



Catalogue no. 11F0019MIE — No. 228

ISSN: 1205-9153

ISBN: 0-662-37802-4

Research Paper

Analytical Studies Branch research paper series

Neighbourhood Inequality, Relative Deprivation and Self-perceived Health Status

By Feng Hou and John Myles

Business and Labour Market Analysis Division
24-F, R.H. Coats Building, Ottawa, K1A 0T6

Telephone: 1 800 263-1136



This paper represents the views of the authors and does not necessarily reflect the opinions of Statistics Canada.



Statistics
Canada

Statistique
Canada

Canada

Neighbourhood Inequality, Relative Deprivation and Self-perceived Health Status

by Feng Hou and John Myles

11F0019MIE No. 228

ISSN: 1205-9153

ISBN: 0-662-37802-4

Business and Labour Market Analysis Division
24-F, R.H. Coats Building, Ottawa, K1A 0T6
Statistics Canada

How to obtain more information :
National inquiries line: 1 800 263-1136
E-Mail inquiries: infostats@statcan.ca

September 2004

We would like to thank James Dunn, Nancy Ross, Blair Wheaton, Russell Wilkins, and Michael Wolfson for their helpful comments and suggestions. This paper represents the views of the authors and does not necessarily reflect the opinions of Statistics Canada.

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2004

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission from Licence Services, Marketing Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

Aussi disponible en français

Table of Contents

1. Introduction	5
2. Data, Measures and Methods	8
2.1 Data	8
2.2 Measures.....	9
2.3 Analytical techniques	12
3. Results	14
3.1 The aggregate (“ecological”) correlation between neighbourhood income inequality and health	14
3.2 The effect of neighbourhood inequality after adjusting for individual characteristics.....	15
3.3 The effect of neighbourhood inequality vs. the effect of neighbourhood income level....	16
3.4 The effect of other neighbourhood characteristics	17
3.5 Do the effects of neighbourhood inequality and neighbourhood income vary among income groups?	19
3.6 The effects of city-level income inequality and economic segregation.....	20
4. Discussion and Conclusion	20
References	29

Abstract

This study examines two theses concerning the relation between individual health status and the socio-economic composition of the neighbourhoods in which they live. In the first variant, the claim is that more unequal communities will not generate the social capital/social cohesion that makes for healthy populations irrespective of whether these communities are rich or poor. At the individual level, the implication is that individuals who live in high inequality neighbourhoods will tend to have poorer health irrespective of their own income levels. The second variant involves testing two competing claims about the health advantages/disadvantages of sharing neighbourhoods with more or less affluent neighbours. On the one hand, sharing neighbourhoods with more affluent families may have negative effects on the health of the less affluent if residential proximity generates invidious social comparisons or competition for scarce resources. On the other hand, the less affluent may derive positive externalities by living with more affluent neighbours because of richer institutional resources and/or “learning effects.” This study combines individual micro-data from Statistics Canada’s 1996/97 National Population Health Survey (NPHS) with neighbourhood-level characteristics estimated from the 1996 Census of Canada 20% sample micro data file. We find that an overall negative association between neighbourhood income inequality and self-assessed health status does not persist once controls are introduced for individual level socio-economic characteristics. However, individuals, regardless of their own income status, derive positive health benefits from sharing neighbourhoods with higher income, better educated, neighbours.

Keywords: Neighbourhood, income inequality, health

1. Introduction

The observation that income inequality is correlated with aggregate measures of population health among affluent nations has stimulated tremendous interest and much debate among health and social science researchers. Among the reasons for this attention is the provocative thesis advanced by Wilkinson (1992; 1996) and others (e.g. Daniels, Kennedy and Kawachi, 1999) to account for the association, namely that the causal mechanisms linking income to health in affluent nations are psycho-social not just material in nature. While there are several variants of Wilkinson's and related theses (see below), roughly speaking the claim is that in affluent societies income differences among individuals generate invidious comparisons that are stress inducing (Wilkinson, 1996) and/or that high levels of income inequality in the population generate low levels of social cohesion or social capital (Kawachi and Kennedy, 1997; Wilkinson, 1996) producing more aggressive and less supportive social environments.

The subsequent inequality/health debate has focused on whether or not the observed statistical association between inequality and average population health is "real" (a direct effect) or a "statistical artefact" (an indirect effect). The fact that at the individual level the relationship between income and health is concave—each additional dollar of income raises a person's health but by ever smaller amounts—ensures that at the aggregate level we will observe an association between average population health and income inequality (Gravelle, 1998; Wolfson et. al., 1999). But the implication of this "statistical artefact" is that the association reflects individual level differences in *absolute* income, not inequality *per se*. As Deaton (2003: 118) observes, the designation of the indirect effect of inequality as a "statistical artefact" is perhaps unfortunate since it suggests there is no real link between inequality and health. This is not a debate about whether inequality "matters" or whether redistribution might improve the health of populations (Deaton, 2003; Gravelle, 1998) but over causal mechanisms. Does the association between inequality and health in affluent nations result from psycho-social processes or the quality of social relations associated with higher and lower levels of inequality or from differences in individual health-related resources and material living standards?

Since the causal mechanisms imputed to account for the relationship are rarely observed directly, debates focus on whether patterns of statistical association are those anticipated by one or other sets of theoretical claims (Wagstaff and van Doorslaer, 2000).¹ In this paper, we use multi-level models to address one of these debates, namely the now conventional finding that the statistical association between inequality and health becomes weaker as one moves from higher (countries, states) to lower (cities, neighbourhoods) levels of aggregation² (Mellor and Milyo, 1999; Soobader and LeClere, 1999; Blakely, Lochner and Kawachi, 2002). This finding is sometimes invoked to refute Wilkinson's thesis on the grounds that if income inequality affects individual health through social comparison it should show a *stronger* effect when it is measured closer to

-
1. The imputation as opposed to direct observation of causal mechanisms is of course hardly idiosyncratic to this literature. For a review of studies that attempt to identify the processes or mechanisms that *account* for contextual effects at the neighbourhood level see Sampson, Morenoff and Gannon-Rowley (2002).
 2. In geographical analysis of aggregate data, this is often referred to as "scale effects". It is different from "zoning effects" that concern the choices of a particular set of boundary definitions at a given geographic scale.

home (Kawachi and Kennedy, 1999; Mellor and Milyo, 2001; Wagstaff and van Doorslaer, 2000). Wilkinson's (1997) explanation for this pattern is that as we move from higher to lower units of aggregation the relevant social comparisons that generate anxiety, stress and a sense of relative deprivation are weakened since the salient heterogeneity is mostly between rather than within smaller geographic units. Within small, relatively homogeneous, neighbourhoods, he suggests, the relevant comparison is based on how affluent the neighbourhood is relative to other neighbourhoods in the wider community. As a result, *average* income in small geographic units such as neighbourhoods matter but the neighbourhood *distribution* of income is less salient.

While there are any number of *good* reasons (see our concluding discussion) to expect that inequality effects may be less likely to appear at lower than at higher levels of aggregation, the assumption of neighbourhood homogeneity is not one of them. The premise that income inequality is always lower at lower levels of aggregation (e.g., neighbourhoods) than at higher levels of aggregation (cities) as a result of high levels of economic segregation is incorrect (see Appendix 1 "Are neighbourhoods more economically homogeneous than cities?" for details). Although economic segregation has been rising in both U.S. (Jargowsky, 1997) and Canadian (Myles, Picot and Pyper, 2000) cities in recent decades, inequality between neighbourhoods ("segregation") still accounts for a small portion of city-level inequality.

In 25 Canadian metropolitan areas, "economic segregation" (the share of total inequality accounted for by between-neighbourhood inequality) accounts for no more than 8 to 12 percent of total city-level income inequality. This is a standard result for U.S. cities as well (Jargowsky 1997; Lobmayer and Wilkinson 2002). As a result, city-level income inequality is almost perfectly correlated with its within-neighbourhood component (see Appendix Table 1.3). This high correlation suggests that any health effect of city-level inequality basically reflects the impact of its *average* within-neighbourhood component *and vice versa* (see Lobmayer and Wilkinson 2002). Most neighbourhoods are economically heterogeneous and, depending on the measure, from 12 to 28 percent of neighbourhoods have higher inequality values than the cities in which they are found. In short, if Wilkinson's and related theses do not hold up at the neighbourhood level, this result cannot be explained away by the absence of sufficient economic heterogeneity and some other mechanism must be at work.

It is important to be precise about the expected statistical relationships implicit in Wilkinson's claims when estimated with individual rather than with aggregate data. As Wagstaff and van Doorslaer (2000) highlight, Wilkinson and others have made several claims about the causal mechanisms relating inequality to health that imply rather different statistical patterns of association when estimated at the individual level.

