Socioeconomic Patterning of Self-Rated Health Trajectories in Canada: A Mixture Latent Markov Model

Maria Hyunji Koh

Department of Epidemiology, Biostatistics and Occupational Health
McGill University, Montreal, Canada
April 2012

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

AIC	Akaike information criterion
BHPS	British Household Panel Survey
BIC	Bayesian information criterion
CAD	Canadian dollar
CATI	Computer assisted telephone interview
CI	Confidence interval
EM	Expectation maximization
FIML	Full information maximum likelihood
GSOEP	German Socio-Economic Panel
LM	Latent Markov
M	Manifest Markov
MI	Multiple imputation
ML	Maximum likelihood
MLM	Mixed latent Markov
NPHS	National Population Health Survey
RRR	Relative risk ratio
SE	Standard error
SEP	Socioeconomic position
SLID	Survey of Labour and Income Dynamics
SRH	Self-rated health
SSABIC	Sample size adjusted Bayesian information criterion

Abstract

This thesis investigates the association between socioeconomic position and self-rated health trajectories among Canadians. Data come from the Survey of Labour and Income Dynamics (SLID), Panel 4 (year 2002 to 2008), conducted by Statistics Canada. These longitudinal data are analyzed using mixed latent Markov model which allows for modeling multiple trajectories of health. Goodness of fit tests showed three trajectories (good health, poor health, and fluctuating health) to provide the best fit to the data. The results show that more than three quarters of Canadians were in the constant good health trajectory whereas 13.95% and 7.99% of Canadians were respectively in the persistent ill health trajectory and fluctuating health trajectory. The relative risk ratios indicate that increasing income and education are independently associated with a greater likelihood of belonging to the persistent good health trajectory rather than the persistent poor health trajectory. Both associations accounted for possible confounders including gender, age, marital status, immigrant status and visible minority status. These results suggest that a socioeconomic gradient exists in the likelihood of belonging to given health trajectories. In addition, the use of mixed latent Markov model is robust in accounting for certain issues inherent to longitudinal analysis. Notably, the Markov chain models the dependency between repeated measurements within the same individual; it allows for the modeling of the latent variables estimate measurement error; the heterogeneity of the population is accounted by finite mixture modeling; and lastly, missing data are dealt with using full information maximum likelihood.

Abrégé

Cette thèse étudie l'association entre la position socioéconomique et les trajectoires de santé perçue parmi la population canadienne. Les données proviennent de l'Enquête sur la dynamique du travail et du revenu (EDTR) de Statistique Canada. Ces données longitudinales couvrant la période 2002-2008 sont analysées à l'aide de chaines de Markov avec variables latentes, qui permettent de modéliser les trajectoires de santé perçue des individus. Les résultats indiquent que plus de trois Canadiens sur quatre appartiennent à la trajectoire de bonne santé stable, alors que 13.95% et 7.99% des Canadiens se trouvent respectivement dans les trajectoires de mauvaise santé persistante et de santé instable. Les ratios de risque indiquent qu'il existe un gradient inverse entre le niveau de revenu et le degré d'instruction et le risque d'appartenir à la trajectoire de mauvaise santé plutôt qu'à celle de bonne santé. Cette association persiste suite à l'ajout des caractéristiques sociodémographiques telles le sexe, l'âge, et les statuts matrimonial, d'immigrant et de minorité visible. Ces résultats établissent la présence d'un gradient socioéconomique dans les trajectoires de santé, démonstration qui n'avait jusqu'à maintenant pas été faite au Canada. Qui plus est, les méthodes utilisées s'avèrent robustes pour l'analyse des données longitudinales et des problèmes qui y sont souvent associés. En effet, les chaines de Markov tiennent explicitement compte de la corrélation entre les réponses fournies à travers le temps par un même individu; l'hétérogénéité dans les trajectoires est prise en compte par un modèle pour un mélange fini de distributions (*finite mixture model*); les erreurs de mesure sont incorporées dans l'estimation des variables latentes; et enfin, les données manquantes sont estimées à l'aide de l'algorithme du maximum de vraisemblance à information complète (*full information maximum likelihood)*.

Acknowledgements

I would like to thank those who gave me an opportunity and to those who have enabled me to grasp those opportunities. To my supervisor, Dr. Amélie Quesnel-Vallée, who took a leap of faith on a student wanting to do research in a field where neither the outcome nor predictors are clearly defined. To Dr. Emilie Renahy, who was there through the trials and tribulations of Mplus and variable definitions. To Danielle Forest, who taught me bootstrapping and weighting - the express version. To the support staff of Mplus, to whom I owe my input codes terminating normally. To my committee member, Dr. Sam Harper, for his thoughtful and constructive comments. To the administrative staff at Purvis who went through the false-starts with my thesis. Lastly, to my family and friends, who kept me grounded and provided unwavering support.

Funding:

- Canadian Institute for Health Research (CIHR) Frederic Banting and Charles Best Canada Graduate Scholarship Award – Master's
- Dr. Amélie Quesnel-Vallée (CIHR operating grant MOP-77800)
- Canadian Institute for Health Research (CIHR) Interdisciplinary Capacity Enhanced Team Grant, Health Disparities: Inter-generational Dynamics and Empirical Assessment Strategies Studentship (HD: IDEAS)
- Population Change and Lifecourse Strategic Knowledge Cluster, Western University
- Quebec Inter-university Centre for Social Statistics (QICSS) matching grant – Master's

Preface

This analysis is based on Statistics Canada's Survey of Labour and Income Dynamics Public Use Microdata, which contains anonymized data collected in the Survey of Labour and Income Dynamics. All computations on the microdata were prepared by Maria Koh. The responsibility for the use and interpretation of these data is entirely that of the author. If the cell counts were below 15, the data cannot be released from Statistics Canada and the information is presented as 'data suppressed'.

Ethics approval was obtained from McGill University's Faculty of Medicine Institutional Review Board on December 13, 2010 (Research project A11-E78-10B).

Chapter 1

INTRODUCTION

- *Why study socioeconomic position and health?*
- *What is the purpose of this study?*

The observation of a social gradient in health outcomes including morbidity, life expectancy, low birth weight, and self-rated health (SRH) has been consistently reported across and within nations over time (Adler & Ostrove, 1999; Feinstien, 1993). The concept of the social gradient in health describes the phenomena where individuals in each subsequent higher step in the social hierarchy have better health (Adler et al., 1994; Marmot, Shipley, & Rose, 1984; Marmot et al., 1991). The gradient implies that socioeconomic inequalities in health cannot solely be explained by material and social conditions associated with severe deprivation. While this association between socioeconomic position (SEP) and health is well-established, establishing causal inference remains a challenge due to limitations frequently plaguing observational research namely measurement error, reverse causation and unobserved heterogeneity (Kawachi, Adler, & Dow, 2010). This thesis constitutes a step forward in this literature by relying on longitudinal data allowing for the temporal ordering of exposure and outcome, using a gold-standard measurement of income taken from administrative tax records, and relying on sophisticated methods incorporating the explicit modeling of measurement error.

In 2002, a Canadian national conference entitled 'Social determinants of health across the life-span: A current accounting and policy implications', highlighted eleven social determinants of health as being most pertinent to Canadians (Raphael, 2004); education, employment status, and income were part of this list and will be examined in this thesis. The majority of the studies establishing this association in Canada have analyzed the data in a cross-sectional manner (Adler & Stewart, 2010; Gunasekara, Carter, & Blakely, 2011; Phipps, 2003). In order to formulate social policies to minimize health inequity, it is central to examine longitudinal evidence concerning the interrelated, dynamic nature of both SEP and health fluctuations over time. This thesis will examine the role of socioeconomic position in determining self-rated health trajectories using longitudinal data from Statistics Canada's Survey of Labour and Income Dynamics (SLID). It will answer the following questions:

(1) How does self-rated health change over time? What proportion of Canadians experience persistent good health, persistent poor health or fluctuate between these two states?
(2) What is the impact of income, education, and employment status on determining which self-rated health trajectory an individual belongs to?

Chapter 2 provides an overview of the research socioeconomic position and health. It pays particular attention to the literature that examines this relationship in its longitudinal nature and theoretical frameworks that propose mediators of this relationship. This section concludes with a discussion on the validity and reliability of the outcome, SRH. Chapter 3 discusses the dataset used for analysis, variable selection and definitions, and describes model building steps and criteria used to arrive at the final model. Chapter 4 contains the descriptive and analytical results from the various extensions of the Markov model. Chapter 5 discusses the implications of the results and compares our findings with previous research. Finally, Chapter 6 examines limitations associated with observational research and research methodology and summarizes the arguments exposed in previous chapters.

Chapter 2

LITERATURE REVIEW

2.1 Socioeconomic Position (SEP)

What is socioeconomic position?

SEP is a measure quantifying a person's social and economic well-being and encompasses two domains: *material and social resources* such as housing and income, and *prestige* which pertains to an individual's rank in the social hierarchy (Bartley, 2004; Krieger, Williams, & Moss, 1997). There have been attempts to directly quantify prestige by the use of scales and rankings (e.g., MacArthur scale of subjective social status), but as a person's prestige is often intertwined with material and social resources, indicators of the two domains have a wide overlap. A plethora of measures have been used to indicate SEP including, at the individual level, absolute income, change in income, occupation, education, employment status, and housing conditions, and at the aggregate level, income inequality and work organization (Galobardes, Shaw, Lawlor, Lynch, & Davey Smith, 2006). There is no perfect indicator of SEP as cohort and period effects may affect the meaning and social impact of an indicator. For example, attaining high school graduation in the 1930s compared to the 1990s would signal different levels of social status attainment. Moreover, the various indicators themselves are often intertwined in complex ways (Figure 1) and their effects are compounded throughout the life-course. For instance, poor housing condition in childhood lead to low achievement in school which in turn affects job prospects and income earned. As such, Galobardes et al., (2006) recommend that the choice of indicators should depend on the research question and the theoretical framework surrounding the outcome of interest. In general, multiple indicators of SEP are preferred over single indicators as multiple indicators account for different components of the relationship (Cairney & Arnold, 1998; Galobardes et al., 2006).

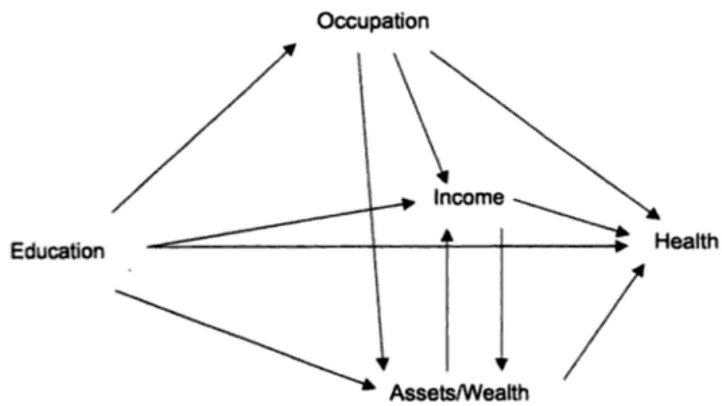


Figure 1: Complex interplay between SEP indicators
Source: House & Williams (2000) in Promoting health: intervention strategies from social and behavioural research. Editors Smedley, B. & Syme, S.

