NOTE TO USERS

This reproduction is the best copy available.

UMI
Individual and neighbourhood socio-economic predictors of chronic health problems and activity limitation
An application of multilevel modelling to 1990 OHS data

by

Jacob Etches

A thesis submitted in conformity with the requirements for the degree of Master of Science
Graduate Department of Community Health
University of Toronto

Copyright ©2001 Jacob Etches.
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.

0-612-58662-6
Individual and neighbourhood socio-economic predictors of chronic health problems and activity limitation

An application of multilevel modelling to 1990 OHS data

A thesis submitted in conformity with the requirements for the degree of Master's of Science
©2001
Jacob Etches
Graduate Department of Community Health
University of Toronto

Whether individual education and income account for between-neighbourhood differences in the average number of chronic health problems and in the average probability of activity limitation is examined. Between-neighbourhood variations in the effects of the individual's education and income on these dimensions of health were estimated. Observed between-neighbourhood variations were tested for their associations with neighbourhood education level in urban and rural areas. Data on individuals are taken from the 1990 Ontario Health Survey, and enumeration area educational data from the 1986 census are used to proxy neighbourhoods. Multilevel models were used to analyze the two distinct levels of analysis in this study, individuals and neighbourhoods. The effects of neighbourhood education are generally small and insignificant, and appear to be much less important than those of individual education. Little evidence is found that neighbourhood education differentially affects individuals according to their education. Methodological limitations and implications for further research are discussed.
Acknowledgements

I would like to thank my supervisor, Dr. Mustard, for his patience, guidance and encouragement, as well as my committee members, Drs. Bondy and Austin, for theirs. I received technical and biostatistical assistance from Dr. Hogg-Johnson, Mr. Manno, Mr. Ibrahim and Ms. Lee at the Institute for Work and Health, and I am indebted to them. I am also thankful for the equipment and support provided by the Institute for Work and Health, without which this analysis would not have been possible. Partial support for this work was provided by the Canadian Population Health Initiative.
# Contents

Abstract ................................................................. ii
Acknowledgements ...................................................... iii
Contents ................................................................. iv
List of Tables ............................................................ vi
List of Figures ........................................................... vii

**Study question**

1 Review of the Literature ........................................... 3
  1.1 Geographic inequalities in health ................................ 4
  1.2 The individual’s socio-economic status ........................ 7
    1.2.1 What is socio-economic status? ............................. 7
    1.2.2 How is socio-economic status measured? .................. 7
    1.2.3 How is socio-economic status related to health? ........ 8
    1.2.4 Causal mechanisms .......................................... 13
  1.3 The analysis of multilevel data ................................. 17
    1.3.1 What is a multi-level model? ............................... 17
    1.3.2 What are the advantages of multilevel models? .......... 18
  1.4 Ecological associations of neighbourhood with population health 19
    1.4.1 What is a neighbourhood? .................................. 20
    1.4.2 Does health vary by neighbourhood? ....................... 21
    1.4.3 How is the social and economic neighbourhood context related to health? 23
    1.4.4 What is the role of urbanism? ............................... 25
    1.4.5 Possible neighbourhood risks ............................... 35
  1.5 Chronic health problems and activity limitation ................ 38
    1.5.1 Chronic health problems .................................... 38
    1.5.2 Activity limitation .......................................... 40
    1.5.3 The former as a risk for the latter ....................... 44
  1.6 Summary of Chapter 1 .............................................. 46

2 This project ............................................................ 47
  2.1 Purpose ............................................................. 47
  2.2 Theoretical Models ............................................... 48
  2.3 Specific hypotheses ............................................... 48
  2.4 Relevance .......................................................... 50
List of Tables

1.1 Geographical units ................................................................. 20
1.2 British neighbourhood effects studies ....................................... 26
1.3 American neighbourhood effects studies ..................................... 28
1.4 Canadian neighbourhood effects studies ..................................... 34
1.5 Neighbourhood characteristics .................................................. 35

3.1 Validation of imputations ......................................................... 53
3.2 Activity restriction variables ..................................................... 56

4.1 Descriptive statistics ................................................................. 64
4.2 Cross-tabulation of individual characteristics ............................... 66
4.3 Attribution of variance in the intercept ....................................... 69
4.4 Neighbourhood effects on chronic health problems ...................... 74
4.5 Neighbourhood effects on self-reported activity limitation ............. 76
4.6 Neighbourhood effects on the HUI ............................................ 78

B.1 Effects of centering ................................................................. 95
# List of Figures

1.1 Context and homicide ................................................................. 4  
1.2 Chicago neighbourhoods: health and crime ...................................... 6  
1.3 Contextual causality schematic ..................................................... 6  
1.4 Social causation schematic .......................................................... 14  
1.5 The environment and activity limitation .......................................... 44  
1.6 Disability models ........................................................................ 45  

2.1 Theoretical models: schematic ....................................................... 48  
2.2 Graphical representation of hypotheses .......................................... 49  

3.1 Neighbourhood education measure distribution .................................. 59  
3.2 Chronic health problems: distribution ............................................. 61  

4.1 Interpreting level 2 parameter estimates .......................................... 68  
4.2 Chronic health problem results graph ............................................. 72  
4.3 Activity limitation results graph .................................................... 73  

5.1 Individual education in urban v. rural EAs ....................................... 81  
5.2 Neighbourhoods and socioeconomic status ...................................... 88  

B.1 Age effects on BINACT .............................................................. 93  
B.2 Age effects on chronic health problems .......................................... 94  
B.3 Age effects on HUI830 ................................................................. 94  
B.4 Age effects on HUI947 ................................................................. 95
Study question

Geographic inequalities in health exist at international, regional and small area scales. There is no question that there are structural causes of these phenomena. But it is not clear, at the scale of neighbourhood inequalities in health within rich countries, whether between-neighbourhood differences in the health of residents are due to compositional differences in individual socio-demographic variables such as age, gender, education and income, or whether adjusting for these socio-demographic variables cannot account for between-neighbourhood health differences. Neither is it clear if education and income have the same predictive value in all neighbourhoods.

In the International Classification of Function and Disability (WHO 1999), it has been proposed that diverse environmental factors can contribute to the development of, and recovery from, biological dysfunction and the limitations such dysfunction can place on the individual’s activities. Geographers, sociologists and epidemiologists are starting to look for evidence that the physical, market, service and social neighbourhood environments have contextual effects on the health of residents, and on the meaning of their education and income for their health.

This study examined whether the age, gender, education and income of residents of neighbourhoods accounted for the differences between Ontario neighbourhoods in their average number of chronic health problems and in their average probability of consequent activity limitation. Simultaneously, between-neighbourhood variation in the effect of individual education and income on these dimensions of health was estimated. The observed between-neighbourhood variations in individual education and income effects and in adjusted neighbourhood health means were tested for their associations with neighbourhood education level, which is hypothesized to represent or predict contextual neighbourhood effects on residents’ health. The effect of neighbourhood education level was allowed to interact with rural status, so that any contextual effects specific to urban or rural areas could be identified.
Health is represented in this study by two related constructs: chronic health problems and activity limitation. Activity limitation is defined as in, or perhaps somewhat exceeding the definition in, the International Classification of Function and Disability (WHO 1999). The count of self-reported chronic health problems is modelled assuming an overdispersed Poisson distribution, and activity limitation, measured first using a self-report of activity limitation, and second using a dichotomized version of the Health Utilities Index Mark III, is modelled logistically. In models of activity limitation, the individual's number of chronic health problems was controlled for in order to distinguish effects on the former from those on the latter. The socio-economic status of the individual is represented primarily by education and secondarily by income. Education was selected as the primary socio-economic indicator since it is less subject to health-selection. When income is included it is in order to provide some evidence that the education effects are not highly confounded by other dimensions of socio-economic status. Neighbourhoods are represented by enumeration areas, which are described as either urban or rural and on a continuum of aggregate education-level.
Chapter 1

Review of the Literature

This is a study of some of the relationships between a group of highly socially constructed variables, namely socio-economic status, neighbourhoods, neighbourhood socio-economic context, chronic health problems, and activity limitation. These constructs require definition before use. There is also much relevant research regarding them that should briefly be reviewed.

The literature on these subjects has not yet been fully synthesized into a coherent conceptual framework. In this chapter, an attempt has been made to order the presentation of the information in such a way that the reader will first appreciate the questions being asked, and then be guided through the literature in an organized manner.

To this end the literature review will begin by introducing the distinction between compositional and contextual explanations of geographic inequalities in health, and then review pertinent aspects of the relationship between socio-economic status and health at the individual level, so that the reader will appreciate why controlling for this pervasive individual-level determinant of health is important for our understanding of geographical inequalities in health. In section 1.4, previous attempts to control for individual socio-economic status in studies of neighbourhood inequalities in health will be reviewed. However, since many of these studies employ multilevel modelling techniques that may not be familiar to the reader, section 1.3 will provide a very brief introduction to the nature and advantages of multilevel modelling. Chapter 1 will end with a brief review of the specific health variables being analysed in the current study, namely chronic health problems and activity limitation.
1.1 Geographic inequalities in health

"[E]xperience shows that each geographic area, each type of society, and each economic group, is characterized by its own pattern of diseases..." - Dubos (1966)

Composition or context?

One of Geoffrey Rose's great contributions to epidemiology was to differentiate between the causes of cases (i.e. the health of individuals) and the causes of incidence (i.e. the health of populations) (Rose 1985). Marmot (1998) reminds us of an unmistakeable example of this by citing homicide rates by age and gender for Chicago and for England and Wales, as shown in figure 1.1.

![Figure 1.1: Rates of homicide in Chicago and England and Wales by age and sex of perpetrator (Marmot 1998).](image)

The remarkable trends by age and gender are completely preserved between locations, so that at first glance these would appear to be the determinants of homicide rates. The astounding additional piece of information is that the scale of the Chicago rate axis is thirty times greater than that for England and Wales. If one wanted to identify high risk individuals, one would ask them their age and gender. If one wanted to promote population health and well-being, one would study the differences between the social context in Chicago and that in England and Wales, and one would attempt to discover the determinants of the homicide behaviour of these populations.

Between-population differences in aggregate health can be due to either compositional or con-
textual differences.¹ Compositional differences in age, for example, can lead to massive differences between populations in their crude mortality rates despite equivalent age-standardized rates, as is the case for the States of Florida and Alaska (Lilienfeld and Stolley 1994). Unlike the Florida-Alaska comparison, between-neighbourhood male life expectancy at birth in Chicago in 1988-93 (with homicide deaths deleted) ranged from 54.3 to 77.4 years (Wilson and Daly 1997). In Ontario, Manuel et al. (2000) calculated Health-Adjusted Life Expectancy (HALE) at age 15 for each Public Health Unit (PHU) in the province, separately by gender, using mortality data and PHU profiles for the Health Utilities Index. As in Chicago, they observed a range between PHUs in HALE of 51.3-58.2 years for men and of 56.6-62.9 years for women, and they identified PHUs that have HALEs that are statistically significantly above and below the Ontario mean.

Neither the Chicago nor the Ontario between-area health differences can be due to differences in the age or gender composition, but there may remain compositional differences between groups. For example, there may be differences in the income and educational compositions, and such compositional differences may explain some or all of the geographic inequalities in health. However, if controlling for the germaine characteristics of individuals does not account for these inequalities, then characteristics of the areas can be tested for their associations with the differences in health. Ecological associations between a contextual characteristic and a properly adjusted population health statistic are candidate contextual determinants of health. In Chicago, the neighbourhood-specific cause-deleted life expectancy was correlated with the neighbourhood homicide rate (men: r=-0.88; women: r=-0.83), as shown in figure 1.2, and was also highly correlated with the median household income and with the Robin Hood index of income inequality (Wilson and Daly 1997). However, without controlling for individual characteristics, and in particular for individual economic security, we do not know the relative contributions of composition and context to this social and geographic trend.

In figure 1.3, our understanding of contextual effects is made slightly more complicated. In addition to independent effects of context (arrow 3) and individual socio-economic status (arrow 4) on individual health, arrow 2 indicates contextual modification of the relationship between individual socio-economic status and health. This is a second kind of contextual effect, and represents an interaction between characteristics of the context and socio-economic status. That is to say,

¹There are many different contexts in which individuals are nested, but the current study focusses on geographic contexts.
Figure 1.2: Neighbourhood-specific homicide rates (per 100 000 population per year) in relation to male life expectancy at birth (with effects of homicide mortality removed) for 77 community areas of Chicago, 1988-93 (Wilson and Daly 1997).

Contexts may be effect modifiers of individual characteristics. Contexts and individuals are often referred to as different 'levels' of analysis, and this type of interaction is therefore referred to as a 'cross-level interaction' (Snijders and Bosker 1999).

In order to analyse such interactions and in order to build models that adjust for the germane individual characteristics before trying to estimate ecological associations, multilevel models should be used. Multilevel models will be described in section 1.3.

Figure 1.3: This schematic illustrates the influence of context on the individual socio-economic characteristics and health. It also illustrates that context might modify the relationship between the individual socio-economic characteristics and health.
1.2 The individual's socio-economic status

Geographic inequalities in health need not represent contextual risks to health if compositional differences can account for them. The individual's socio-economic status is a pervasive determinant of social inequalities in health, and individuals are recognized to be spatially clustered according to their socio-economic characteristics. This section will briefly define socio-economic status, its measurement, its associations with health, and the causal mechanisms thought to underlie this association.

1.2.1 What is socio-economic status?

The two main schools of thought regarding the distribution of power and resources within and between communities are the Marxian and the Weberian. Marx wrote first, and had a single, all-important dimension to what he termed class, namely one's relationship to the means of production. According to Marx, one either owns capital and exploits labour, or does not own capital and sells one's labour on an exploitative market (Liberatos et al. 1988; Krieger et al. 1997; Lynch and Kaplan 2000). Weber wrote second, and argued for a more complex view of the distribution of power that has come to dominate social epidemiology and which informs the current study.\(^2\) Weber (1946) defined the terms legal order, power, honour (i.e. status), social order, economic order, class (and class situation), status group (and status situation), and party. The reader is referred to appendix A for brief definitions of these seminal concepts. The term 'socio-economic status' is an amalgamation of Weber's notion of class with his notion of status.

1.2.2 How is socio-economic status measured?

There are numerous measures of individual socio-economic status, and these have been reviewed elsewhere (Liberatos et al. 1988; Krieger et al. 1997; Lynch and Kaplan 2000).\(^3\) Lynch and Kaplan divide these into the categories 'occupation,' 'income,' 'education' and 'wealth,' although some of the measures in the occupation category are actually composite measures, a heading which appears in the Liberatos et al. enumeration.

Income and wealth map well onto class, while occupation and education map to both class and status. Party is not assessed by these indicators. Class and status are distinct but interrelated by

\(^2\)Marxian analysis still has its proponents. See, for example, the recent paper by Wright (2000).

\(^3\)Lynch provides some international scales, but all three articles are dominated by American measures.
definition, so that ‘clear mapping’ is to some degree a misnomer. Because the theoretical constructs of class and status are so difficult to dissect empirically, the construct of ‘socio-economic status’ is popular among epidemiologists.

The different measures of socio-economic position are recognized to have different properties that can be useful to the researcher (Liberatos et al. 1988; House et al. 1990; Hertzman et al. 1994; Krieger and Fee 1994; Marmot et al. 1995; Williams and Collins 1995; Krieger et al. 1997; Mustard et al. 1997; Lynch and Kaplan 2000). Education is usually acquired early in life, and is therefore less subject to modification by later ill-health. For this reason, education is the main indicator of socio-economic status in the current study. Income is much more easily modelled as a continuous measure than is occupation. This property makes it preferable as a second socio-economic indicator with which to test the stability of education effects in the current study. Income is also highly dynamic, and has the potential to indicate unstable or evolving class situations in a way that education cannot. This means that a study of activity limitation in cross-section would be expected to confute socio-economic patterns in the incidence of activity limitation with the impact of work-related activity limitation on income. For this reason the use of income in this study is merely to test the stability of any observed education effects.

1.2.3 How is socio-economic status related to health?

The current study aims to examine the interactions between neighbourhood characteristics and individual socio-economic status in the determination of chronic health problems and activity limitation. A brief review of the pertinent empirical work on the relationship of socio-economic status and health is therefore required in order to properly specify the individual level models and in order to attempt to interpret the results.

Which measures of health status are predicted by socio-economic status?

It is widely recognized that the health conditions that do not increase in prevalence with decreasing socio-economic status are the exceptions to the rule. The following Canadian examples demonstrate the generality of the association.

Wilkins et al. (1991), using census tract income quintiles as a proxy measure of individual income, list the income-related potential years of life lost (PYLL; years prior to age 75) by age group, by gender and for both 1971 and 1986 in Canada. The causes of death most responsible
for the socio-economic mortality gradient were not stable in time, place, by gender or by age-group. The 1986 combined-gender and all-age (0-74) causes of income-related PYLL are accidents (24.9%), circulatory (17.1%), perinatal (15.4%), respiratory (8.4%), neoplasms (7.6%), congenital (7.0%), digestive (6.9%) and other (12.7%). Mustard et al. (1997) described the social inequalities in morbidity and mortality in Manitoba and found that even between the most and the second most advantaged education or income quartiles that there was a modest general decrease in risk of morbidity and mortality with increased social advantage. The odds ratios between the least advantaged and the most advantaged quartiles were very significant for both mortality and co-morbidity. Humphries and van Doorslaer (2000) observe significant inequality in both self-assessed health and functional status among Canadians, complementing the work by Badley and Ibanez (1994), who showed steep educational and income-related gradients in disability in the general Canadian population. Wilkins and Adams (1983) observed large income-related differences in quality adjusted life-expectancy in Canada in the late 1970's. Millar and Wigle (1986) demonstrated socioeconomic gradients in health behaviours and their sequelae. In Canada there appears to be an income gradient in both the prevalence and the episode duration of depression (Stephens 1999), as well as educational gradients in self-esteem, mastery, happiness, distress, and cognitive impairment (Stephens et al. 1999), always favouring the more socially advantaged.

Given that socio-economic inequalities in health appear to date from at least the mid-nineteenth century (Krieger and Fee 1994), when the disease profile was significantly different from today, and given the marked instability and diversity of the causes of differential mortality by income in Wilkins et al. (1991), the argument that socio-economic status is an enduring determinant of health generally, while the specific expressions of this inequity are transient and therefore less important, is persuasive (Link and Phelan 1995). For this reason, the findings of studies of neighbourhood effects on other health variables are relevant to the current study (these are reviewed in section 1.4), and the findings of the current study may allow some speculation regarding area effects in other health domains.
Which indicators of socio-economic status are predictive?

The simple answer to this question is 'most.'\(^4\) Occupational hierarchies,\(^5\) occupational mobility,\(^6\) employment status,\(^7\) income levels (personal or household),\(^8\) income dynamics,\(^9\) total assets,\(^10\) liquid assets,\(^11\) ownership of one's house,\(^12\) difficulty paying one's mortgage,\(^13\) crowding within the house,\(^14\) possession of a car/second car,\(^15\) gravestone height,\(^16\) class of berth on the Titanic,\(^17\) educational achievement,\(^18\) and composite measures\(^19\) all demonstrate gradients of various strengths for the vast majority of health status indicators to which they have been applied.

Is the socio-economic gradient in health an artifact, non-causal, the result of health-selection or of social causation?

The Black Report (Black et al. 1988) lists these as the four possible explanations for the observed relationship between socio-economic status and health, and there has been no argument about the list itself. It is important that the assumption of social causation be firmly rooted in empirical evidence, since the interpretation of regressions on cross-sectional data, such as those presented

\(^4\)The following lists of studies are not exhaustive, either of measures or of studies using measures, but are intended to display the range and approximate frequency of use of different measures of socio-economic status.


\(^6\)(Hart et al. 1998)

\(^7\)(Badley and Ibanez 1994; Lahelma et al. 1997; Ostrove and Feldman 1999; Birch et al. 2000)


\(^9\)(Lynch et al. 1997; McDonough et al. 1997)

\(^10\)(Smith and Kington 1997; Kington and Smith 1997; Muntaner et al. 1998; Ostrove and Feldman 1999)

\(^11\)(Robert and House 1996)

\(^12\)(Marmot et al. 1995; Robert and House 1996; Wannamethee and Sharper 1997)

\(^13\)(Nettleton and Burrows 1998)

\(^14\)(Bobak et al. 1999)

\(^15\)(Davey Smith et al. 1990; Marmot et al. 1995; Wannamethee and Sharper 1997; Bobak et al. 1999)

\(^16\)(Smith et al. 1992)

\(^17\)(Lord 1955)


\(^19\)(House et al. 1990; Chen and Matthews 1999; Harley and Mortimer 1999)
here, can provide little evidence of causation on their own.

There is nearly unanimous agreement that socio-economic status predicts health far more than health predicts socio-economic status and that the relationship is generally not artifactual or non-causal (Macintyre 1986; Evans et al. 1994; House et al. 1994; Adler et al. 1994; Williams and Collins 1995; Link and Phelan 1995; Wilkinson 1996; Mustard et al. 1997). The purely artifactual explanation of the association is not plausible in light of the nearly complete irrelevance of the specific measures of health and socio-economic status to the observed association.

The non-causal dismissal requires that a third variable (or a set of variables) be found that is associated with both socio-economic status and with health, but which is causally unrelated to socio-economic status. Blaxter (1997) finds that the argument that socio-economic gradients are merely gradients in health behaviours, along with the false assumption that the behaviours of individuals are unconditioned by their class and status situations, is the prevalent lay understanding of the determinants of social inequalities in health. Perhaps because of the theoretical breadth of class and status, no properly formulated candidates for a ‘third variable’ have been proposed.

The significant theoretical and empirical question is to what degree (1) healthy individuals are upwardly mobile and unhealthy ones downwardly so (health selection), and/or (2) lower status individuals become unhealthy more often or earlier in the lifecourse than their higher status counterparts (social causation). Illsley (1986) skillfully examines the dangers of making causal conclusions regarding health inequalities based on a series of cross-sectional occupation-based studies, but the effects of parental social class on children’s health and risk exposures, neonatally, in childhood, adolescence or as adults does not allow for downward mobility as an explanation (Power et al. 1986; Power 1991; Hertzman 1994; Brunner et al. 1996; Dahl and Birkelund 1997; Lundberg 1997; Lynch et al. 1997; Power and Matthews 1997; Wadsworth 1997; Call and Nonnemaker 1999; Chen and Matthews 1999; Goldstein et al. 1999; Pensola and Valkonen 2000). Similarly, it is unlikely that mobility can explain educational gradients in health conditions whose onset occurs much later, although Koivusilta et al. (1998) do observe that health-promoting behaviours often predate high educational achievement. Longitudinal studies that measured health at baseline have provided further evidence for the social causation hypothesis (Power et al. 1990; House et al. 1994; Lynch et al. 1997; McDonough et al. 1997). The strongest evidence comes from the OPCS Longitudinal, Whitehall and Alameda County cohorts, which tracked mortality by baseline socio-economic status.
status, and demonstrated that the cumulative mortality experiences for each socio-economic group diverge smoothly over extended periods of time (Fox et al. 1985; Marmot et al. 1995; Kaplan and Keil 1993). If ongoing social mobility according to health status were a major force in preserving the gradient in health by occupational status or income group, one would expect baseline socio-economic status to be decreasingly predictive, and this does not seem to occur. The Alameda county cohort is particularly important to this point since it is a general adult sample rather than a male-only sample.