For the income inequality (IIH) variant (Wilkinson, 1996; 1999), it is inequality *per se* that matters: more unequal communities will not generate the social capital/social cohesion that makes for healthy populations irrespective of whether these communities are rich or poor. At the individual level this means that health will be a positive function of individual income but *inversely* related to neighbourhood inequality. As shown later, high inequality is characteristic of neighbourhoods with concentrations of disadvantaged populations and also of neighbourhoods with concentrations of advantaged (high income, well educated) populations. This empirical fact

provides an opportunity to test the claim that income inequality *per se* rather than material conditions accounts for health outcomes. If inequality is the culprit it should affect the rich as well as the poor.

The main challenge to the IIH thesis is the counter-claim that the effects of income inequality are indirect, mediated by differences in individual living standards that are correlated with inequality. The absolute income (AIH) hypothesis implies that the association will disappear when one correctly identifies and controls for individual level differences in socio-economic resources. Both claims may be true at the same time, of course, but they do lead to very different predictions about the likely effects of rising inequality. Inequality can rise because of increases in the incomes of the rich or because of declines in the incomes of the poor, that is, by changes in *either* tail of the distribution. For the Wilkinson and related theses, however, differences of this sort are irrelevant. What matters is the average economic “distance” among individuals, irrespective of how that “distance” is produced, and that is precisely what valid inequality measures capture (Cowell, 1995).

In contrast, if one believes health outcomes are mainly the result of differences in individual living standards (the AIH hypothesis), one would predict very different outcomes when inequality rises because the rich are getting richer than when inequality rises because the poor are getting poorer. Under the assumption of declining marginal returns, one would predict small marginal gains in population health when the rich are getting richer (the rich will be a little healthier). In contrast, when the poor are getting poorer, one would predict somewhat more substantial declines in average population health.

The relative income (RIH) hypothesis (Wilkinson, 1998) leads to a related but rather different debate. The question is whether lower income people derive positive externalities (RIH⁺), negative externalities (RIH⁻) or none at all from sharing environments with higher income families? The sociological debate on this topic was stimulated by William Julius Wilson’s (1987) *The Truly Disadvantaged*. Wilson’s thesis is that, *ceteris paribus*, the poor derive positive externalities from sharing neighbourhoods with more affluent families as a result of richer institutional resources and/or “learning effects.”

In their classic review of the topic, however, Jencks and Mayer (1990) identify equally plausible theoretical reasons for assuming the opposite may be true. Anticipating Wilkinson’s thesis (RIH⁻), they suggest that the less affluent may experience greater levels of stress, anxiety and relative deprivation as well as sharper competition for scarce institutional resources (grades in school, access to health care facilities) when faced with competition from better educated, more affluent, neighbours (see also Deaton, 2003: 123). Though not always recognized as such (Subramian, Kawachi and Kennedy, 2001), the appropriate statistical model for testing this claim in multi-level models with individual level data is one that tests for an effect of average (or median) community income controlling for differences in individual socio-economic resources (Wagstaff and van Doorslaer, 2000: 557). Empirically, the implication of the relative deprivation (RIH⁻) thesis is that holding individual income constant, *average neighbourhood income* (not income inequality) will be *inversely* related to health: less affluent individuals will have better outcomes when they share neighbourhoods with their economic peers than when they live near

more affluent families. Wilson's counter-claim (RIH⁺) predicts the opposite. Since policy-makers have considerable influence on the economic composition of neighbourhoods through zoning laws and other instruments, the answers to these questions can potentially have non-trivial public health implications.

The evidence that exists provides little support for the negative variant of the RIH hypothesis (for reviews see Brooks-Gunn, Duncan and Aber, 1997; Deaton, 2003; Wagstaff and van Doorslaer 2000) at the neighbourhood level and for the most part contradicts it. As Wheaton and Clarke (2003) highlight, however, there is no inherent reason why we should expect individual and contextual effects to be additive. It may well be that middle class and more affluent individuals do derive positive externalities from living with advantaged neighbours while the personally disadvantaged derive no, or even negative, externalities from association with more affluent neighbours. If such an interaction is present, it suggests that earlier conclusions are based on misspecification of the association between health outcomes and the social context in which people live.

Although not the main focus of our analysis, we conclude with model specifications that incorporate *city-level* measures of income inequality, economic segregation and median income in part to clarify the difference between city and tract-level measures of "neighbourhood" characteristics and to control for city-level differences in our estimates of neighbourhood effects. Based on previous research on Canadian cities (Ross et al., 2000), we do not anticipate finding significant associations between health status and *city-level* indicators of income inequality, a point to which we return below.

2. Data, Measures and Methods

2.1 Data

This analysis is based on cross-sectional household data from Statistics Canada's 1996/97 National Population Health Survey (NPHS) and the 1996 Census 20% sample micro data. There were 81,804 respondents to the questions on in-depth health information. This study is based on 34,613 NPHS respondents aged 12 or older who were residing in one of Canada's 25 Census Metropolitan Areas (CMAs) that are delineated around urban areas (*urbanized core*) with a population of at least 100,000. In this analysis, the census tract represents the basic neighbourhood unit.³ We excluded 13 census tracts with total population less than 500 to ensure reliable estimation of neighbourhood characteristics. Thus our final sample included 34,592 individuals living in 3044 census tracts.

Neighbourhood and city level income, income inequality and other aggregate characteristics were calculated directly from the 1996 Census micro-data file allowing us to make adjustments and calculate measures of neighbourhood inequality and related measures unavailable to most

3. Census tracts have carefully designed attributes, contain a wide range of demographic and socio-economic information, and allow for national and historical statistical comparisons. Studies have shown that contextual variables derived from census tracts and "natural" neighbourhoods have similar health effects (Ross, Tremblay and Graham 2003).

researchers. These neighbourhood attributes were then matched to the records of each NPHS respondent.

2.2 Measures

Health Status: The health outcome used in this analysis is self-perceived health. Self-perceived health refers to an individual’s assessment of his or her general health. Many American and European studies have demonstrated that self-perceived health is an important predictor of the onset of disability and mortality, independent of other medical conditions and psychosocial states (Idler and Benyamini, 1997; Ferraro, Farmer and Wybraniec, 1997). Respondents were asked: “In general, would you say your health is excellent, very good, good, fair or poor?” Scores range from 1 to 5, with a higher score indicating better self-perceived health. In our multivariate analyses, we examined the sensitivity of our results to treating the self-perceived health measure as an ordinal, continuous and categorical variable (see details later in the section on *Analytical techniques*).

Inequality measures: Although some studies suggest that the choice of inequality measures does not change the inequality-health association (Kawachi and Kennedy, 1997), others report that measures reflecting income dispersion between particular economic strata do matter (Daly, Duncan, Kaplan, and Lynch, 1998; Weich, Lewis and Jenkins, 2002). In this study, we use six income inequality measures that satisfy the standard criteria for valid inequality measures (Cowell, 1995). Unlike the majority of studies, index values for census tracts and cities (CMAs) are estimated directly with the underlying census micro-data after adjusting economic family income with an equivalence scale and assigning the equivalized income to individuals (see Deaton 2003: 135-136).⁴

Among the six indexes, the mean logarithmic deviation (I_0),⁵ the Theil index (I_1),⁶ and the squared coefficient of variation (I_2)⁷ all belong to the Generalized Entropy class $GE(\alpha)$. The measure I_0 , I_1 , I_2 corresponds to $\alpha=0$, 1, and 2. We also include the Gini index (G_i),⁸ median

4. The equivalence scale for adjusting family size assigns a weight of 1 to the first person and 0.4 to each additional person.

5. $I_0 = \frac{1}{n} \sum_{i=1}^n \log \frac{\bar{y}}{y_i}$, where n is the number of individuals, y_i is the income of individuals, and $\bar{y} = \frac{1}{n} \sum y_i$

6. $I_1 = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{\bar{y}} \log \frac{y_i}{\bar{y}}$.

7. See footnote 9 for the computation of the coefficient of variation.

8. $Gini = \frac{1}{2n^2 \bar{y}} \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|$.

share (MS),⁹ and the coefficient of variation (CV).¹⁰ The various indices are sensitive to transfers at different income levels. The mean logarithmic deviation and the Theil index are sensitive to transfers at the lower end of income distribution. The Gini index and median share are sensitive to transfers in the middle of the distribution while the coefficient of variation and the squared CV are more sensitive to transfers among high-income earners (Allison, 1978).