What are the theoretical pathways between SEP and health?

There are four main theoretical pathways between SEP and health (Bartley, 2004; Kawachi et al., 2010; Raphael, 2004): the behavioral and cultural explanation, the material pathway, the neo-materialist approach and the social comparison approach (also known as psychosocial comparison). The behavioural and cultural explanation posits that the gradient exists partially due to differences in health behavior influenced by a given social group's culture and norms. Concrete examples of how different economic and social conditions pattern health behavior are evident in Canada: smoking, physical inactivity, and alcohol binging all tend to follow an inverse social gradient (Canadian Institute for Health Information, 2008). Yet, the social gradient in health still persists after controlling for these health-related behaviors, indicating the existence of other factors that are at play in mediating the association between SEP and health.

The material pathway is perhaps the most intuitive: income provides differential access to material goods such as adequate housing and nutritious food. More diffusely, the neo-material approach suggests that there also exist societal factors and public infrastructure, such as health care and social services, which may shape the material conditions of individuals beyond their individual income.

Finally, the social comparison approach hypothesizes that health is a product of a person's perception of their standing in a social hierarchy. For instance, people who perceive themselves as being low in the social hierarchy may suffer from stress, which leads them to find solace and immediate gratification in deleterious health behaviours such as alcohol and tobacco use. Although this coping phenomenon may seem less intuitive, Khaneman & Tversky (1974) have shown that human behavior is prone to biases and use of heuristics can lead to behavior that may appear irrational according to medical rationality (e.g. smoking in the face of known health risks). Additionally, upon qualitative examination, the importance of lay knowledge and its contextualized rationality must also be recognized (Popay et al., 2003). Thus the social comparison approach relies on the fact that a person's perception of his/her social standing can have a powerful impact on their behavior, and in turn, on their health.

A review of Canadian literature on how income is conceptualized when examining its effect on health found that the majority of the studies took the materialist approach followed by neo-materialist and social comparison (Macdonald et al., 2009). In practical terms, however, it is impossible to tease apart the difference between the material pathway and the social comparison approach as a change in material pathway inevitably affects the social comparison approach (Kawachi et al., 2010). That is, a raise in individual's income would improve health as it provides greater access to goods and services (material pathway), but the very increase in income potentially also signals an increase in the person's prestige (social comparison approach). A similar analogy can be used for education and employment: a higher achievement in education standing concedes both materialistic goods and prestige. Therefore, this thesis examines an individual's absolute income, education and employment status on health, all of which are conceptualized to directly affect the material pathway and, indirectly, the social comparison approach.

What do we know about the causal link between SEP and health through longitudinal and biological studies?

Studies of physiological pathways in humans lend biological plausibility to SEP and health. As the brain serves as the central coordinating organ, many biological studies focus on neurological pathways. Biological mediators such as glucocorticoids in blood and adrenal steroid hormones from the adrenal gland are but two examples used to map the pathway linking stress and health in the context of SEP. A process coined allostatic load describes the physical consequences from chronic deregulation of stress mechanisms that are normally used as a short-term adaptation strategy. A person of lower SES may be more likely to be exposed to stressful life events at individual, family and residential level (McLeod & Kessler, 1990) which may alter an individual's coping profile (McEwen, 2007). Prolonged stress has been shown to adversely affect health in many ways including cardiovascular function and reproduction (McEwen & Gianaros, 2010). Observational studies have shown increases in secretion of cortisol among men who report low levels of control at work (Kunz-Ebrecht, Kirschbaum, & Steptoe, 2004).

Although longitudinal studies are less robust than randomized control trials in causal inference because of such limitations as unobserved confounders and limited generalizability due to sample selection, they provide a valuable opportunity to treat SEP and health as dynamic variables that change over time. Yet, the dearth of studies that have examined changes in SRH over time has previously been noted in the literature (Bailis, Segall, & Chipperfield, 2003; Rodin & McAvay, 1992). A recent systematic review by Gunasekara et al. (2011) examined literature on repeated measure of income and SRH in the adult population. The review found only 13 studies that met the inclusion criteria: four used the British Household Panel Survey (BHPS, U.K.), three used the Panel Study of Income Dynamics (PSID, U.S.), one used both the BHPS and PSID, three used the German Socio-Economic Panel (GSOEP, Germany), one used the Health and Retirement Study (HRS, U.S), and one used the Household, Income and Labor

Dynamics, Australia (HILDA, Australia). Of these thirteen, ten found a statistically significant positive association between income and health. Of the three studies that did not find a significant association, one study (Fischer & Sousa-Poza, 2009) limited the study sample to those who were employed, and the other two studies did find a significant association with mental health status but not SRH. In addition to the small number of studies, the authors also discuss the lack of studies that have accounted for measurement error in SRH. One of the studies that did account for measurement error in SRH is by McDonough, Worts and Sacker (2010) who contrasted the BHPS in the U.K. and the PSID in the U.S. with the mixed latent Markov model and found an effect of income, education, unemployment and occupation on which health trajectory a person belonged to. None of the longitudinal studies were conducted using Canadian data.

In a much broader review, Macdonald et al., (2009) examined studies in Canada that looked at indicators of SEP (income, income distribution, educational attainment or poverty status) and a wide range of health outcomes including SRH, morbidity, mortality, child development and access to resources. With the wider scope than the review done by Gunasekara et al., (2011), the authors identified 241 Canadian studies between 1995 and 2002. Of the 241 studies, 39 (17%) employed a longitudinal analysis. This scarcity of Canadian longitudinal research may be due to the fact that longitudinal data such as the National Longitudinal Survey of Children and Youth, the National Population Health Survey, and the SLID were only broadly made available to researchers outside of Statistics Canada over the past decade or so, which is late when compared with American and European counterpart (Dunn, 2009; Phipps, 2003).

What indicators are used to measure SEP?

Income has been measured in a number of ways: absolute level of income, pre and post taxes and transfers, not to mention the modeling decisions of looking at change in income, income inequality and temporary versus persistent poverty. Regardless of how income is measured, the association between income and variety of health outcomes is persistent and robust (Duncan, 1996; Krieger et al., 1997; Link & Phelan, 1995; Wagstaff & van Doorslaer, 2000; Wilkinson, 1996). As income largely dictates the extent to which a person attains material goods and achieves social standing (Lynch & Kaplan, 2000), it has been found to be a stronger predictor of mortality compared to other measures of SEP such as education and occupation (Duncan, Daly, McDonough, & Williams, 2002).

Yet, there are many issues surrounding measurement of income. First, there is the question of how income is conceptualized as there are assumptions made for each measure. For example, with measuring change in income, the permanent-income hypothesis proposes that the consumption behavior does not change despite fluctuations in income as people tend to 'smooth-out' fluctuations by saving in periods of high income and spend proportionately more when earning less income. With regards to measuring income inequality, using absolute versus relative measures can depict contradictory findings (Harper et al., 2010). A second issue is capturing true income without error. Ideally, income should be captured in its heterogenic and dynamic nature (Krieger et al., 1997) and disposable income should account for different sources of income (e.g., wages, investment, pensions) as well as transfers (e.g., tax credits, guaranteed income supplements). However, income is rarely captured in its heterogenic nature (Duncan et al., 2002), as recall bias impairs the accurate accounting of all sources of income, especially for those whose income is not salary based, as well as of deductions. In Canada, in 2008, government transfers were, on average, $13,500 CAD annually for families of two people or more in the lowest income quintile (Statistics Canada, 2011b). In addition, self-reported income (particularly from all sources) is fraught with recall bias, social desirability bias (Krieger et al., 1997), and is considered to be a socially sensitive item (Turrell, 2000). Indeed, a study done by Kim et al. (2007) found those who were missing income information to have systematically different socioeconomic characteristics from those who do report income information; those with missing information were less educated, reside in poor neighborhoods, and are younger. Third, there is the lesser challenge of deciding at which level income should be measured: individual, household, or ecological. Household income can take into account household composition and size and thus provide a more accurate representation of purchasing power especially for single earner families (Galobardes et al., 2006). Using household income, however, assumes that there is an equitable distribution amongst the members of the same household. Fourth, geographical variation in purchasing power in access to goods and services exists which further complicates measuring the value of income in different areas (Macintyre, Maciver, & Sooman, 1993; Zenk et al., 2005).

Education is a popular measure of SEP for several reasons: ease of measurement, stability during adulthood, high response rates, predictive power for mortality and morbidity, and wide generalizability as it is inclusive of those who are not active in the labour force (Galobardes et al., 2006; Krieger et al., 1997; Ross & Wu, 1995). Despite its popularity, education, by itself, may only account for a partial component of the true SEP as it does not perfectly overlap with income for instance. This discordance between education and true SEP further emphasizes the need to have multiple measures of SEP rather than one measure. Another drawback related to education is its shifting meaning and value over time and place of attainment. This is of particular concern in Canada, where immigrants compose a significant part of the population, and education obtained in another country may not be recognized by some employers contributing to the disjunction between education level and economic advantage. Thirdly, cohort effects are present as the percent of the population with university degrees nearly doubled between 1999 and 2009 in Canada (Statistics Canada, 2009). The temporal trend of a greater proportion of people attaining higher level of education speaks to cohort differences in employment prospects with a certain educational attainment. In addition, it suggests that there may be cross-cohort differences in the relative position and labor force impact of a given absolute level of education.

The association of education with health outcomes is thought to be a reflection of a person’s knowledge and characteristics that are associated with health such as self-direction and the ability to delay gratification (Fuchs, 1979). It is also a predictor for other indicators of SEP; namely, employment and income. Furthermore, some researchers argue that achieving a certain education level facilitates the attainment of meaningful and more rewarding work conditions and the development of gratifying interpersonal relationships (Mirowsky & Ross, 2003; Ross & Wu, 1995).

Education can be measured continuously or categorically. Continuous measurement of education assumes that each additional year of education contributes to health in a dose-response mechanism whereas a categorical measure of education emphasizes the major milestones such as obtaining a high school diploma or bachelor's degree as critical to later socio-economic achievement, and therefore on health, in a cumulative process. Hence, Krieger (1997) recommends measuring education in terms of attainment rather than continuously because achieving certain credentials has implications for employment which therefore implies that the differential benefit gained for each year of education is not equidistant. On the other hand, a twin study in the US by Fujiwara and Kawachi (2009) found each additional year of schooling resulted in a lower prevalence of rates of smoking, which tends to support a dose-response mechanism.