Who bears the burden most?

Care must be taken in the current study to express any curvilinearities in the income effect and to be wary of the assumptions imposed on the relationship of educational achievement to the dependent variables. If the relationship between socio-economic status and health were merely a threshold effect, then some large part of the population would not be disadvantaged relative to anyone else. This is not the case. Income is the only routinely used socio-economic indicator for which the shape of the curve can be discussed, and only then for interval measures of health. Ecob and Davey Smith (1999) found height, waist-hip ratio, respiratory function (as a proportion of expected FEV1), malaise (number of symptoms), and probability of limiting long-term illness to have a linear relationship to the log of income, so that an equivalent health gain could be achieved by a doubling of income, within the central 80% of the distribution. Wolfson et al. (1999) also used a log of income to describe its relationship to the relative risk of death. For non-continuous socio-economic indicators, it is usually observed that effect sizes are disproportionately larger for the most disadvantaged groups than for groups of intermediate advantage, relative to the most advantaged, but the Whitehall studies of British civil servants are famous for having large effects between each occupational grade for most health measures (Davey Smith et al. 1990; Marmot et al. 1991; Marmot et al. 1995; Rossum et al. 2000). It has been observed, as well, that the Whitehall civil servants may have unusually collinear class and status situations (Marmot et al. 1991), and that, unlike the general population, the entire Whitehall sample is employed.

Is the effect of socio-economic status constant throughout the lifecourse?

Socio-economic status does not predict morbidity or mortality equally at all ages. There are three aspects of this to consider: (1) there may be a series of qualitatively different windows of
susceptibility to class and status during the lifecourse, (2) there may be condition-specific windows for the expression of accumulated socio-economic exposure as compromised health, and (3) there may be a survivorship effect within age cohorts of progressively greater magnitude that distorts the composition of each class in a manner relevant to health. It is beyond the scope of the current review to summarize the findings of the rapidly growing body of research which is addressing these points. The common cross-sectional finding is that socio-economic status predicts health significantly in infancy, to a small degree in childhood, variably in young adulthood according to the health measure in question, greatly in middle-age and early old age and perhaps moderately among the elderly (Pincus et al. 1987; House et al. 1990; Wilkins et al. 1991; House et al. 1994; Kunst and Mackenbach 1994; Ford et al. 1994; Mustard et al. 1997). This consensus indicates that complex interactions of socio-economic status with age may exist that the current study will be unable to model. Limitations in model complexity will be discussed further in section 3.4.2.

1.2.4 Causal mechanisms in the social determination of health

Notwithstanding the baneful luxuries in which the European rich indulge themselves, and the disorders of repletion, inactivity and vice to which they are subject, the mean duration of their lives exceeds about ten years that of their inferiors, whom excessive fatigue has contributed to wear out before their time; whom poverty has deprived of the means of proportional comfort and subsistence. (1793)\(^{20}\)

The mechanisms by which socio-economic indicators are thought to affect health are summarized and synthesized in figure 1.4. This scheme draws especially on the work of Antonovsky (1980), Adler et al. (1994), Evans et al. (1994), Mustard et al. (1997) and McEwen (McEwen 1998; McEwen and Seeman 1999), but is not necessarily representative of their views. Some of the elements of figure 1.4 will now be described briefly.

Socio-cultural and historical context

The impact of the individual's class and status situations on his or her health should be examined with the following two points in mind: social and economic orders are not all alike, nor are they inherently stable. Regarding the first point, Amartya Sen distinguishes between support and growth mediated development (Sen 1999a; Sen 1999b). He reminds us that, after controlling for investment in health and the presence of poverty, there is no relationship between gross domestic product per capita and population health, even among low-income countries. Furthermore, Kerala State (India),

Figure 1.4: Domains are separated by vertical lines. If read left-to-right, assuming that the elements of each domain cause the elements of the domains to its right as indicated by the arrows connecting the domain titles, the schematic describes a particularly linear understanding of the social causation argument. While causal arrows could have been used to link the domains instead of vertical lines, the resulting causal web would be too complex to be informative. Health behaviours include hygiene (including sexual), diet (including alcohol), tobacco and illicit drug use, physical activity, sleeping habits, social integration and participation, spiritual and intellectual activity, and the use of accessible health services. Material and information resources include transportation, medical, grocery, clothing and fashion, entertainment and recreation. Insults could be biological, chemical or physical, mild or severe, and acute or chronic. The accumulation of impairments is inevitable, but limited by allostatic systems, which both create additional insults (sometimes greater than the one to which they respond) and also are subject to impairment, such that impairment accumulates more rapidly over time, eventually compromising vital functions.

China and Sri Lanka, by having excellent population health in the midst of extreme poverty (Sen 1999a; Sen 1999b), prove that national wealth is not a necessary cause of population health, and that population health is not a sufficient cause of economic growth. Regarding the second point, the most rapid improvements in British life expectancy of the twentieth century occurred during the
two 'World Wars,' and this has been attributed to wartime redistributive measures and perhaps to a public sense of solidarity and purpose (Winter 1988; Wilkinson 1996). Similarly, Eastern Europeans and the people of the former Soviet Union experienced precipitous declines in population health following the collapse of their command economies and governance mechanisms (Hertzman and Marmot 1996; Walberg et al. 1998).

**Health behaviours, psychosocial environments, and access to health care**

Lalonde (1974) identified heavy drinking, over-eating, inactivity, the consumption of diets high in animal fat or sugar or low in essential nutrients, smoking, coffee drinking, working in high-stress jobs, not wearing seat-belts and careless driving, promiscuity, abuse of and addiction to drugs, and living in remote areas as the 'lifestyle choices' that create 'populations at risk.' This report has been criticized as accusatory, and for failing to investigate the determinants of these behaviours, but most of the behaviours cited remain on the risk list (Adler et al. 1994). Others have been added. Unsafe sexual behaviour (Sionean and Zimmerman 1999), chronic sleep curtailment and reduced opportunities to recover from sleep loss (Cauter and Spiegel 1999) may characterize low-status groups. Status appears to express itself through most of these specific modern behaviours that are well-acknowledged determinants of health (Marmot et al. 1995; Lynch et al. 1997). Kogevinas et al. (1991) suggest that delay in seeking care may also contribute to class gradients in survival from cancers of good prognosis, since there is evidence that high socio-economic groups make more use of preventive health services. Huntington et al. (1999) show that some of the social gradient in oral health can be attributed to preventive dental care. Katz (1998) has shown that self-care activities among persons with rheumatoid arthritis are related to years of education. The demand and control characteristics of one's daily work (Karasek and Theorell 1990), as well as, and as modifiers of physical, chemical and biological exposures, mediate some of the status gradient (Marmot and Theorell 1988).

**Things that money can buy**

Quality housing (Brunner et al. 1996; Dunn and Hayes 2000), optimal diets, clothing that gives one confidence, many urban entertainments that maintain social networks and reduce psychological stress, cars and transportation, and uninsured health services, for example, all cost money and are therefore unequally distributed according to the economic order. Predatory and exploitative
advertising campaigns, such as the recently exposed targeting of youth by tobacco companies, likely compromise the rational use of limited resources in a socially-graded manner. The success of advertising in modifying behaviour can be appreciated by the U.S. Office of Drug Policy's recently acquired ability to approve television scripts (Gostin 2000).

Status dependent discounting of delayed benefits

'Why do poor people behave poorly?', asks Lynch et al. (1997). Drawing on theories in evolutionary psychology, "long-termism" was proposed by Charlton (1996) "to be a status-dependent ability to defer current satisfaction and invest resources to produce greater gains in the future." Similarly, Wilson and Daly (1997) refers us to research showing "that young adults, poor people, and criminal offenders all tend to discount the future relatively steeply... Steep discounting may be a 'rational' response to information that indicates an uncertain or low probability of surviving to reap delayed benefits, for example, and 'reckless' risk taking can be optimal when the expected profits from safer courses of action are negligible."

Psychosomatic pathways: allostasis

There is also a great deal of evidence that mood, affect and cognition, in addition to their motivational and capacity-related roles in determining behaviour, can trigger allostatic systems such that they cause the organism significant harm (McEwen 1998; McEwen and Seeman 1999). "Allostatic systems enable us to respond to our physical states (e.g., awake, asleep, supine, standing, exercising) and to cope with noise, crowding, isolation, hunger, extremes of temperature, danger, and microbial or parasitic infection." (McEwen 1998). The allostatic systems are active when called upon, and then the body recovers, returning to a resting state. The resting state is biologically

---

21 Several of the biological theories referred to in this section have been, and continue to be, investigated in animal models such as mice, rats and non-human primates. In fact, leaders in these fields are actively promoting the use of animal models of disease (Moynihan and Ader 1996). The author concedes the arguments that such use of animals has scientific advantages, such as simplicity, cost and speed. However, the author wishes to distance himself from this type of research, and requests that the reader consider the following argument. Either animal experimenters are inflicting psychological suffering without consent on sentient beings relevantly similar to human beings in the way they experience pain and psychological hardship, or these animals are not valid models of human pain and psychological hardship. That is to say, either these models are completely unethical or they are not valid. No ethically consistent argument has ever been advanced that would allow an ethical distinction to be made between species per se. Arbitrary ethical preference is nothing other than discrimination, and in this light animal experimentation is no better than the use of orphans, the mentally ill, prisoners of war or the racially marginalized, as occurred so recently. Whether it is a blow to science or not, and many argue the latter, sentient organisms should not be used without their free and informed consent for research where likely harm exceeds likely benefit to the individual organism. Where such consent is not possible, neither should be the research.
preferable in that it involves less wear and tear, so the amount of time spent in states of stress-induced arousal should be minimized. In addition to the more familiar role of hormones, and in particular the hypothalamic-pituitary-adrenal (HPA) axis, the immune system plays an important role. "All immunoregulatory processes take place within a neuroendocrine environment that is sensitive to the influence of the individual's perception of and response to events in the external world... Psychosocial factors [have been implicated] in the predisposition to and initiation and progression of various pathophysiological states, including infectious, bacterial, allergic, autoimmune, and neoplastic diseases that involve alterations in immunological defence mechanisms" (Ader et al. 1995). McEwen calls the amount of punishment that the allostatic systems inflict on the organism 'allostatic load,' and he describes four types: (1) frequent stressors, (2) perception of stress in non-threatening situations, (3) delayed recovery from arousal, and (4) dysregulation of allostatic mechanisms.

1.3 The analysis of multilevel data

The pervasive effects of socio-economic status reviewed in the previous section provide a plausible compositional explanation of geographic inequalities in health. Many of the papers reviewed in section 1.4 use a relatively new analytic technique, known as multilevel modelling, to test the hypothesis that there is no residual variation in health between neighbourhoods after controlling for the individual's socio-economic status. Due to the novelty and complexity of multilevel models, a brief summary of multilevel modelling is provided here. More technical information will be presented in section 3.4.1.

1.3.1 What is a multi-level model?

Not all users of 'multilevel models,' use the same vocabulary. Some prefer to describe the structure of the data, emphasizing that observations at one level are logically nested within the next higher level, such as individuals in neighbourhoods. These researchers refer to multilevel models or to hierarchical models. Others prefer to describe the statistical model in a descriptive rather than prescriptive manner, and so they refer to random coefficient models or mixed effects models.

Whatever they are called, multilevel models are simply a generalization of multiple regression models for clustered data. Whereas in a classical multiple regression model the coefficients (both
the intercept and the parameter estimates) are assumed to remain constant for all observations, a multilevel model allows all or some of the coefficients to vary randomly or systematically across clusters. Classical multiple regression is simply the special case where there is assumed to be no variation between clusters.

1.3.2 What are the advantages of multilevel models?

Geoffrey Rose made an important distinction between the causes of cases and the causes of means, which he referred to as the difference between individual and population health, and the difference between medicine and epidemiology (Rose 1985). Multilevel models explicitly make this differentiation by allowing the means, as well as the individual observations, to be modelled. Group means can be predicted with aggregate or with ecological measurements.\footnote{In the case of neighbourhood contextual analysis, aggregate measures are those that summarise the attributes of individuals, and include all measures based on census data. Ecological measures are those that have no meaning at the individual level, such as the presence of graffiti, the crime rate, and the density of goods and services.} For example, the social gradient in health is an individual phenomenon, while aggregate and/or ecological characteristics of groups may determine the mean around which health is socially distributed.

Multilevel models have another scientifically relevant property: differences in slopes, which are coefficients just like intercepts in a multiple regression model, can be modelled as well. A slopes-as-outcomes model could investigate whether the slope of the social gradient varies in place or time, or between subpopulations, such as ethnic groups or labour market sectors. Marmot’s interpretation that class inequalities are growing greater in Britain (Marmot 1997) could be tested in this way, as can neighbourhood effects on the effect of education on chronic health problems and activity restriction.

To recapitulate, multilevel models allow for the simultaneous estimation of (1) the individual-level effects (controlling for between-group variation), (2) the determinants of the group mean, and (3) the determinants of the group slopes, or effect sizes, of any individual-level effects.

Multiple variance terms and the attribution of variance

Why not just put the neighbourhood-level predictors in a classical multiple regression model as though they were attributes of the individual? One of the assumptions of classical multiple regression is that all of the predictors share the same variance, which is why there is only one error term. If, as in this study, one has roughly 2000 clusters with 40,000 respondents distributed between
them, there would be 2000 observations at the neighbourhood level and 40,000 observations at the individual level. Assuming equal variances will artificially improve the confidence associated with estimates of neighbourhood effects. A multilevel model has a level 1 error term, 23 and one error term for each random coefficient, so that the number of observations, as well as the nature of the variables, is respected.

Multiple variance terms also allow the variance to be attributed to one of the two levels. 24 This allows the researcher to assess how much variation there is at the neighbourhood level before and after the introduction of level 2 predictors.

**Shrinkage estimation of cluster means and slopes**

As explained by Longford (1993), the traditional route for estimating cluster means is to stratify by cluster, but this results in a loss of power. The use of multilevel models makes allowances for the cluster effect and at the same time uses all of the observations, thus improving the estimation of individual effects, potentially their cluster-specific effects, and cluster means. The result is slightly biased towards the mean, but is a much more efficient estimation of these parameters.

**Clear thinking**

Perhaps the most important feature of multilevel models is their influence on the theoretical understanding of issues involving complex variance structures by researchers. Multilevel models promote Rose's distinction of means and cases, and allow for explicit and transparent cross-level interactions.


### 1.4 Ecological associations of neighbourhood with population health

In this section, previous work on neighbourhood effects adjusted for individual characteristics is reviewed. The important issue of neighbourhood definition is only briefly addressed, while more time

---

23 In a two-level model of area effects, level 1 refers to the individual data and effects, and level 2 refers to the area data and effects.

24 Statistical research in this area is improving, but currently the quantification of level 2 variance is possible only for linear and logistic models.
is spent (1) reviewing evidence of between-neighbourhood variations in health after adjustment for demographic and socio-economic composition, and (2) reviewing evidence that the socio-economic characteristics of neighbourhoods predict the residual between-neighbourhood differences in health. Evidence from the UK, the US and Canada are reviewed separately in an attempt to respect the political and cultural contexts in which neighbourhoods have been forged in these jurisdictions.

1.4.1 What is a neighbourhood?

Hillary (1955), cited by Boyle and Lipman (1998), abstracted three common features of neighbourhoods from ninety-four definitions that he found in the literature: shared identity, shared social relations, and shared locality. In practice, theoretically justified neighbourhood or community boundaries have rarely been operationalized in health surveys. Instead, administrative boundaries, such as electoral, statistical or postal boundaries, are used as stages in multi-stage sampling designs, and these boundaries are then used as proxy definitions of neighbourhoods. The available administrative boundaries and their sizes vary nationally. The commonly used administrative units in Canada, the United Kingdom and the United States, and their approximate population sizes, are summarized in table 1.1.25

<table>
<thead>
<tr>
<th>Nation or Region</th>
<th>Small Area</th>
<th>Larger Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>Enumeration Area</td>
<td>600 Census Tract</td>
</tr>
<tr>
<td>United States</td>
<td>Block Group</td>
<td>1000 Census Tract</td>
</tr>
<tr>
<td>England and Wales</td>
<td>Enumeration District</td>
<td>400 Ward</td>
</tr>
<tr>
<td>Scotland</td>
<td>Enumeration District</td>
<td>400 Post-code sector</td>
</tr>
</tbody>
</table>

Table 1.1: The Anglo-American administrative geographical units commonly used in health surveys as neighbourhoods or local communities, their average population sizes and international equivalents.

In the UK, most of the available data has been nested wards and postcode-sectors, and in the US most data is nested in census tracts. In Canada this type of analysis has used enumeration areas exclusively.

In the following sections, published research findings from these three jurisdictions will be reviewed. The articles were selected by thorough, if unsystematic, searches of Pubmed, and through the acquired bibliographies. No attempt was made to find unpublished papers, though some were

acquired serendipitously.

1.4.2 Does health vary by neighbourhood when controlling for individual socio-economic status?

This is an important first test, since it allows the researcher to distinguish between inter-area differences in health experiences and the explanatory power of area characteristics, which will be reviewed in section 1.4.3. Unfortunately, most multilevel publications, perhaps because of the statistical complexity of this argument, fail to fully report and discuss the variance components along with the fixed effect estimates. Let us quickly review the evidence regarding the amount of variance in various health indicators by nation. In multilevel models, the intra-class correlation coefficient (\( \rho \)) represents the percentage of the variance in the intercept that occurs across areas in random intercept models. The following section makes reference to all of the reviewed papers that addressed this preliminary question.

...in the UK?

Humphreys and Carr-Hill (1991) report that only one of their dependent variables (perceived health) displayed between-ward variation after controlling for individual occupational status, and that for this model \( \rho = 3\% \).\(^{26}\) Duncan et al. (1993) report that approximately 5% of variance in smoking status and 0.7% of variance in units of alcohol consumed per week is at the ward level after controlling for individual social and demographic variables. Duncan et al. (1996) find that there is no significant area variance in the number of cigarettes smoked by smokers and a small amount of area variance in the log-odds of smoking, but they either could not or did not report \( \rho \). Ecob (1996) reported \( \rho \) unadjusted for individual socio-economic status for waist-hip ratio (4.7%), reaction time (2.6%), limiting long-standing illness (2.0%) and self-rated health (0.0%). One can assume these low values would be lower after his adjustment for social class and deprivation. 1.5% of drinking behaviour could be attributed to area by Rice et al. (1998), although this may have been diminished by the inclusion of potential mediators of the neighbourhood effect (perceived social support, stress, and physical activities). Jones et al. (2000) found that no variance in mortality could be attributed to the ward using either logistic or proportional hazards multilevel models. Finally, Mitchell et al. (2000) find that 11% of the prevalence of reporting five or more symptoms (from a list not unlike

\(^{26}\) Jones and Duncan (1995) state that the algorithms used likely underestimated area effects.
the chronic conditions list in this study) occurred at the ward level.

In the UK, there are small amounts of between-area variance in disability, death, health behaviours and central fat, and perhaps slightly more between-area variance in having multiple ailments. This finding of statistically significant small-area clustering is of uncertain consequence. None of the studies reviewed discussed the amount of variance that could be attributed to cross-level interactions involving individual socio-economic variables.

... in the US?

None of the American publications included in this review address the issue of whether there are between-area differences in health separately from discussion of predictors of that variance. In the absence of this debate, one is left with only the effects of the area characteristics to describe the amount of area variance. Since it is unlikely that the area characteristics in the models account for all of the clustering of health outcomes, this is an incomplete picture of neighbourhood inequalities in health.

... in Canada?

There are two Canadian studies that report estimates of the between-area differences in health after adjustment for individual socio-economic indicators. Boyle and Lipman (1998) found that approximately 5% in each of their measures of child behaviour problems in a national sample was attributable to the enumeration area after adjusting for parental and family characteristics. Boyle and Willms (1999) found essentially no variation in the HUI, well-being or family functioning, and only 3.6% of the variance in health problems could be attributed to the enumeration area using data from the 1990 Ontario Health Survey.

Summary

Statistically significant area variation in the individually-adjusted level of many health variables has been observed in these studies. However, it is unusual for this variance to represent more than 5% of the total variance. Since individual characteristics are imperfectly described, this amount of variance could easily represent unmeasured compositional differences between areas. The one possible exception is the finding by Mitchell, Gleave, Bartley, Wiggins, and Joshi (2000) of 11% area variance in experiencing multiple health symptoms. This is particularly relevant to the current
project since the measure of multiple health symptoms is quite similar to the measure of chronic health problems used in the current study. Very few studies have reported any investigation of area variation in the effects of individual characteristics, such as socio-economic status, and none have discussed area variation separately from the significance of the interaction coefficients.

1.4.3 How is the social and economic neighbourhood context related to health?

Structured summaries of the available British, American and Canadian studies have been tabulated by nation in tables 1.2, 1.3 and 1.4. In Britain the development of widely used area proxies for individual socioeconomic status sparked discussion and research regarding their real meaning. In general, the small amounts of area variance in health not accounted for by individual socio-economic characteristics are partially attributable to aggregate measures of area deprivation. This is not, in itself, evidence of an ecological effect, since area measures could easily be detecting unmeasured compositional differences. Consider Jones and Duncan (1995) as an example of a study that makes an above-average attempt to fully control for individual socio-economic characteristics. They include occupational class (I/II,III*III*,IV/V,other), tenure (local authority rental, private rental, owner occupied, missing) and income (low, middle, high, missing). While this level of detail is sufficient to demonstrate social gradients in health using classical regression techniques, and is complex enough to cause statistical difficulties, it cannot hope to fully capture individual class and status situations.

Davey Smith et al. (1998) present the following argument in favour of an ecological interpretation: if area-based measures were simply proxies for socio-economic status at the individual level, then they should be decreasingly predictive with increasing area size due to increasing misclassification, and yet the evidence suggests that for a range of small area sizes this is not the case. But increasing area size need not immediately increase misattribution. If the geography of socio-economic composition is relatively constant over a larger area than the larger of the two areas being used for proxy purposes (ward and enumeration district), then the characteristics of any small-area within this 'neighbourhood' would be an equivalent measure of area characteristics. In ecological studies in Glasgow, Oakland, Chicago and Harlem, neighbourhoods were built from multiple post-code sectors/census tracts for sociological reasons (Sooman and Macintyre 1995; Haan et al. 1987; Sampson et al. 1997; McCord and Freeman 1990), and in a housing study in Vancouver a
census tract was chosen from *within* each of two 'neighbourhoods' (Dunn and Hayes 2000). The irrelevance of the area used as proxy is less surprising given these neighbourhood definitions. It is also possible that the boundaries of census tracts/wards/post-code sectors are fairly consistent with any sudden spatial changes in socio-economic composition.

However, if the magnitude of the area effect is large or if the sign is opposite to that of the individual indicators, then one would have stronger evidence of an ecologic effect. The only such finding among the studies in table 1.2 is the inconclusive observation that particularly advantaged groups experience fewer of the symptoms of heart disease in more deprived wards in the UK (Jones and Duncan 1995).