One desirable attribute of $GE(\alpha)$ measures is that inequality at the city level (I_T) can be decomposed into between- and within-neighbourhood components: $I_T = I_B + I_W$ (Jenkins, 1995) where I_W is the average of neighbourhood inequality weighted by each neighbourhood's share in the city population. The ratio I_B / I_T —the share of total inequality accounted for by between neighbourhood inequality—is often seen as indexing the level of economic “segregation” but should not be confused with traditional segregation indices such as the index of dissimilarity that measure a rather different concept.¹¹ Jargowsky's (1996) neighbourhood sorting index is simply a specific case of I_B / I_T that can be calculated with a variety of decomposable inequality measures. For two cities with the *same* value of I_T , a higher value of I_W indicates less economic segregation at the city level. Decomposable inequality measures of this sort have been used in previous research (Lobmayer and Wilkinson, 2002) to examine the effects of city-level inequality on health outcomes. Note, however, that like I_T , I_B , I_W , and segregation (I_B / I_T) are city-level, not neighbourhood, attributes.

For each of the six inequality measures shown here, higher values indicate greater inequality. In our multivariate analyses, each of the six inequality measures is calculated at the neighbourhood level and neighbourhoods are grouped into five quintiles based on the level of inequality in each neighbourhood: (Q1) low inequality neighbourhoods; (Q2) lower-middle; (Q3) middle; (Q4) upper middle; and (Q5) high inequality neighbourhoods. Neighbourhood inequality coefficients are estimated with dummy variables with low inequality neighbourhoods serving as the reference category to capture possible non-linearities in the relation between income inequality and health.

Neighbourhood economic conditions: *Neighbourhood median income* is calculated for all individuals in a tract based on their equivalized family income, i.e., after adjusting for the economies of scale associated with family size (see footnote 5 above).

9. The median share measure (MS) is defined here as the proportion of total family income belonging to the more affluent 50 percent of families within a geographical area. The larger value of the measure, the higher level of inequality. This measure is just in the opposite direction of the one used by Kaplan et al., (1996) (let's call LMS). LMS is defined as the proportion of total family income belonging to the less affluent 50 percent of families within a geographical area. $MS = 1 - LMS$.

10.
$$CV = \frac{1}{y} \left[\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right]^{1/2}.$$

11. Traditional segregation measures of economic segregation such as the index of dissimilarity begin by assigning persons to fixed income classes (people are distributed in “income space”) whereas inequality measures distribute income in “people space” (the share of income going to fixed percentage of the population). In the inequality literature these two approaches are referred to as “polarization” and “inequality” measures, respectively and over time they may move in opposite directions (see Wolfson, 1997).

Other neighbourhood variables: In order to examine the correlates of neighbourhood income inequality that might account for or mediate the relation between neighbourhood inequality and health, we select five other neighbourhood characteristics: education, age, family type, immigration, and race. For each of these characteristics, we measure the prevalence of a particular category. We choose the percentage of adults (age 15 years and over) with a university degree, seniors (aged 65 years and over), single-parent families, recent immigrants (living in Canada 10 years or less), and percent non-white.

Individual level control variables: Family income measured with a five-category measure provided in the survey data is based on a measure of income level adjusted for household size (Appendix Table 1). The bottom three categories (Very Low, Low and Lower-Middle) capture the bottom half of the income distribution and the top two (upper-middle and high income) the top half. Five dummy variables were created and the “high income” category used as the reference group. Since a large percentage of respondents did not report their income, another dichotomous variable (income missing = 1, others = 0) was created to incorporate them in the analyses.

Other variables included for analysis are age and sex, education, immigrant status, and racial minority status. These variables are often associated with individual level income. The inclusion of these variables in the model may, to a large extent, capture the effect of permanent income or income potential since current income is often not a reliable indicator of economic well-being.

Table 1: Descriptive statistics for individual level variables

Variable	Sample size	Weighted frequency distribution or mean (standard deviation)
General health		
Poor	922	2.10%
Fair	2,654	6.80%
Good	8,988	26.70%
Very good	13,289	38.30%
Excellent	8,739	26.10%
Age	34,592	41.3(18.2)
Sex		
Male	16,182	49.20%
Female	18,410	50.90%
Immigration status		
Immigrants	8,062	25.30%
Non-immigrants	26,530	74.70%
Visible minority status		
Non-whites	3,631	14.20%
Whites	30,961	85.80%
Education		
< High school	8,727	27.10%
High school graduation	5,837	15.80%
Some postsecondary	13,857	39.60%
University	6,171	17.60%
Income status*		
Very Low	1,315	3.80%
Low	2,742	8.40%
Lower-middle	7,116	22.30%
Upper-middle	10,615	32.70%
High	4,962	15.00%
Income not reported	7,842	17.80%

Note: see Appendix Table 1 for the definitions of income categories.

Age is a continuous variable, coded by single year, ranging from 12 to 99; sex was coded as females = 1, males = 0. Immigrant status was coded as immigrants = 1, non-immigrants = 0. Race is coded with Statistics Canada's indicator of "visible minority status" and assigned values of one for non-white, non-Aboriginal and zero otherwise.

Education was coded into four categories: less than high school (the reference category), completed high-school, some post-secondary education, and completed university.

2.3 Analytical techniques

In our analyses, health outcomes are predicted by both individual-level and neighbourhood-level variables. Since all people in the same neighbourhood have the same value on the neighbourhood variables, there is a problem of independence (or lack of it) for observations drawn from the same neighbourhood. To address this issue, we use hierarchical linear models (HLM) and the HLM 5 program (Bryk and Raudenbush, 1992; Raudenbush, Bryk and Cheong, 2000).

Since self-reported health is an ordinal level measure, the ordered logit model based on the cumulative distribution probabilities of the response categories is the appropriate estimation method. Unfortunately, while the current version of HLM can incorporate population weights for multi-level models with a continuous dependent variable this is not possible with ordered logit models. To determine the effect of alternative estimation techniques, we first established that the *unweighted* multi-level ordered logit models produced the same results as the *unweighted* multi-level models in which health status is treated as a continuous measure both in terms of the significance of variables and the direction of their effects. Second, we compared weighted and unweighted results with health measured as a continuous variable and concluded that *unweighted* estimates biased our findings in *favour* of our main conclusions. Since the unweighted estimates produced smaller standard errors than the weighted multi-level models, we were much more likely to get statistically significant coefficients with the former than with the latter. All three sets of estimates, however, lead to the same qualitative conclusions.

We also explored two ways of creating a dichotomous variable for self-perceived health. In the first, we contrasted "poor" (fair or poor) with "good" (excellent, very good, good) health. In the second, we contrasted "healthy" (excellent or very good) with others (good, fair or poor). The comparison between "healthy" and others and the continuous measure of health status yield the same results with respect to the significance of associations between the outcome variable and neighbourhood explanatory variables. The comparison between "poor" vs. "good" often produced non-significant results largely because it suppresses most of the measured variance in self-perceived health status. Less than 10% of respondents reported poor (2.1%) or fair (6.8%), while more than half of the respondents reported either very good (38.3%) or excellent health (26.1%).

Here, we only present weighted multi-level modeling results treating self-perceived health as a continuous variable although the results based on the ordinal and categorical outcomes are available upon request. With a continuous dependent variable we could also perform ANOVA and ANCOVA to decompose the variance in the dependent variable into between- and within-

neighbourhood components allowing us to estimate what proportion of the variance in the outcome can be explained by neighbourhood variables.

The presentation of findings is organized as follows. First, a means-as-outcome model (Model 1) is used to regress the neighbourhood average of the health outcome on neighbourhood income inequality. This model is equivalent to estimating the simple (“ecological”) correlation between average self-perceived health in a neighbourhood and neighbourhood inequality.

Model 2 adds individual family income and other individual level variables to Model 1. The results indicate whether neighbourhood income inequality has a unique effect on individual health over and above the effect of income measured at the individual level. The concern here is whether the association between income inequality and health may actually reflect neighbourhood differences in the composition of income groups as the AIH hypothesis suggests.

Model 3 adds neighbourhood median income to Model 2 to assess the Relative Income (RIH) hypothesis and to determine whether the association between health status and neighbourhood inequality exists independently of the effect of neighbourhood income level.

In Model 4, we determine whether other neighbourhood characteristics can further explain the association between health and neighbourhood income inequality. To avoid problems of multicollinearity, we test the selected neighbourhood characteristics one at a time and only report the results if the inclusion of one of the characteristics changes the effect of neighbourhood income inequality.

In Model 5, we test whether the effect of neighbourhood income inequality and absolute income level differs across income groups. Specifically, we test whether people in lower income groups are more subject to the effects of neighbourhood economic conditions than higher income individuals.

Finally, we add measures of city-level income inequality and economic segregation to Model 4 to see whether inequality measured at a higher geographic level has any independent effect on individual health after controlling for neighbourhood income inequality and individual level characteristics. We report results from adding higher order, city level, effects for city-level inequality (I_T), city-level economic segregation (I_b/I_T), the within-neighbourhood component (I_W), the between-neighbourhood component (I_b) and median city income. Since the city level Gini, median share, and the coefficient of variation cannot be decomposed into between- and within-neighbourhood components, we use Theils’ two components for city-level estimates that include neighbourhood level values for Gini and median share as the inequality measure, and squared CV’s two components for the model using CV as the neighbourhood inequality measure.