Occupation is thought to affect health through material resources from income derived from work (Ferrie, Martikainen, Shipley, & Marmot, 2005), as well as prestige and social standing related to the occupation (Bosma et al., 1997), psychosocial benefits (e.g., social support and employment security) (Virtanen, Vahtera, Kivimaki, Pentti, & Ferrie, 2002) and work-related physical exposures (Power, Manor, & Matthews, 1999). Measures that have been developed emphasize the different pathways in which occupation is thought to influence health. For example, the Erikson-Goldthorpe Schema focuses on characteristics of the workplace but do not have an implicit hierarchical rank, and the British Occupational-Based Social Class classifies occupation based on prestige. A limitation with these indicators is that they do not have universal applicability, either within or across jurisdictions. Within a given jurisdiction, certain groups such as the retired, students, and those out of the labor force or who are working in the informal sector are often excluded in these measures. There is also the concern of classifying individuals based on their occupational groups e.g., farmers, construction workers as the link between the kind of employment and employment conditions may vary. Therefore, classifying the current labour force based on a person's occupation group may lead to misclassification bias in capturing the factors that contributes to health.

To overcome the limitations of classifying occupations based on prestige, indicators of occupation such as job security and level of control and autonomy exist. These indicators capture the degree of precariousness in jobs. Capturing the concept of precariousness in occupation is especially relevant as the labour market has restructured to include greater variability and intensity in types of employment such as contract and service jobs based rather than unionized manufacturing jobs (Benach, Mutaner, & Santana, 2006; Quinlan, Mayhew, & Bohle, 2001). These non-standard types of work (working poor, precarious work) have been associated with negative impact on health (Benavides, Benach, Diez-Roux, & Roman, 2000; De Witte, 1999; Quesnel-Vallee, DeHaney, & Ciampi, 2010; Sparks, Cooper, Fried, & Shirom, 1997), and this impact seems to be disproportionately larger in economically marginalized groups (Kalleberg, 2009). In addition, there are increasingly variable definition of 'active labour force' where some are including informal sectors such as volunteers and caregivers (World Health Organization, 2011). With this post-industrial society, the traditional classification systems may be inadequate in capturing the current workforce. In light of these trends, employment status can be used as an indicator to capture exclusion from the workforce and has the advantage of being able to capture variety of statuses.

Regardless of how SEP is measured and combined, the association between SEP and a variety of health outcomes is clear. Nonetheless, despite decades of research, the best indicator of SEP that is most amenable to social policies and also balances the need for ease of measurement has not been established, leading to a plethora of research that measures income, education and occupation in different ways that seem to be driven by data availability rather than theory. Moreover, the strength of the association varies due to measurement of SEP, research design, population and countries (with different social policies) examined, and how covariates were chosen and included. In order to formulate policies to minimize health inequalities due to SEP, the first step would be to gather evidence on the strength of the SEP-health association that utilizes accurately measured covariates and also takes into account temporality.

2.2 Self-Rated Health (SRH)

- *Self-rated health: reliability and validity*
- *Is SRH better represented as categorical or binary?*
- *What are the issues with using self rated health as an outcome for studying socioeconomic differences?*

> **Interviewer:** Is it hard for you to compare your own health with that of other people of your own age, would you say it is…
> **Respondent (85-year-old):** Well most of them are dead, aren't they?
> (Jylhä, 1994)

The World Health Organization takes a holistic approach to defining health, as 'a state of complete physical, mental and social well-being, and not merely the absence of disease or infirmity' (World Health Organization, 1948). SRH provides an avenue to assess health that encompasses the multi-dimensional aspect of health beyond objective measures. Objective measures are designed to capture the presence of disease and symptoms and not necessarily its impact on quality of life – or as Litva & Eyles (2002) argues, *health and being healthy are not the same*. Kelinman (1989), in the anthropological narrative "The Illness Narratives" illustrates the limitations with taking an objective view of illness and health and describes this view as being inadequate in capturing the effect of the *illness* that comes with the *disease.* A qualitative study by Manderbacka (1998) echoes a similar notion; it was found that middle aged Scandinavians consider a wide range of aspects of health when determining their own rating of SRH including the impact of the disease on their lifestyle. Others have hypothesized that assessment of SRH acts like a 'crystal ball' by influencing future health behaviours that in turn affect mortality and morbidity (DeSalvo, Bloser, Reynolds, He, & Muntner, 2006).

Self-rated health (SRH) is widely used in population surveys as an indicator of health because it is relatively simple to administer and it has been shown to be a powerful predictor of a range of health outcomes. Specifically, SRH has been shown to predict use of healthcare services (Hansen, Fink, Frydenberg, & Oxhoj, 2002; Miilunpalo, Vuori, Oja, Pasanen, & Urponen, 1997; Wolinsky, Culler, Callahan, & Johnson, 1994), disability (Idler & Kasl, 1995), recovery from illness (Wilcox, Kasl, & Idler, 1996), and mortality (Idler & Benyamini, 1997; Kawada, 2003; van Doorslaer & Gerdtham, 2003). Furthermore, studies in multiple countries have used SRH as an outcome and have confirmed the association between health and SEP (Benzeval, Taylor, & Judge, 2000; Contoyannis, Jones, & Rice, 2004; Ettner, 1996). This association of SRH with numerous health outcomes in multiple countries lends to the validity of SRH. A meta-analysis by De Salvo (2006) revealed that this association is attenuated when covariates are accounted for but still remained significant and this is in agreement with previous research (Benyamini, Leventhal, & Leventhal, 1999; Idler & Benyamini, 1997).

SRH is ascertained by asking respondents about their health on an ordinal ranking of their perceived health status. There are different ways SRH can be asked; both the wording of the question and response categories can differ. For example, the question can ask about health in general ("how would you rate your health?") or relative to others ("how would you rate your health compared to others your age?"). Response categories are similarly varied, for instance, those from the United States PSID are "excellent, very good, good, fair, or poor" whereas in the BHPS from the United Kingdom, they are "excellent, good, fair, poor or very poor". To assess the impact of different wording and response options, Eriksson (2001) compared the results of different SRH questions and found high correlation in the responses regardless of whether the question was asked in relative or absolute terms and whether the person was presented with five or seven response categories. Moreover, regardless of the wording and response categories, reviews have consistently shown a link between SRH and health (Idler & Benyamini, 1997).

Analyzing longitudinal trends in SRH patterns imposes the additional challenge of temporality on top of the existing constraints of having the outcome be reliable over populations, socioeconomic groups, gender and age. The following will describe the limitations associated with analyzing SRH.

Due to its subjective nature, one of the major concerns associated with measuring SRH is determining the effect cultural and social environments have on an individual's assessment of health. Murray (2001) gave the pointed example of how aboriginals in Australia report better health than the general Australian population despite the fact that mortality and morbidity rates are higher in the aboriginal population. As far back as 1950s, researchers have conceptualized cultural differences in experiencing and reporting indicators of health such as pain (Zola, 1966). How closely an individual's own self-perception of health is related to his/her true health and has been referred to as 'response category cut-point shift' (Lindeboom & Van Doorslaer, 2004), 'scale of reference bias' (Groot, 2000) and 'positional objectivity' (Sen, 1993), among other terms. All of these concepts highlight the substantial assumption made when analyzing SRH across heterogeneous populations: that people in the study have similar comprehension and interpretation of the question asked, are equally affected by the environment, have similar values and are using similar criteria to evaluate their health. The concerns associated with heterogeneity can be minimized by observing a more homogenous population or limiting analysis of SRH to within one country. Nonetheless, in a country like Canada, where there is a large immigrant population with strong cultural ties to their place of birth, it is difficult to assume homogeneity. Studies have shown differences in interpretation and reference level of health across cultures (Desesquelles, Egidi, & Salvatore, 2009; Jurges, 2007; Jylha, Guralnik, Ferrucci, Jokela, & Heikkinen, 1998). These differences are hypothesized to be a reflection of social inequalities and differential access to various social services such as health care and education (McFadden et al., 2009). If SRH is a reflection, at least partially, of one's own physical condition (Jylha, 2009), sub-groups of people who have better access to health care should theoretically have better awareness of their own health status, consequently influencing their SRH ratings. Nonetheless, a qualitative study exploring how individuals arrive at their SRH, found that despite cultural differences in Western European countries, most sub-populations evaluate their health using a similar weighting of different dimensions of health (Jylha et al., 1998). Evidence of accuracy of SRH in different cultures and ethnicity is inconsistent, but it is clear that despite the difference in interpretation, the link to morbidity and mortality has been consistently found in different countries (Idler & Benyamini, 1997). Furthermore, SRH is recommended by the WHO as one of the measures of health in national studies (de Bruin, Picavet, & Noosikov, 1996).

The reliability of SRH and health across SEP has mixed evidence. Some studies show no variation between low and high social class (Burstrom & Fredlund, 2001) while others indicate that the association between SRH and health varies depending on the measures of SEP (Dowd & Zajacova, 2007; Singh-Manoux et al., 2007). Humphries and van Doorslaer (2000) have shown that subjective measures of health have a stronger association with income than objective measures of health (McMaster's Health Utility Index Mark 3 (HUI 3)) in Canada. Additional studies based in other countries including France (Etile & Milcent, 2006; Singh-Manoux et al., 2007) and the Netherlands (Huisman, Van lenthe, & Mackenbach, 2007) provide evidence consistent with Humphries and van Doorslaer's findings. Even though some studies show differences in the predictive ability of SRH on mortality between different SEP strata, it should be noted that these studies are not always sufficient evidence for significant bias: Huisman (2007) found a difference in predictive power only for the highest educational group in men (RR 1.33 (1.07-1.66)). Furthermore, a substantial number of other studies have not found a difference in mortality by social class (Burstrom & Fredlund, 2001; McFadden et al., 2009; van Doorslaer & Gerdtham, 2003).

In addition to SEP and cultural difference, SRH ratings show effect modification by age and gender. The frame of reference for SRH varies by age with younger people typically focusing on health behaviors and older adults focusing on health problems (Krause & Jay, 1994). The validity of SRH has been confirmed in working age groups (Miilunpalo et al., 1997; Wannamethee & Shaper, 1991). The relationship between SRH and mortality is modified by gender (Idler, 2003; Lindeboom & Van Doorslaer, 2004), with the association being stronger in men than women. Indeed, a review by Idler & Benyamini (1997) found SRH to be a stronger predictor of mortality in men than women in ten of the sixteen studies, with the largest difference existing in studies based on people of working age (Miilunpalo et al., 1997; Strawbridge & Wallhagen, 1999). Idler (2003) provides a comprehensive review of possible explanations for these gender differences. First, it may be that women tend to consider dimensions of heath that are unrelated to mortality or are non-fatal more often than men when determining their SRH (Deeg & Kriegsman, 2003). So as men base their SRH rating on conditions that are more closely tied to mortality, the SRH is a stronger predictor of mortality in men than women. Another interpretation is the differential level of knowledge about health between women and men. Women are more likely to be aware of their health conditions as they utilize health services more frequently and health information is often targeted preferentially towards women. As a result of this heightened awareness, women are more likely to give accurate answers in other health status information besides SRH. When this more accurate health information is taken into account in a multivariate analysis, the association between SRH and health is attenuated to greater degree in women than in men. Idler (2003) illustrates with three empirical examples where the adjusted hazard ratio shows greater attenuation in women than men after accounting for other health information (e.g., physical activity). Thus, gender and age differences in SRH ratings shed light on the processes whereby individuals establish a SRH rating, and may not constitute biases.