There are several American studies (see table 1.3) that show changes of sign for the effects of neighbourhood deprivation, such that the poor have worse health in rich areas. Medicaid may only be a predictor of low birth weight among mothers living in rich communities (O’Campo et al. 1997). Yen and Kaplan (1998) observe that non-poverty area residence may benefit the rich and the non-black and hinder the poor and the black as regards leisure-time physical activity. They subsequently reported a graded increase in mortality with decreasing census tract ‘quality’ for the rich, and rapidly decreasing mortality for the poor, so that rich and poor fare equally in ‘poor’ quality census tracts, while the poor have their worst health experiences in the most ‘advantaged’ areas (Yen and Kaplan 1999a). Otherwise, as in the United Kingdom, American research shows small amounts of risk associated with decreasing neighbourhood levels of socio-economic advantage.

Finally, the Canadian research does not differ from the previous findings (see table 1.4). Small amounts of variance are observed after controlling for individual characteristics. Boyle and Lipman (1998), however, found that neighbourhood disadvantage predicted lower levels of parentally-reported child behaviour problems. They speculate that this might be a contextually determined reporting bias, such that in poor areas one’s children’s behaviour seems appropriate by comparison. While the effect is on perception rather than health, this would nonetheless be an ecological effect.
1.4.4 What is the role of urbanism?

In Britain, as in Canada, the binary definition of urban v. rural is highly arbitrary and therefore contentious (Haynes and Gale 1999). Population density, ethnic composition, anonymity, labour market sectors, the accessibility of goods and services, the form and function of public spaces, educational opportunity, pollution, public transportation, traffic and pedestrian patterns, housing type/quality/cost, and recreation are all elements of urbanism, and a binary classification is a particularly crude representation of this complexity. Many of these are genuinely ecological constructs for which quality data does not commonly exist in surveys and censuses focused on individuals.

Senior et al. (2000) briefly review the British studies of age and sex-standardized differences in average health statistics for rural and urban areas, and they conclude that all major disease classes are more common in urban areas. Heart disease, strokes, most cancers, and respiratory diseases are particularly related to urban living, and the only causes of death more common in rural areas are suicide and traffic deaths. Senior et al. (2000) then attempt to control for area socio-economic characteristics using ward-level census data to determine if any residual effect of urbanism can be observed in all-cause, circulatory, cancer, and respiratory mortality. They divided wards into six categories of urbanism: deep rural (>10% employed in agriculture, forestry and fishing), rural (<10% so employed), rural town (1-5k), small town (5-20k), large town (20-100k), and city (100k+). They found, after adjusting for the components of the Carstairs index and for age and sex, that there was a small amount of excess all-cause mortality mostly in small and large towns, due mostly to respiratory diseases, but also to cardiovascular ones. They further adjusted for the proportion employed in manufacturing and mining, which left the same pattern but with yet smaller effect sizes, since both of these industries are significant predictors of area all-cause, circulatory and respiratory mortality.

The study by Senior et al. (2000) suggests that while rural areas appear to confer a contextual health advantage, urban areas are more heterogeneous. Rather than increasing urbanism being correlated with decreasing population health, small and large towns seem to be at a disadvantage while cities of 100,000 or more people display health profiles quite similar to those in rural areas. The heterogeneity of urban areas will not be modelled in the current study, which uses a dichotomous variable.
Table 1.2: British neighbourhood effects studies

<table>
<thead>
<tr>
<th>Aggregation</th>
<th>Dependent Variables</th>
<th>Explanatory Variables</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carstairs and Morris (1989a)</td>
<td>Post-code sector</td>
<td>Mortality</td>
<td>IND: age, class; AREA: deprivation</td>
<td>1981 Census and</td>
<td>Class and deprivation mortality gradients are similar and largely independent</td>
</tr>
<tr>
<td></td>
<td>(Scotland)</td>
<td></td>
<td></td>
<td>1980-2 vital statistics, men aged 20-64</td>
<td></td>
</tr>
<tr>
<td>Humphreys and Carr-Hill (1991)</td>
<td>Wards</td>
<td>Self-assessment, limiting long-standing illness/disability, symptom score, respiratory function</td>
<td>IND: age, sex, ethnicity, overcrowding, tenure, garden, class, education, alcohol and tobacco consumption, exercise; AREA: deprivation</td>
<td>Multilevel and cross-tabular</td>
<td>A significant area effect likely related to aggregate socio-economic conditions is observed</td>
</tr>
<tr>
<td>Duncan, Jones, and Moon (1993)</td>
<td>Wards</td>
<td>Smoking (y/n) and drinking (units/wk)</td>
<td>IND: sex, class, age, employment, house ownership, marital status; WARD; REGION</td>
<td>HALS</td>
<td>4.5% of smoking variance and &lt; 1% of drinking variance can be attributed to the ward level</td>
</tr>
<tr>
<td>Jones and Duncan (1995)</td>
<td>Wards</td>
<td>Self-assessed health, cardiovascular conditions, FEV1</td>
<td>IND: Age, height, sex, class, employment, tenure, income, smoking, alcohol, eating, exercise, WARD: urban (defined unusually), deprivation, CONSTITUENCY: mean income</td>
<td>Multilevel</td>
<td>Significant area variation; ward deprivation independently predictive; urban non-smokers have lowered FEV1; Class III and very sensitive to deprivation effects on self-assessed health; some evidence that rich individuals suffer increased cardiovascular morbidity in advantaged wards, opposite to the general trend</td>
</tr>
<tr>
<td>Sooman and Macintyre (1995)</td>
<td>Neighbourhoods (Four in Glasgow, each containing 1-4 socially similar post-code sectors)</td>
<td>Anxiety, depression, self-assessed health</td>
<td>Perceived amenities, crime, neighbourhoodness, area reputation, satisfaction; NEIGHBOURHOOD</td>
<td>Classical regression</td>
<td>Perceived neighbourhood quality predicts anxiety and self-assessed health when controlling for social class</td>
</tr>
<tr>
<td>Duncan, Jones, and Moon (1996)</td>
<td>Wards</td>
<td>Smoking and amount smoked</td>
<td>IND: age, gender, social class, marital status; WARD</td>
<td>HALS</td>
<td>Odds of being a smoker and of quantity smoked demonstrate statistically significant area variation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Multilevel (mixed, multivariate)</td>
<td></td>
</tr>
</tbody>
</table>

continued on next page
<table>
<thead>
<tr>
<th>Aggregation</th>
<th>Dependent Variables</th>
<th>Explanatory Variables</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecob (1996)</td>
<td>Self-rated health, limiting long-standing illness, height, FEV1, waist-hip ratio, reaction time</td>
<td>IND: Age, sex, class, material deprivation; AREA: deprivation</td>
<td>West of Scotland Twenty-07 Study (baseline, '87/8)</td>
<td>Multilevel</td>
<td>Little variance in the health measures was observed across areas, and observed variance was inconsistent across age and gender strata</td>
</tr>
<tr>
<td>Ellaway and Macintyre (1998)</td>
<td>Long-standing illness, limiting long-standing illness, anxiety, depression</td>
<td>IND: age, sex, income, housing stressors, tenure, area assessment, housing type; NEIGHBOURHOOD</td>
<td>West of Scotland Twenty-07 Study (1992 - elderly only)</td>
<td>Cross-tabulation and poorly documented/reported multiple regressions</td>
<td>Net of controls, there were neighbourhood differences in all health measures, and area assessment was predictive of most measures, particularly anxiety and depression</td>
</tr>
<tr>
<td>Macintyre and Ellaway (1998)</td>
<td>Health behaviours (at all, locally)</td>
<td>IND: age, sex, social class; NEIGHBOURHOOD</td>
<td>West of Scotland Twenty-07 Study (1992 - elderly only)</td>
<td>Classical logistic regression</td>
<td>Neighbourhood populations differ in their participation in health promoting behaviours generally and locally</td>
</tr>
<tr>
<td>Rice, Carr-Hill, Dixon, and Sutton (1998)</td>
<td>Units of alcohol consumed per week</td>
<td>IND: age, gender, marital status, perceived social support, stress, activities, education, socio-economic, social class, income, gender, age, household size, car ownership, household effect; AREA</td>
<td>1993 Health Survey for England</td>
<td>Multi-level</td>
<td>Approximately 2% of variance in the dependent variable was between districts (between-level interactions not examined)</td>
</tr>
<tr>
<td>Davey Smith, Hart, Watt, Hole, and Hawthorne (1998)</td>
<td>Diastolic BP, cholesterol, height, body mass index, FEV score, current smoking, former smoking, MRC bronchitis, angina, ECG ischaemia, all cause mortality</td>
<td>IND: age, sex, social class (manual/non-manual), morbidity (see DV list); AREA: deprivation</td>
<td>Renfrew and Paisley Study, 1972-6</td>
<td>Standardized relative rates/risks and tests for trend</td>
<td>Post-code sector and enumeration area (data not published) performed similarly; deprivation and social class make significant and independent contributions to mortality, with the effect of deprivation non-significantly concentrated in the non-manual classes</td>
</tr>
<tr>
<td>Mitchell, Gleave, Bartley, Wiggins, and Joshi (2000)</td>
<td>Self-report of 5+ health symptoms (binary)</td>
<td>IND: age, sex, work status, social class, &quot;Do you feel part of the community?&quot; (y/n); AREA: deindustrialization</td>
<td>HALS</td>
<td>Multilevel</td>
<td>Feeling part of the community is protective, while residing in a ward that is in transition from high to lower levels of industrial employment is a hazard; controlling for deindustrialization does not diminish individual risk estimates</td>
</tr>
</tbody>
</table>

continued on next page
<table>
<thead>
<tr>
<th>Aggregation</th>
<th>Dependent Variables</th>
<th>Explanatory Variables</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haan, Kaplan, and Camacho (1987)</td>
<td>Federal poverty</td>
<td>Mortality</td>
<td>Alameda county study, 1965 and</td>
<td>Logistic regression</td>
<td>No differences in cause of death exist between areas; poverty area residence has a significant effect while controlling for individual socio-economic measures individually</td>
</tr>
<tr>
<td></td>
<td>area (37 contiguous census tracts, 151k people) v. non-poverty area (65 census tracts)</td>
<td>IND: age, sex, race, baseline chronic conditions, income adequacy, education (yrs), employment, health insurance, regular physicians visits, smoking (pack-yrs), BMI, weekly alcohol consumption, sleep patterns, physical activity, marital status, close friends, relatives, group membership, church membership, depression, personal uncertainty; AREA: poverty</td>
<td>1974 waves, n=1811 (351 deaths), ages 35+ in '65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCord and Freeman (1990)</td>
<td>Harlem v. US</td>
<td>Mortality</td>
<td>Vital Statistics for the United States, 1980; Files from the New York City Health Department</td>
<td>Sex-specific, age-standardized mortality rates, age-specific mortality rates, survival curves</td>
<td>Harlem's population crashed from 233k in 1960 to 228k in 1980; median 1980 family income was half the national non-white value, 40% of the New York value and 30% of the national value; despite decreasing child mortality and a strong downward national trend, adult mortality rose between 1960 and 1980; in 1980 the life expectancy was lower for men in Harlem than for men in Bangladesh; this study does not dissect the effects of individual and neighbourhood socio-economic characteristics</td>
</tr>
<tr>
<td>Aggregation</td>
<td>Dependent Variables</td>
<td>Explanatory Variables</td>
<td>Data</td>
<td>Methodology</td>
<td>Findings</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Kaplan (1996)</td>
<td>Mortality</td>
<td>IND: age, sex, health, income, race, married, smoking, activity, alcohol, BMI, isolation, depression; 4 AREA scales: lower (female blue-collar, male blue-collar, separated/divorced females, separated/divorced males, deteriorating housing units, Black), house (dilapidated housing units, deteriorating housing units, housing units with no heat, housing units with no/shared bathroom, housing units with &gt;1 person/room), old-down (males &gt;65 yrs old, housing units with no/shared bathroom, housing units with no heat, widowed males), upper (same residence as 5 yrs ago, some college education, owner-occupied housing units, drive car to work, employed males, median income)</td>
<td>Alameda county study, 1965, 1984 waves</td>
<td>Cox PH regression</td>
<td>High v. low scores for tract predict increased hazard of age and sex-adjusted morality, and controlling separately for the other individual characteristics has little impact on the hazard estimate (confidence intervals not reported)</td>
</tr>
<tr>
<td>Anderson, Sorlie, Backlund, Johnson, and Kaplan (1997)</td>
<td>Mortality</td>
<td>IND: family income (3 categories) AREA; median family income (3 categories)</td>
<td>National Longitudinal Mortality Study, n=234k, ages 25+ (ie. multiple Current Population Surveys linked with the National Death Index); 1980 Census</td>
<td>Age, sex and racially stratified Cox PH regression with robust covariance matrix</td>
<td>Area median income predicts death in a graded manner among men and women aged 25-64 when controlling for family income, with larger effect sizes for blacks than for whites; note that standard regression treatment of area characteristics likely artificially narrowed their confidence intervals</td>
</tr>
<tr>
<td>Diez-Roux, Nieto, Muntaner, Tyroler, Comstock, Shahar, Cooper, Watson, and Szkl (1997)</td>
<td>CHD</td>
<td>IND: family income, education level, occupation (6 level hierarchy), smoking, serum cholesterol, HDL and fibrinogen, systolic BP, diabetes status, physical activity, Keys score (for cholesterol elevating potential of current diet), use of antihypertensives AREA: % aged 25+ without high school, median household income, % in occupational classes II-VI, median house value</td>
<td>ARIC Study (baseline: 1987-9); n=12,6k ind. in 567 block-groups (Nested within 3 predominantly white counties and one entirely black one: Jackson); ages 45-64; 1990 Census</td>
<td>Multilevel</td>
<td>Odds of CHD increase with increasing area deprivation while controlling for socio-economic and biomedical risk factors, especially for women, except in Jackson, where the center of the deprivation distribution may place men at increased risk; effects of individual socio-economic indicators did not vary by neighbourhood context</td>
</tr>
<tr>
<td>Aggregation</td>
<td>Dependent Variables</td>
<td>Explanatory Variables</td>
<td>Data</td>
<td>Methodology</td>
<td>Findings</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
<td>-----------------------</td>
<td>------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>O'Campo, Xue, Wang, and Caughy (1997)</td>
<td>Low-birthweight</td>
<td>IND: maternal age, education, health insurance status, trimester of prenatal care initiation; AREA: ration of home owners:renters, number of community groups, % unemployed, rate of housing violations, per capita crime rate, log of mean wealth, per capita income (4 categories)</td>
<td>Computerized birth certificates 1985-9; Baltimore Department of Planning: census tract files; Maryland Department of Employment and Economic Development: 1988-9 unemployment statistics; Claritas/NDPC: per income and household wealth estimates (200-900 observations/CT, 180 CTs)</td>
<td>Two-stage multi-level regression</td>
<td>Low per capita income is a weak predictor of low birthweight directly, and it interacts with Medicaid insurance status so that the latter predicts risk only in richer neighbourhoods; area unemployment weakly amplifies the risk associated with maternal age and reduces the apparent value of early prenatal care (itself perhaps a social, as well as a medical, variable); increasing average wealth also reduces the value of early prenatal care; per capita crime amplifies the risk associated with low maternal education</td>
</tr>
<tr>
<td>Robert (1998)</td>
<td>Chronic conditions, functional limitations, self-rated health</td>
<td>IND: age, race, sex, education (yrs), income (logged), non-home assets &lt;$10k (y/n/missing); AREA: % on public assistance, % families with $30k+ income, % adult unemployment, community disadvantage index (summary of the previous three)</td>
<td>Americans' Changing Lives Study 1986, n=3.6k (randomly dispersed in the US): 1980 Census</td>
<td>OLS with serial correlation of parameter variances adjusted by a Taylor series linearization method</td>
<td>Area measures were marginally significant when controlling for individual socio-economic indicators for chronic conditions; % on public assistance predicts self-rated health independently of individual characteristics; note that OLS treatment of area characteristics likely artificially narrowed their confidence intervals</td>
</tr>
</tbody>
</table>

*continued on next page*
<table>
<thead>
<tr>
<th>Aggregation</th>
<th>Dependent Variables</th>
<th>Explanatory Variables</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wasserman, Shaw, Selvin, Gould, and Syme (1998) Block-group</td>
<td>neural tube defect</td>
<td>IND (both mother and father): high school education (y/n), employment, labourer/operator occupation (y/n), pre-tax family income &lt; $10,000 (y/n) (maternal); race, age, BMI, periconceptional vitamin use, first-trimester fever, gravidity, periconceptional recreational drug use, alcohol use, cigarette smoking, solvent exposure, nutritional status</td>
<td>California Birth Defects Monitoring Program 1989-91 (479 cases); 1987-88 birth cohort files (498 controls); interviews with cases and controls; 1990 Census</td>
<td>Adjusted Odds-ratios</td>
<td>All neighbourhood measures are associated with increased odds of neural tube defect after controlling for maternal and family socio-economic indicators, race, age, vitamin use, BMI and fever</td>
</tr>
<tr>
<td>Yen and Kaplan (1998) Federal poverty area v. non-poverty area</td>
<td>Changing physical activity level 1965-74</td>
<td>IND: family income adequacy (adj. for family size), education level, race, smoking (current/former/never), BMI (underweight, normal, obese), alcohol consumption (abstain, moderate, heavy); AREA: poverty</td>
<td>Alameda County Study</td>
<td>Classical linear regression</td>
<td>Poverty area residence was associated with decreased physical activity independently of individual socio-economic indicators, though perhaps not of race; non-poverty area residence is a benefit to the rich and the non-black, and a hindrance to the poor and the black</td>
</tr>
<tr>
<td>Diez-Roux, Nieto, Caulfield, Tyroler, Watson, and Szko (1999) Block group</td>
<td>Consumption of vegetables, fruits, meats, fish, saturated fat, cholesterol, polyunsaturated fat; Keys score</td>
<td>IND: age, sex, energy intake, family income, race; AREA: median family income</td>
<td>ARIC Study (baseline: 1987-9), n=13k, ages 45-64; 1990 Census</td>
<td>Multilevel and classical logistic regression</td>
<td>Socio-economic differences in diet appear to be accounted for by individual income, such that neighbourhood is generally insignificant or inconsistent in its effects</td>
</tr>
<tr>
<td>Robert (1999) Census tract (enumeration district &amp; block numbering area)</td>
<td>Self-rated health</td>
<td>IND: age, race, sex, education (yrs), income (logged), non-home assets &lt; $10k (y/n/missing), BMI, physical activity, current smoking, mastery, vulnerability, number of life events, informal integration, formal integration, married; AREA: % on public assistance</td>
<td>Americans' Changing Lives Study 1986; n=3.6k (randomly dispersed in the US); 1980 Census</td>
<td>OLS with serial correlation of parameter variance adjusted by a Taylor series linearization method</td>
<td>half of the individual education effect and half the area (% on social assistance) effect appear to be mediated by BMI, physical activity, smoking, mastery, vulnerability, life events and married status (a risk), so that the area effect loses significance; the individual income effect was not affected by the mediators</td>
</tr>
</tbody>
</table>

continued on next page
Yen and Kaplan (1999a)
Census tract Mortality IND: sex, age, income, race, education, smoking, perceived health, BMI, monthly alcohol consumption; AREA: 1st v. 2nd and 3rd tertiles of factor analysis summary measure (and 3 sub-measures) of: commercial stores (grocery, supermarket, beauty/barber, laundry/dry-cleaner, pharmacy), motor vehicle crashes, parks, % white-collar, % renters, % single-family dwelling, crowding, % black, per capita income, population, geographic area
Alameda county Study 1983 wave + 11 yr follow-up (n=1129, ages 36-96); 1980 census; 1980,83,84 Yellow Pages; Parks and Recreation departmental records; Department of California Highway Patrol computer records Multilevel logistic analysis
Residence in a tract scoring in the bottom third of the area score is a significant predictor of death independent of all individual characteristics; the presence of commercial stores predicts death, poor housing/environment predicts death only among the poor, increasing area SES predicts health for the rich and death for the poor in a graded manner, rich and poor fare equally in the most deprived areas

Yen and Kaplan (1999b)
Poverty area residence v. non-poor area Depression, perceived health status IND: sex, age, income adequacy, race, education level, smoking status, BMI, alcohol consumption, chronic condition count (0-2+); AREA: poverty Alameda county study, 1965 and 1974 waves, n=370 (poverty), 1125 (non-poor) Classical logistic regression
Poverty area residence is an age and sex-adjusted risk for depressive symptoms when controlling for other individual characteristics individually, but is not significant when they are simultaneously present; poverty area residence is an independent risk for fair/poor perceived health

Census tract Motor-vehicle, homicide, suicide, other-injury mortality IND: age, sex, race, marital status, income adequacy, educational attainment, employment, occupational status; AREA: % blue-collar, median family income, % poverty, % < high school, median housing value, % >1 person/room, % 5yr immobility, % unemployment, % vacant housing, % female headship, % divorced, % black/hispanic, % housing in S+ unit buildings, urban residence
National Health Interview Survey 1987-94, n=472k (1.2k deaths), ages 18-64; National Death Index 1987-95; 1990 Census Cox PH regression using SUDAAN v7.11 to adjust for design effects
urban/multi-unit living is protective except regarding homicide, where it is a hazard; low area socio-economic characteristics generally predict homicide and traffic deaths in a graded manner; racially mixed neighbourhoods may predict fewer traffic and 'other' deaths; predominantly black neighbourhoods contain excess homicide; female headship is a strong predictor of homicide, and, when poor, they predict increased traffic deaths; concentrations of divorcees predict homicide; note that standard regression treatment of area characteristics likely artificially narrowed their confidence intervals

continued on next page
<table>
<thead>
<tr>
<th>Aggregation</th>
<th>Dependent Variables</th>
<th>Explanatory Variables</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jackson, Anderson, Johnson, and Sorlie (2000)</td>
<td>Mortality</td>
<td>IND: age (25-44,45-64,65+), sex, race, family income (7 categories); AREA: racial segregation (% black)</td>
<td>National Longitudinal Study, n=239k; 1980 Census</td>
<td>Cox PH regression</td>
<td>Middle-aged (45-64) white women and young (25-44) black men experience a graded increase in mortality hazard with increasing proportion of blacks in the census tracts; young women, regardless of colour, are at slightly increased risk when skin colours are approximately equally represented; all other demographic groups showed impressively little trend; note that standard regression treatment of area characteristics likely artificially narrowed their confidence intervals</td>
</tr>
</tbody>
</table>

| Ross (2000) | Walking, exercising, smoking | IND: age, sex, race, ethnicity, marital status, education (yrs), household income, poverty; AREA: % poverty, % aged 25+ with college degree, % black (non-hispanic), % hispanic, | Survey of Community Crime and Health 1995, n=2482, ages 18+ (representative of Illinois, note: 66% of tracts contain a single observation); 1990 Census | Lagrange Multiplier (LM) test, OLS, logistic regression | the analytic approach cannot establish that there is variation to be explained at the neighbourhood level, and the LM test suggests that there may not be, therefore area measures may simply be standing proxy for unmeasured individual-level risks; there may be more walking in poor neighbourhoods and men may be more likely to smoke 7+ cigarettes/wk in poor neighbourhoods |
Table 1.4: Canadian neighbourhood effects studies

<table>
<thead>
<tr>
<th>Aggregation</th>
<th>Dependent Variables</th>
<th>Explanatory Variables</th>
<th>Data</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beland, Stoddart, and Birch (1998)</td>
<td>Perceived health, experience of illness, functional status, well-being</td>
<td>IND: Gender, age, unemployment (self/household), perceived stress, occupational status, income, schooling, social network, perceived social support, locus of control, experience of illness, functional status, perceived health; AREA: age distribution, labour force participation, unemployment, occupation, income level, education level, family size, single parents</td>
<td>1987 Sante Quebec; 1986 Census</td>
<td>Multilevel</td>
<td>Low cluster density, possible over specification and difficult presentation of output make the results of this study difficult to interpret; while the authors claim that there are some multilevel interactions, they appear to be weak ones, and the area variance appears to be minimal; note that in these models income and education, when included, are never predictive</td>
</tr>
<tr>
<td>Boyle and Lipman (1998)</td>
<td>Conduct problems, hyperactivity, emotional problems</td>
<td>IND(parent): sex, age, born outside Canada; FAMILY: # siblings, single parent, &lt; poverty line, SES (unweighted mean of standardized household income and the education and occupational statuses of both parents); AREA: % lone parent, %&lt; poverty line, neighbourhood disadvantage (unweighted mean of standardized % neighbourhood income from government transfer payments, % 15 yrs+ without high school, % 15 yrs+ with university (reversed), mean household income (reversed), % 15 yrs+ unemployed), neighbourhood disadvantage²</td>
<td>National Longitudinal Survey of Children and Youth, n=11.5k, ages 4-11</td>
<td>3-level Multilevel (children in 7799 families in 5325 enumeration areas)</td>
<td>5, 5 &amp; 6% percent of variance in the behavioural problems respectively was attributable to the enumeration area after adjusting for child and family characteristics; neighbourhood characteristics (mostly % lone parents (a risk) and neighbourhood disadvantage (protective)) explain 0.6, 0.2 and 0.4% of this variance; there was no significant variance in the effect of SES across areas</td>
</tr>
<tr>
<td>Boyle and Willms (1999)</td>
<td>Health problems (chronic or acute), health utilities index, well-being, family functioning</td>
<td>IND: Gender, age, age², education, marital status, immigrant status, household income, income²; FAMILY; ENUMERATION AREA; PUBLIC HEALTH UNIT</td>
<td>1991 Ontario Health Survey, n=48.5k, ages 12+</td>
<td>4-level Multilevel model (Individuals in families in enumeration areas in public health units)</td>
<td>After adjusting for individual characteristics, variance at the neighbourhood level equaled 0.5% (family dysfunction), 0.6% (general well-being), 1.5% (HUI), and 3.6% (health problems); note that the distributional assumptions regarding the dependent variables were probably violated, and no evidence was presented justifying the reported linear treatment</td>
</tr>
</tbody>
</table>
Table 1.5: Neighbourhood characteristics implicated as candidate contextual determinants of health, and the authors who suggested them, originally or by reference.