The nature of our design clearly precludes the possibility of making strong claims about causality (Kawachi and Berkman, 2003). The most obvious problem is that the influences between income and health run in both directions and our cross-sectional design offers no solution to this problem. The second problem is selection bias. People have limited choice over their nation of residence as a result of immigration barriers but they can and do make choices about both cities and neighbourhoods based on their health conditions. Third, while there is little evidence in the

literature, we cannot preclude the possibility that individual level differences such as income are not themselves an “effect” of neighbourhood context. Finally, current health outcomes may well be the product of earlier social contexts (Wheaton and Clarke, 2003) that will not be captured by point-in-time estimates for the geographically mobile. Our use of the term “effect” or “effects” therefore should always be read with implied quotation marks.

3. Results

3.1 The aggregate (“ecological”) correlation between neighbourhood income inequality and health

In Canada, as elsewhere, individual level health has been shown to rise with income but at a decelerating rate (“declining marginal returns”) as income rises (Wolfson, et. al., 1993). Given this fact, it is perhaps surprising that previous research (Ross et al., 2000) has not found the expected “statistical artefact” (Gravelle, 1998)—a negative association between average health outcomes and income inequality—when measured at the aggregate level with census metropolitan areas (see also section 5 below). As shown in Table 2, however, we do find the expected negative association between inequality and self-perceived health at the neighbourhood level.

In short, unlike U.S. findings, Canadian results show a *stronger* (“ecological”) association between inequality and health at lower (neighbourhoods) than at higher (cities) levels of aggregation. The reasons for these differences, we suspect, have to do with the number of observations (there are few cities in Canada but many neighbourhoods), the lower levels and restricted range of variation in city-level inequality in Canada compared to the U.S. (Ross et al., 2000), and the greater variation in neighbourhood- than in city-level inequality.

The six income inequality measures¹² are similar in that they are all significantly associated with individual health in the same direction but differ with respect to where the differences are located and the strength of the association. In general, the coefficients show a shallow U shape pattern with health levels first declining and then rising as neighbourhood inequality rises. For Theil, Gini, and median share, the inflexion point is the upper-middle category while for the mean log deviation and CV squared the inflexion occurs in the middle of the neighbourhood inequality distribution. To anticipate later results, we will see this U-shaped pattern results from the fact that high inequality is characteristic of both the most advantaged and the least advantaged neighbourhoods (i.e., the result of compositional differences between neighbourhoods characterized by higher and lower levels of inequality).

12. Since the results for CV and CV squared are identical, the result for CV is not presented in this and subsequent tables.

Table 2. Hierarchical linear models for individual self-perceived health and neighbourhood inequality

	Log deviation		Theil		Gini		Median share		CV squared	
	B	SE	B	SE	B	SE	B	SE	B	SE
Intercept	3.843 ***	0.020	3.834 ***	0.020	3.840 ***	0.017	3.852 ***	0.017	3.832 ***	0.017
Neighbourhood income inequality										
Lower-middle	-0.036	0.024	-0.031	0.025	-0.038	0.025	-0.023	0.023	-0.060 *	0.026
Middle	-0.089 **	0.026	-0.082 **	0.026	-0.072 **	0.026	-0.100 ***	0.027	-0.090 **	0.026
Upper-middle	-0.066 *	0.028	-0.097 **	0.027	-0.099 ***	0.027	-0.116 ***	0.027	-0.059 *	0.027
Highest	-0.071 *	0.027	-0.006	0.028	-0.039	0.029	-0.069 *	0.029	0.006	0.027
Neighbourhood variance explained										
	0.45%		0.99%		0.67%		1.27%		0.61%	

Note: * significant at $p < .05$, ** $p < .01$, *** $p < .001$.

In terms of the strength of the association, the median share explains 1.3% of the *neighbourhood* variance (i.e., the between-neighbourhood component of the total variance) in the outcome, about twice as large as the variance explained by the mean log deviation, CV squared, and Gini respectively. Thus, although statistically significant, the association between neighbourhood income inequality and health is rather weak. Since the between-neighbourhood variance accounts for about 15% of the total variance in outcomes, neighbourhood inequality measured by the median share explains only about 0.2% ($1.27\% \times 15\%$) of the total variance in the outcome. The “effect” of neighbourhood income inequality as measured by the coefficients is also small. The largest difference is between neighbourhoods with upper-middle inequality and those with the lowest inequality when median share is used as the inequality measure. This difference is 0.116, about one tenth of the standard deviation of the outcome.

3.2 The effect of neighbourhood inequality after adjusting for individual characteristics

After controlling for individual level socioeconomic characteristics, the association between neighbourhood income inequality and the health outcome is not statistically significant (Table 3). This result implies that the observed negative health effect of neighbourhood income inequality in Table 2 can be attributed to neighbourhood differences in the composition of people with different family incomes and other socioeconomic characteristics. Individuals in the lowest income group are about half a point lower in their scores of self-perceived health than those in the highest income group. The difference in health status between people with university education and those with less than high-school education is close to the difference between those in the highest and lowest income groups. Women tend to have poorer perceived health than men, and older people tend to have poorer perceived health than younger ones. Immigrant status and visible minority status are generally not significantly associated with individuals’ self-perceived health.

Table 3. Hierarchical linear models for individual self-perceived health and neighbourhood inequality controlling for individual level variables

	Log deviation		Theil		Gini		Median share		CV squared	
	B	SE	B	SE	B	SE	B	SE	B	SE
Intercept	3.789 ***	0.016	3.779 ***	0.016	3.781 ***	0.016	3.803 ***	0.020	3.780 ***	0.016
Individual Level Variables										
Age	-0.013 ***	0.001	-0.013 ***	0.001	-0.013 ***	0.001	-0.013 ***	0.001	-0.013 ***	0.001
Female	-0.044 **	0.017	-0.045 **	0.017	-0.045 **	0.017	-0.049 *	0.019	-0.045 **	0.017
Immigrants	-0.025	0.024	-0.026	0.024	-0.026	0.024	-0.034	0.027	-0.026	0.024
Non-white	-0.057	0.030	-0.057	0.030	-0.057	0.030	-0.073 *	0.033	-0.057	0.030
Education										
High school graduate	0.231 ***	0.027	0.231 ***	0.027	0.231 ***	0.027	0.211 ***	0.029	0.231 ***	0.027
Some post secondary	0.256 ***	0.024	0.256 ***	0.024	0.257 ***	0.024	0.249 ***	0.026	0.256 ***	0.024
With university degree	0.407 ***	0.027	0.405 ***	0.027	0.406 ***	0.027	0.425 ***	0.031	0.404 ***	0.027
Family income										
Very Low	-0.476 ***	0.061	-0.478 ***	0.061	-0.478 ***	0.061	-0.477 ***	0.066	-0.478 ***	0.061
Low	-0.411 ***	0.045	-0.411 ***	0.045	-0.411 ***	0.045	-0.400 ***	0.048	-0.411 ***	0.045
Lower-Middle	-0.204 ***	0.030	-0.203 ***	0.030	-0.203 ***	0.030	-0.207 ***	0.032	-0.203 ***	0.030
Upper-Middle	-0.109 ***	0.027	-0.108 ***	0.027	-0.109 ***	0.027	-0.123 ***	0.030	-0.108 ***	0.028
Income missing	-0.139 ***	0.029	-0.138 ***	0.029	-0.139 ***	0.029	-0.163 ***	0.032	-0.138 ***	0.029
Neighbourhood income inequality										
Lower-middle	-0.001	0.022	0.007	0.023	0.005	0.023	0.017	0.021	-0.019	0.024
Middle	-0.024	0.025	-0.014	0.024	-0.003	0.024	-0.035	0.025	-0.020	0.024
Upper-middle	-0.007	0.026	-0.019	0.025	-0.021	0.025	-0.031	0.025	0.007	0.026
Highest	-0.001	0.026	0.045	0.026	0.028	0.027	0.003	0.027	0.043	0.024

Note: * significant at p<.05, ** p<.01, *** p<.001.

3.3 *The effect of neighbourhood inequality vs. the effect of neighbourhood income level*

Up to this point, we have not included neighbourhood income level in our regression models. This is an important omission since the sign and size of this coefficient bear on three distinct claims. First, the relative income (social comparison) hypothesis predicts that, controlling for individual income, individual health will be inversely related to neighbourhood income. Failure to find such an association could, as Wilkinson argues, be a result of the fact that there is insufficient heterogeneity at the neighbourhood level for the relevant social comparisons to be salient. If this were the case, however, the association between neighbourhood income level and health should disappear once we control for individual income. Economic homogeneity at the neighbourhood level is not the only reason one might anticipate the absence of an association of health with neighbourhood income but such a result is consistent with such a claim. Finally, Wilson's (1987), and related claims concerning the positive externalities individuals derive from sharing neighbourhoods with more affluent individuals lead us to expect a positive coefficient.