With respect to time, there is even less certainty concerning its effects on the association between SEP and SRH. The time dimension introduces two major issues: within-person serial correlation and construct validity. The within-person serial correlation is accounted by analysis. However, there is the question of reliability and validity over time: does the interpretation of the SRH question stay stagnant over time? Are the changes in health status a true reflection of change in health status? Do individuals adapt to their disability over time? There are at least some studies that point towards changes in physical status predicting changes in SRH (Benyamini, Idler, Leventhal, & Laventhal, 2000; Rodin & McAvay, 1992).

Quantifying the scale of these measurement biases is difficult. The mixed results of how the environment shapes the ratings of SRH and the direction of change provides evidence of this difficulty. Despite the variation in the interpretation and assessment of SRH across different sub-groups, the predictive capacity for future health outcomes across dimensions of gender, race, ethnicity and time is consistent and cannot be discounted (Chandola & Jenkinson, 2000; Jylha, 2009).

Chapter 3

METHODS

3.1 Data Source

The Survey of Labour and Income Dynamics (SLID) is a longitudinal survey initiated in 1993 by Statistics Canada. The main objective of the survey is to measure changes in economic well-being of Canadians. It is a household panel survey with an overlapping cohort design where each cohort remains for six consecutive years and a new cohort is added every three years to maintain sample size and representativeness (Figure 2). A household is defined as those who have been selected in the reference year of the panel and these individuals are followed for six years regardless of whether they move away from the household. Those who join the household after the reference year are considered to be cohabitants and are also interviewed as long as they live with any of the original longitudinal respondents. Each panel counts approximately 17 000 households and 34 000 respondents.

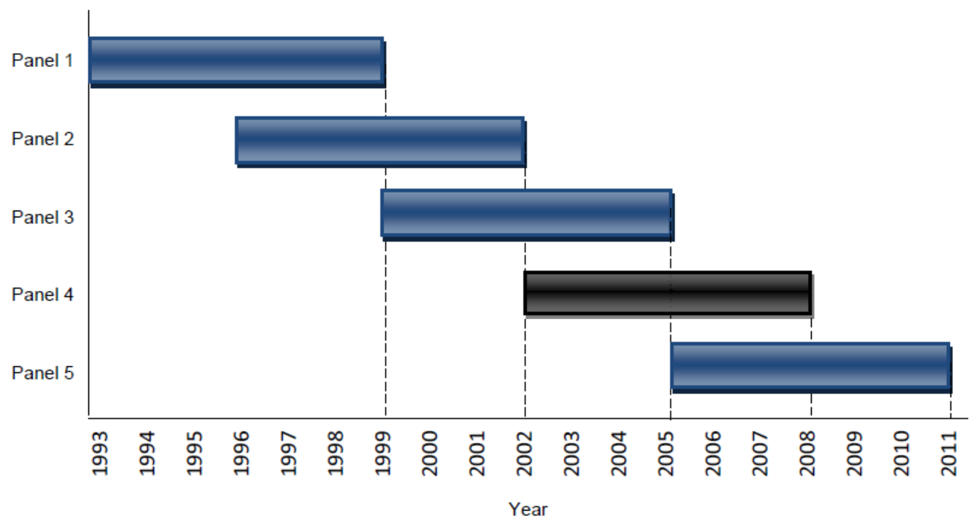


Figure 2: Overlapping survey design of SLID

Prior to 2004, there were two surveys per year. In January, households were interviewed using computer assisted telephone interviews (CATI) to collect demographic information and labour data. CATI reduces response error as it reminds the respondents of the information they had provided in previous interviews. In May of the same year, households who declined to have their income tax records linked were contacted again to collect information on income for every person in the household 16 years or older. Starting in 2004 and onwards, income data along with demographics and labour data were collected concurrently in January. That is, respondents who did not give Statistics Canada permission to access their income tax records were asked of their income information in January. For the period examined in this study, over 70% of the respondents gave permission for Statistics Canada to link their income tax file from Revenue Canada. As income measurement is fraught with recall bias and social desirability bias, information linked to an administrative income tax records is considered to be the gold standard in income measurement. The timing of the surveys implies that the income information collected pertains to the year before health data is collected, lending credence to the temporal ordering of events leading from income to health. Proxy response was allowed; one household member may answer questions on behalf of all members of the same household on the condition that the person is knowledgeable enough to do so. For the purpose of this study, we will be using the most recently completed panel, panel four, which includes 42 232 individuals and sampled between 2002 and 2007 (the black bar of Figure 2).

3.2 Sample

The target population is all persons living in Canada except for those who are living in the territories (Yukon, Northwest and Nunavut), reserves, or institutions, and military personnel living in barracks. The sample is drawn from a stratified, multi-stage design with all members of the household being included. As the sample is not drawn at random with independent equal probabilities, appropriate steps were taken to address this complex survey data including sample weights to reflect population distribution and bootstrapping for variance estimation. Bootstrap refers to repeatedly taking a large number (a minimum of 200 replicates are recommended) of random subsamples of the primary sampling unit (the original dataset) to obtain estimates from each replicated sample then taking the variance of the mean of all replicate estimates (Lee & Forthofer, 2006; Rust & Rao, 1996). The bootstrap method was applied with 1000 replicates that were provided by Statistics Canada. Cross-sectional response rates ranged from 69.6% in 2007 to 79% in 2002 with an average of 75.2% over the six years (Table 1).

Table 1: Cross-sectional response rates for Panel 4 of SLID

Year	2002	2003	2004	2005	2006	2007
%	69.6	74.9	74.7	74.7	78.3	79.0

Longitudinal studies are subject to attrition. Table 2 lists the reasons why respondents were lost. Those who have been lost to attrition include those who have moved to the Territories, emigrated, institutionalized and deceased individuals and this constitutes 36.8% of the total sample (15524/42232). These dropouts may not be randomly distributed and may introduce selection bias. To mitigate the effect of attrition and to ensure that estimates are representative of the initial population, longitudinal weights provided by Statistics Canada were applied. Longitudinal weights are designed to generate estimates that are representative of the population of Canada's ten provinces at the time the longitudinal sample is selected. Longitudinal weights account for (1) nonresponse (2) influential values and (3) post-stratification. Please see the report by Levesque & Franklin (2000) for detailed information on how the longitudinal weights were derived for SLID.

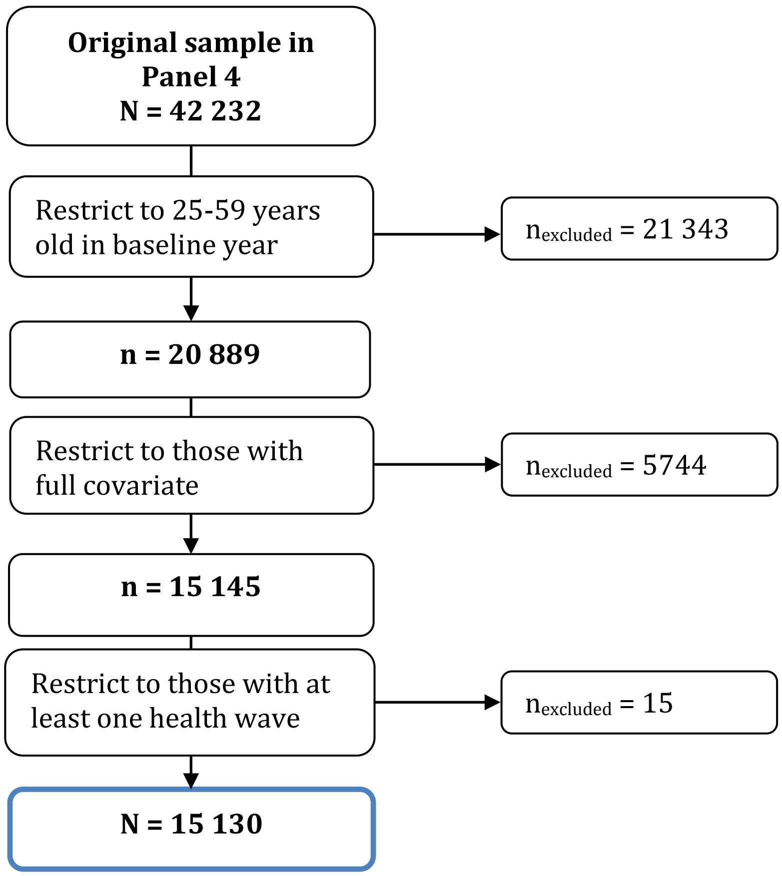
Table 2: Status of respondents in 2007

Reason for attrition	n
Non-respondents	3248
Moved to territories	9
Moved outside of Canada	323
Institutionalized	657
Deceased	1391
Removed from sample	9859
Error	37
Total	15524

In Figure 3 we present the flowchart that describes our sample selection with regards to age as well as exclusion due to missing values. To ensure that the sample contained those who had completed their education and are most likely to be active participants in the labour force, the sample was restricted to 25 to 59 years old at baseline year (Y2002). This led us to exclude 21 343 of the 42 232 participants, and left us with a sample of 20 889 respondents. The upper limit was imposed to ensure that the oldest individuals would be nearing or reaching age 65, at which they become eligible for full benefits from the federal pension system, by the end of the 6-year observation period. The lower limit was chosen to maximize the number of respondents having completed their schooling. This lower limit cut-point appears to be appropriate, as only 2.27% of 25 to 59 year-old respondents reported being a student in 2002, our baseline year. Furthermore, this proportion steadily decreased over time in the sample selected for analysis.).

In addition, as with any survey, SLID's sample is subject to unit and item non-response. While unit (household and individual) non-response is dealt with by weighting, item non-response remains an issue (discussed below in the missing data section). As shown in Figure 3, we performed a listwise deletion of cases with any missing values on the covariates (taken at baseline), which led us to exclude 5744 cases. Finally, while the analytic method we use (full information maximum likelihood, described in greater detail below) can accommodate some amount of missing data in estimating latent trajectories, it nevertheless needs some observed information to do so, and hence we further restricted our sample to those respondents with at least one measurement of SRH. With the exclusion of the 15 respondents lacking at least one wave of SRH data, we were left with a final analytic sample of 15 130 respondents.

Figure 3: Flowchart of sample restriction



3.3 Variable Definition and Selection

The independent predictors of SRH are income, education and employment status. Confounders, chosen a priori on the basis of our literature review, are gender, age, minority status, visible minority status and marital status. For the purpose of this analysis, only baseline predictors and confounders will be used. However, the dynamic nature of the outcome variable, SRH, will be modeled with trajectories. The following statistics refer to the analysis sample (n=15 130).