<table>
<thead>
<tr>
<th>Physical environment</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>Macintyre et al. (1993)</td>
</tr>
<tr>
<td>Climate</td>
<td>Macintyre et al. (1993), Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Cleanliness, beauty, nature, tranquility</td>
<td>Macintyre et al. (1993), Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Radiation</td>
<td>Lindheim and Syme (1983)</td>
</tr>
<tr>
<td>Recreational space, opportunities (e.g. bike paths, pleasant walkways, community tennis courts and pools; several others gave no examples, and may have had something different in mind)</td>
<td>Macintyre et al. (1993), Diez-Roux et al. (1997), Taylor et al. (1997), Boyle and Lipman (1998), Curtis and Jones (1998), Davey Smith et al. (1998), Yen and Kaplan (1998), Ross (2000)</td>
</tr>
<tr>
<td>Lighting</td>
<td>Macintyre et al. (1993), Yen and Kaplan (1998)</td>
</tr>
<tr>
<td>Risk of accidental injury or death</td>
<td>Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Abandoned housing</td>
<td>Boyle and Lipman (1998)</td>
</tr>
<tr>
<td>Vehicular traffic</td>
<td>Boyle and Lipman (1998)</td>
</tr>
</tbody>
</table>

continued on next page

1.4.5 Possible neighbourhood risks

It is widely acknowledged throughout the literature on area effects that the field is ‘under-theorized’ (Curtis and Jones 1998). Table 1.5 is an attempt to list the types of neighbourhood causal factors that have been implicated. The categories physical environment, service environment, and social environment were used by Robert (1998), while market environment, and ‘other’ have been added as bins to contain factors she did not discuss. Special attention should be paid to ‘mobility due to or prevented by health,’ since this is a selective rather than a causal proposition.

Currently there is little evidence of contextual effects for the vast majority of the causal factors. The factors are variously hypothesized to act as immediate physical hazards, by modifying
### Market environment

<table>
<thead>
<tr>
<th>Factor</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affordable and nutritious food</td>
<td>Macintyre et al. (1993), Diez-Roux et al. (1997), Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Non-specific availability of goods or retail facilities</td>
<td>Jones and Duncan (1995), Taylor et al. (1997), Davey Smith et al. (1998)</td>
</tr>
<tr>
<td>Behaviour elicited through economic pressures</td>
<td>Anderson et al. (1997)</td>
</tr>
<tr>
<td>Housing prices</td>
<td>Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Banks</td>
<td>Taylor et al. (1997)</td>
</tr>
<tr>
<td>Deindustrialization</td>
<td>Mitchell et al. (2000)</td>
</tr>
<tr>
<td>Secure and non-hazardous labour market</td>
<td>Macintyre et al. (1993), Curtis and Jones (1998)</td>
</tr>
</tbody>
</table>

### Service environment

<table>
<thead>
<tr>
<th>Service area</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>Macintyre et al. (1993), Boyle and Lipman (1998), Ross (2000)</td>
</tr>
<tr>
<td>Street cleaning or sanitation</td>
<td>Macintyre et al. (1993), Robert (1998)</td>
</tr>
<tr>
<td>Welfare services</td>
<td>Macintyre et al. (1993)</td>
</tr>
<tr>
<td>Fire services</td>
<td>Robert (1998)</td>
</tr>
<tr>
<td>Congregate meals, senior centers, family services</td>
<td>Robert (1998)</td>
</tr>
</tbody>
</table>

*continued on next page*

behaviour (through norms, opportunities and constraints) and through psychosomatic pathways, as described in section 1.2.4. Some authors proposed factors with specific behaviours or health outcomes in mind. Many authors mention housing quality as a neighbourhood characteristic, but this is a household characteristic. A household with sufficient resources will choose between building a new home, renovating a substandard home, or buying a good home within commuting distance of its members' workplaces or schools. The real estate market, however, is an area characteristic, but
### Social environment

<table>
<thead>
<tr>
<th>Category</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnic history and composition, discrimination</td>
<td>Macintyre et al. (1993), Taylor et al. (1997), Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Churches, other community organizations</td>
<td>Macintyre et al. (1993)</td>
</tr>
<tr>
<td>Existence of networks of community support</td>
<td>Macintyre et al. (1993)</td>
</tr>
<tr>
<td>Facilitation of social networks</td>
<td>Boyle and Lipman (1998)</td>
</tr>
<tr>
<td>Social disorganization</td>
<td>Davey Smith et al. (1998)</td>
</tr>
<tr>
<td>Social capital (especially the supervision of adolescents)</td>
<td>Taylor et al. (1997)</td>
</tr>
<tr>
<td>Political, economic, religious history</td>
<td>Macintyre et al. (1993)</td>
</tr>
<tr>
<td>Norms and values</td>
<td>Macintyre et al. (1993), Anderson et al. (1997), Diez-Roux et al. (1997)</td>
</tr>
<tr>
<td>Reputation</td>
<td>Macintyre et al. (1993)</td>
</tr>
</tbody>
</table>

### Other

<table>
<thead>
<tr>
<th>Category</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban overload (social and information; after Milgram), or &quot;stress&quot;</td>
<td>Lindheim and Syme (1983), Anderson et al. (1997), Diez-Roux et al. (1997), Taylor et al. (1997), Boyle and Lipman (1998)</td>
</tr>
<tr>
<td>Urbanism</td>
<td>Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Publicity for cigarettes</td>
<td>Diez-Roux et al. (1997)</td>
</tr>
<tr>
<td>Mobility due to or prevented by health</td>
<td>Jones and Duncan (1995), Curtis and Jones (1998)</td>
</tr>
<tr>
<td>Immunity and exposure levels regarding infectious disease</td>
<td>Curtis and Jones (1998)</td>
</tr>
</tbody>
</table>

was mentioned only by Curtis and Jones (1998).
1.5 Chronic health problems and activity limitation

The studies reviewed in the previous section addressed diverse dimensions of health and health behaviours. The current study examines only two health constructs as defined by Badley et al. (1993). They derived a count of chronic health problems, and a measure of their potential impact on daily life, which they refer to as long-term disability. In this document, long-term disability will be referred to as activity limitation in keeping with the revisions being made to the International Classification of Functioning and Disability\(^{27}\) (WHO 1999). In this section, the natures and prevalence of similar measures used in the literature will be summarized, and the relationship between the two constructs will be discussed.

1.5.1 Chronic health problems

What qualifies as a problem?

The stated purpose of the question on chronic health problems in the OHS User's Guide (1990) was "to ensure that all long-term, permanent or recurring physical\(^{28}\) health problems (such as asthma, allergies, periodic migraine headaches, etc.) are identified."\(^{29}\)

Chronic mental illnesses were included in previous tools that were components of the Canada Health Survey and the Quebec Health Survey, but mental health was not considered a chronic health problem for the OHS since it was assessed separately\(^{30}\) (MOH 1990), nor were mental health problems included in the chronic health problem tool included in the National Population Health Surveys (NPHS).\(^{31}\) This separate treatment is unfortunate because mental illness is a prominent, and soon to be the leading, cause of activity limitation worldwide (Dewa and Lin 2000).

Migraines and Alzheimer's disease were mentioned in publications as being missing from the OHS survey instrument (To and Wu 1995; Badley et al. 1993). The first omission is mysterious, while the second might be due to it being defined as a mental illness. The National Population Health Survey was established in 1994, and its format and content were very similar to the 1990

---

\(^{27}\)This document is referred to here and elsewhere by the anachronistic abbreviation ICIDH-2.

\(^{28}\)Emphasis added.

\(^{29}\)Migraine headaches, despite this documentation, were not on the OHS list. Please refer to page 54 for the list of chronic health problems that were included on the 1990 OHS instrument.

\(^{30}\)There was an Ontario Health Supplement on mental health that was administered to a subsample of the last two quarterly replicates of the Ontario Health Survey (Dewa and Lin 2000)

\(^{31}\)In the NPHS, chronic health problems were referred to as chronic conditions, which seems to be a cosmetic or irrelevant change, and so is ignored in this discussion.
OHS. Nonetheless, in the 1994 NPHS chronic health problems list, they divided ‘allergies’ into ‘food’ and ‘other,’ and ‘eye problems’ into ‘cataracts’ and ‘glaucoma;’ they dropped ‘other digestive disorders,’ ‘other serious problems with joints and bones,’ ‘circulatory problems,’ ‘goitre or thyroid trouble’ and the ‘kidney disease’ component of urinary problems; and they added ‘migraine headache,’ ‘sinusitis,’ ‘Alzheimer’s disease or other dementia,’ and ‘acne requiring medication (for respondents aged less than 30).’ In 1996 the NPHS reinstated ‘a thyroid condition,’ and created the item ‘a bowel disease such as Crohn’s disease or colitis’ to replace ‘other digestive disorder,’ which had previously been deleted. In 1998 the NPHS made no further changes, perhaps indicating minimal critical feedback from the research community.

Lists of this type date at least to the 1983 screening tool administered during cycles of the Labour Force Survey to identify children for The Canadian Health and Disability Survey. On that tool some impairments and diseases missing from more recent tools were included, and some of the currently included health problems were missing. This seems to indicate that there are theoretically valid chronic health problems that are excluded for reasons specific to each survey. The reason does not appear to be related to the target population, since children are not the group most prone to uncorrectable hearing and vision loss, yet they alone, using the tools mentioned, have been surveyed for these impairments. The number of problems included in each tool is relatively constant, which indicates a likely pressure to achieve brevity rather than comprehensiveness, but the debate surrounding the priorities for inclusion on these tools is missing from the published literature.

Clearly survey lists of chronic health problems are intended to capture, in broad categories understood by most people, the types of chronic (and mostly physical) health problems that burden the population and consequently publicly financed health care. Close examination of the included problems and consideration of the revisions made between surveys illustrate the contentiousness of certain categories, a bias towards diseases, and the probable modest sensitivity of such survey instruments. The typically modest sensitivity of survey instruments for measuring population morbidity is well recognized and has inspired the metaphor ‘iceberg of morbidity,’ with an ‘iceberg

---

\(^{32}\text{prevalence} = 7.8\% \) (Mittmann et al. 1999)

\(^{33}\text{prevalence} = 5.3\% \) (Ibid.)

\(^{34}\text{prevalence} = 0.1\% \) (Ibid.)

\(^{35}\text{prevalence} = 2.7\% \) (Ibid.)
of disability' proposed by analogy (Verbrugge 1990). Nonetheless, research using such measures demonstrates that they can meaningfully distinguish groups of people based on the degree to which their health is imperfect (Badley et al. 1993), in such a way that imperfection reflects, directly and by proxy, elements of both disease and impairment (see figure 1.6).

The prevalence of chronic health problems

In the United States in 1987, 45% of the non-institutionalized population had at least one, and 20% had more than one, chronic health problem (Hoffman et al. 1996).36 In the 1994/5 National Population Health Survey (NPHS), among persons aged 12 or greater responding to a similar chronic health problems tool as was used in the OHS, these prevalences were 54.9% and 28.3% (Mittmann et al. 1999). Among adults aged 25 or greater in the OHS, these prevalences were 57.5% and 28.4%. There is some evidence to suggest that there is little educational or income-related inequality in the probability of reporting a chronic health problem at ages 25-34 or shortly after retirement, but that during midlife social inequalities in health appear, and then disappear (Badley et al. 1993). In the OHS is has been shown that five of the chronic health problems are the most common: hay fever or other allergies (17.7%), arthritis or rheumatism (14.0%), high blood pressure or hypertension (9.8%), serious trouble with back pain (9.2%), skin allergies or other skin diseases (8.4%). Of these, only hypertension and arthritis/rheumatism increase notably in prevalence with age.

1.5.2 Activity limitation

Activity limitation is defined in the ICIDH-2 as “difficulties an individual may have in the performance of activities,” and an activity is defined and explained as “the performance of a task or action by an individual...Activities represent the integrated use of body functions, often in a purposeful manner...The Activity dimension deals with the actual performance of the individual; it does not refer to an aptitude, potential, capacity or what an individual might do...The use of assistive devices or personal assistance does not eliminate the impairment but may remove limitations on activity in specific domains...” (p. 18)

36The definition of chronic health problem used in this study is not necessarily compatible with that in the OHS.
Operationalizing the activity limitation construct

Too often the theoretical models underlying disability and health status instruments are tacit rather than explicit, as was discussed above regarding chronic health problems. Ziebland et al. (1993) analysed these tacit models and enumerated their theoretical differences. Instruments can focus on physicians' assessments or self-reports, which affects scope, objectivity and proxy effects; can be restricted to physical function or can include physical, emotional, pain and social functions and experiences; can focus on actual performance or potential capacity; can focus on necessary activities such as working and self-care, or on usual activities, where usual could be in the recent or distant past, and usual for the individual or for the average person; and functional reduction could be in kind, degree or could be absolute, where kind includes dependence on people and/or devices. These theoretical differences can dramatically alter instrument characteristics. Measuring capacity rather than performance, for example, has been estimated by Anderson et al. to reduce prevalence estimates by 15-20% (Ziebland et al. 1993).

Activity limitation is a notoriously difficult construct to represent in a single question. A review of such questions by the International Network on Health Expectancy recommended their inclusion in population health surveys and concluded that they tend to have good coverage of specific disabilities and that they are distinct from self-rated health (Verbrugge et al. 1999). The Network's concept of disability\footnote{It is only in this sense that the word disability, which has too many meanings, will be used in this document.} included both the performance of activities and of roles, which exceeds the ICIDH-2 activity limitation construct, including also the ICIDH-2 participation restriction construct. They include in their survey of candidate questions the 1994 NPHS activity limitation screening question as an example of the relatively invariant Canadian attempts at single-question disability prevalence assessment, and note no major flaws.

Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) were combined with the activity limitation measure used by Badley et al. (1993). First proposed by Katz (Katz et al. 1963; Katz and Akpom 1976; Spector et al. 1987), this approach to measuring disability focusses on dependence on others in the necessary activities of the elderly, namely bathing, dressing, going to toilet, transferring, continence, and feeding (ADL), and shopping, transportation and housekeeping (IADL). These constructs map onto either or both of the ICIDH-2 concepts of activity limitation and participation restriction. Since these are not being distinguished by Badley...
et al. (1993) or in the current study, this non-specificity is not important for either study. Due to the detail of the ADL/IADL questions, combining this information with the self-report of activity limitation in the OHS most likely makes a small improvement in sensitivity at little or no cost to specificity.

Due to the limitations of single questions about the impacts of health problems, multi-item survey instruments have been developed, such as the Health Utilities Index Mark III (HUI), which was developed at McMaster University (Feeny et al. 1994; Boyle et al. 1995; Feeny et al. 1995; Feeny et al. 1998). The HUI is a multi-attribute utility measure. The multiple attributes are vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain, each of which is rated on a 5-6 item severity scale. Preference scores for levels of severity within each attribute and preference scores for health states comprised of one level each for all eight attributes are obtained empirically in separate survey work from the instrument’s administration as a part of population surveys. Preferences are generated by standard gamble and visual analog scale methods. A description of these methods is provided in Feeny et al. (1994). The summary score is bounded by 0, which represents desirability equal to death, and 1, which represents perfect health. The HUI aims to measure capacity rather than performance, and “adopt[s] a ‘within the skin’ definition of health status, i.e., the extent to which one’s health status permits one to engage in the activities one chooses to engage in or the extent to which health status inhibits such engagement.” (Feeny et al. 1995; Feeny et al. 1998) Limitations in the HUI attributes (except the emotion attribute) are graded in severity according to the individual’s reliance on aids and personal assistance, access to which is assumed. Due to these theoretical decisions in the design of the instrument it is not possible to map the HUI cleanly onto ICIDH-2 constructs. As part of the General Social Survey Reliability Study, the test-retest reliability of the HUI was examined and found to be acceptable (Boyle et al. 1995).

**The prevalence and distribution of activity limitation**

The prevalence is 9.7% in the Ontario population (all ages) according to Badley et al. (1993), using the same instruments and data as are used in this study. The Ontario figure is 16.0% according to the 1991 Health and Activity Limitation Survey (HALS),\(^{38}\) and 13.9% according to the same

\(^{38}\)The HALS uses a disability screening question that is virtually identical to that in the OHS.
survey in 1986.\textsuperscript{39} The HALS estimates for Ontario are higher than the national averages of 15.5% in 1991 and 13.2% in 1986 (Statistics Canada 1990; Statistics Canada 1995). 5% of the 1986 HALS 'no' responses on the screening question changed to 'yes' on follow-up, which may explain part of the discrepancy between the OHS and the HALS (Badley et al. 1995).

The prevalence of activity limitation definitely increases with age, although perhaps not for the poor or poorly educated, whose rates at all ages approximate those of the elderly (Badley et al. 1993). The 1991 HALS prevalence estimates by age group are 7.0% (<15yrs), 8.0% (15-34yrs), 14.0% (35-54yrs), 27.1% (55-64yrs), 37% (65-74yrs), 57% (75-84yrs) and 84% (85+yrs). The all ages increase over 1986 was concentrated in all age groups less than 55 years of age. (Statistics Canada 1995). While the prevalence of activity restriction increases rapidly with age, the number of people affected peaks in the 35-44 age group and the 75+ group accounts for only 8.8% more cases than the 25-34 year old group in Ontario, due to the taper of our population pyramid (Badley et al. 1993).

\textbf{The role of the environment regarding activity limitation}

The ICIDH-2 describes environmental factors as either barriers or facilitators of activity and participation in life situations, and divides the environment into broad domains: (1) products and technology, (2) natural environment and human made changes to environment, (3) support and relationships, (4) attitudes, values and beliefs, (5) services, and (6) systems and policies.

Brandt and Pope (1997) argue that environmental modification could eliminate much activity limitation and participation restriction. By 'environment' they refer to both physical space and to social structures at the family, community and societal levels. This notion of environment includes, but is clearly larger than that of neighbourhood. In the current study neighbourhoods are described by their educational level and by their urbanism, and so this model of the role of the environment is only tested insofar as the barriers and facilitators are located within the neighbourhood, and to the extent that they are collinear with either education level or urbanism.

\textsuperscript{39}The HALS estimates include institutionalized persons. The Ontario household numerator over the total denominator, however, only reduces the provincial estimate to 12.9% in 1986.
Figure 1.5: The man in the diagram represents the person, and the dark box the environment. Arrow A represents the disabling process, in which the needs of the individual grow and exceed those supported by the environment. Arrows B and C show the two types of enabling processes, where B represents environmental modification and C represents functional restoration. Modified from Brandt and Pope (1997).

1.5.3 The influence of chronic health problems on activity limitation

An important theoretical assumption here is that chronic health problems are necessary, though not always sufficient, causes of activity limitation. This has been asserted in the disability literature using varying terminology (WHO 1980; Verbrugge 1990; Pope and Tarlov 1991; Badley 1995; Fitzpatrick and Badley 1996) and assumed in prior work on the Ontario Health Survey (Badley et al. 1993). Verbrugge et al. (1999), using the term disability to refer “to the impacts health problems have on people’s social functioning, that is, their ability to perform roles and activities,” go on to say that “[b]ecause research and policy interests are typically on long-term dysfunctions associated with chronic conditions,40 that is the focus here.” Graphical models of this relationship

---

40Emphasis in the original.
Model 1

Disease or Disorder \(\rightarrow\) Impairment \(\rightarrow\) Activity Limitation \(\rightarrow\) Participation Restriction

Model 2

Chronic Health Problems (Disease, disorder & impairment) \(\rightarrow\) Activity Limitation & spill-over of survey item into Participation Restriction

Figure 1.6: Model 1 is a figure appearing in both the ICIDH and the Nagi traditions of disablement studies, which represents the two main models of disablement (WHO 1980; Verbrugge 1990; Pope and Tarlov 1991; Badley 1995). While model 1 uses ICIDH-2 language, the causal flow from left to right through constructs of increasing physical and social scale is common between these schools of thought. Model 2 is a modification of this figure to represent the assumptions expressed in Badley et al. (1993), and in the current study. Since the ICIDH-2 claims not to propose models of disablement, but merely to provide a language for this pursuit, no official ICIDH-2 model is provided.

are summarized in figure 1.6.