Table 4. Hierarchical linear models for individual self-perceived health and neighbourhood inequality controlling for individual level variables and neighbourhood median income

	Log deviation		Theil		Gini		Median share		CV squared	
	B	SE	B	SE	B	SE	B	SE	B	SE
Intercept	3.678 ***	0.030	3.679 ***	0.030	3.679 ***	0.031	3.684 ***	0.031	3.757 ***	0.016
Neighbourhood income inequality										
Lower-middle	0.025	0.022	0.035	0.023	0.034	0.023	0.048 *	0.022	0.006	0.024
Middle	0.016	0.025	0.026	0.025	0.038	0.025	0.008	0.026	0.015	0.025
Upper-middle	0.047	0.027	0.026	0.026	0.030	0.026	0.024	0.026	0.041	0.026
Highest	0.057 *	0.027	0.076 **	0.026	0.071 *	0.027	0.057 *	0.028	0.052 *	0.024
Neighbourhood median income										
Lower-middle	0.033	0.029	0.032	0.029	0.033	0.029	0.031	0.030	0.032	0.029
Middle	0.058	0.031	0.054	0.031	0.057	0.031	0.051	0.032	0.054	0.031
Upper-middle	0.099 **	0.031	0.093 **	0.030	0.097 **	0.031	0.091 **	0.032	0.089 **	0.030
Highest	0.177 ***	0.029	0.167 ***	0.029	0.172 ***	0.029	0.170 ***	0.030	0.161 ***	0.029

Note: * significant at $p < .05$, ** $p < .01$, *** $p < .001$.

Note: Individual level variables are included in each model, but not presented here.

The results of models that control for individual level characteristics and include both median neighbourhood income and neighbourhood income inequality are shown in Table 4. Neighbourhood income level has an independent, positive, association with individual health over and above the effects of individual level characteristics. People living in neighbourhoods with upper-middle and the high-income levels have significantly better self-perceived health. People in high-income neighbourhoods, on average, have scores of self-perceived health about 0.17, or about 5% higher than those living in neighbourhoods with the lowest median income. Hence, while the gains only appear for individuals in upper-middle and higher income neighbourhoods (the association with neighbourhood income is exponential), we cannot reject Wilson's (1987) hypothesis that individuals obtain positive externalities from sharing neighbourhoods with more affluent families. It is worth noting that the unweighted regressions from both the ordered logit and from models with health status measured as a continuous variable produced significant coefficients for lower-middle and middle income neighbourhoods as well and a linear rather than exponential pattern of association.

Paradoxically, once we control for neighbourhood income, neighbourhood inequality shows a significant, positive, although not strong, association with individual health. People living in neighbourhoods with the highest income inequality tend to have *better* self-perceived health than those living in neighbourhoods with the lowest income inequality and this result is consistent across different inequality measures. Why so?

3.4 The effect of other neighbourhood characteristics

To understand both the U-shaped aggregate correlation between neighbourhood inequality and average neighbourhood health status (Table 2) and the change in sign (from negative to positive) in the association after we control for median neighbourhood income (Table 4), it is instructive to consider the neighbourhood correlates of neighbourhood inequality: What sorts of neighbourhoods are characterized by high/low levels of income inequality?

We address this question by examining the correlations between the six measures of income inequality with median neighbourhood income and the proportions of people at the low and high ends of the income distribution (see Appendix 2: Alternative Inequality Measures and Their Correlates, for details). The low-income rate is defined as the proportion of families in the neighbourhood with incomes less than half of the city median income and the high-income rate is the proportion of families in the neighbourhood with income twice the city median. The median share, the mean log deviation and Gini are moderately associated with neighbourhood median income (Pearson r ranges from -0.26 to -0.36). By comparison, Theil, CV and squared CV have a very weak correlation with neighbourhood median income (r : -0.11 to 0.04).

These moderate to weak associations are a result of the fact that the association between neighbourhood inequality and neighbourhood income tends to be U-shaped: higher inequality is characteristic of both affluent and disadvantaged neighbourhoods.¹³ This pattern is apparent in the fact that neighbourhood inequality is positively associated with both the neighbourhood low-income rate *and* the neighbourhood high-income rate although the size and pattern of association differs across inequality measures. In sum, high inequality is characteristic of both advantaged neighbourhoods and disadvantaged neighbourhoods.

The conclusion that high neighbourhood inequality is in part a proxy for concentrations of *both* disadvantaged and advantaged populations is also evident in the correlations with other neighbourhood characteristics. Neighbourhood inequality is positively correlated with the percent of university educated adults (for all inequality measures) and the percentage of recent immigrants and non-whites (for three inequality measures). Strikingly, the partial correlation between inequality indices and the percentage of persons with a university degree, after controlling for neighbourhood median income, becomes even larger. When each of these variables is separately added to the previous model (as in Table 4), the percent of seniors, recent immigrants and visible minorities are not significant. The percent of lone-parent families is significantly and negatively associated with health, but its inclusion in the model has no impact on the positive association with neighbourhood income inequality.

In contrast, when the percent of people with university degrees is included (Table 5), the association is positive and significant and the positive association with neighbourhood inequality disappears. Again, this result is consistent across different inequality measures. In the model with mean log deviation as the inequality measure, the difference in average health between a neighbourhood with 5% of its adult population with university degrees and a neighbourhood with 40% of its adult population with university degrees is 0.105 (from $0.003 \times 40 - 0.003 \times 5$), which is close to the size of the coefficient for the highest neighbourhood median income category. This result implies that neighbourhood inequality reflects socio-economic conditions (well educated residents) that are not fully captured by either neighbourhood income level or individual level characteristics.

13. For example, measured with the Theil index, average neighbourhood inequality in Toronto's poorest neighbourhoods is 0.26, falling to 0.19 in middle-class neighbourhoods and rising to 0.39 in the most affluent neighbourhoods.

Table 5. Hierarchical linear models for individual self-perceived health and neighbourhood inequality controlling for individual level variables, neighbourhood median income and percent of people with university degrees

	Log deviation		Theil		Gini		Median share		CV squared	
	B	SE	B	SE	B	SE	B	SE	B	SE
Intercept	3.731 ***	0.033	3.732 ***	0.034	3.739 ***	0.036	3.766 ***	0.037	3.780 ***	0.018
Neighbourhood income inequality										
Lower-middle	0.017	0.022	0.027	0.023	0.026	0.023	0.037	0.022	-0.002	0.024
Middle	-0.001	0.026	0.006	0.026	0.017	0.026	-0.018	0.027	-0.004	0.025
Upper-middle	0.016	0.028	-0.006	0.028	-0.011	0.029	-0.026	0.030	0.009	0.027
Highest	0.011	0.030	0.023	0.031	0.006	0.033	-0.021	0.034	0.005	0.027
Neighbourhood median income										
Lower-middle	0.015	0.030	0.013	0.030	0.008	0.030	0.001	0.030	0.017	0.029
Middle	0.031	0.032	0.026	0.032	0.021	0.032	0.007	0.033	0.031	0.031
Upper-middle	0.059	0.033	0.054	0.032	0.048	0.034	0.028	0.035	0.058	0.031
Highest	0.104 **	0.036	0.099 **	0.035	0.091 *	0.037	0.068	0.039	0.102 **	0.033
% with university degree										
	0.003 **	0.001	0.003 **	0.001	0.004 **	0.001	0.004 ***	0.001	0.003 **	0.001

Note: * significant at $p < .05$, ** $p < .01$, *** $p < .001$.

Note: Individual level variables included in each model, but not presented here.

We interpret these results with caution since there are so many possible interpretations available. A contextual explanation would imply that people enjoy positive health externalities from living among more affluent and better educated neighbours irrespective of their own income and educational attainments. The causal mechanism could be “material” (better neighbourhood services, amenities, less exposure to environmental risk factors) or social (e.g., “learning effects” from being exposed to people with healthier life styles), or both and possibly by other mechanisms. Alternatively, the statistical association between health status and *neighbourhood* income and education levels may reflect unmeasured differences in individual characteristics or selection bias.