3.3.1 Predictors

Income

The majority of the income information was derived from administrative income tax records (Canada Revenue Agency's income tax records) rather than self-reported, and this source of information constitutes one of the strengths of the SLID's income data. In any given year, more than 70% of respondents chose to link their records. This greatly improves the accuracy of the income information in providing a close estimate of personal disposable income as it takes into account income from multiple sources including transfers from government (e.g., old age security and employment insurance) and progressive tax deductions. However, there are also limitations from using income in this format. First, it cannot account for distribution of resources (or lack thereof) within the family. Second, the information collected does not include assets such as owning a house which can buffer effects of negative income dynamics on health. After tax income is defined as follows:

Market income
- Earnings
  - Wages, salaries and commission
  - Self-employment income
- Investment income
- Retirement pensions
- Other income

(plus) **Government transfers**

Child tax benefits
Old Age Security and Guaranteed Income Supplement/Spouse's Allowance
Canada Pension Plan/Quebec Pension Plan benefits
Employment Insurance benefits
Social Assistance
Workers' compensation benefits
GST/HST Credit
Provincial/territorial tax credits
Other government transfers

(equals) **Total Income**
(minus) **Income tax**
(equals) **After-tax Income**

Please see appendix A for full definition of each item used to calculate after-tax income.

It is also important to account for household size and composition as income adequacy changes with household size. Indeed, economies of scale on fixed expenditures that occur as household size (and potential number of additional earners) increases can partially offset the food and clothing costs incurred by these additional members. There is no agreed upon method to adjust for household size and composition (Atkinson, Rainwater, & Smeeding, 1995; Buhmann, Rainwater, Schmaus, & Smeeding, 1988; Lanjouw & Ravallion, 1995), but there is some consensus around the fact that household expenditures are non-linearly related to household size. Thus, the equivalence scale used by Statistics Canada is as follows: the oldest person in the family is assigned a weight of 1.0; the second oldest is assigned 0.4; other members age 16 and over are assigned 0.4; and members under the age of 16 are assigned 0.3. Family income is then divided by the sum of the weights assigned to the family members for each year therefore capturing dynamics in family composition. This was done prior to dropping those who were under age 25. The equivalency scale is applied homogeneously across households, although the validity may not hold across all households – there is evidence to suggest gender differences in resource allocation (Thomas, 1994) and mothers of poor families tend to prioritize their spouse and children before their own needs (Krieger et al., 1997).

Income was categorized into 11 categories starting at $0-$10000 in increments of $10000 in order to account for the non-linear relationship between income and health (Backlund, Sorlie, & Johnson, 1996).

As mentioned earlier, income information was derived from one of two sources: consenting respondents' income were derived from T1 Income Tax and Benefits Return file from Canada Revenue Agency or the respondents had the choice to self-report their income. Due to the varying sources of financial information, a t-test[1] was done to compare the difference between self-reported and tax-file derived for the baseline year. The t-test revealed that there is a statistically significant difference between the mean income of a tax derived information and self-reported income information ($t= -3.59$, $p =0.0003$). Those who self-reported their income had a higher mean income compared to those whose income information came from tax files. However, with the majority of income information being derived from income tax files (ranging from 70.08%-88.18%), bias from differential reporting should be minimized.

Education:

The original question for education classified individuals into one of the following categories: no schooling, elementary, some secondary, secondary school graduation, other beyond high school, some trade school, some community college, some university, diploma or certificate trade school, diploma or certificate of community college, bachelor degree, master's or MD or PhD degree.

This variable was recoded into four categories: less than high school, graduated high school, non-university postsecondary certificate, and university degree or certificate. Education was captured as a categorical variable with milestones as cutoffs rather than continuous years of schooling as we interpreted the relationship between education and health to be a reflection of milestones achieved. As demonstrated in the literature

[1] Bartlett's test for equal variance was performed to check for equal variance at the $p=0.05$ level. Chi square statistic 0.31 ($p=0.58$) indicated equal variance of both groups.

review, this measure of credentialing also provides a better estimation of the potential returns to education in the labor force.

Employment:

The Statistics Canada derived variable for working status asked for the person's main activity at the end of the reference year. The response categories were working at a job or business, looking for work, going to school, keeping house, caring for other family members including young children, retired, long term illness or disability, doing volunteer work, no main activity, and other. These categories were collapsed into employed (working at a job or business), unemployed (looking for work), student (going to school), out of labour force (caring for other family members including young children, retired, long term illness or disability), and other (doing volunteer work, no main activity, other).

3.3.2 Outcome

Self-rated health:

The question pertaining to SRH asks, "In general, how would you describe your state of health? Would you say it is…". It is self-reported and the five response categories were excellent, very good, good, fair and poor. The same question was administered in all years. The health status question can be answered by proxy. In 2002, 26.8% (4034/15042) responded to the January questionnaire (which contains the SRH question) by proxy. To assess the impact of proxy respondents, a chi square test[2] (x^2) was conducted to compare the proportions of responses between proxy and non-proxy respondents. Only the baseline line year (Y2002) had significantly different proportions (x^2=12.22, df=4, p=0.016), and all other years were non-significant at α= 0.05 level (Appendix B). To categorize SRH into a binary variable the original five categories were grouped into 'good health' which included excellent, very good and good and 'poor health' which included fair and poor. Many studies also dichotomize SRH into good and less than good health (Blakely, Kennedy, Glass, & Kawachi, 2000; Bobak, Pikhart, Rose,

[2] The expected value of each cell was far greater than 5.

Hertzman, & Marmot, 2000; Newbold, 2005) due to small number of respondents in one category (Benyamini, Blumstein, Lusky, & Modan, 2003) or to maintain consistency between different surveys (Burstrom & Fredlund, 2001). Manor (2000) tested the justifiability of collapsing the variable by comparing the results from a binary collapsed variable using logistic regression versus ordered categorical variable using polytomous regression, cumulative odds, continuation ratio and adjacent categories. The authors find that, in a large sample size, analyses based on binary SRH responses typically have similar results as those based on categorical analyses.

3.3.3 Confounders

Visible minority and immigrant status:

Measurement of race and ethnicity varies by jurisdiction and social context. While some countries, like the U.S., espouse relatively consensual conceptual frameworks, this has not historically been the case in Canada. Hence, our approach here draws from the social exclusion literature to capture populations that were found in a Canadian assessment of social determinants of health to unduly and disproportionately suffer from exclusion, namely visible minorities and immigrants (Bertrand & Mullainathan, 2004; Raphael, 2004). More specifically, the term visible minority has been described as a 'group that carries with it the connotation of the unequal relationships among groups within society, in which groups are subjected to greater prejudice and discrimination' (Sue, 1991). And further, immigrant status has been shown to play a role in determining one's position within the social structure (Angel & Angel, 2006).

The Employment Equity Act defines visible minorities as "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour" (Employment Equity Act, SC 1995, c 44). This includes those who are Chinese, South Asian, Black, Arab, West Asian, Filipino, Southeast Asian, Latin American, Japanese and Korean. This information was collected at baseline year by self-report. The respondents who reported themselves as Black, South Asian (Indo-Pakistan), Chinese, Korean, Japanese, South

East Asian, Filipino, Oceanic (other pacific islands), West Asian & North African (Arab), and Latin American were considered to be a visible minority and captured as a binary variable (visible minority and non-visible minority). Immigration status was ascertained with the question “Did you immigrate to Canada?”. While the two concepts sometimes refer to the same populations, they do not perfectly overlap, notably due to variations in the source countries of Canadian immigration, as well as with the birth of second and higher order generations of members from visible minority communities in Canada. Indeed, when weighted, 8.79% of our respondents listed themselves as an immigrant but not as a visible minority. Conversely, 1.66% categorized themselves as a visible minority but not as an immigrant (e.g., second or higher order generation children).

Ethnicity has been conceptualized as an enduring, fundamental aspect of self that includes a sense of membership in an ethnic group and attitudes and feelings associated with that membership (Phinney, 1996). Thus, identification as a visible minority may allow the identification of groups at greater risk of experiencing systematic biases of discrimination, as most Americans and Europeans do not see ethnicity as part of their identity (Alba, 1990). Yet there are drawbacks from capturing this concept as a binary variable, as the impact of minority status is not uniform across individuals nor groups. For example, Harrison and Thomas (2009) documented a gradation in the experience of job discrimination by African Americans in relation to the variation of their skin color. Furthermore, personality characteristics such as self-esteem may influence the perception of discrimination and introduce individual variation in its health consequences (Phinney, 1996).

Similarly, subsuming all immigrant groups under the same category masks the substantial heterogeneity stemming from variations in the country of origin, level of development, and ethnicity status within that country (Setia, 2011). In addition, our measure also obscures the fact that the effect of immigrant status may be mitigated by the length of time that they have resided in Canada. For example, in 2010, the unemployment rate for immigrants who landed less than five years ago was 15.8%, while those who immigrants who landed 5-10 years or more than 10 years ago respectively experienced unemployment rates of 11.1% and 8.2% (Statistics Canada, 2011c). One should note however, that these measures of time spent in Canada also conflate the potential impact of secular changes in the source countries of immigration, from those of European descent to primarily Asian and South Asian populations. Thus, our use of the visible minority indicator along with that of immigrant may capture some of that variation (Citizenship and Immigration Canada, 2009).

Marital status:

Marital status of the respondent was measured at baseline. Dummy variables were created for each category: married, common law, separated (including from a common-law union), divorced, widowed and single, with never married as the reference category.

Sex:

Sex was measured at baseline, with male as the indicator and female as the reference category.

Age:

Age at baseline (in 2002) was recoded into 5 year group dummy variables as follows: 25-29, 30-34 (reference group), 35-39, 40-45, 45-49, 50-54, 55-59. This categorization is consistent with previous studies.

3.4 Model Selection

3.4.1 Markov Models and Extensions

In longitudinal data, where variables are measured repeatedly over time within the same individual, the problem of dependency between the repeated measurements arises. The correlation between measures can be modeled using Markov models (Markov, 1907). Markov models have been extensively used to account for state dependency; where the current response is influenced by past response in a non-random manner. Three different types of models will be fitted and the need for each will be assessed: Manifest Markov model (M), Latent Markov model (LM), and Mixed Latent

Markov model (MLM). Comparison between the M and the LM will allow testing for the necessity of accounting for measurement error. Fitting of the MLM will test the assumption of heterogeneity of populations – that is, multiple latent Markov chains. Within these models different types of restrictions will be placed including time invariant transition probabilities and time invariant response probabilities where applicable.

Manifest Markov model (M):

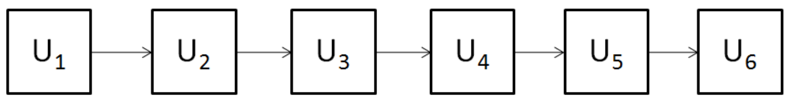


Figure 4: Graphical model of first order manifest Markov chain. Observed variables are represented by squares.