Despite this perspective being widely accepted, some research makes a persuasive case for feedback effects (Verbrugge and Jette 1993; Simeonsson et al. 2000), an example of which would be the development of chronic ulceration due to immobility and the resulting chronic pressure. The language proposed is often 'secondary impairment,' meaning that an activity limitation resulting from an impairment has caused another one to develop. There remains an assumption that chronic health problems are necessary causes for activity limitations. The ICIDH-2 avoids providing guidance on this and related issues of causality. It will be assumed in the current study that activity limitation is a very minor contributing cause of chronic health problems.

The chronic health problems that account for the greatest proportions of limited activities are musculoskeletal disorders, which alone accounts for half of activity limitation, followed by circulatory and respiratory diseases (Badley et al. 1993; Badley et al. 1994). Previous work on the OHS has shown that 93% of those with activity limitation report a chronic health problem, and that 70% of this overlap is causally attributed to a reported chronic health problem. For every health problem the unadjusted odds were shown to be significant for experiencing activity limitation, with
odds ratios ranging from 1.1 for hay fever and other allergies to 7.3 for serious trouble with back pain. This evidence agrees well with that provided by Verbrugge et al. (1999). It would appear that the impact of chronic health problems on activity limitation is highly conditional on co-morbidity. The same research also showed that the number of chronic problems greatly increases the risk of activity limitation (Badley et al. 1993).

1.6 Summary of Chapter 1

There are between-neighbourhood inequalities in diverse dimensions of health. These inequalities are not entirely explained by the statistical descriptions of neighbourhood socio-economic and demographic composition that have been used to date. The large, pervasive and multidimensional impact of individual socio-economic status could account for many of the putatively contextual effects reported in the literature, although a small number of studies have observed effects in directions that would not be predicted under the causal assumption of unmeasured socio-economic compositional differences. The differential impact of context on socio-economic groups has rarely been examined. None of the Canadian studies, and very few studies internationally, have distinguished between urban and rural neighbourhood status when examining contextual effects.

Chronic health problems and activity limitation are prevalent and costly types of human suffering, and have been theorized to be susceptible to contextual determination and modification. Of the dimensions of health examined to date, levels of co-morbidity may exhibit the greatest unexplained neighbourhood inequalities. In Canada, very little research has attempted to control for compositional differences that might explain neighbourhood levels of chronic health problems and activity limitation.
Chapter 2

This project

2.1 Purpose

This study examined whether the age, gender, education and income of residents of neighbourhoods accounted for the differences between Ontario neighbourhoods in their average number of chronic health problems and in their average probability of consequent activity limitation. Simultaneously, between-neighbourhood variation in the effect of individual education and income on these dimensions of health was estimated. The observed between-neighbourhood variations in individual education and income effects and in adjusted neighbourhood health means were tested for their associations with neighbourhood education level, which is hypothesized to represent or predict contextual neighbourhood effects on residents’ health. The effect of neighbourhood education level was allowed to interact with rural status, so that any contextual effects specific to urban or rural areas could be identified.

Health is represented in this study by two related constructs: chronic health problems and activity limitation, as defined in, or perhaps somewhat exceeding the definition in, the International Classification of Function and Disability (WHO 1999). The count of self-reported chronic health conditions is modelled assuming an overdispersed Poisson distribution, and activity limitation, measured first using a self-report of activity limitation, and second using a dichotomized version of the Health Utilities Index Mark III, is modelled logistically. In models of activity limitation, the individual’s number of chronic health problems was controlled for to distinguish effects on the former from those on the latter. The socio-economic status of the individual is represented primarily by education and secondarily by income. Education was selected as the primary socio-economic indicator since it is less subject to health-selection. When income is included it is in order to
provide some evidence that the education effects are not highly confounded by other dimensions of socio-economic status. Neighbourhoods are represented by enumeration areas, which are described as either urban or rural and on a continuum of aggregate education-level.

2.2 Theoretical Models

The theoretical models in this study are summarized in figure 2.1. These are the causal schematics on which the multilevel models are designed. The top half of the schematic represents characteristics of the neighbourhood, while the bottom half represents individual characteristics, in this case health and socio-economic status.

![Diagram](image)

Figure 2.1: This schematic adapts the general case described in figure 1.3 to the current study. Context has become urbanism and neighbourhood education level, and health is represented by the related constructs of chronic health problems and activity limitation.

The multi-level nature of the models is evident in figure 2.1. Area variation in health (arrows 3) is distinguished from the effect of individual socio-economic indicators (arrow 4) and neighbourhood modification (arrow 2) of the individual socio-economic effect on health (arrow 4) is hypothesized. Any significant variance regarding arrows 2 (significant variation in the individual level socio-economic effects) and 3 (significant variation in the intercept) can be tested to see if urbanism or area education level are predictors of between-neighbourhood differences.

2.3 Specific hypotheses

(1) It was hypothesized that controlling for individual-level socio-demographic characteristics would reduce between-neighbourhood variation in the mean number of chronic health prob-
lems and the mean probability of activity limitation, but that both health constructs would exhibit significant residual area variation.

(2) It was expected that the effect of the individual's education and income on both health constructs would vary by neighbourhood.

(3) Lower levels of neighbourhood education were expected to predict a greater mean number of chronic health problems and a higher prevalence of activity limitation.

(4) The effects of neighbourhood education level was expected to be greater among persons with lower levels of individual education.

(5) All neighbourhood effects were expected to be greater in urban than in rural areas.

These hypotheses are represented graphically in figure 2.2.

Figure 2.2: LEGEND:
A: The ideal.
B: Highly educated people in neighbourhoods with high levels of education.
C: Highly educated people in neighbourhoods with low levels of education.
D: Less educated people in neighbourhoods with high levels of education.
E: Less educated people in neighbourhoods with low levels of education.

These two drawings illustrate the hypothesized influence of neighbourhood context on the individual characteristics SES and health. The hypothesised survival curve on the left is the longitudinal trend that is thought to underlie the cross-sectional data investigated here. The squaring of the health-adjusted survival curve as a phenomenon of the epidemiological transition has been proposed by Fries (Fries 1980; Vita et al. 1998) and the socially-graded nature of this process has been suggested by House et al. (1994). The version presented here adds the role of neighbourhood context to House's proposal. The drawing on the right mimics the graphs of the results from multilevel models in a simplified form, and introduces the urban v. rural distinction. Note that the role of neighbourhood education was expected to be greater in urban areas, but the primary effect of urban v. rural status was not originally considered, and so the placement of the urban and rural means relative to one another is a refinement of the hypotheses at the time of writing.
2.4 Relevance

There are three possible conclusions to this study, each with its own messages for social epidemiological theory:

(1) If it is found, having controlled for individual age and gender, that there is little or no variance in these health indicators across enumeration areas, then we can conclude that the only differences between neighbourhoods in chronic health problem co-morbidity and activity restriction are demographic in nature, and that we have no moral mandate to study further or to intervene.

(2) If it is found that individual socio-economic status explains all age and sex standardized area differences in health, then the question will become a sociological one: to what extent do people have access to, or make choices regarding, neighbourhoods based on their socio-economic status, as opposed to neighbourhoods determining the socio-economic characteristics of individuals. Because of the importance of socio-economic status for the health of individuals, the latter causal explanation would be of relevance to public health.

(3) If it is found, having controlled for individual socio-economic and demographic composition, that there are random differences between neighbourhoods in health or in the meaning of individual socio-economic status for health, then there is a need to describe the exposures, be they ecological or individual, that account for these between-neighbourhood differences. If the area descriptors used in this study are predictive, they may suggest avenues for further research.
Chapter 3

Methods

3.1 Data sources

3.1.1 The 1990 Ontario Health Survey

Sample design

The 1990 OHS sampling scheme was stratified by Public Health Unit (PHU) and by Urban/Rural status according to the 1986 Census. EAs were randomly selected from PHUs, and households were randomly selected from EAs. One individual per household answered on behalf of all household members. An average of 46 EAs per PHU were selected, and roughly 15 and 20 dwellings were sampled from each urban and rural EA respectively, such that the total number of respondents is 61,239 (living in 35,479 dwellings), of which 38,818 were aged 25+ and completed the entire survey. Within each household a knowledgeable member replied to the interviewer-completed portion of the survey on behalf of each household member (from which all variables for this study are drawn). The target population consisted of all residents of private dwellings in Ontario during the survey period January through December 1990 (though no interviews took place in July or December, as response rates are typically poor during these months). Residents of Indian reserves, inmates of institutions, foreign service personnel and residents of remote areas were excluded.

Response rates

The household response rate for the interviewer-administered questionnaire, from which all responses were taken, was 87.5% (5.9% not at home, 4.2% refusal, 2.4% language/other) and was lower in urban areas (86%) than in rural areas (90%). Unfortunately, the public use file contains data only for individuals who also completed the self-administered questionnaire, and this response
rate is 77.2

Missing health data

Values for dependent variables were missing for 658 individuals (1.7% of the sample). These respondents were excluded from the study sample. All other missing data were imputed by the ministry or myself, as discussed below.

Ministry imputations

The Ontario Ministry of Health performed imputations on missing values for age and sex, since these were required for the weighting process. Sex was imputed by hot deck\(^1\) with separate decks for values of Public Health Unit (PHU). The values for age were also imputed by hot deck, with separate decks for values of PHU, marital status, and main activity (Question 96, Form 4). It was not reported how many records were imputed (MOH 1990).

Imputations completed as part of this project

Missing values for education and income were imputed by hot deck with separate decks for gender and age. Age was divided into only two groups: 25–64 and 65+. This was performed using the SAS hot deck method developed by Stiller and Dalzell at the U.S. Census Bureau (1998). 0.77% of education records were imputed. The incomes on records with missing values were very similar to incomes on complete records (see figure 3.1). 14.71% of income records were imputed. The educational attainments stated on records with missing values were very similar to the educational attainments stated on complete records (see figure 3.1).

Data limitations

Since the survey is representative of the Ontario population, it would be appropriate to use sample weights in order to maximize the external validity (generalizability) of the results. However, the software used to run the hierarchical models does not permit the use of sample weights in non-linear

---

\(^1\)Hot deck imputation involves stratifying the population into theoretically valid categories, then for each stratum the records are treated as cards in a deck and cards with missing values are randomly assigned the value from a complete card in the same deck. The advantages of hot deck imputation are (1) that the records with missing values for one variable are not lost from the sample, and other responses on the record can be used, and (2) that the imputed value is selected from its estimated distribution, thus maintaining a smooth distribution in the final data set. The selection of the stratifying variables and the proportion of records imputed are key discretionary elements of this technique. Despite its popularity, the literature on the theoretical properties of the method are very sparse. (Little and Rubin 1987)
Table 3.1: Imputations on education and income are validated by analysis of the unimputed distribution of the alternate socio-economic indicator for the complete and imputed records. This is the complete set of records prior to the listwise deletion of records for missing values on the dependent variables. All numbers are percentages. Income categories are in thousands of dollars.

<table>
<thead>
<tr>
<th>Imputations for education</th>
<th>Imputations for income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>Complete</td>
</tr>
<tr>
<td>0-12</td>
<td>6.1</td>
</tr>
<tr>
<td>12-20</td>
<td>10.4</td>
</tr>
<tr>
<td>20-30</td>
<td>13.4</td>
</tr>
<tr>
<td>30-40</td>
<td>15.6</td>
</tr>
<tr>
<td>40-50</td>
<td>15.0</td>
</tr>
<tr>
<td>50-60</td>
<td>12.9</td>
</tr>
<tr>
<td>60-70</td>
<td>9.0</td>
</tr>
<tr>
<td>70-80</td>
<td>6.1</td>
</tr>
<tr>
<td>80+</td>
<td>11.5</td>
</tr>
</tbody>
</table>

models. The survey also interviews a single household member who self-reports and also acts as proxy for the other household members, which may introduce a mixed proxy effect. Investigation of this possibility revealed that for more objective phenomena (e.g., chronic health problems, labour force participation) the reliability of proxy reports was acceptable, while for more subjective phenomena (e.g., emotion, pain) it was biased towards under-reporting of suffering. This effect was mainly observed for the interviewer-administered questionnaire (Grootendorst et al. 1994), from which all the data for this study are taken. This effect likely reduces the predictive value of individual characteristics, but unless the proxy rate or effect varies by enumeration area, there is not likely to be any reduction or bias in random intercept models, and only a bias toward the null regarding random-slope models.

3.1.2 The 1986 Canadian Census

1986 census information was used rather than 1991 because the definition of enumeration area boundaries shifts between censuses, and the OHS respondents are coded to 1986 definitions, such that it would have been much more difficult to match individuals to the level 2 clusters using 1991 data.

The long-form education question

The long-form, which contains the question regarding the highest level of formal schooling achieved by the individual, is administered to 20% of the population aged 15 or more on Census day.
Electronic versions of the basic summary tables for Canadian Enumeration Areas were acquired from the University of Toronto. For each enumeration area the 15-24 yr old values were subtracted from the 15+ yr values to generate estimates of the educational profile for each enumeration area among adults aged 25+. Six education categories to parallel the individual level were available in the census for 1986. Primary education or less is defined by the long-form as 'less than grade 9,' and university education as 'university degree.'

Data limitations

The Census data is four years older than the individual data, but this is not expected to introduce much error since it has been observed in other settings that the socio-economic ranking of small areas is slow to change (Dolan et al. 1995).

3.2 Dependent variables

3.2.1 Chronic health problems

The OHS allows binary responses to each part of the following compound question (Question 77, Form 4):

Do/Does _______ has/have:
(a) Skin allergies or other skin diseases?
(c) Serious trouble with back pain?
(e) Other serious problems with the joints or the bones?
(g) Asthma?
(i) Epilepsy?
(k) Circulatory problems?
(m) Diabetes?
(o) Stomach ulcer?
(q) Goiter or thyroid trouble?
(s) Cancer

(b) Hay fever or other allergies?
(d) Arthritis or rheumatism?
(f) Paralysis or speech problem due to stroke?
(h) Emphysema or chronic bronchitis or persistent cough?
(j) High blood pressure or hypertension?
(l) Heart disease?
(n) Urinary problems?
(p) Other digestive problems?
(r) Eye problems, for example, glaucoma, cataract?

Question 79 on the same form asks “Do/does _______ have any other long term health problem?” Up to eight conditions could be identified. For conditions such as Alzheimer’s disease and migraine headache, this was the only question that could identify such a problem. This format, however, places a considerable burden on the respondent/proxy to identify any conditions that might be valid members of the set described to them in question 77. This format likely biases the ‘other’ category towards more serious cases (To and Wu 1995), and a higher than normal socio-economic reporting bias is possible. Because of the unknown number of conditions that a positive
response indicates, and because of the possibility of a confounding bias, question 79 was not used.

There appear to be a small number of individuals who claim to have most chronic health problems, and yet appear healthy on other health status measures, such as the HUI. It is also the case that although there are 19 chronic health problems, and that there are people who claim to have 18 and 19 of them (however, not 13-17), 99.9% of the sample had 8 or fewer chronic health conditions. In order to minimize the effect of these potential outliers, the count variable has been truncated to the nine values of 0 through 8+.

To count or not to count

Counting chronic health problems ignores their relative prognostic severity and functional significance. Mitchell et al. (2000), provided with similar data in the UK, chose to dichotomize the count on the grounds that the items were not of equal severity and that integer treatment was therefore artificially precise. It should also be recognized that the prognostic severity and functional significance of a chronic problem may depend very much on the other chronic problems experienced by the individual in question. To the extent that this is true, the assumption of additivity (which is distinct from equivalence) implicit in the arguments presented by Mitchell et al. (2000) may not be valid.

3.2.2 Health Utilities Index III

It was originally intended to use the Health Utilities Index (HUI) as a second measure of impairment. This seemed justified given the intended ‘within the skin’ definition of health status that the HUI was designed to measure (Feeny et al. 1998). However, after careful consideration of the HUI question items (Questions 35-65, Form 4) and the ICIDH-2 framework, it was felt that the HUI measured neither impairment nor activity restriction, but rather represented elements of both. Since the study requires the impairment measure to be at least logically prior to the activity measure, it was decided to examine the HUI to determine whether it had relationships to the other dependent variables that could validate its position in the ICIDH-2 framework.

The HUI produces a score that is bounded by 0 and 1 and, due to ceiling effects, has a general population mean in excess of 0.9. This score is not only not normally distributed, it is extremely difficult to transform. (Kopec et al. 2000) resolved this problem to some extent by establishing and validating two cut points for dichotomization on the HUI scale. The higher cut point of 0.947
<table>
<thead>
<tr>
<th>BINACT</th>
<th>HUI830</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>30229</td>
<td>78.9</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3715</td>
<td>9.7</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1392</td>
<td>3.6</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2973</td>
<td>7.8</td>
</tr>
</tbody>
</table>

Mean: 0.115 0.175

Table 3.2: The interrelationship of the two activity restriction variables, BINACT and HUI830. Spearman Correlation Coefficient = 0.478, p < 0.0001.

(we will refer to this variable as HUI947) separates ‘healthy’ from ‘not healthy’ individuals, such that healthy individuals can have as complaints only correctable vision problems. The lower cut point of 0.830 (we will refer to this variable as HUI830) separates ‘disabled’ from ‘not disabled’ individuals, such that disabled individuals have a utility score less than that scored by someone who required a cane or walker, but has no other limitations. Preliminary investigations revealed that the relationship of HUI947 to the explanatory variables, particularly age (the dominant effect), did not justify its use as a measure of impairment (see figures B.1, B.2, B.3 and B.4 in appendix B). HUI830, however, performed very much like the BINACT variable (described below), although, as shown in table 3.2, HUI830 appears to be identifying more individuals as disabled. Other investigators have used the 0.830 cut point (Williams et al. 1998), but it should be remembered that such dichotomization is fundamentally arbitrary, and that other cut-points have been used, such as 0.800 (Roberge et al. 1995).

Given these theoretical decisions and empirical findings, HUI947 will not be used in this study, and HUI830 will be used as a second measure of activity restriction.

3.2.3 BINACT

BINACT (short for binary activity restriction) is the name of the main activity restriction variable used in this study, and is referred to sometimes as ‘self-reported activity restriction.’ This variable has a value of 1 only if the respondent answered yes to the question:

Compared to other people of the same age in good health, are/is ...... limited in the kind or amount of activity ...... can do because of a long-term physical or mental condition or health problem?
(Question 70a, Form 4)

or to the activities of daily living (ADL) questions:

Do/does ...... need the help of another person with personal care such as eating, bathing, dressing or getting around inside the house, because of any impairment or health problem?
(Question 66, Form 4)
Do/Does ..... need the help of another person in looking after personal affairs, doing everyday household chores, going shopping or getting around outside the house, because of any impairment or health problem?
(Question 67, Form 4)

As seen in table 3.2, the prevalence of activity restriction according to this variable is 11.5%. BINACT positive values are generated primarily by question 70a. 1.0% of the sample answered 'no' to question 70a and 'yes' to questions 66 or 67. These individuals are thought to be more likely disabled than not, given that question 70a is much less specific, and so they are included as positives in BINACT.

3.3 Independent variables

3.3.1 Individual

Age

The OHS contains continuous integer age data. In order to create stable multilevel models, the age effect could not be modelled categorically, and it was necessary to treat it as a centered continuous variable, and to make assumptions about the relationship of age to the dependent variables. Age was centered on its mean of 48 years.

For each dependent variable age was first modelled as 5 year categories up to 80+, with 25-29 as the reference category, so that the relationship of age to the dependent variables (controlling for sex, income and education) could be assessed visually as a histogram of the coefficients for each age variable. These analyses were performed using PROC LOGISTIC and PROC GENMOD, since the likelihood estimators in HLM (Hierarchical Linear Modelling 5.0) cannot be used for likelihood ratio tests. The suspected linear, quadratic, or cubic expressions were then compared as nested models by likelihood ratio tests (Zelterman 1999).

For activity limitation, which is modelled as two binary variables (BINACT, HUI830), the relationship is strongly cubic. For chronic health problems, which are modelled as a count variable assuming a poisson distribution, a quadratic expression best characterizes the relationship. The linear effect of age was centered on the grand mean, and the squared and cubic age variables are based on this centered variable.
Female

Sex is treated as a binary variable that takes a value of 1 for women. In all models, for uncentered independent variables the reference group was made to be the healthiest group, in this case men, in order to facilitate interpretation of the intercept.

Education

Preliminary investigation of the education effect revealed it to have a non-ordinal effect on the dependent variables, so a continuous approximation of education was not attempted, and its effects were tested categorically. The education categories used are (1) primary education only, (2) some secondary education, (3) secondary education, (4) some post-secondary education, (5) completed college diploma, and (6) completed university degree. The sixth category was used as the reference group.

Income

Information on annual household income was available in the following categories: $0k, 0-6, 6-12, 12-20, 20-30, 30-40, 40-50, 50-60, 60-70, 70-80 and 80k+.

Due to small numbers and erratic effects, the first three categories were combined to form a $0-12k group. As was done for age, the income effects were visualized, controlling for age, sex and education, and quadratic expressions of income were selected by testing nested models by likelihood ratio test. In multilevel analyses the income variables were centered on the central group. Since the nine groups are of nearly equivalent sizes, centering on the central group ($40-50k) was equivalent to centering on zero. The categories were therefore coded as -4, -3, -2, -1, 0, 1, 2, 3 and 4, where -4 represents $0-12k and 4 represents $80k+.

Chronic health problems

Please refer to section 3.2.1 for a description of the variable construction. The count of chronic health problems is used as a predictor of activity limitation. As was done for age and income, the effects were visualized, then a quadratic expression for the effect of chronic health problems was selected by testing nested models by likelihood ratio test. In multilevel analyses the chronic health problem variables were not centered on the grand mean when used as predictors, so that 0 chronic conditions would be the reference group.

\(^2\)k is used to refer to thousands, as in $80 thousand +.
Figure 3.1: The neighbourhood education measure is the log of the ratio of the proportion of adults aged 25+ with primary education or less to the proportion with a university degree. Higher values indicate a lower aggregate level of education for the enumeration area. The mean, error and range are reported in table 4.1. This figure represents the uncentered data.

3.3.2 Enumeration area

Education level (EA-education)

There is no standard ecological description of enumeration area educational level, so for this study three summaries were compared. Median educational category, expressed as a categorical and as a continuous variable resulted in models with excessive multi-collinearity. The proportion with primary or less than primary education and the proportion with university education, used separately to represent the tails of the EA educational profile, also resulted in inflation of the standard error of the level 2 parameter estimates. The log of the ratio of the primary or less proportion to the university proportion was used to combine the information on the two tails and, when centered, could be included in stable models. In the construction of the ratio measure, a value of half a person was added to proportions equalling zero to avoid dividing by zero or taking the natural log of zero. The frequency distribution of the logged ratio measure is displayed in figure 3.1.

See table B in appendix B for documentation of this decision to center the EA-education variable.
Rural

The rural variable takes a value of 1 if, according to the 1986 census, the EA was considered not urban by Statistics Canada. Urban areas are defined as “continuously built-up areas with a population concentration of 1,000 or more and a population density of 400 or more per square kilometer based on the previous census.” (Statistics Canada 1987) While centering the urban-rural dummy on zero reduces the collinearity among the level 2 predictors, this was found to be unnecessary (see table B in appendix B.).

3.4 Multilevel analysis

All multilevel models were estimated using HLM 5 (Hierarchical Linear Modelling), distributed by Scientific Software International (http://www.ssicentral.com/), on a SUN Enterprise 3000 server. All other calculations were performed using SAS 6.12.