3.5 Do the effects of neighbourhood inequality and neighbourhood income vary among income groups?

Our analysis, thus far, is based on the assumption that the effects of individual and contextual socio-economic resources are purely additive. These effects, however, may well differ across specific sub-populations. While in its most general form, Wilkinson’s theory of the health-inequality linkage assumes that high inequality and relative deprivation is bad for the health status of the rich as well as the poor, it is scarcely the only one available. It is equally plausible that the contextual and individual aspects of socio-economic status interact in predicting individual health. Wheaton and Clarke (2003), for example, suggest that the model most consistent with the relative deprivation thesis is a “compound disadvantage” model that would predict that the greatest health disadvantage occurs for the personally disadvantaged with disadvantaged neighbours while the personally advantaged derive positive health benefits (compound advantage) from sharing neighbourhoods with advantaged neighbours. If such an interaction exists, our models until now are misspecified.

To test this possibility within a multi-level model framework, we allow the coefficients of individual income categories to vary across neighbourhoods and then regress these coefficients on neighbourhood income inequality and absolute income level. This procedure is equivalent to incorporating interaction terms for individual incomes and neighbourhood variables in the usual OLS models. The results (data not presented, but available upon request) show that the effects of neighbourhood income inequality and neighbourhood absolute income level do not vary across income groups. The poor derive the same health benefit from living with more affluent neighbours as other income groups.

3.6 The effects of city-level income inequality and economic segregation

Finally, we add a third level model to the two-level models to test whether city-level measures of inequality, economic segregation, or median city income are associated with individual health after controlling for neighbourhood and individual-level characteristics. In the third level model, we first include a city-level inequality measure. We did not find a significant effect (data not presented, but available upon request). Then, instead of using the overall city-level inequality measure, we include in the third level model its two components—economic segregation and within-neighbourhood inequality. Based on the sign of the coefficients for these variables, economic segregation is negatively associated with self-perceived health, while within-neighbourhood inequality is positively associated with self-perceived health. However, only the within-neighbourhood inequality coefficient for the mean log deviation was marginally significant ($p=.04$). Similarly the coefficient for median city income was insignificant.

4. Discussion and Conclusion

There are many *good* reasons why we might expect no (direct) association between inequality and health at the neighbourhood level. The strongest U.S. evidence for a (direct) relationship between inequality and health, net of individual differences in socio-economic resources, is at the state and metropolitan level and the effects are mainly experienced by the poorest sections of the population, pointing toward policies toward the poor that are correlated with income inequality as a likely culprit (Wagstaff and Van Doorslaer, 2000: 564). If politics and policies are the mediator, (direct) inequality effects are unlikely to show up at lower levels of aggregation than those at which the policies are set. In more centralized polities, one would expect no, or only weak, “effects” to appear among sub-national units and such effects are least likely to appear at the level of neighbourhoods. The conclusion, of course, is not that income inequality and redistribution are irrelevant for the health of populations. The fact that the effect of income inequality on health is “indirect,” mediated by differences in individual living standards, does not diminish its salience.

Our neighbourhood level results are decidedly more salient for the RIH variant of Wilkinson’s thesis. Do the less affluent experience poorer health as a result of greater stress or competition for resources when they share environments with the more affluent? Or, as Wilson claims, do they generally benefit from such environments? If most neighbourhoods are highly segregated by income, of course, the debate is moot but this is generally not the case. Our cross-sectional estimates of the implications of co-residence with more affluent and better educated individuals are likely inflated by selection bias and unmeasured individual differences but attenuated as a

result of failing to control for the cumulative effects of past neighbourhood contexts (Wheaton and Clarke, 2003). Within these limitations, however, we cannot reject Wilson's counter-hypothesis that, on average, the less affluent who share neighbourhoods with better educated and more affluent neighbours experience better health outcomes. These results suggest that populations may benefit from housing and zoning strategies that encourage economically "mixed" neighbourhoods and discourage high levels of residential segregation by income level.

Appendix Table 1: Family Income Categories in the National Population Health Survey, 1996/97

Income status	Income	Household size
Very Low	< \$10,000	1 to 4 persons
	< \$15,000	5 or more
Low	\$10,000-14,999	1 or 2
	\$10,000-19,999	3 or 4
	\$15,000-29,999	5 or more
Lower-Middle	\$15,000-29,999	1 or 2
	\$20,000-39,999	3 or 4
	\$30,000-59,999	5 or more
Upper-Middle	\$30,000-59,999	1 or 2
	\$40,000-79,999	3 or 4
	\$60,000-79,999	5 or more
High	\$60,000 or more	1 or 2
	\$80,000 or more	3 or more
Income not reported	Note stated	Not applicable

Appendix 1. Are Neighbourhoods More Economically Homogeneous Than Cities?

A standard conclusion that has found widespread consensus in the literature is that the strength of the relationship between inequality and health outcomes is influenced by the geographic scales being compared (Mellor and Milyo, 1999; Soobader and LeClere, 1999; Blakely, Lochner and Kawachi, 2002). The size of the correlation between income inequality and mortality, for example, declines as one shifts the units of aggregation from countries to states to cities or from larger to smaller counties. Wilkinson (1997), for whom the relationship between inequality and health outcomes is psycho-social in origin, argues that as we move from higher to lower units of aggregation the relevant social comparisons that generate anxiety, stress and a sense of relative deprivation are weakened since the salient heterogeneity is mostly between rather than within smaller geographic units. Within small, relatively homogeneous, neighbourhoods, he suggests, the relevant comparison is based on how affluent the neighbourhood is relative to other neighbourhoods in the wider community. As a result, *average* income in small geographic units such as neighbourhoods matter but the neighbourhood *distribution* of income is less salient.

While we do not question the claim that the size of the correlation between inequality and health is influenced by the size of the geographic units, we do question the usual explanation for this result. The premise that income inequality is always lower at lower levels of aggregation (e.g., neighbourhoods) than at higher levels of aggregation (cities), however, is incorrect.

Appendix Table 1.1 Descriptive statistics of income inequality indices at the census metropolitan area (CMA) and neighbourhood level

	Mean	Standard deviation	Minimum	Maximum
City level (n=25)				
Mean log deviation (I_0)	0.32	0.03	0.27	0.37
Theil (I_1)	0.25	0.03	0.20	0.31
Gini (Gi)	0.37	0.02	0.33	0.40
Median share (MS)	0.75	0.01	0.73	0.77
Coefficient of variation (CV)	0.86	0.13	0.68	1.13
Squared CV (I_2)	0.75	0.24	0.46	1.28
Census tract level (n=3,044)				
Mean log deviation (I_0)	0.30	0.10	0.07	1.01
Theil (I_1)	0.22	0.09	0.05	1.47
Gini (Gi)	0.35	0.05	0.18	0.69
Median share (MS)	0.74	0.03	0.62	0.89
Coefficient of variation (CV)	0.73	0.26	0.33	4.57
Squared CV (I_2)	0.60	0.76	0.11	20.85

Data Source: calculated from the 1996 Canada Census 20% sample micro-data.

As shown in Appendix Table 1.1, all inequality measures at the city level have larger *average* values than their corresponding average values at the level of the neighbourhood (census tract). This result, however, does not justify the claim that incomes are always more homogenous within neighbourhoods than within cities. Although, on *average*, income inequality is lower in the sub-areas, group averages for higher levels of aggregation (e.g., for a city) are the product of sub-areas with both lower and higher levels of inequality than for the city as a whole unless economic segregation between areas is high and this is generally not the case. The between-neighbourhood component accounts for a very small portion of the city-level inequality.

Appendix Table 1.2 CMA-level income inequality indices and their between- and within-neighbourhood components

CMAs	Mean log deviation (I_0)			Theil (I_1)			Squared CV (I_2)		
	Total	Between As % of Total	Within	Total	Between As % of Total	Within	Total	Between As % of Total	Within
St. Johns	0.32	7%	93%	0.25	10%	90%	0.62	8%	92%
Halifax	0.31	7%	93%	0.23	9%	91%	0.58	7%	93%
Saint John	0.33	11%	89%	0.26	14%	86%	1.14	7%	93%
Chicoutimi	0.35	4%	96%	0.23	6%	94%	0.48	5%	95%
Quebec	0.32	10%	90%	0.23	13%	87%	0.54	12%	88%
Sherbrooke	0.34	9%	91%	0.25	11%	89%	0.65	8%	92%
Trois-Rivières	0.35	8%	92%	0.25	10%	90%	0.59	9%	91%
Montreal	0.37	13%	87%	0.29	17%	83%	1.28	9%	91%
Ottawa–Hull	0.32	11%	89%	0.24	14%	86%	0.74	10%	90%
Oshawa	0.27	7%	93%	0.20	9%	91%	0.54	6%	94%
Toronto	0.37	14%	86%	0.31	17%	83%	1.13	11%	89%
Hamilton	0.30	12%	88%	0.25	14%	86%	0.96	8%	92%
St Catharines	0.28	7%	93%	0.22	9%	91%	0.56	7%	93%
Kitchener	0.28	8%	92%	0.24	9%	91%	0.91	5%	95%
London	0.31	10%	90%	0.24	13%	87%	0.71	9%	91%
Windsor	0.32	9%	91%	0.25	12%	88%	0.74	8%	92%
Sudbury	0.30	8%	92%	0.22	10%	90%	0.50	9%	91%
Thunder Bay	0.27	6%	94%	0.20	8%	92%	0.46	7%	93%
Winnipeg	0.31	14%	86%	0.24	18%	82%	0.73	13%	87%
Regina	0.30	13%	87%	0.23	17%	83%	0.60	13%	87%
Saskatoon	0.33	8%	92%	0.25	10%	90%	0.69	7%	93%
Calgary	0.33	11%	89%	0.29	14%	86%	1.03	8%	92%
Edmonton	0.32	10%	90%	0.25	13%	87%	1.00	7%	93%
Vancouver	0.36	9%	91%	0.29	12%	88%	1.05	7%	93%
Victoria	0.30	7%	93%	0.23	10%	90%	0.63	8%	92%
Average	0.32	9%	91%	0.25	12%	88%	0.75	8%	92%