The diagram in Figure 4 depicts a first-order manifest Markov model where the response at time2, U_2, is dependent on response at time1, U_1. The first order refers to a process where the state in time t only depends on the state at t-1 *only* and not on earlier points in time. Higher order Markov chains, where the current state can depend on more than one previous time period, can be modeled but require more degrees of freedom to be identified.

The following introduces the statistical notation for a manifest Markov model for a repeated observation of one item over six time periods:

$$P_{ijklmn} = \delta_{i}^{1} \tau_{j|i}^{12} \tau_{k|j}^{23} \tau_{l|k}^{34} \tau_{m|l}^{45} \tau_{n|m}^{56}$$

The subscripts refer to response categories and the superscripts refer to time points. The frequency of response pattern is represented by P_{ijklmn} where the subscripts i, j, k, l, m, n corresponds to the categories of the manifest variable (I, J, K, L, M, N) at Time 1 to 6 respectively. In this study, there are two possible categorical responses possible for i, j, k, l, m, n: good and poor health. The δ is the unconditional proportion of individuals who have good health or poor health at time 1. The τ are the transitional probabilities at time t+1 conditional on the response category at time t. For example, $\tau^{12}_{j|i}$ is the probability that an individual will be in category j at time 2, given that she or he was in category i at time 1.

The transition probabilities were fixed to be time homogenous (stationary Markov chain). That is, $\tau^{12}_{j|i} = \tau^{23}_{k|j} = \tau^{34}_{l|k} = \tau^{45}_{m|l} = \tau^{56}_{n|m}$, and the validity of this assumption was evaluated by comparing the model fit with and without homogenous transition probabilities. It was necessary to impose this restriction for identification purposes in further extensions of the Markov models discussed below (Vermunt, Langeheine, & Bockenholt, 1999).

Latent Markov model (LM):

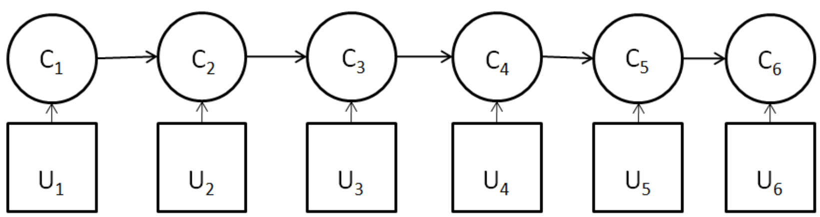


Figure 5: Graphical representation of latent Markov model. Observed variables are represented by squares and the circles represent the latent variables.

One of the limitations of the manifest Markov model is that it cannot differentiate between true and spurious change arising from measurement error (Engel & Reinecke, 1996; Hagenaars, 1992). Thus, an extension of the manifest Markov model is to include latent variables in order to account for the measurement error associated with the observed measurement (Wiggins, 1973). The latent Markov chain is composed of two parts: (1) the measurement component, represented in the diagram as U1→C1, U2→C2…U6→C6, which contains the relationship between latent variables (C) and the manifest indicator variables (U) and (2) the structural component which is the change process between the latent variable modeled by Markov chain (C1→C2→C3→C4→C5→C6). As it can be seen from the graphical model, the transition structure is between the latent responses rather than the manifest variables. LM is also known as a hidden Markov model and has been used extensively in the fields of speech recognition (Rabiner, 1989), education (Kaplan, 2008), and econometrics (Hamilton, 1989).

Latent variables are used in cases where the construct being measured is not directly observable. In this case, the true underlying concept is health, and it is captured by an indicator of health, SRH. Misclassification of health can result from social desirability bias, misinterpretation of the question, data entry errors, and "all factors other than the A (*true status of health)* that determine the value of A*(*self-rated health status)*" (Hernan & Cole, 2009). A way to quantify the measurement error, especially in self-reported data, is to use latent variable models (Biemer & Wieson, 2002; Engel & Reinecke, 1996). As opposed to latent class analysis which requires multiple indicators, latent Markov model can be estimated using only one indicator if there are multiple waves (McCutcheon, 1996; Vermunt et al., 1999). An explanation of measurement error in latent Markov model is provided in Appendix C.

Given six time points, where ijklmn denotes the response of the manifest variable for time 1 to 6 respectively, and abcdef represents responses to the latent variables ABCDEF, the LM is written as:

$$P_{ijklmn} = \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{c=1}^{C} \sum_{d=1}^{D} \sum_{e=1}^{E} \sum_{f=1}^{F} \delta_a^1\, \rho_{i|a}^1 \tau_{b|a}^{12}\, \rho_{j|b}^2 \tau_{c|b}^{23}\, \rho_{k|c}^3 \tau_{d|c}^{34}\, \rho_{l|d}^4 \tau_{e|d}^{45}\, \rho_{m|e}^5 \tau_{f|e}^{56}\, \rho_{n|f}^6$$

where:

P_{ijklmn} is the expected proportion of individuals in cell ijklmn.

δ_a^1 proportion of individuals at time 1 having 'a' latent health state

$\rho_{i|a}^1$ response probability of observing value i, given membership in latent state a.

$\tau_{b|a}^{12}$ transition probability from time 1 to time 2 for those in latent state b given that they were in latent state a at time 1.

The response probabilities (ρ) correspond to the measurement component of the model and the transition probabilities (τ) describe the structural component. By combining the markov model with the latent class model, the resulting model allows for analysis of change and account for measurement error. This equation reduces to the manifest Markov model if the measurement error is non-existent ($\rho_{i|a}^1 = \rho_{j|b}^2 = \rho_{k|c}^3 = \rho_{l|d}^4 = \rho_{m|e}^5 = \rho_{n|f}^6 = 1$). In other words, if SRH had perfect reliability, the latent self rated classes will have the same distribution as the observed SRH status and the manifest Markov model would suffice in describing the data.

A necessary condition for the identifiability of a model is to have degrees of freedom that are greater than zero. Due to the computation burden of the model, restrictions were imposed in order to reduce the number of parameters that were estimated in order to have more degrees of freedom. Multiple authors have also pointed out the need for additional restrictions that must be placed for identifiability (Van de Pol & Langeheine, 1990; Vermunt, Tran, & Magidson, 2008; Vermunt et al., 1999). In line with this literature, measurement errors across all six time periods were held equal.

To evaluate the optimal number of latent classes, different number of classes were fit to the six waves of health data e.g., 2 classes corresponding to 'good' and 'poor' health, three classes corresponding to 'very good', 'good' and 'poor'. The choice of number of latent states usually equals the number of categories of the observed response variable (Vermunt et al., 2008).

Mixed latent Markov model (MLM)

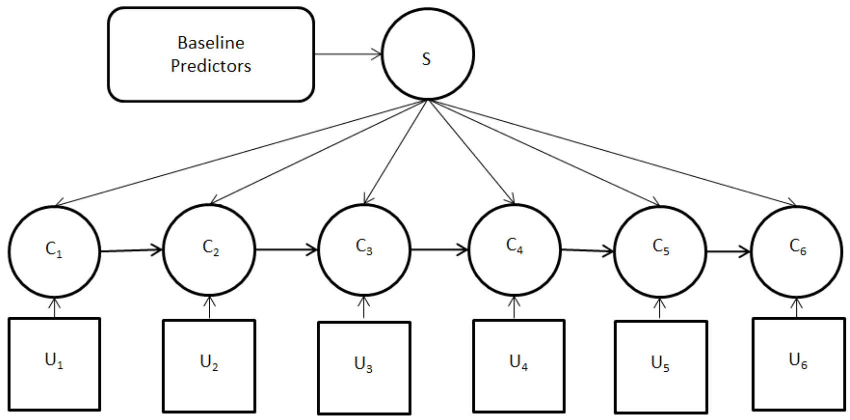


Figure 6: Graphical representation of mixed latent Markov model

A limitation with both latent Markov and the manifest Markov models is in assuming that the process is the same for all individuals. In other words, individuals are from a single population. It is plausible, however, to assume that the population is composed of multiple subpopulations. For example, those who have constant good health, constant poor health and those who fluctuate between good and poor health. Moreover, it has been shown that latent Markov models overestimate the size of measurement error probabilities (Magidson, Vermunt, & Tran, 2007). Mixture modeling therefore provides an opportunity to postulate multiple subpopulations that will be inferred from the data (McLachlan & Peel, 2000). Mixture modeling allows for several trajectories at the latent level, and having multiple latent Markov chains allows for each chain to have its own set of parameters. A consequence of partitioning the heterogeneous population into more homogeneous sub-population is that no inter-individual variation is allowed within each latent trajectory (Langeheine, Stern, & van de Pol, 1994). Therefore, all subjects within one chain are characterized by one set of parameters.

The following formula is an extension of the LM model to account for multiple latent Markov chains:

$$P_{\mathrm{ijklmn}} = \sum_{s=1}^{S}\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\sum_{d=1}^{D}\sum_{e=1}^{E}\sum_{f=1}^{F} \pi_s \delta^{1}_{\mathrm{a|s}}\, \rho^{1}_{\mathrm{i|as}}\tau^{12}_{\mathrm{b|as}}\, \rho^{2}_{\mathrm{j|bs}}\tau^{23}_{\mathrm{b|as}}\, \rho^{3}_{\mathrm{k|cs}}\tau^{34}_{\mathrm{d|cs}}\, \rho^{4}_{\mathrm{l|ds}}\tau^{45}_{\mathrm{e|ds}}\, \rho^{5}_{\mathrm{m|es}}\tau^{56}_{\mathrm{f|es}}\, \rho^{6}_{\mathrm{n|fs}}$$

where,

S represents latent Markov chains

π_s represents the proportion of the population in each S chain

The initial probability ($\delta^{1}_{\mathrm{a|s}}$), conditional response probabilities associating the latent variables to their respective indicator variables ($\rho^{1}_{\mathrm{i|as}}$, $\rho^{2}_{\mathrm{j|bs}}$, $\rho^{3}_{\mathrm{k|cs}}$, $\rho^{4}_{\mathrm{l|ds}}$, $\rho^{5}_{\mathrm{m|es}}$, $\rho^{6}_{\mathrm{n|fs}}$), and the transition probabilities between latent measurements at t and t+1 ($\tau^{12}_{\mathrm{b|as}}$, $\tau^{23}_{\mathrm{b|as}}$, $\tau^{34}_{\mathrm{d|cs}}$, $\tau^{45}_{\mathrm{e|ds}}$, $\tau^{56}_{\mathrm{f|es}}$) are all conditional on chain membership. As with the LM being deduced to M when there is no measurement error, the MLM equation reduces to the LM when there is only one chain. Comparing the fit between LM and MLM will evaluate the need to account for a heterogeneous population.

The MLM requires specification of the number of latent Markov chains (S). There are different variations of the MLM depending on the number of chains specified. One of these variations is a mover-stayer model originally conceived by Blumen et al. (1955). In this model, one trajectory is fixed to have a probability of 0 of transitioning to another state (stayers) and the other chain has a non-zero probability of moving (movers). Once the number of chains is specified, covariates and independent predictors at baseline can be incorporated into the MLM models by multinomial logistic regression to evaluate its effect on trajectory membership.