3.4.1 Models used in this study

The following is the notation that will be used to discuss multilevel models throughout this document. Equation 3.1, which we will call level 1, is some function of the individual-level outcome variable (i.e. dependent variable), regressed on individual characteristics (i.e. $X_{ij}$). Equations 3.2 and 3.3, which we will call level 2, are the adjusted neighbourhood intercept and slope regressed on characteristics of the neighbourhood (i.e. $W_j$ or $Z_j$).

\[
\begin{align*}
  f(Y_{ij}) &= \beta_0j + \beta_{1j}X_{ij} + r_{ij} \\
  \beta_0j &= \gamma_{00} + \gamma_{01}W_j + u_{0j} \\
  \beta_{1j} &= \gamma_{10} + \gamma_{11}Z_j + u_{1j}
\end{align*}
\]

Where:

- $i$ are individuals in groups $j$;
- $f(Y_{ij})$ are the values of the dependent variable transformed by the function $f$ for individuals $ij$;
- $X_{ij}$ are the values of the independent variable $X$ for individuals $ij$;
- $W_j$ are the values of the independent variable $W$ for groups $j$;

\footnote{This notation is based on that used by Bryk and Raudenbush (1992), with some modification for linearized dependent variables.}
Figure 3.2: Observed counts of chronic health problems in the study sample compared to the Poisson distribution with a mean of 1.1 demonstrate that the Poisson distribution is a valid assumption for this data.

$Z_j$ are the values of the independent variable $Z$ for groups $j$;

$\beta_{0j}$ are the intercepts for groups $j$;

$\beta_{1j}$ are the coefficients associated with $X_{ij}$ for groups $j$;

$\gamma_{00}$ is the intercept for $\beta_{0j}$;

$\gamma_{01}$ is the coefficient associated with $W_j$;

$\gamma_{10}$ is the intercept for $\beta_{1j}$;

$\gamma_{11}$ is the coefficient for associated with $Z_j$;

$r_{ij}$ is the level-1 error term, or unique effect of person $ij$;

$u_{0j}$ is the error term for the level-1 intercept, or the unique effect of group $j$ on $\beta_{0j}$; and

$u_{1j}$ is the error term for the level-1 coefficient (slope), or the unique effect of group $j$ on $\beta_{1j}$.

For the count of chronic health problems (CC8), the distribution of the data suggests an overdispersed Poisson (see figure 3.2), or log linear, model, while the dichotomous variables for activity restriction (BINACT and HUI830) are modelled with Bernouilli, or logit, models. Therefore the following link functions$^5$ ($f(Y_{ij})$) can be specified:

\[ f(CC8_{ij}) = \ln(CC8_{ij}) \]  

$^5$Link function: A particular function of the expected value of the response variable used in modelling as a linear combination of the explanatory variables. For example, logistic regression uses a logit link function rather than
\[ f(BINACT_{ij}) = \logit(BINACT_{ij}) = \ln \left( \frac{BINACT}{1 - BINACT} \right) \] 
\[ f(HUI830_{ij}) = \logit(HUI830_{ij}) = \ln \left( \frac{HUI830}{1 - HUI830} \right) \] 

Technically speaking, it is the probability that BINACT and HUI830 equal 1, and not the values themselves that should appear in the logit functions. If this were not the case the odds would be undefined whenever the variables take a value of 1. For simplicity’s sake, the link functions have been written ignoring this technicality.

### 3.4.2 Novel challenges for model building

#### Centering

Centering variables on zero by subtracting the mean from each value does not alter the variance structure of the variable and has two convenient properties. First, since zero is defined as the mean, the intercept for the model refers to the average person for that variable. While zero sometimes has a useful theoretical meaning, such as zero chronic conditions, other variables, such as age, if uncentered, make the value of the intercept nonsensical. When variance across clusters in the level 1 intercept is being modelled, a sensible definition of the intercept is essential. Second, the correlation between \( x \) and \( x^2 \), and between \( x \) and \( xz \) are dependent on the scale of \( x \). If \( x \) is uncentered, these pairs will be highly correlated, but if \( x \) is centered both pairs will be orthogonal if \( x \) and \( z \) are normally distributed. For non-normal variables the collinearity is greatly reduced (Aiken and West 1991). This reduction in collinearity is clearly desirable. In a multilevel context, centering becomes a more important factor than in fixed-effects regression, because of the potential for a high correlation between the random intercept and random slopes. This correlation can also be eliminated or greatly reduced by centering raw scores on zero (Kreft et al. 1995).

#### Level 1 interactions and random slopes: constraints on model complexity

The computational requirements and the difficulty of achieving convergence in multilevel models depends a great deal on the number of random coefficients in the model. For this project the raw expected values of the response variable; and Poisson regression uses the log link function. The response variable values are predicted from a linear combination of explanatory variables, which are connected to the response variable via one of these link functions. In the case of the general linear model for a single response variable (a special case of the generalized linear model), the response variable has a normal distribution and the link function is a simple identity function (i.e., the linear combination of values for the predictor variable(s) is not transformed). From: *Common Concepts in Statistics*, M. Tevfik Dorak (http://dorakmt.tripod.com/mtd/glosstat.html).
most complex models were specified with a random intercept, with five random dummy variables for individual education, and with two random coefficients for the quadratic expression of income, giving a total of eight random coefficients. Some of these models required in excess of one-hundred hours of CPU time, as compared to a minute or so for models with only a random intercept. These full models were also extremely sensitive to multicollinearity among the level 2 predictors. If an additional goal had been to control for the age-specific effects of individual education, then the interaction of age and indicators of socio-economic status would have increased the number of random coefficients to fifteen or twenty-two, depending on whether age was being expressed quadratically or cubically. If such models could be made to converge, the CPU hours required would be unjustifiable. This example serves to illustrate the constraints on model complexity that are encountered in multilevel modelling. If one wishes to know, as we do here, whether the effect of socio-economic variables depends on the characteristics of the neighbourhood, the only practical route is to ignore the level 1 interactions involving socio-economic variables so that the effects of each is represented by a single coefficient. This does, however, allow known but unrepresented level 1 compositional effects to manifest themselves as a contextual (level 2) effect. For this reason conclusions from slopes as outcomes models should be interpreted with caution.
Chapter 4

Results

4.1 The study population

In table 4.1 the mean, standard error and range of each variable, and a reminder of how many individuals and enumeration areas these apply to, are provided. Centered variables are easily identified by their mean of zero. Refer to sections 3.2 and 3.3 for details regarding variable definitions.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>N</th>
<th>Unweighted</th>
<th>Weighted¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Level 1 - dependent variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC8</td>
<td>38160</td>
<td>1.10</td>
<td>1.34</td>
</tr>
<tr>
<td>BINACT⁵</td>
<td>38160</td>
<td>0.11</td>
<td>0.32</td>
</tr>
<tr>
<td>HUI830</td>
<td>38160</td>
<td>0.18</td>
<td>0.38</td>
</tr>
<tr>
<td><strong>Level 1 - predictor variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>38160</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>Age</td>
<td>38160</td>
<td>0.00</td>
<td>15.74</td>
</tr>
<tr>
<td>Primary</td>
<td>38160</td>
<td>0.14</td>
<td>0.35</td>
</tr>
<tr>
<td>Some Secondary</td>
<td>38160</td>
<td>0.24</td>
<td>0.43</td>
</tr>
<tr>
<td>Secondary</td>
<td>38160</td>
<td>0.25</td>
<td>0.43</td>
</tr>
<tr>
<td>Some Post-Secondary</td>
<td>38160</td>
<td>0.11</td>
<td>0.31</td>
</tr>
<tr>
<td>College</td>
<td>38160</td>
<td>0.13</td>
<td>0.33</td>
</tr>
<tr>
<td>University</td>
<td>38160</td>
<td>0.13</td>
<td>0.33</td>
</tr>
<tr>
<td>Income</td>
<td>38160</td>
<td>-0.06</td>
<td>2.35</td>
</tr>
<tr>
<td><strong>Level 2 - predictor variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA-education</td>
<td>1925</td>
<td>0.00</td>
<td>1.68</td>
</tr>
<tr>
<td>Rural</td>
<td>1925</td>
<td>0.21</td>
<td>0.41</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>1925</td>
<td>0.16</td>
<td>0.83</td>
</tr>
</tbody>
</table>

¹Weights were not used due to software limitations. Weighted descriptive statistics based on the same observations are provided for comparison purposes only.

⁵The denominator for the variance was the sum of the weights less one.

Table 4.1: Descriptive statistics for OHS and Census-derived variables. Note by the means that age and EA-education have been centered.
Weighted estimates of individual variables are provided to demonstrate the effects of the sample weights provided with the OHS data files, since weights could not be used due to software limitations. The weighted estimates show that relative to the Ontario population this sample has a slight excess of women and persons whose highest academic achievement is 'some secondary' education, and that university graduates, rich people and young people are slightly under-represented. Since this study is not generating population estimates, but rather is testing predictive associations, such small deviations from representativeness are not expected to affect the results. For a multi-tabulation of frequency counts by individual education, gender and age in urban and rural areas, see table 4.2.1.

4.2 Multilevel analysis

4.2.1 How to read the output tables

Summary tables (tables 4.4, 4.5 and 4.6) of key output from the multilevel analyses will be used to describe the presence of neighbourhood variance in random coefficients, as well as the magnitude, direction and statistical significance of the parameter estimates for individual and neighbourhood effects. Since multilevel analysis may be unfamiliar to the reader, and since the terminology and presentation used here may not be identical to that familiar to users of other software packages, a brief explanation of the tables seems warranted.

The multilevel regression tables are simply an adaptation of summary tables of classical regression analyses. The candidate variables are listed vertically at the left, and each column (these are titled with model numbers) provides the parameter estimates for a set of these variables, each adjusted for the others. After scanning each column (i.e. model) to see which variables have been included, the parameter estimates can then be read horizontally, such that the changes represent the effects of controlling for a slightly different set of confounders and interactions.

At the bottom of each table is a section that may be unfamiliar to some readers titled 'random effects.' Here the variance component attributable to each random coefficient (i.e. residual neighbourhood variation in the specified level 1 coefficient) is quantified and tested for significance on a $\chi^2$ distribution. When there are no level 2 predictors, this is a test for neighbourhood variation in the coefficient, adjusted for whichever level 1 (individual level) descriptors are included in that model. Each time new explanatory variables are added to the model, the significance of the variance
<table>
<thead>
<tr>
<th>Sex</th>
<th>Age</th>
<th>Urban</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Primary</td>
<td>Secondary</td>
</tr>
<tr>
<td>Male</td>
<td>25-29</td>
<td>38</td>
<td>273</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>40</td>
<td>345</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>40</td>
<td>310</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>41</td>
<td>320</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>113</td>
<td>307</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>100</td>
<td>279</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>218</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>248</td>
<td>258</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>237</td>
<td>219</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>165</td>
<td>193</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>149</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>80+</td>
<td>111</td>
<td>56</td>
</tr>
<tr>
<td>All</td>
<td>1631</td>
<td>2880</td>
<td>2920</td>
</tr>
</tbody>
</table>

Table 4.2: Frequency counts for both sexes by educational achievement and age in urban and rural areas.
components in the random effects output allows one to see if any neighbourhood variation remains to be explained in a model.

It is expected that the term 'level 2 intercept' may be confusing. Recall equations 3.2 and 3.3 on page 60:

\[ \beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j} \]
\[ \beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \]

\( \gamma_{00} \) and \( \gamma_{10} \) are the intercepts of these level 2 regressions of level 1 coefficients on the level 2 predictor variables. In a model with no level 2 predictors, the level 2 intercept represents the average level 1 coefficient, adjusted only for random variation. In a model with level 2 predictors, the level 2 intercept refers to the average level 1 coefficient for the reference neighbourhood type. In this study the reference neighbourhood type has been defined as an urban neighbourhood with an average level of education (average for all neighbourhoods, not just urban ones.) The meaning of level 2 parameter estimates is presented diagramatically in figure 4.1.

The presentation of the results will take the form of a sequential appraisal of the hypotheses presented in section 2.3.

4.2.2 Hypothesis tests

**Hypothesis 1:**

It was hypothesized that controlling for individual-level socio-demographic characteristics would reduce between-neighbourhood variation in the mean number of chronic health problems and the mean probability of activity limitation, but that both health constructs would exhibit significant residual area variation.

Although it was not possible in log-linearized models to calculate proportions of variance attributable to the neighbourhood level,\(^1\) this has been done for the logistic activity restriction variables, BINACT and HUI830, for models with only a random intercept.\(^2\) Results of these analyses are presented in table 4.3. These models allow the appraisal of the roles of the various individual and neighbourhood predictors in explaining the observed between-neighbourhood variation in the mean of the logit of the activity restriction measures.

The intra-class correlation coefficient (ICC, or \( \rho \)) represents the proportion of variance attributable to level 2. It is calculated by dividing the variance component for the intercept by the

---

\(^1\)The calculation, if it is possible, has not yet been developed.
\(^2\)This calculation cannot be applied to models with random slopes.
Random level 1 intercept:

$$\beta_{0j} = \gamma_0 + \gamma_{01}\text{Rural}_j + \gamma_{02}\text{EA-education}_j + \gamma_{03}\text{Rural}_j \times \text{EA-education}_j + u_{0j}$$

Random level 1 slope:

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\text{Rural}_j + \gamma_{12}\text{EA-education}_j + \gamma_{13}\text{Rural}_j \times \text{EA-education}_j + u_{1j}$$

Figure 4.1: These two schematics represent the study hypotheses and display what the level 2 parameters in tables 4.4, 4.5 and 4.6 are actually describing. There are two kinds of random coefficients in this study: level 1 intercepts and level 1 slopes. This diagram uses the logistic model of HUI830 as an example. The level 1 intercept ($\beta_{0j}$) for logit(HUI830) could be referred to as the adjusted neighbourhood mean, where 'adjusted' refers to the fact that the effects of individual-level socio-demographic characteristics have been controlled for statistically. Likewise, the level 1 slope ($\beta_{1j}$) for the effect of primary education on logit(HUI830) could be referred to as the adjusted neighbourhood slope for the primary education effect. Both the intercept, or adjusted neighbourhood mean, and the effect of primary education are expected to increase in magnitude with decreasing EA-education level. A rural advantage has been added to the hypotheses in order to separate the points where the rural and urban estimates cross the y axis so that the level 2 intercept of both random coefficients can be clearly seen to refer to the reference group, in each case urban EAs. The important thing to understand from this diagram about multilevel models is that the level 2 parameters do not describe $f(Y_{ij})$ directly, but instead describe one of the individual level coefficients on which $f(Y_{ij})$ is being regressed.

sum of this variance component and the level 1 variance component, which in logistic models is assumed to be $\pi^2/3$. In table 4.3, one can easily see that BINACT is a more clustered indicator of activity restriction than its counterpart, HUI830. That is to say, there are greater between-neighbourhood differences in BINACT prevalence than in HUI830 prevalence (with and without adjusting for individual and neighbourhood characteristics). The unadjusted model for BINACT
Table 4.3: Intra-class Correlation Coefficients (ICC) for logistic models of activity restriction with random intercepts. The phrase ‘EA-education level by urbanism,’ which appears below, is meant to succinctly refer to EA-education, rural status, and the interaction term between these two. Together these variables allow the effect of EA-education to be assessed separately ‘by urbanism,’ that is, for each stratum: urban and rural.

<table>
<thead>
<tr>
<th>Model</th>
<th>Controlling for:</th>
<th>% Variance at area level</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unadjusted</td>
<td>BINACT 10.2   HUI830 5.9</td>
</tr>
<tr>
<td>1</td>
<td>EA-education by urbanism</td>
<td>9.3 5.1</td>
</tr>
<tr>
<td>2</td>
<td>Age and sex</td>
<td>9.6 5.5</td>
</tr>
<tr>
<td>3</td>
<td>2 + chronic condition count</td>
<td>9.2 5.0</td>
</tr>
<tr>
<td>4</td>
<td>3 + education</td>
<td>8.9 4.4</td>
</tr>
<tr>
<td>5</td>
<td>4 + income</td>
<td>8.8 4.3</td>
</tr>
<tr>
<td>6</td>
<td>4 + EA-education by urbanism</td>
<td>8.7 4.2</td>
</tr>
<tr>
<td>7</td>
<td>5 + EA-education by urbanism</td>
<td>8.6 4.2</td>
</tr>
</tbody>
</table>

This model, referred to by Bryk and Raudenbush (1992) as a ‘one-way anova with random effects,’ tests whether there is any clustering of health at all. Clearly the majority of variation in both measures remains at the individual level. In models 2 through 5 the roles of the individual level predictors can be seen cumulatively. All of these variance components are significantly non-zero (p=0.000 by $\chi^2$ test), supporting hypothesis 1.

In tables 4.4, 4.5 and 4.6, simultaneous tests for variation in the intercept and in the socio-economic effects on the number of chronic health problems, BINACT and HUI830 are carried out in models 1. Note that the reference group in all analyses is men aged 48 with a university education and a household income of $40-50,000. At the bottom of each table, in the section labelled ‘Random Effects,’ the variance component for the intercept, and a $\chi^2$ test of $h_0$: variance component = 0 is reported. In models with and without household income, there is statistically significant between-neighbourhood variation in the log of the chronic health problem count after controlling for age, sex and education (p for variance component = 0.022 without income; 0.004 with income). The same is true for BINACT and HUI830 after controlling for age, sex, number of chronic health problems and education. These findings also support hypothesis 1.

$logit(BINACT_{ij}) = \beta_{0j} + r_{ij}$

$\beta_{0j} = \gamma_{00} + u_{0j}$
Hypothesis 2:
It was expected that the effect of the individual's education and income on both health constructs would vary by neighbourhood.

In tables 4.4, 4.5 and 4.6, tests for between-neighbourhood variation in individual socio-economic effects on the log of the number of chronic health problems, and on the logit of BINACT and of HUI830 are carried out in models 1. In the section labelled ‘Random Effects,’ the variance component for the slope associated with each socio-economic variable, and a \( \chi^2 \) test of \( h_0: \) variance component = 0 is reported. In models without income, there is marginally significant between-neighbourhood variation in the effect of 'some secondary' education on the log of the chronic health problem count after controlling for age and sex (\( p=0.077 \)). In models with income, there is marginally significant between-neighbourhood variation in the effect of each education category on the log of the chronic health problem count after controlling for age and sex (\( p=0.093, \) 0.040 some secondary; 0.080 some secondary; 0.028 some post-secondary; 0.074 college). The same is not true for BINACT, where no neighbourhood variation in the education effect was observed.

For HUI830, in the absence of income, and after controlling for age, sex and number of chronic health problems, the effect of some secondary education showed marginally significant between-neighbourhood variation (\( p=0.078 \)). The addition of income made this variation fully significant (\( p=0.001 \)), and resulted in the effect of primary education also appearing to vary between neighbourhoods (\( p=0.021 \)). The effects of income on all dependent variables displays more significant between-neighbourhood variation than do education variables. These findings support hypothesis 2 for chronic health problems, but present mixed evidence regarding activity restriction.

Hypothesis 3:
Lower levels of neighbourhood education were expected to predict a greater mean number of chronic health problems and a higher prevalence of activity limitation.

and Hypothesis 5:
All neighbourhood effects were expected to be greater in urban than in rural areas.

Model 1 in table 4.3 (page 69) is adjusted for the neighbourhood predictors in the absence of individual level controls. If some of the unadjusted model's neighbourhood variance were explained by this model, but subsequently accounted for by individual level controls, then most of the apparent effects of urbanism and neighbourhood education level (EA-education) are caused or mediated by individual characteristics. When these neighbourhood characteristics are re-introduced to models with individual level controls (models 6 and 7 in table 4.3), they explain less variance than they did initially, suggesting a causal explanation of unknown sequence and complexity.
In tables 4.4, 4.5 and 4.6, tests for the association of neighbourhood education level (EA-education) with the number of chronic health problems, BINACT and HUI830 in urban and rural EAs are carried out in models 2. In the section labelled 'Fixed Effects,' these parameter estimates are reported with their standard errors and a p-value for a t-ratio test of significance. These show that there is a non-significant risk of increased chronic health problems with decreasing neighbourhood education level in urban areas only (p=0.135 without income, see figure 4.2). For activity limitation, there is a marginally significant risk associated with decreasing neighbourhood education level only in the absence of income (p= 0.087 BINACT; 0.046 HUI830). In rural areas the role of neighbourhood education is less consistent (see figures 4.2 and 4.3, and tables 4.4, 4.5 and 4.6).

These findings do not support hypothesis 3 or 5, but do demonstrate small and non-significant effects in the expected directions for each hypothesis.

Hypothesis 4:
The effects of neighbourhood education level was expected to be greater among persons with lower levels of individual education.

and Hypothesis 5:
All neighbourhood effects were expected to be greater in urban than in rural areas.

In tables 4.4, 4.5 and 4.6, estimates of the association of EA-education in urban and rural areas with between-neighbourhood variation in socio-economic effects on the number of chronic health problems, BINACT and HUI830 are carried out in models 2. In the section labelled 'Fixed Effects,' these parameter estimates are reported with their standard errors and a p-value for a t-ratio test of significance. In general, the contextual effect of neighbourhood education did not vary by individual educational achievement in urban areas for chronic health problems or for either measure of activity limitation. In rural areas the pattern is less consistent. These slopes can be inspected visually in figures 4.2 and 4.3. In no models did the area characteristics explain any of the between-neighbourhood variation in the effect of individual income.

These findings do not support hypothesis 4 or 5.

Model 2 (education only) results for chronic health problems are presented graphically in figure 4.2, and those for activity limitation (BINACT and HUI830) are presented in figure 4.3.
Figure 4.2: This graph shows the effect of area education level (EA-education) on the adjusted
eighbourhood mean number of chronic health problems, stratified by individual educational
achievement and by urban (U) v. rural (R) neighbourhood status. These figures are adjusted
for age and sex, the reference group being men aged 48. EA-education below the mean indicates an
increasing proportion of university-educated residents relative to the number with primary education only, and values greater than the mean indicate the reverse trend. All education groups except 'secondary' suffered significantly more chronic health problems than did the university educated. Rural area residence predicts improved health (p=0.000). The effect of EA-education was not significant for any group in urban areas, but relative to the role of EA-education in rural areas for the university educated, there was a significantly opposite slope for the college-educated (p=0.004), and marginally significantly flatter slope for those with primary (p=0.045), some secondary (p=0.057), and some post-secondary education (p=0.093). The contextual effect of EA-education did not vary by individual educational achievement in urban areas.