Data Source: calculated from the 1996 Canada Census 20% sample micro-data

As shown in Appendix Table 1.2, in Canada's 25 metropolitan areas, economic segregation (between-neighbourhood component as a percentage of the city-level total income inequality) on average accounts for 8% (I_2) to 12% (I_0) of city-level income inequality. This is a standard result for U.S. (Jargowsky, 1996; Lobmayer and Wilkinson, 2002) as well for Canadian cities.

City-level income inequality is primarily determined by the within-neighbourhood component and most neighbourhoods are economically heterogeneous. As a result, city-level income inequality is almost perfectly correlated with its within-neighbourhood component (Appendix Table 1.3). This high correlation suggests that any health effect of city-level inequality basically reflects the impact of its *average* within-neighbourhood component *and vice versa*. The correlation between city-level inequality and the between-neighbourhood component, in contrast, is decidedly more modest and the correlation with economic segregation is lower still implying that city-level economic segregation may have additional effects beyond those that may be associated with overall inequality.

Appendix Table 1.3. Pearson R correlations between CMA-level income inequality indexes and their components

	Mean log deviation (I_0)			Theil (I_1)			Squared CV (I_2)		
	Total	Within	Between	Total	Within	Between	Total	Within	Between
Within neighbourhoods	0.94***			0.94**			0.99***		
Between neighbourhoods	0.54**	0.23		0.79***	0.53**		0.81***	0.76***	
Economic segregation	0.31	-0.01	0.97***	0.60**	0.29	0.96***	0.01	-0.05	0.58**

Data Source: calculated from the 1996 Canada Census 20% sample micro-data

Note: n=25. ** significant at $p < .01$; *** $p < .001$.

More importantly, neighbourhood income inequality has a much larger variation than the city-level counterpart. The standard deviations of income inequality indexes at the census tract level are about three times larger than the corresponding indexes at the city level (see Appendix Table 1.1, above). Thus, even though on average, neighbourhood-level inequality is somewhat lower than city-level inequality, some neighbourhoods have much higher and others much lower levels of inequality than the city as a whole.

In our data, about 28% of neighbourhoods have higher values on the mean log deviation than their corresponding city values; about 21% of neighbourhoods have higher values on the Theil index than their corresponding city values; and about 12% of neighbourhoods have higher value on the squared coefficient of variation than their corresponding city values. Quite simply, the standard conclusion that lower average inequality at lower levels of aggregation is always a result of a process of economic segregation that “generates homogenous neighbourhoods that vary greatly from one another” (Soobader and LeClere, 1999: 738) is incorrect.

Appendix 2: Alternative Inequality Measures and Their Correlates

Students of income inequality typically deploy a family of inequality measures in their investigations since different measures capture different features of the income distribution. Atkinson (1970), for example, finds that developing nations tend to show relatively less inequality in measures that are more sensitive in the lower range of incomes since developing nations tend to have a large, homogeneous low-income population together with great inequality among the rich. The opposite is true for developed nations.

Here we examine how similar or different empirically the selected six measures are in measuring neighbourhood income inequality (Appendix Table 2.1). Correlations among measures tend to be high between indices with similar sensitivity to a particular range of incomes, in particular, between Gini and median share and between CV and squared CV but tend to be low between indices that are sensitive to different ranges of incomes. The exception is the Theil index that is highly correlated with both the Gini index and CV.

Appendix Table 2.1. Pearson correlations among income inequality indices at the neighbourhood level

	Mean log deviation (I_0)	Theil (I_1)	Gini (G_i)	Median share (MS)	Coefficient of variation (CV)
Mean log deviation (I_0)					
Theil (I_1)	0.83				
Gini (G_i)	0.90	0.92			
Median share (MS)	0.90	0.83	0.98		
Coefficient of variation (CV)	0.64	0.93	0.76	0.64	
Squared CV (I_2)	0.48	0.81	0.55	0.44	0.92

Data Source: calculated from the 1996 Canada Census 20% sample micro-data.

Note: all coefficients are significant at $p < .001$

The differences among the six measures can be better illustrated from their correlations with neighbourhood income level, the proportion of people at the low and high end of the income distribution, as well as the income dispersion at both ends of the income distribution. Here we use neighbourhood median income to measure neighbourhood income level. We use low- and high-income rates to measure proportions of people at the low and high end of the income distribution. The low-income rate is defined as the proportion of families in the neighbourhood with incomes less than half of the city median income and the high-income rate is the proportion of families in the neighbourhood with income twice the city median. A standard assumption in the inequality-health debate is that inequality is, in part, a proxy for differences in absolute living standards and/or poverty levels in the population. At the neighbourhood level, this is true for some inequality measures but not for all.

As shown in Appendix Table 2.2, median share, the mean log deviation and Gini are moderately associated with neighbourhood median income. By comparison, Theil, CV and squared CV have

a very weak correlation with neighbourhood median income. Median share, Gini, and mean log deviation are correlated more strongly with the neighbourhood low-income rate than with the high-income rate. The correlations between the Theil index and both the low- and high-income rates are virtually identical. By comparison, CV and squared CV are correlated more strongly with the high-income rate than with the low-income rate. However, their correlations with the dispersions at both ends are much weaker for CV and squared CV than for the other four indexes.

Appendix Table 2.2 Pearson correlations between neighbourhood income inequality indices and other neighbourhood income characteristics (n=3,045)

	Mean log deviation (I ₀)	Theil (I ₁)	Gini (Gi)	Median share (MS)	Coefficient of variation (CV)	Squared CV (I ₂)
Neighbourhood median income (log)	-0.34	-0.11	-0.26	-0.36	0.04*	0.06
Low-income rate	0.55	0.31	0.49	0.58	0.12	0.04
High-income rate	0.07	0.28	0.22	0.12	0.33	0.23

Data Source: calculated from the 1996 Canada Census 20% sample micro-data.

Note: * significant at p<.05, all other coefficients are significant at p<.001

The reason for these patterns is shown in Appendix Table 2.3 where we present the average low- and high-income rates by level of neighbourhood inequality. For all measures, the association between the neighbourhood low-income rate and neighbourhood inequality is monotonic: the percent of low-income individuals rises as one moves from low inequality to high inequality neighbourhoods. For four of the five measures of inequality, however, the association between neighbourhood inequality and the high-income rate is U-shaped: concentrations of very affluent families live in *both* low inequality and high inequality neighbourhoods. In sum, high inequality neighbourhoods are characterized by concentrations of both high- and low-income families, a result that has important implications for our subsequent analyses.

Appendix Table 2.3. Low-income and high income rates by neighbourhood inequality group

Neighbourhood income inequality	% low income		% high income		% low income		% high income		% low income		% high income	
	Log deviation		Theil		Gini		Median share		CV			
Lowest	1.9	22.0	2.0	19.8	2.0	20.2	1.9	20.8	2.0	18.3		
Lower-middle	2.9	19.3	3.0	18.3	3.2	18.5	3.0	19.7	3.8	17.4		
Middle	4.8	16.3	5.2	16.2	4.4	16.9	4.3	16.5	5.6	15.5		
Upper-middle	5.5	16.7	5.8	15.6	6.4	15.6	5.9	15.7	5.7	16.1		
Highest	8.6	16.9	7.9	22.0	8.0	20.6	8.7	18.7	6.5	24.8		

Data sources: the 1996 census 20% sample micro-data

The conclusion that high neighbourhood inequality is in part a proxy for concentrations of *both* disadvantaged and advantaged populations is also evident in the correlations with other neighbourhood characteristics (Appendix Table 2.4). Neighbourhood inequality is positively correlated with both the percentage of university educated (for all inequality measures) and the percentage of recent immigrants and non-whites (for three inequality measures). After controlling for neighbourhood median income, the partial correlation between inequality indices and the percentage of persons with a university degree becomes even larger. This partial correlation is rather instrumental in explaining the aggregate association between health and neighbourhood inequality. The more general lesson is that simple assumptions concerning the likely correlates of any measure of neighbourhood inequality (e.g., high inequality is correlated with higher levels of poverty) are made at one's peril.