3.4.2 Summary of Markov models and extensions

- Manifest Markov (M) model, latent Markov (LM) model and mixed latent Markov (MLM) model will be evaluated for their fit.
- MLM takes into account unobserved heterogeneity, autocorrelation and measurement error.
- Different variations of the MLM model will be fitted.
- Assumptions made are homogenous transition probabilities and homogenous response probability.

3.5 Model Identification & Estimation

- *How were the model parameters estimated?*
- *And how were problems of non-convergence and local solutions dealt with?*

Maximum Likelihood (ML) & Expectation Maximization (EM)

The model parameters were estimated using maximum likelihood estimates (ML) through an expectation maximization algorithm (EM). ML produces estimates that maximize the likelihood (largest likelihood is chosen) of the data given the specified model (Myung, 2002). One of the advantages of EM over alternatives, namely the Newton Raphson algorithm, is in its robustness in finding maxima even if the initial starting values are far from the final estimates (McCutcheon, 2002). Nonetheless, different sets of starting values were specified in the model until the same value was found (the global maximum) and had large number of iterations (see Appendix D for number of iterations, random start values, and the minimum convergence criteria specified for each model). There are other advantages of using ML, especially in a large sample: consistent convergence to the true parameter values, standard errors are smaller than other methods (better efficiency), and the sampling distribution is approximately normal.

Identification

As mixed latent Markov modeling with number of chains is computationally heavy, sometimes the parameters are not uniquely identified. Identification refers to unique solutions of the parameters that are determined by the sample. In other words, an unidentified model refers to estimating parameters that take on no value or multiple values. Identification of the model was checked by re-running the model with different start values and comparing the likelihood value and the estimated expected frequency. A model was deemed to be unidentified when the same value of likelihood was found but the estimated expected frequencies and the parameter estimates were different.

3.6 Missing Data

Despite rigorous data collection and methods used by Statistics Canada to minimize occurrence of missing data, attrition and item non-response inevitably introduces missing data. While attrition is dealt with weights, one method to account for item non-response requires studying *why* the data are missing. Table 4 shows the extent of missing data for SRH for each wave and Table 3 shows the proportion of respondents who have completed each wave.

Table 3: Longitudinal missing data pattern for SRH (not weighted, n=15 130)

Number of waves with complete SRH information	n	%
1	1261	8.33
2	778	5.14
3	833	5.51
4	1073	7.09
5	1886	12.47
6	9299	61.46

Table 4: Cross-sectional missing data pattern for SRH (not weighted, n=15 130)

Year	n	%
2002	53	0.35
2003	2016	13.32

2004	2659	17.57
2005	3264	21.57
2006	3941	26.05
2007	4015	26.54

As shown in Table 3, the majority of the participants (61.5%) had a full six waves of data. Cross sectional examination of missing data shows growing proportions of missing data due to attrition.

Missing data can arise from three mechanisms: missing completely at random (MCAR), missing at random (MAR) and not missing at random (NMAR). If the data are MCAR, the power to detect differences is minimized but estimates are not biased. Missing values on a given variable are considered to be MAR when their distribution is unrelated to responses on the variable itself after controlling for other variables in the dataset. For example, the probability of missing data on SRH is correlated with education status but not to the respondent's SRH. Unfortunately, the previous scenario is plausible, as when individuals with poor SRH are more likely to be lost to attrition because of the increased risk of mortality. The last mechanism is NMAR, where the dependency remains even after taking into account measured covariates in the study. A person with poor health may not report their health status because they are uncomfortable with reporting this information. In this thesis, the statistical method used to deal with missing data assumes that the data is either MAR or MCAR. However, the method that was chosen has shown to be robust in obtaining unbiased parameter estimates even when the data may not be MAR (Enders, 2001).

There are a variety of methods for dealing with missing data: pairwise deletion, listwise deletion, multiple imputation (MI) and full information maximum likelihood (FIML). MI and FIML produce similar results under certain assumptions, namely that the number of imputations in MI is sufficiently large and the imputation model is equivalent to the model that is used for analysis in the maximum likelihood in FIML (Schafer & Graham, 2002). Full information maximum likelihood (FIML) was introduced in the 1950s by Koopmans, Rubin and Leipnik and was chosen here to account for missing data. All variables (variables correlated with outcome and missingness, variables correlated with outcome but not with missingness) were included in predicting the missing data as Collins, Schafer and Kam's simulation (2001) demonstrated that employing an inclusive strategy had larger gains for efficiency with minor costs to bias (given moderate missing data).

3.7 Model Fit Statistics

Four model evaluation criteria are often used for Markov chains and its latent variable extensions: Pearson chi square (X^2), likelihood ratio chi square (L^2), Akaike information criteria (AIC), and Bayesian information criteria (BIC) (Kaplan, 2008). Model fit statistics are used to evaluate how well a statistical model fits the data by comparing the expected cell counts under the model hypothesis to the original observed cell counts. The replication power of the hypothesized model is used to determine the choice of a model: if the observed cell frequency under the hypothesized model is too far from the observed cell frequency, the model is deemed to be implausible. Multiple competing models are fitted to determine the best model. One caveat to model selection is that saturated models tend to have better fit. Although the information criteria (AIC and BIC) penalizes for increased number of parameters being estimated, relying solely on empirical information for model selection may lead to the wrong conclusion as it can be the result of capitalization on chance (Cudeck & Henly, 2003; Rindskopf, 2003). The challenge of choosing a model therefore lies in balancing parsimony, acceptable fit, interpretability and theory. The following describes two chi square statistics (Pearson chi square and Likelihood ratio chi square) and two descriptive fit indexes (AIC and BIC) that will be used here to determine statistical fit.

Pearson chi square (X^2)

This statistic tends to be conservative when sample size is large – it is difficult to reject significance. The likelihood ratio chi square is preferred over chi square (Langeheine et al., 1994) as the test statistic distribution is not well approximated by a chi square distribution. As the p-values are not being used in evaluating fit of the model, the

Pearson chi square evaluation, in conjunction with other goodness of fit criteria, provides valuable information.

Likelihood ratio chi square (L^2)

A perfect model fit is indicated when the L^2 statistic equals zero. The further away it is from zero, the greater the proportion of the data that remains unexplained by the model.

Akaike Information Criteria (AIC)

AIC penalizes total number of parameters required for model estimation. AIC can be used to compare non-nested models which means that specification of each model can differ. Models with lower AIC are preferred.

Bayesian Information Criteria (BIC)

BIC penalizes for both the number of parameters and total sample size. In contrast to the AIC, the BIC penalizes for additional parameters more severely than the AIC. Models with lower BIC are preferred.

The BIC and L^2 usually lead to similar conclusion (McLachlan & Peel, 2000), but the AIC and BIC may not always agree as the AIC tends to overfit and the BIC tends to underfit. The BIC is preferred over the AIC in contingency table analysis (Langeheine et al., 1994). In latent Markov literature, the BIC has been utilized extensively for model selection by Langeheine, Stern & Van de Pol (1994), Vermunt,Tran, and Magidson, (2008), Bartolucci, Pennoni, & Francis (2007). In sum, the BIC and AIC are relative fit measures which are useful in comparing competing models. These criteria will be used to guide in determining number of latent states and in selection of model.

3.8 Software

SLID retrieval software (SLIDRET) was used to extract variables of interest. Data preparation and descriptive statistics were done with Stata MP 10 and Mplus 6.1 on a 64 bit computer on multiple servers was used for data analysis.

Chapter 4

RESULTS

4.1 Descriptive Results

4.1.1 Characteristics of the study population

Table 5 presents weighted descriptive data on predictors, outcome and confounders of the participants from year 2002 to 2007 (n= 15 130, n=13 615 622 after weighting).

Table 5: Study participant characteristics for year 2002 to 2007 (n= 15 130, n=13 615 622 after weighting) (% unless otherwise stated)

	Year					
Characteristic	2002	2003	2004	2005	2006	2007
Male	49.16	49.18	49.43	49.35	49.55	49.55
SRH						
Excellent	30.6	27.87	27.3	24.5	22.00	21.6
Very good	35.22	38.03	37.12	37.57	38.18	38.1
Good	24.5	24.26	24.96	27.4	27.61	28.28
Fair	6.86	6.77	7.51	6.99	8.5	8.36
Poor	2.81	3.08	3.11	3.54	3.72	3.66
Income (adjusted for family size)						
Mean (SE)	34757.41 (440.08)	35813.13 (605.79)	37488.21 (536.38)	39410.33 (464.77)	40960.81 (412.16)	43469.29 (403.85)
Education						
< High school	14.92	14.86	14.66	14.46	14.5	13.95
High school	28.57	27.64	26.89	26.77	25.97	25.82
Non-university post-secondary	33.49	34.17	34.6	35.01	35.39	35.77
University	23.01	23.33	23.85	23.75	24.15	24.46
Employment status						
Employed	73.59	75.25	75.25	75.4	75.18	73.91
Unemployed: looking for work	4.08	3.73	3.32	2.61	2.14	2.46
Student	2.27	2.08	1.44	1.29	1.15	0.97
Out of labour force	18.07	17.38	18.04	19.18	19.68	20.49
Other	2.00	1.56	1.95	1.52	1.85	2.17
Immigrant	19.85	-	-	-	-	-
Visible minority status	12.5	-	-	-	-	-
Cohabitation status						
Married	71.53	72.48	72.58	72.58	73.47	73.00
Single	17.48	15.7	14.67	13.92	12.73	12.31
Other	10.99	11.82	12.75	13.5	13.8	14.69
January Proxy	26.8	31.69	31.06	31.11	29.79	29.91
May Proxy	29.91	30.76	-	-	-	-
Age						
Mean (95%CI)	41.80 (41.63-42.00)	42.78 (42.62-42.94)	43.76 (43.60-43.92)	44.74 (44.58-44.91)	45.73 (45.57-45.89)	46.71 (46.55-46.87)
Income info source						
Self-reported	13.64	6.97	5.34	3.7	3.04	3.02
Tax	70.08	82.68	85.01	87.27	87.59	88.18
Neither	16.28	10.36	9.64	9.02	9.37	8.8

Outcome

Between 88 and 90% of the participants rated themselves as having at least good health throughout six years. The proportion of people who rated themselves as being in excellent health declined by 9% by the end of the study period (30.6% to 21.6%). Conversely, there was a slight increase in proportion of people in all other health categories with an increase of 2.88% for very good, 3.78% for good, 1.5% for fair and 0.85% for poor status. This decrease in excellent health and increases in the other health categories is congruent with an aging-related decline in general health in this population. Once SRH was dichotomized into good (excellent, very good, good) and poor health (fair and poor), the temporal trend of decline in the highest category of health is attenuated (Table 6).