4.2.3 Other observations

The most pronounced neighbourhood effect in table 4.4 and figure 4.2 is that of rural status. Rural
status is associated with fewer chronic health problems. This effect is sizable, significant, and
independent of household income.
Figure 4.3: This graph shows the effect of area education level (EA-education) on the adjusted neighbourhood mean probability of activity limitation for each of the two activity limitation measures, BINACT and HUI830. The results are stratified by individual educational achievement and by urban (U) v. rural (R) neighbourhood status. These figures are adjusted for age and sex, the reference group being men aged 48. EA-education below the mean indicates an increasing proportion of university-educated residents relative to the number with primary education only, and values greater than the mean indicate the reverse trend. Note that despite the different adjusted prevalence of the two activity limitation variables, the roles of individual and neighbourhood education and of urbanism for each are extremely consistent. Compared to the university educated, all other groups are significantly more likely to experience activity limitation, but for the college-educated this effect is marginal (p=0.084) and is reduced to insignificance in the model controlling for household income. The slope for the EA-education effect is marginally significant (BINACT: p=0.087; HUI830: p=0.046), but is insignificant in models adjusting for household income. The test for a different effect of EA-education in rural areas was marginally significant for BINACT only (BINACT: p=0.048; HUI830: p=0.888). In the main, contextual effects affected all education groups equally.
Table 4.4: Neighbourhood effects on the log of the chronic health problems count, controlling for the socio-demographic characteristics of individuals. Fixed effects are reported as \("\text{effect (standard error ; p-value for t-ratio)}\) and random effects as \("\text{variance component (chi-square ; df ; p-value for chi-square)}\)."

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Level 1</th>
<th>Education only</th>
<th>Education and Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 2</td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Fixed Effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.165 (0.0200;0.000)</td>
<td>-0.119 (0.0251;0.000)</td>
<td>-0.135 (0.0213;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.217 (0.0652;0.000)</td>
<td>.</td>
<td>-0.225 (0.0657;0.000)</td>
</tr>
<tr>
<td>EA-education</td>
<td>0.019 (0.0128;0.135)</td>
<td>.</td>
<td>0.007 (0.0139;0.598)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.051 (0.0325;0.115)</td>
<td>.</td>
<td>0.062 (0.0344;0.070)</td>
</tr>
<tr>
<td>Primary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.206 (0.0236;0.000)</td>
<td>0.187 (0.0295;0.000)</td>
<td>0.119 (0.0240;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.104 (0.0642;0.105)</td>
<td>.</td>
<td>0.127 (0.0650;0.053)</td>
</tr>
<tr>
<td>EA-education</td>
<td>0.005 (0.0165;0.759)</td>
<td>.</td>
<td>-0.008 (0.0172;0.638)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.074 (0.0370;0.045)</td>
<td>.</td>
<td>-0.072 (0.0380;0.057)</td>
</tr>
<tr>
<td>Some Secondary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.180 (0.0222;0.000)</td>
<td>0.165 (0.0278;0.000)</td>
<td>0.114 (0.0227;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.065 (0.0624;0.301)</td>
<td>.</td>
<td>0.084 (0.0638;0.187)</td>
</tr>
<tr>
<td>EA-education</td>
<td>0.022 (0.0157;0.161)</td>
<td>.</td>
<td>0.014 (0.0161;0.378)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.070 (0.0390;0.057)</td>
<td>.</td>
<td>-0.069 (0.0373;0.062)</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.041 (0.0225;0.072)</td>
<td>0.021 (0.0282;0.457)</td>
<td>0.0041 (0.0226;0.986)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.107 (0.0639;0.094)</td>
<td>.</td>
<td>0.115 (0.0638;0.072)</td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.001 (0.0159;0.933)</td>
<td>.</td>
<td>-0.008 (0.0169;0.623)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.054 (0.0373;0.151)</td>
<td>.</td>
<td>-0.052 (0.0372;0.164)</td>
</tr>
<tr>
<td>Some Post-Secondary</td>
<td>0.145 (0.0259;0.000)</td>
<td>0.152 (0.0318;0.000)</td>
<td>0.107 (0.0259;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.001 (0.0724;0.990)</td>
<td>.</td>
<td>0.026 (0.0721;0.721)</td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.004 (0.0180;0.832)</td>
<td>.</td>
<td>-0.011 (0.0180;0.524)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.071 (0.0425;0.093)</td>
<td>.</td>
<td>-0.073 (0.0423;0.084)</td>
</tr>
<tr>
<td>College</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.088 (0.0256;0.001)</td>
<td>0.087 (0.0317;0.006)</td>
<td>0.060 (0.0256;0.020)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.097 (0.0695;0.163)</td>
<td>.</td>
<td>0.105 (0.0693;0.128)</td>
</tr>
<tr>
<td>EA-education</td>
<td>0.021 (0.0185;0.256)</td>
<td>.</td>
<td>0.017 (0.0185;0.362)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.123 (0.0419;0.004)</td>
<td>.</td>
<td>-0.124 (0.0418;0.003)</td>
</tr>
<tr>
<td>University: Reference category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.</td>
<td>-0.046 (0.0030;0.000)</td>
<td>-0.047 (0.0036;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>.</td>
<td>.</td>
<td>0.006 (0.0074;0.415)</td>
</tr>
<tr>
<td>EA-education</td>
<td>.</td>
<td>.</td>
<td>-0.003 (0.0021;0.203)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>.</td>
<td>.</td>
<td>0.001 (0.0041;0.860)</td>
</tr>
<tr>
<td>Income Squared</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.</td>
<td>0.006 (0.0011;0.000)</td>
<td>0.006 (0.0013;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>.</td>
<td>.</td>
<td>-0.000 (0.0028;0.888)</td>
</tr>
<tr>
<td>EA-education</td>
<td>.</td>
<td>.</td>
<td>0.001 (0.0089;0.362)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>.</td>
<td>.</td>
<td>-0.001 (0.0016;0.689)</td>
</tr>
</tbody>
</table>

continued on next page
<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Age</th>
<th>Age squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.166 (0.0113;0.000)</td>
<td>0.166 (0.0113;0.000)</td>
<td>0.147 (0.0112;0.000)</td>
</tr>
<tr>
<td></td>
<td>0.026 (0.0005;0.000)</td>
<td>0.026 (0.0005;0.000)</td>
<td>0.024 (0.0005;0.000)</td>
</tr>
<tr>
<td></td>
<td>-0.0002 (2e-5;0.000)</td>
<td>-0.0002 (2e-5;0.000)</td>
<td>-0.0002 (2e-5;0.000)</td>
</tr>
<tr>
<td>Random Effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.1759 (839,759;0.022)</td>
<td>0.0280 (822,756;0.047)</td>
<td>0.0406 (861,752;0.004)</td>
</tr>
<tr>
<td>Primary</td>
<td>0.1368 (778,759;0.307)</td>
<td>0.0207 (770,756;0.352)</td>
<td>0.0193 (804,752;0.093)</td>
</tr>
<tr>
<td>Some Secondary</td>
<td>0.2041 (815,759;0.077)</td>
<td>0.0420 (809,756;0.090)</td>
<td>0.0568 (821,752;0.040)</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.2101 (809,759;0.103)</td>
<td>0.0401 (800,756;0.131)</td>
<td>0.0376 (807,752;0.080)</td>
</tr>
<tr>
<td>Some Post-Secondary</td>
<td>0.2230 (805,759;0.119)</td>
<td>0.0524 (797,756;0.145)</td>
<td>0.0474 (828,752;0.028)</td>
</tr>
<tr>
<td>College</td>
<td>0.1724 (796,759;0.173)</td>
<td>0.0288 (789,756;0.193)</td>
<td>0.0383 (800,752;0.074)</td>
</tr>
<tr>
<td>University: Reference category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Income Squared</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>LEVEL 1</td>
<td>1.2521</td>
<td>1.2533</td>
<td>1.1975</td>
</tr>
</tbody>
</table>

The $\chi^2$ statistics reported above are based on a subset of the 1985 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.
Table 4.5: Neighbourhood effects on the logit of self-reported activity limitation (BINACT), controlling for the socio-demographic characteristics of individuals. Fixed effects are reported as “effect (standard error ; p-value for t-ratio)” and random effects as “variance component (chi-square ; df ; p-value for chi-square).”

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Education only</th>
<th>Education and Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td><strong>Fixed Effects:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-4.286</td>
<td>-4.155</td>
<td>-4.669</td>
<td>-4.539</td>
</tr>
<tr>
<td></td>
<td>(0.0827;0.000)</td>
<td>(0.0967;0.000)</td>
<td>(0.0981;0.000)</td>
<td>(0.1143;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.362</td>
<td>-0.411</td>
<td>-0.461</td>
<td>-0.461</td>
</tr>
<tr>
<td></td>
<td>(0.2183;0.096)</td>
<td>(0.2621;0.110)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA-education</td>
<td>0.088</td>
<td>0.062</td>
<td>0.088</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>(0.0517;0.087)</td>
<td>(0.0637;0.331)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.268</td>
<td>-0.324</td>
<td>-0.268</td>
<td>-0.324</td>
</tr>
<tr>
<td></td>
<td>(0.1358;0.048)</td>
<td>(0.1651;0.049)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.085</td>
<td>1.019</td>
<td>1.019</td>
<td>1.019</td>
</tr>
<tr>
<td></td>
<td>(0.0903;0.000)</td>
<td>(0.1064;0.000)</td>
<td>(0.1064;0.000)</td>
<td>(0.1266;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.359</td>
<td>0.376</td>
<td>0.359</td>
<td>0.376</td>
</tr>
<tr>
<td></td>
<td>(0.2480;0.148)</td>
<td>(0.2934;0.200)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.048</td>
<td>-0.014</td>
<td>-0.048</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.0631;0.445)</td>
<td>(0.0760;0.856)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.164</td>
<td>0.164</td>
<td>0.164</td>
<td>0.164</td>
</tr>
<tr>
<td></td>
<td>(0.1506;0.277)</td>
<td>(0.1803;0.363)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Some Secondary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.852</td>
<td>0.863</td>
<td>0.863</td>
<td>0.863</td>
</tr>
<tr>
<td></td>
<td>(0.0849;0.000)</td>
<td>(0.1008;0.000)</td>
<td>(0.1008;0.000)</td>
<td>(0.1168;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.285</td>
<td>0.289</td>
<td>0.285</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>(0.2378;0.231)</td>
<td>(0.2758;0.295)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.075</td>
<td>-0.051</td>
<td>-0.075</td>
<td>-0.051</td>
</tr>
<tr>
<td></td>
<td>(0.0589;0.206)</td>
<td>(0.0694;0.460)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.216</td>
<td>0.241</td>
<td>0.216</td>
<td>0.241</td>
</tr>
<tr>
<td></td>
<td>(0.1402;0.140)</td>
<td>(0.1722;0.162)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.384</td>
<td>0.365</td>
<td>0.365</td>
<td>0.365</td>
</tr>
<tr>
<td></td>
<td>(0.0874;0.000)</td>
<td>(0.1006;0.001)</td>
<td>(0.1006;0.001)</td>
<td>(0.1206;0.008)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.140</td>
<td>0.124</td>
<td>0.140</td>
<td>0.124</td>
</tr>
<tr>
<td></td>
<td>(0.2455;0.551)</td>
<td>(0.2818;0.658)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.054</td>
<td>-0.050</td>
<td>-0.054</td>
<td>-0.050</td>
</tr>
<tr>
<td></td>
<td>(0.0603;0.373)</td>
<td>(0.0701;0.474)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.207</td>
<td>0.266</td>
<td>0.207</td>
<td>0.266</td>
</tr>
<tr>
<td></td>
<td>(0.1406;0.167)</td>
<td>(0.1747;0.128)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Some Post-Secondary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.395</td>
<td>0.352</td>
<td>0.352</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>(0.0987;0.000)</td>
<td>(0.1132;0.002)</td>
<td>(0.1132;0.002)</td>
<td>(0.1335;0.061)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.144</td>
<td>0.198</td>
<td>0.144</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>(0.2752;0.600)</td>
<td>(0.3150;0.530)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.102</td>
<td>-0.112</td>
<td>-0.102</td>
<td>-0.112</td>
</tr>
<tr>
<td></td>
<td>(0.0677;0.133)</td>
<td>(0.0780;0.155)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.357</td>
<td>0.435</td>
<td>0.357</td>
<td>0.435</td>
</tr>
<tr>
<td></td>
<td>(0.1607;0.031)</td>
<td>(0.1934;0.024)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>College</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.461</td>
<td>0.457</td>
<td>0.457</td>
<td>0.457</td>
</tr>
<tr>
<td></td>
<td>(0.0952;0.000)</td>
<td>(0.1096;0.000)</td>
<td>(0.1096;0.000)</td>
<td>(0.1297;0.002)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.081</td>
<td>0.098</td>
<td>0.081</td>
<td>0.098</td>
</tr>
<tr>
<td></td>
<td>(0.2619;0.756)</td>
<td>(0.2998;0.743)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.022</td>
<td>-0.009</td>
<td>-0.022</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>(0.0678;0.745)</td>
<td>(0.0783;0.904)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.233</td>
<td>0.285</td>
<td>0.233</td>
<td>0.285</td>
</tr>
<tr>
<td></td>
<td>(0.1622;0.151)</td>
<td>(0.1886;0.130)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>University: Reference category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>-0.114</td>
<td>-0.114</td>
<td>-0.114</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0121;0.000)</td>
<td>(0.0141;0.000)</td>
<td>(0.0141;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td></td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0302;0.871)</td>
<td>(0.0302;0.871)</td>
<td>(0.0302;0.871)</td>
</tr>
<tr>
<td>EA-education</td>
<td></td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0086;0.433)</td>
<td>(0.0086;0.433)</td>
<td>(0.0086;0.433)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td></td>
<td>-0.004</td>
<td>-0.004</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0170;0.800)</td>
<td>(0.0170;0.800)</td>
<td>(0.0170;0.800)</td>
</tr>
<tr>
<td>Income Squared</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0048;0.320)</td>
<td>(0.0055;0.480)</td>
<td>(0.0055;0.480)</td>
</tr>
<tr>
<td>Rural</td>
<td></td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0122;0.704)</td>
<td>(0.0122;0.704)</td>
<td>(0.0122;0.704)</td>
</tr>
<tr>
<td>EA-education</td>
<td></td>
<td>-0.000</td>
<td>-0.000</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0033;0.918)</td>
<td>(0.0033;0.918)</td>
<td>(0.0033;0.918)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td></td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0068;0.763)</td>
<td>(0.0068;0.763)</td>
<td>(0.0068;0.763)</td>
</tr>
</tbody>
</table>

*continued on next page*
<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Age</th>
<th>Age squared</th>
<th>Age cubed</th>
<th>CC8</th>
<th>CC8 squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.207 (0.0298;0.000)</td>
<td>-0.208 (0.0299;0.000)</td>
<td>-0.275 (0.0296;0.000)</td>
<td>-0.276 (0.0296;0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.010 (0.0018;0.000)</td>
<td>-0.011 (0.0018;0.000)</td>
<td>-0.017 (0.0019;0.000)</td>
<td>-0.017 (0.0019;0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age squared</td>
<td>-0.001 (0.0001;0.000)</td>
<td>-0.001 (0.0001;0.000)</td>
<td>-0.001 (0.0001;0.000)</td>
<td>-0.001 (0.0001;0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age cubed</td>
<td>$4 \times 10^{-5}$ (3e$^{-6}$;0.000)</td>
<td>$4 \times 10^{-5}$ (3e$^{-6}$;0.000)</td>
<td>$5 \times 10^{-5}$ (3e$^{-6}$;0.000)</td>
<td>$5 \times 10^{-5}$ (3e$^{-6}$;0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC8</td>
<td>1.217 (0.0261;0.000)</td>
<td>1.217 (0.0261;0.000)</td>
<td>1.335 (0.0265;0.000)</td>
<td>1.335 (0.0266;0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC8 squared</td>
<td>-0.084 (0.0042;0.000)</td>
<td>-0.084 (0.0042;0.000)</td>
<td>-0.093 (0.0044;0.000)</td>
<td>-0.093 (0.0044;0.000)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Random Effects:

<table>
<thead>
<tr>
<th></th>
<th>Intercept 2.43 (959;756;&gt;0.5)</th>
<th>Primary 3.53 (709;756;&gt;0.5)</th>
<th>Some Secondary 3.34 (755;756;&gt;0.5)</th>
<th>Secondary 3.30 (592;756;&gt;0.5)</th>
<th>Some Post-Secondary 4.01 (503;756;&gt;0.5)</th>
<th>College 3.42 (519;756;&gt;0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income: Reference category</td>
<td>Incoome 0.097 (140;729;0.000)</td>
<td>Income Squared 0.017 (1439;729;0.000)</td>
<td>LEVEL 1 0.526</td>
<td>LEVEL 1 0.527</td>
<td>LEVEL 1 0.406</td>
<td>LEVEL 1 0.406</td>
</tr>
</tbody>
</table>

The $\chi^2$ statistics reported above are based on a subset of the 1925 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.
Table 4.6: Neighbourhood effects on the logit of the HUI, controlling for the socio-demographic characteristics of individuals. Fixed effects are reported as “effect (standard error ; p-value for t-ratio)” and random effects as “variance component (chi-square ; df ; p-value for chi-square).”

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Education only</th>
<th>Education and Income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Fixed Effects:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-3.041 (0.0636;0.000)</td>
<td>-2.952 (0.0751;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.203 (0.1621;0.212)</td>
<td>-0.052 (0.0476;0.275)</td>
</tr>
<tr>
<td>EA-education</td>
<td>0.077 (0.0384;0.046)</td>
<td>0.013 (0.0910;0.888)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.013 (0.0910;0.888)</td>
<td>0.013 (0.0910;0.888)</td>
</tr>
<tr>
<td>Primary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.983 (0.0687;0.000)</td>
<td>0.940 (0.0834;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.108 (0.1856;0.560)</td>
<td>0.052 (0.0476;0.275)</td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.052 (0.0476;0.275)</td>
<td>-0.052 (0.0476;0.275)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.095 (0.1033;0.350)</td>
<td>-0.095 (0.1033;0.350)</td>
</tr>
<tr>
<td>Some Secondary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.676 (0.0643;0.000)</td>
<td>0.607 (0.0780;0.000)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.127 (0.1780;0.477)</td>
<td>0.044 (0.0447;0.322)</td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.044 (0.0447;0.322)</td>
<td>-0.044 (0.0447;0.322)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>-0.030 (0.1002;0.765)</td>
<td>-0.030 (0.1002;0.765)</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.339 (0.0643;0.000)</td>
<td>0.274 (0.0781;0.001)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.056 (0.1797;0.755)</td>
<td>0.056 (0.1797;0.755)</td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.073 (0.0445;0.100)</td>
<td>-0.073 (0.0445;0.100)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.072 (0.1006;0.471)</td>
<td>0.072 (0.1006;0.471)</td>
</tr>
<tr>
<td>Some Post-Secondary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.363 (0.0734;0.000)</td>
<td>0.287 (0.0882;0.002)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.172 (0.2004;0.395)</td>
<td>0.172 (0.2004;0.395)</td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.050 (0.0504;0.323)</td>
<td>-0.050 (0.0504;0.323)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.013 (0.1123;0.911)</td>
<td>0.013 (0.1123;0.911)</td>
</tr>
<tr>
<td>College</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.161 (0.0774;0.038)</td>
<td>0.160 (0.0929;0.084)</td>
</tr>
<tr>
<td>Rural</td>
<td>-0.164 (0.2116;0.438)</td>
<td>-0.164 (0.2116;0.438)</td>
</tr>
<tr>
<td>EA-education</td>
<td>-0.037 (0.0546;0.494)</td>
<td>-0.037 (0.0546;0.494)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.050 (0.1206;0.678)</td>
<td>0.050 (0.1206;0.678)</td>
</tr>
<tr>
<td>University: Reference category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.000 (0.003;0.956)</td>
<td>0.000 (0.003;0.956)</td>
</tr>
<tr>
<td>Rural</td>
<td>0.000 (0.003;0.956)</td>
<td>0.000 (0.003;0.956)</td>
</tr>
<tr>
<td>EA-education</td>
<td>0.000 (0.003;0.956)</td>
<td>0.000 (0.003;0.956)</td>
</tr>
<tr>
<td>Rural*EA-education</td>
<td>0.000 (0.003;0.956)</td>
<td>0.000 (0.003;0.956)</td>
</tr>
</tbody>
</table>

continued on next page
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-0.098 (0.0278;0.001)</td>
<td>-0.100 (0.0279;0.001)</td>
<td>-0.121 (0.0268;0.000)</td>
<td>-0.124 (0.0269;0.000)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.002 (0.0017;0.246)</td>
<td>0.002 (0.0017;0.237)</td>
<td>-0.001 (0.0017;0.786)</td>
<td>-0.000 (0.0017;0.772)</td>
<td></td>
</tr>
<tr>
<td>Age squared</td>
<td>-0.001 (9.1e-5;0.000)</td>
<td>-0.001 (9.1e-5;0.000)</td>
<td>-0.001 (9.1e-5;0.000)</td>
<td>-0.001 (9.1e-5;0.000)</td>
<td></td>
</tr>
<tr>
<td>Age cubed</td>
<td>3.4e-5 (3e-6;0.000)</td>
<td>3.5e-5 (3e-6;0.000)</td>
<td>3.9e-5 (3e-6;0.000)</td>
<td>3.9e-5 (3e-6;0.000)</td>
<td></td>
</tr>
<tr>
<td>CC8</td>
<td>0.871 (0.0240;0.000)</td>
<td>0.870 (0.0240;0.000)</td>
<td>0.895 (0.0234;0.000)</td>
<td>0.894 (0.0234;0.000)</td>
<td></td>
</tr>
<tr>
<td>CC8 squared</td>
<td>-0.047 (0.0043;0.000)</td>
<td>-0.047 (0.0043;0.000)</td>
<td>-0.048 (0.0043;0.000)</td>
<td>-0.048 (0.0043;0.000)</td>
<td></td>
</tr>
</tbody>
</table>

**Random Effects:**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.98 (867;750;0.004)</td>
<td>0.96 (851;756;0.009)</td>
<td>1.65 (943;752;0.000)</td>
<td>1.63 (937;749;0.000)</td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>1.32 (798;759;0.158)</td>
<td>1.31 (785;756;0.222)</td>
<td>2.06 (836;752;0.018)</td>
<td>2.04 (830;749;0.021)</td>
<td></td>
</tr>
<tr>
<td>Some Secondary</td>
<td>1.23 (820;759;0.060)</td>
<td>1.23 (812;756;0.078)</td>
<td>1.96 (885;752;0.001)</td>
<td>1.95 (881;749;0.001)</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>0.96 (713;759;&gt;0.5)</td>
<td>0.95 (705;756;&gt;0.5)</td>
<td>1.73 (738;752;&gt;0.5)</td>
<td>1.70 (735;749;&gt;0.5)</td>
<td></td>
</tr>
<tr>
<td>Some Post-Secondary</td>
<td>1.26 (641;759;&gt;0.5)</td>
<td>1.24 (637;756;&gt;0.5)</td>
<td>2.38 (648;752;&gt;0.5)</td>
<td>2.36 (647;749;&gt;0.5)</td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>2.08 (679;759;&gt;0.5)</td>
<td>2.08 (669;756;&gt;0.5)</td>
<td>2.84 (665;752;&gt;0.5)</td>
<td>2.81 (663;749;&gt;0.5)</td>
<td></td>
</tr>
</tbody>
</table>

**University:** Reference category

- Income
  - .
  - .
  - 0.03 (1194;752;0.000)  | 0.03 (1189;749;0.000)
- Income Squared
  - .
  - .
  - 0.01 (1214;762;0.000)  | 0.01 (1208;749;0.000)
- LEVEL
  - 0.772
  - 0.773
  - 0.638
  - 0.639

*The χ² statistics reported above are based on a subset of the 1925 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.*
Chapter 5

Discussion

5.1 Conclusions

In section 2.4 three possible conclusions to this study were presented. The first conclusion was that between-neighbourhood differences in health were entirely due to compositional differences in age and sex, and this demographic explanation has been shown not to be the case. The second conclusion was that socio-economic status explains all of the observed non-demographic between-neighbourhood differences, and while this extreme case is not true, the socio-economic composition of neighbourhoods does explain some of the observed between-neighbourhood health differences. The sociological implications of this are discussed below. The third possible conclusion was that there were significant between-neighbourhood differences in health that were not explained by demographic and socio-economic composition, and that we would have a mandate to discover the individual or contextual effects responsible for this pattern. Small amounts of variance in health were observed between neighbourhoods after controlling for individual socio-economic and demographic factors. The effect size and significance of the effect of rural status on chronic health problems demonstrates that even this small amount of variance can represent important geographic inequalities in health (see figure 5.1 for a graph of this effect in isolation). However, the effects of EA-education are generally small and insignificant, and appear to be much less important than the individual-level educational differences in chronic health problems and activity restriction. Very little evidence is found that neighbourhood education level differentially affects individuals according to their individual educational achievements.
Figure 5.1: This graph is a simplified version of figure 4.2 for EAs of average education level. In this version it is easier to visually compare the slope of the individual-level education gradient in rural and urban areas (p=0.105 primary; 0.301 some secondary; 0.094 secondary; 0.990 some post-secondary; 0.163 college). In the left-hand panel the linear trend has been added to the joined point estimates. The effect sizes and significance for the effect of rural status on all educational groups relative to the main effect of rural status on the university group increases when household income is included in the model (see table 4.4). The main effect of rural status is highly significant (p=0.000).