Appendix Table 2.4 Pearson correlations between neighbourhood income inequality indices and other neighbourhood characteristics (n=3,044)

	Mean log deviation (I ₀)	Theil (I ₁)	Gini (G _i)	Median share (MS)	Coefficient of variation (CV)	Squared CV (I ₂)	
<u>Raw correlation</u>							
Percent of university educated	0.27	0.41	0.41	0.35	0.38	0.26	
Percent of seniors	0.14	0.21	0.28	0.29	0.19	0.10	
Percent of lone-parent families	0.13	-0.04 *	0.05 *	0.13	-0.13	-0.10	
Percent of recent immigrants	0.36	0.22	0.30	0.34	0.12	0.07	
Percent of visible minorities	0.31	0.16	0.23	0.26	0.07	0.03	ns
<u>Controlling for median income</u>							
Percent of university educated	0.55	0.54	0.66	0.66	0.42	0.27	
Percent of seniors	0.10	0.20	0.25	0.25	0.19	0.12	
Percent of lone-parent families	-0.06	-0.12	-0.11	-0.08	-0.13	-0.08	
Percent of recent immigrants	0.31	0.20	0.25	0.28	0.13	0.09	
Percent of visible minorities	0.26	0.14	0.18	0.21	0.08	0.04	*

Data Source: calculated from the 1996 Canada Census 20% sample micro-data.

Note: ns- not significant, * significant at p<.05, all other coefficients are significant at p<.001

References

- Allison, P.D. 1978. "Measures of inequality." *American Sociological Review*, 43(6): 865-880.
- Atkinson, A. 1970. "On the measurement of inequality." *Journal of Economic Theory*, 2: 244-263.
- Blakely, T., K. Lochner and I. Kawachi. 2002. Metropolitan area income inequality and self-rated health—a multi-level study. *Social Science & Medicine* 54(1): 65-77.
- Brooks-Gunn, J., G. Duncan and J.L. Aber. (Eds.). 1997. *Neighborhood Poverty: Context and Consequences for Children*. New York: Russell Sage Foundation.
- Bryk, A.S. and A.W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, California: Sage.
- Cowell, F. A. 1995. *Measuring Inequality*. New York: Prentice Hall/Harvester Wheatsheaf.
- Daly, M., G. Duncan, G. Kaplan and J. Lynch. 1998. "Macro-to-micro links in the relation between income inequality and mortality." *The Milbank Quarterly*, 76(3): 315-339.
- Daniels, N., B.P. Kennedy and I. Kawachi. 1999. "Why justice is good for our health: the social determinants of health inequalities." *Daedalus*, 128(4): 215-251.
- Deaton, A. 2003. "Health, inequality and economic development." *Journal of Economic Literature* 41(1): 113-158.
- Ferraro, K.F., M.M. Farmer and J.A. Wybraniec. 1997. Health trajectories: long-term dynamics among black and white adults. *Journal of Health and Social Behavior*, 38(1): 38-54.
- Gravelle, H. 1998. "How much of the relation between population mortality and unequal distribution of income is a statistical artefact?" *British Medical Journal*, 316: 382-385.
- Idler, E.L. and Y. Benyamini. 1997. "Self-rated health and mortality: a review of twenty-seven community studies." *Journal of Health and Social Behavior*, 38(1): 21-37.
- Jargowsky, P.A. 1996. "Take the Money and Run: Economic Segregation in U.S. Metropolitan Areas." *American Sociological Review*, 61(6): 984-998.
- . 1997. *Poverty and Place: Ghettos, Barrios, and the American city*. New York: Russell Sage Foundation.
- Jencks, C. and S. Mayer. 1990. The social consequences of growing up in a poor neighborhood. In L.E. Lynn Jr. and M.H. McGeary (eds.), *Inner-City Poverty in America* (pp. 111-185). Washington, D.C.: National Academy Press.

- Jenkins, S.P. 1995. "Accounting for inequality trends: decomposition analysis for the U.K., 1971-86." *Economica*, 62(245): 29-63.
- Kaplan, G.A., E.R. Pamuk, J.W. Lynch, R.D. Cohen, and J.L. Balfour. 1996. "Inequality in income and mortality in the United States: Analysis of mortality and potential pathways." *British Medical Journal*, 312(7037): 999-1003.
- Kawachi, I. and L.F. Berkman. 2003. *Neighborhoods and Health*. New York: Oxford University Press.
- Kawachi, I. and B.P. Kennedy. 1997. The relationship of income inequality to mortality: Does the choice of indicator matter? *Social Science & Medicine*, 45(7):1121-1127.
- . 1999. "Income inequality and health: pathways and mechanisms." *Health Services Research*, 34(1) Part II: 215-227.
- Lobmayer, P. and R.G. Wilkinson. 2002. "Inequality, residential segregation by income, and mortality in US cities." *Journal of Epidemiological and Community Health*, 56(3): 183-187.
- Mellor, J. and J. Milyo. 2001. "Reexamining the evidence of an ecological association between income inequality and health." *Journal of Health Politics, Policy and Law*, 26(3):487-522.
- Mellor, J. and J. Milyo. 1999. *Income Inequality and Health Status in the United States: Evidence from the Current Population Survey*. Princeton NJ: Robert Wood Johnson Foundation.
- Myles, J., G. Picot and W. Pyper. 2000. *Neighbourhood inequality in Canadian cities*. Analytical Studies Branch Research Paper Series, No. 160. Catalogue No. 11F0019MIE. Ottawa: Statistics Canada
- Raudenbush, S.W., A. Bryk, Y.F. Cheong and R. Congdon. 2000. *HLMTM 5: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: Scientific Software International, Inc.
- Ross, N.A., M.C. Wolfson, J.R. Dunn, J.M. Berthelot, G.A. Kaplan and J.W. Lynch. 2000. Relation between income inequality and mortality in Canada and in the United States: cross sectional assessment using census data and vital statistics. *British Medical Journal*, 320(7239): 898-902.
- Ross, N., S. Tremblay and K. Graham. 2003. Neighbourhood influences on health in Montreal, Canada. Paper under review for publication.
- Sampson, R., J. Morenoff and T. Gannon-Rowley. 2002. Assessing 'neighbourhood effects': social processes and new directions in research. *Annual Review of Sociology*, 28: 443-478.

- Soobader, M., and F. LeClere. 1999. "Aggregation and the measurement of income inequality: effects on morbidity." *Social Science & Medicine*, 48(6): 733-744.
- Subramanian, S.V., I. Kawachi and B.P. Kennedy. 2001. "Does the state you live in make a difference? Multilevel analysis of self-rated health in the US." *Social Science & Medicine*, 53(1): 9-19.
- Wagstaff, A. and E. van Doorslaer. 2000. "Income inequality and health: what does the literature tell us?" *Annu Rev Public Health*, 21: 543-67.
- Weich, S., G. Lewis and S.P. Jenkins. 2002. "Income inequality and self rated health in Britain." *Journal of Epidemiology and Community Health*, 56(6): 436-441.
- Wheaton, B. and P. Clarke. 2003. "Space meets time: combining temporal and contextual influences." *American Sociological Review*, 68(5): 680-707.
- Wilkinson, R.G. 1992. "National mortality rates: the impact of inequality?" *American Journal of Public Health*, 82(8): 1082-1084.
- . 1996. *Unhealthy Societies: the Afflictions of Inequality*. London: Routledge.
- . 1997. "Commentary: income inequality summarises the health burden of individual relative deprivation." *British Medical Journal*, 314: 1727-28.
- . 1998. "Low relative income affects mortality." *British Medical Journal*, 316: 1611.
- . 1999. "Income inequality, social cohesion, and health: clarifying the theory--a reply to Muntaner and Lynch." *International Journal of Health Services*, 29(3): 525-543.
- Wilson, W.J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Wolfson, M. 1997. "Divergent inequalities: theory and empirical results." *Review of Income and Wealth*, 43: 401-421.
- Wolfson, M., G. Kaplan, J. Lynch, N. Ross and E. Bachlund. H. Gravelle and R.G. Wilkinson. 1999. "Relation between income inequality and mortality: empirical demonstration." *British Medical Journal*, 319: 953-7.
- Wolfson, M., G. Rowe, J.F. Gentlemen and M. Tomiak 1993. "Career earnings and death: a longitudinal analysis of older Canadian men." *J Gerontol Soc Sci.*, 48: S167-S179.