal of Contemporary Ethnography 1992;20:4:417-441.

18. Caliper Maptitude Users's Guide: Version 4.0 for use with Microsoft Windows. In:. Newton, MA: Caliper Corporation; 1996.

19. Census of Population and Housing, 1980 [United States]: Summary Tape File 3A: United States Department of Commerce, Bureau of the Census; 1980.

20. Census of Population and Housing, 1990 [United States]: Summary Tape File 3A: United States Department of Commerce. Bureau of the Census; 1990.

21. Geographic Areas Reference Manual. Washington: US Bureau of the Census; 1994 1994.

22. Ricketts E, Sawhill I. Defining and Measuring the Underclass. Journal of Policy Analysis and Management 1988;7(2):316-325.

23. StataCorp. Stata Statistical Software: Release 6.0. College Station, TX: Stata Corporation; 1999.

24. Geronimus A, Bound J, Neider L. On the Validity of Using Census Geocode characteristics to Proxy Individual Socioeconomic Characteristics. JASA 1996;91(434):529-537.

24. Geronimus A, Bound J, Neider L. On the Validity of Using Census Geocode characteristics to Proxy Individual Socioeconomic Characteristics. JASA 1996;91(434):529-537.