5.2 Contribution to the literature

This study found somewhat more between-neighbourhood variation in health than did Boyle and Willms (1999), which is the only published Canadian work using similar dependent variables. Nonetheless, the amount of variance observed was relatively small, as has been found in all studies of the contextual effects of neighbourhood socio-economic level within rich countries. This is the first Canadian work to examine the differential effects of neighbourhood context on socio-economic groups within neighbourhoods. Unlike the American finding by Yen and Kaplan (1999a) that neighbourhood socio-economic context had large and opposite effects on the mortality of different groups according to their individual socio-economic status,¹ this study found that neighbourhood socio-economic context has little effect on health in urban areas and generally less in rural ones. This study also found that neighbourhood education level affects individuals of differing socio-economic

¹Note that Yen and Kaplan (1999a) do not provide enough evidence for the significance of this finding to be evaluated.
status equally in urban areas, and largely equally in rural areas. Given the much greater magnitude of the individual socio-economic effects, and likelihood of significant unmeasured socio-economic compositional differences between neighbourhoods, these small effects doubtfully represent important contextual determinants of chronic health problems and activity restriction.

It has long been observed that there is an apparent health advantage to rural life (Law and Morris 1998; Senior et al. 2000). This study provides a particularly clear demonstration of this advantage since the effects of both the individual and neighbourhood socio-economic characteristics have been controlled. No publication could be found of a comparable approach to examining the rural advantage. It is interesting to note that the salutary role of rural residence is much more marked for chronic health problems than for consequent activity restriction. This is not the pattern that health-related mobility towards well-serviced areas would predict, but it could represent under-reporting of non-activity-restricting chronic health problems in rural areas. However, rural mortality advantages, such as those observed by Law and Morris (1998) and Senior et al. (2000) in the UK, provide some evidence that a reporting bias is not a complete explanation. It should also be noted that mortality studies suggest a U-shaped risk related to urbanism, such that cities/metropolitan areas experience advantages similar to deep rural ones (Barnett et al. 2000; Senior et al. 2000). The current study was not designed to examine this phenomenon, and a dichotomous measure of urbanism does not allow for detailed speculation. This is an area that deserves further research.

5.3 Study limitations

The most serious limitations of this study regard its ability to measure individual and aggregate socio-economic situations. Education is relied upon at the individual level because it allows for tentative causal interpretation of cross-sectional associations. Education does not, however, fully describe the class and status situation of the individual. It should be noted that at the individual level the correlations between the most common indicators of socio-economic status is lower than is often assumed. The correlation between education and income is typically in the range of 0.2–0.4 in the general population in both the US (Winkleby et al. 1992) and in Canada (preliminary analyses for the current project using data from the 1990 OHS and the 1994 NPHS), depending on gender and the measures of income (household v. individual) and of education (milestones v. number of
years). The same sources show the correlations between income and occupation to be 0.3 – 0.5 and for that between education and occupation to be stronger at 0.5 – 0.7. These correlations are highly significant statistically. In Canada there does not seem to be any correlation between household ownership and either education or occupation, and only a modest correlation between household ownership and income (0.3). These correlations demonstrate that socio-economic status cannot be equivalently measured by any one of these survey items, regardless of the comparability of the gradients in health that the models may or may not reveal. This conclusion is supported by the general finding that multiple ‘markers of socio-economic status’ often have independent effects in predictive models (Liberatos et al. 1988; House et al. 1990; Badley and Ibanez 1994; House et al. 1994; Marmot 1996; Mustard et al. 1997; Wannamethee and Sharper 1997; Lynch and Kaplan 2000). The different behaviour of household income from individual education in the current study’s models highlights the likely complexity of the relationship between individual and contextual characteristics in the determination of the health and well-being of individuals.

At the neighbourhood level, a continuous measure of neighbourhood educational composition was created, but it incorporated only the proportions with primary and university education (for variable definitions see section 3.3). This measure may not fully capture the ‘neighbourhood socio-economic context.’ However, evidence from American factorial ecological work, using census information at the census tract level, shows that proportions of the adult population with post-secondary education load on factors or scales representing aggregate advantage and disadvantage. Kaplan (1996) found that residential mobility, owner-occupied housing, driving to work, employed males, median income, and levels of ‘some-college’ education formed one of four neighbourhood scales. In James (2000) a two factor scheme was presented for 1980 and 1990 census data in which levels of post-secondary education loaded quite negatively on the first factor (driven by the proportion of households headed by women, male long-term unemployment, the poverty rate, the welfare rate, the high-school drop-out rate, and the proportion in crowded housing) and quite positively on the second factor (driven by the proportion with post-secondary education, managerial/professional positions and the proportion of college students). It is interesting to note that in this study the rate of owner occupancy loaded negatively on both factors. In Manitoba a regional socio-economic index was developed by Frohlich and Mustard (1994, Mustard and Frohlich 1995). At the municipal level it was observed that the proportion of adults aged 25-34 who had completed high school correlated
significantly ($p \leq 0.001$) with four of the five other elements of the index: the unemployment rate among those aged 15-24 ($r=-0.450$) and those aged 45-54 ($r=-0.405$), the labour force participation rate among women aged 15+ ($r=0.531$), and the mean value of dwellings ($r=0.312$). The proportion of the population aged 35-44 and 45-54 with a high school diploma were not included in the index, but were significantly correlated to the index ($r=-0.568,-0.535$; both $p \leq 0.001$) at the municipal level. These studies indicate that the measure of neighbourhood education level used in the current study was likely highly correlated to many other dimensions of the neighbourhood socio-economic context.

The other source of error regarding the measurement of neighbourhood socio-economic status regards the definition of neighbourhoods as enumeration areas, which was done for reasons of data availability. This may have led to the misattribution of individuals relative to a more meaningful definition of neighbourhood, and to measurement error regarding the educational composition of neighbourhoods. However, as was discussed in section 1.4.3, sociologically meaningful neighbourhoods are likely larger than enumeration areas, and it has been the experience of those using census descriptions of small areas as proxy measures of individual socio-economic status that the size of the censal unit used makes little difference to the predictive potential. It seems unlikely that measurement error in the neighbourhood socio-economic characteristics would greatly exceed that in the individual socio-economic characteristics, and so the relative effect sizes of these are likely meaningful.

There is a need in Canada for more sophisticated and theorized measures of the socio-economic ranking of enumeration areas and census tracts and for the evaluation of the effects of boundary definitions, both for investigations of neighbourhood context (Boyle and Willms 1999) and also for use as a proxy measure of individual socio-economic status. This proxy use of neighbourhood socio-economic context has obvious empirical and theoretical faults, but at least in part overcomes the absence of socio-economic information that can easily be linked to most medical data (Mustard et al. 1999).

The dependent variables were also limited. Chronic health problems have not been explicitly defined, but merely defined by analogy. We cannot deduce why obesity, emaciation, poor muscle tone, missing limbs, hearing loss, multiple sclerosis, hepatitis, hepatic cirrhosis, thalassemia, haemophilia, cystic fibrosis, HIV seropositivity, sterility, impotence, dysmenorrhea, amenorrhea,
speech pathology or muscle twitches, to name a few examples at random, are not included in recent Canadian lists of this type. Why is hypertension usually included and hypercholesterolemia usually excluded? Some of these do not seem more rare, less severe, or more stigmatized than some of the other 'long-term, permanent or recurring health problems' that are typically assessed. Clearly rarity, severity and stigma were not essential criteria, and the OHS chronic health problems list had questionable content validity. It would be useful to construct a more comprehensive list of chronic health problems based on an explicit definition.

The two measures of activity limitation performed extremely similarly despite significant differences in prevalence. This observation demonstrates their relative construct validity, but also demonstrates that the dichotomous treatment of activity limitation is a simplification, as theory would suggest. The challenges of measuring activity limitation in general populations are well acknowledged and not easily resolved.

This study also has methodologic limitations. Despite the complexity of multilevel models and their ability to correctly estimate theoretically important interactions between individual and contextual characteristics, there remain limits to what single models can attempt to explain due to problems of algorithm convergence, computational intensiveness, power and interpretability. In this study no interactions between individual-level characteristics were specified for these reasons.

5.4 Questions arising

Variance in the determinant

It may be that Ontario neighbourhoods are sufficiently equitable that no effect can be demonstrated. This would not be a study limitation, but rather would be a good thing. Note that in section 1.1 evidence from Chicago was used in order to arouse interest in the study question. It may be that nowhere in Ontario is there a community as dysfunctional as is Chicago. Hajnal (1995), from the University of Chicago, investigated the levels of concentrated urban poverty in Canada and the United States, and observed that Canada had proportionately more people living in urban census tracts in which more than 40% of the population lived in poverty. Hajnal also found that in Canada the regional poverty rate is not predictive of its concentration, and that ethnically heterogeneous cities (e.g. Toronto, Kitchener) had less concentrated poverty than ethnically homogeneous ones (e.g. Montreal, Chicoutimi), despite a national trend for ethnic minorities to live in concentrated
poverty. However, he noted that Quebec accounted for a great deal of Canadian concentrated urban poverty, with Montreal harbouring 115 of Canada's 225 poverty tracts, and that Ontario had less than its share. If we continue to assume a high degree of collinearity between different dimensions of neighbourhood socio-economic context, then it may be that Ontario simply does not have a sufficient range in neighbourhood socio-economic context for its effect on health to be significant in statistical models. However, the results of the current study are not particularly different from those in other jurisdictions, which suggests that if sufficient variance is the 'problem,' it is also not common.

**The role of neighbourhood socio-economic context**

Context determines health in many ways, and recent calls for contextual analysis in epidemiology should be acted upon (Susser and Susser 1996; Kaplan and Lynch 1997; Shy 1997; Mackenbach 1998; Susser 1998; McMichael 1999; Schwartz et al. 1999). The historical and international differences in the health of populations are perhaps the clearest examples of the contextual determination. However, and notwithstanding the study limitations listed above, which are common to most studies on this subject, both the literature review and the empirical findings of this study indicate that much of the excitement over neighbourhood socio-economic context as an independent determinant of health has been overstated. This conclusion has been reached by others (Duncan et al. 1993; Boyle and Willms 1999). There is the danger here of usurping research and policy attention from the individual socio-economic determinants of health and from the determinants of individual socio-economic status, which this and other studies have demonstrated to be much more significant determinants of the health of populations. Individual socio-economic determinants of health are too frequently ignored in policy discourses. In his detailed review of legal and public policy interventions to promote population health, Gostin (2000) acknowledges that tort action by the state on businesses can have negative population health consequences when the tort costs to the businesses are passed on to the employees and to customers. However, it is remarkable that he fails to acknowledge that state policies that directly or indirectly affect the distribution of economic security are a powerful tool for the promotion of the health of the population. This omission is doubly remarkable since he acknowledges that the economic order is a determinant of population health, and the book in which his paper appears focusses on equity in the social and economic
determinants of health. We have inherited a long tradition of looking for non-redistributive ways to address public health problems that are fundamentally rooted in the social and economic orders; this tradition can be traced back at least to the sanitary efforts of Lords Lincoln and Morpeth and of Edwin Chadwick (Hamlin 1995; Hamlin and Sheard 1998).

Researchers sometimes suggest factors as possible causes of neighbourhood contextual effects on health that are actually contextual determinants of socio-economic status or socio-economic composition. For example, Robert (1998) suggests that the availability of healthy workplaces, and Ross (2000) implies that educational opportunities, are possible contextual determinants of health independent of socio-economic status. Robert also cites the work of Wallace and Wallace as suggesting that the withdrawal or denial of municipal services can affect the health and safety of residents, yet their proposed mechanism for the main effect of these service denials was the emigration of the socially advantaged (Wallace 1990; Wallace and Wallace 1998). In their discussion of the 450,000 'brownfield' properties in American cities, Greenberg et d. (1998) acknowledge that the abandoned and degraded land in brownfields (which are typically found in poor, predominantly black neighbourhoods) represent not only a physical threat, but also a threat to the property value of local landowners. They also argue that by providing an incentive for the advantaged to emigrate, by providing a convenient site for illegal activity, and by undermining the local tax base, these brownfield properties create a downward spiral in economic activity and in municipally funded services. Strong cases can be made for contextual determination of socio-economic status and composition.

Multilevel modelling techniques as they have been used here were established in the investigation of school effects, where it was found that schools had a very significant and independent contextual effect on students' educational trajectories (Willms and Raudenbush 1989). Due to the connection between educational achievement and subsequent health, such school effects are very relevant to population health. School effects and local labour force opportunities are examples of neighbourhood determination of individual socio-economic status. Recall that in figure 2.1 arrow 1 is two-headed. There are two pathways to consider, and these are illustrated in figure 5.2.

The process of social sorting need not involve any neighbourhood effects on health. Differences in neighbourhood composition might be traceable to conscious efforts at segregation of the vulnerable by the powerful, but could also be traceable to voluntary segregation, such as the formation of
Figure 5.2: Neighbourhoods and socio-economic status may be related in two distinct, and not exclusive ways. On the left the process of social sorting, which may or may not have consequences for health is illustrated. On the right the process whereby local opportunities shape the class and status situations of individuals is illustrated.

Linguistic and ethnic communities. The ecological relationship between aggregate neighbourhood advantage and individual health could be completely confounded by the individual's class and status situations if this is the only process at work.

However, social sorting could lead to, or be achieved by, a reduction in local opportunities, such as educational and employment opportunities. This would lead to the determination of the individual's socio-economic status by neighbourhood characteristics. Here again, the neighbourhood need not have direct effects on health or be a modifier of the socio-economic conditioning of the individual's health. In this scenario the neighbourhood is 'upstream' of socio-economic status. We saw in table 4.3 that the predictive association of neighbourhood characteristics with individual health was greater in the absence of individual controls than in their presence. This does not distinguish between the two pathways in figure 5.2. These could only be distinguished with longitudinal data gathered to examine the determinants of the individual's socio-economic status, such as the school effects literature mentioned above. In such research the class-specific roles of mobility and transportation should be investigated, because one can sometimes relocate towards opportunity, and can sometimes commute out of one's neighbourhood to work or study.
Neighbourhood inequalities in opportunities and constraints and in psychological stressors could arise in three distinct and non-exclusive ways: (1) the differential ability of people to live in neighbourhoods with good qualities (class-dependent mobility of the individual), (2) the market- or local tax-driven creation of neighbourhood qualities (class-based contingency of neighbourhood qualities), and (3) residents lobby for, organize to promote, or have better representation regarding improvements in their neighbourhood's quality (one example of Weber's notion of party). The first two mechanisms are completely dependent on the economic order, and so redistribution would be an effective response. The third mechanism would also likely be limited by the redistribution of economic and status power, since the neighbourhood advantages in party representation are likely rooted in the class and status situations of the neighbourhood residents. It is logically necessary that the geographic concentration of power will be limited by its equitable distribution among the people.

Macro-economic trends

Both urbanism, as measured in this study, and de-industrialization, as measured by Mitchell et al. (2000), despite being legitimate neighbourhood characteristics from a methodological point of view, have geographical patterns that greatly exceed neighbourhoods in scale. These patterns have complex implications for the physical and social environments as well as for the labour force opportunities available to the residents of large clusters of neighbourhoods. These are contextual differences that deserve further research. Few predict a future decrease or reversal in urbanization, but it may be that advantages of rural life, if they exist and are identified, can be strategically reinvented in urban settings. Likewise, de-industrialization, despite creating immediate decreases in the health status of affected residents, may offer long-term opportunities for net and sustainable population health promotion. The health implications of these broad social and economic patterns should be understood so that policy, to the extent that it has influence over such processes, can be created in the long-term interests of the public.

5.5 Summary

Neighbourhood socio-economic context does not appear to be a significant determinant of health independently of individual socio-economic status, nor does it appear to modify the health sequelae
of individual socio-economic status. However, neighbourhood socio-economic context may be an important determinant of individual class and status situations. Due to the profound effects of individual class and status on diverse dimensions of health, neighbourhood context and other determinants of the economic and social orders should be carefully examined for their potential modification in the interests of the public's health. Rural residence is associated with fewer chronic health problems, and the reasons for this should be investigated further.
Appendix A

Definitions of Weberian terminology

The following definitions are drawn from Weber (1946).

*Legal order* is maintained by a staff of persons who will use physical or psychical compulsion with the intension of obtaining conformity with that order. The structure of the legal order can directly enhance the chance to hold power and honour within its respective community (state or other), but is not normally their primary source.

*Power* is the chance of individuals or groups to realize their own will in a communal action, even against the resistance of others who are participating in the action. Power can be valued 'for its own sake.' Power comes in many forms, including political and economic. Economic power (especially 'naked' money power) can, but often does not, confer honour.

*Honour*, or status honour, or social honour, or prestige, can form the basis of power, and

*social order* is the way in which honour is distributed in a community. It has a similar relationship to the legal order as does the economic order.

*Economic order* is (merely) the way economic goods and services are distributed and used in a community. It is not identical to social order, but affects and is affected by it.

*Classes, status groups and parties* are phenomena of the distribution of power within a community.

*Classes* are not communities, but represent possible and frequent bases for communal action. Classes are groups of people who have the same class situation (or market situation), that is to say, that in the context of labour and commodity markets they have in common a specific causal component of their life chances, insofar as this component is represented exclusively by economic interests in the possession of goods and opportunities for income. Property and its lack are the basic categories of class situation, but each is further differentiated by the amount and kind of property possessed by the propertied, and by the kind of services that can be offered on the market by the propertyless. Persons not permitted to take part in the market, such as slaves and dependent children, cannot be a class, but rather are status groups.

*Status groups*, unlike classes, are normally communities, although often of an amorphous kind. The 'status situation' is defined by every typical component of one's life fate that is determined by a specific, positive or negative, social estimation of *honour*. Honour may be gained from class, but class is not a necessary component of status, which is often opposed to the
pretensions of sheer property. Despite this opposition, property is a status qualification in the long run. This purely conventional situation becomes legal privilege, positive or negative, when the social order has been 'lived in' and has achieved stability by virtue of a stable distribution of economic power. When the bases of the acquisition and distribution of goods are relatively stable, stratification by status is favoured. Technological change and economic transformations threaten stratification by status and redistribute power according to the economic order.

Parties are oriented toward the acquisition of social power, that is to say, toward influencing a communal action no matter what its content may be. Parties may represent interests determined through class or status situations, through both or through neither. Parties aim to influence the staff of persons charged with enforcing the legal order through naked violence, canvassing for votes, money, social influence, the force of speech, suggestion, clumsy hoax, and so on to the rougher or more artful tactics of obstruction in parliamentary bodies.
Appendix B

Model building

Prior to the multilevel modelling, exploratory work was done to determine how best to treat the variables. Important aspects of this work are summarized here in graphs and tables.

Figure B.1: In order to compare BINACT to dichotomized versions of the HUI, the betas for the dominant effect, age, were graphed. These models control for sex, education and income, with and without controlling for the count of chronic conditions as a quadratic.
In order to compare the count of chronic health problems to dichotomized versions of the HUI, the betas for the dominant effect, age, were graphed. This model controls for sex, education and income.

Figure B.2: In order to compare the count of chronic health problems to dichotomized versions of the HUI, the betas for the dominant effect, age, were graphed. This model controls for sex, education and income.

In order to compare HUI830 to the count of chronic health problems and to BINACT, the betas for the dominant effect, age, were graphed. These models control for sex, education and income, with and without controlling for the count of chronic conditions as a quadratic. Note that these curves are very similar in shape to those for BINACT, and not similar to that for chronic health problems.

Figure B.3: In order to compare HUI830 to the count of chronic health problems and to BINACT, the betas for the dominant effect, age, were graphed. These models control for sex, education and income, with and without controlling for the count of chronic conditions as a quadratic. Note that these curves are very similar in shape to those for BINACT, and not similar to that for chronic health problems.
Figure B.4: In order to compare the count of chronic conditions to dichotomized versions of the HUI, the betas for the dominant effect, age, were graphed. These models control for sex, education and income, with and without controlling for the count of chronic conditions as a quadratic. Note that these curves are neither similar to those for BINACT or to that for chronic health problems.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>EA-education</td>
<td>0.79</td>
<td>0.26</td>
<td>0.26</td>
<td>0.06</td>
</tr>
<tr>
<td>Urban/rural</td>
<td>0.44</td>
<td>0.44</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>EA-education*Urban/rural</td>
<td>0.79</td>
<td>0.56</td>
<td>0.33</td>
<td>0.10</td>
</tr>
</tbody>
</table>

1 EA-education not centered; urban=1
2 EA-education not centered; rural=1
3 EA-education centered; rural=1
4 EA-education centered; urban/rural centered

Table B.1: This table illustrates the effects of centering on multicollinearity. For each combination of variable definitions, each variable has been regressed on the other two, and the $R^2$ has been listed. Note that redefining the urban/rural dummy so that rural=1 reduces collinearity, as does centering the EA-education variable. Centering the urban/rural dichotomous variable further reduces the collinearity, but this treatment can be confusing and somewhat cumbersome to interpret as a predictor in multilevel analysis. Since the model 3 definitions resulted in stable standard errors among level 2 parameter estimates, and there was no apparent advantage to using model 4 definitions, the latter were not used.
Bibliography


Frohlich, N. and C. Mustard (1994, January). *Socio-Economic Characteristics*. Winnipeg: Manitoba Center for Health Policy and Evaluation, Department of Community Health Sciences, Faculty of Medicine, University of Manitoba.


Pincus, T., L. F. Callahan, and R. V. Burkhauser (1987). Most chronic diseases are reported more frequently by individuals with fewer than 12 years of formal education in the age 18-64 United States population. Journal of Chronic Diseases 40(9), 865–74.