Table 6: Dichotomized SRH from 2002 to 2007 (n= 15 130, n=13 615 622 after weighting) (%)

	2002	**2003**	**2004**	**2005**	**2006**	**2007**
Good	90.32	90.15	89.39	89.47	87.79	87.98
Poor	9.68	9.85	10.61	10.53	12.21	12.02

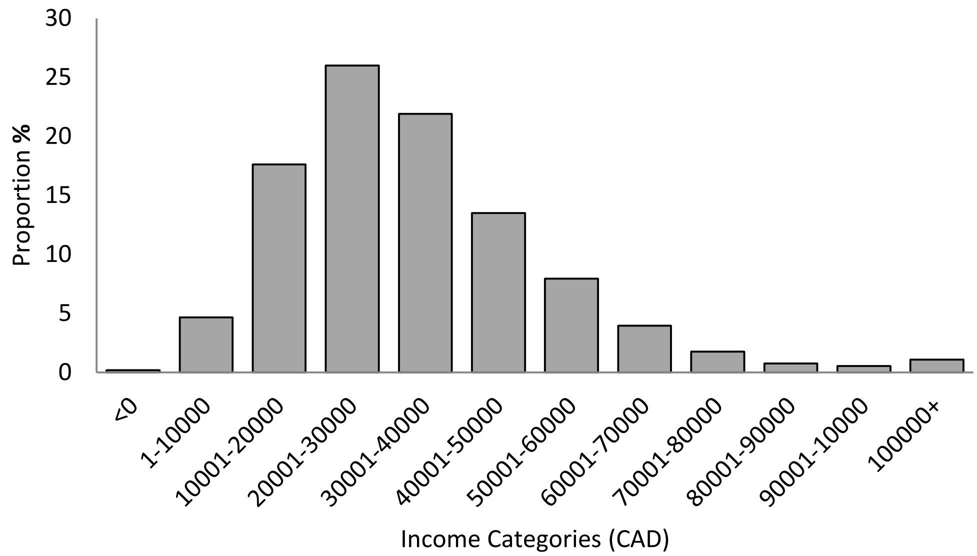
Note: Good includes the excellent, very good, good categories and poor health includes the fair and poor categories.

Predictors

In order to accurately capture purchasing power, income is adjusted to constant dollars (Table 7). This adjustment will allow us to examine temporal trends, while accounting for inflation. To calculate constant income, current dollars are multiplied by a ratio of baseline year rate: rate of year of interest. 2002 was chosen as the baseline year and set to 100%. The inflation rates were gathered from Statistics Canada (Statistics Canada, 2011a). For example, for 2003, constant income is calculated as $35813.13×(100/102.8) = $34837.68. As seen in Table 7, the average income adjusted for inflation rose by approximately $4000 over the six years.

Table 7: Mean income adjusted for inflation and family size

	2002	2003	2004	2005	2006	2007
Current income	34757.41	35813.13	37488.21	39410.33	40960.81	43469.29
Inflation rate	100%	102.8%	104.7%	107.0%	109.1%	111.5%
Constant income	34757.41	34837.67	35805.36	36832.08	37549.8	38985.91



The descriptive data on employment status and education status reflects our restriction of the sample to only those of working age. The highest level of education achieved stayed relatively stable over time with around 15% having less than high school, 26% having a high school degree, 35% having a non-university post-secondary degree and the remainder having a university degree or greater. With regards to employment status, about 75% of the respondents were employed at all waves and less than 3% were students. The proportion of people who were unemployed ranged from 2.14% (Y2006) to 4.08% (Y2002) and exhibited a general declining trend from 2002 to 2007. Close to 20% of the population were out of the labour force during this period which was an amalgamation of keeping house, caring for other family members including young children, retired, or have a long term illness or disability. The remainder (~2%) belonged to "other" category which included those who listed themselves as volunteers, had no main activity or classified themselves as having 'other activities'.

Confounders

The number of respondents who identify themselves as immigrant (19.85%) was greater than those who identified as a visible minority (12.5%). When weighted, 8.79% of the people listed themselves as an immigrant but not as a visible minority. Additionally, 1.66% categorized themselves as a visible minority but not as an immigrant (e.g., second generation children). The discordance between these two variables further confirms the importance of including both variables in the analysis. Immigration status and visible minority status were only collected at baseline. For cohabitation status, the majority of the respondents (~75%) were married, followed by single and other (common-law, widowed, and separated). Close to 30% of participants responded via proxy for both January and May surveys. The difference in income and SRH for those who opted to report via proxy is shown in the Methods section (section 3.3.2). As the May survey was combined with the January survey starting in 2004, the proxy information is not available after that year. The average age of responders in 2002 was 41.8 and increased by one year at each subsequent wave as expected. See Figure 9 for the distribution at baseline.

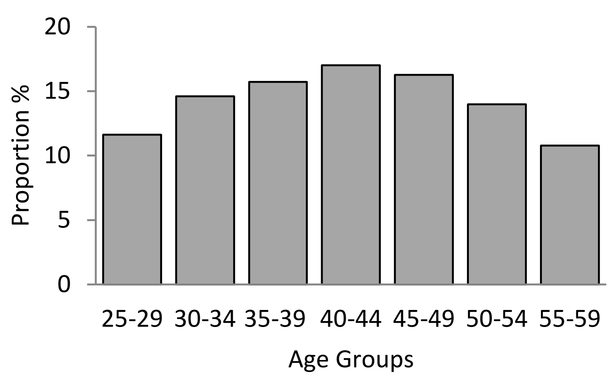


Figure 7: Age distribution

In 2002, 70% of the participants gave permission for Statistics Canada to link their data to their tax files, a proportion that increased to nearly 90% in 2007. The difference in income between self-reported and tax derived was found to be significant, as discussed previously in the Methods section (section 3.3.1).

4.1.2 Self-Rated Health Validation

SRH was validated within the sample by comparing its distribution with another health indicator, self-reported functional limitation. The functional limitation indicator was constructed using the following questions:

1. Does a physical condition or mental condition or health problem reduce the amount or the kind of activity you can do in *other activities*, for example, transportation or leisure?
2. Does a physical condition or mental condition or health problem reduce the amount or the kind of activity you can do at a *job or business or at school*?
3. Does a physical condition or mental condition or health problem reduce the amount or the kind of activity you can do at *work*?
4. Does a physical condition or mental condition or health problem reduce the amount or the kind of activity you can do at *home*?
5. Do you have any difficulty hearing, seeing, communicating, walking, climbing stairs, bending, learning or doing any similar activities?

The response categories for all five of these questions were: yes, sometimes; yes, often; and no. If the respondent answered 'yes, sometimes' or 'yes, often' to any of these questions, the person was flagged as having a functional limitation.

Cross tabulation of SRH status with functional limitation is shown below for year 2003 (Table 8). The baseline year (2002) is not shown as the counts in some cells were below 15 and is thereby restricted from release from Statistics Canada.

Table 8: Distribution of self-rated health by functional limitation, Y2003 (not weighted)

	Functional limitation % (n)		
SRH	Yes	No	Total
Excellent	6.6 (235)	93.4 (3312)	3547
Very good	13.7 (689)	86.3 (4357)	5046
Good	29.0 (919)	71.0 (2249)	3168
Fair	72.5 (689)	27.5 (262)	951
Poor	92.8 (361)	6.4 (25)	389
Total	2893	10205	13098

The first column shows an increasing proportion of respondents having functional limitation as the health status worsens. For instance, among those who categorized themselves as having poor health status, 92.8% also reported a functional limitation. Conversely, of the people who rated themselves as having excellent health, the majority (93.4%) reported to be free from a functional limitation. A similar pattern of association is found for all years (Appendix D).

4.1.3 Distribution of covariates by SRH status

The baseline year, 2002, was analyzed to examine the distribution of covariates in those who had good health versus poor health (Table 9). The purpose was to confirm the association of confounders with health and also to examine the general direction of the association.

Table 9: Relation between variables and dichotomized SRH for the baseline year (Y2002) (n= 15 130, N=13 615 622 after weighting)

	Good Health (%)	Poor Health (%)
Sample	90.32	9.68
Male	49.99	44.74
Income		
Mean (95%CI)	35332.32 (34747.49-35917.15)	25633.24 (24148.12-27118.36)
Education		
< High school	13.07	31.94
High school	28.09	26.33
Non university post-secondary	24.1	30.51
University	24.75	11.21
Employment status		
Employed	77.14	38.79
Unemployed: looking for work	3.87	5.4
Student	2.49	1.45
Out of labour force	14.63	51.62
Other	1.87	2.74
Immigrant	19.13	20.72
Visible minority status	12.28	11.58
Cohabitation status		
Married	73.69	61.18
Single	16.67	18.27
Other	9.63	20.55
January Proxy- Yes	27.52	23.05
January Proxy – No	72.48	76.95
May Proxy- Yes	32.55	30.08
May Proxy- No	67.45	69.92
Age		
Mean (95%CI)	41.33 (41.16-41.50)	46.24 (45.50-46.97)

The majority of the respondents were healthy (90.32%) with only 9.68% reporting poor health. The group of respondents in good health was almost even split between males and females, while the group in poor health counted a slight majority of women. . Income, education and employment status were all associated with health in the expected direction. Those in poor health had a lower average income ($\bar{x}$= \$25 633.24), counted a greater proportion of those who had less than a high school education, and were more likely to belong in the out of the labour force category compared to those who categorized themselves as being in good health.

While visible minority status should be associated with a poorer health status, in this sample, it tended to be associated with better health. Proxy answers tended to report better health, and more so in January than in May. As expected, those who rated themselves as having poor health status had a higher mean age.

4.2 Analytical Results

4.2.1 Manifest Markov Model (M)

The transition probabilities from time x to time x+1 are shown in Table 10.

Table 10: Heterogeneous transition probability for M model (n= 15 130, n=13 615 622 after weighting)

		t+1	
		Transition probabilities (%)	
t	Health Category	good	poor
1	good	95.5	4.5
	poor	39.7	60.3
2	good	95.1	4.9
	poor	36.0	64.0
3	good	95.5	4.5
	poor	37.0	63.0
4	good	94.4	5.60
	poor	32.8	67.2
5	good	94.7	5.3
	poor	37.7	62.3

For the transition from time 1 to time 2, 95.5% of those who were in good health at time 1 remained in good health at time 2, while 4.5% transitioned to poor health at time 2. In contrast, of those who started in poor health, 39.7% transitioned into good health while 60.3% remained in poor health. Each transition probability row must be equal to 1 as the probabilities are conditional on the previous category that the person belonged to (e.g., $\tau_{j|i}^{12}$). The same pattern of interpretation can be used for transition periods 2→3, 3→4, 4→5 and 5→6.

The transition probabilities are similar for all five transition matrixes across waves. The proportion of those who stay in good health ranges from 94.4% to 95.5%, while the proportion of those who remain in poor health ranges from 60.3% to 67.2%. The same narrow range is observed with those who make a transition, as the proportion of people who transition from poor health to good health ranges from 32.8% to 39.7% and the proportion of those who transition from good health to poor health ranges from 4.5% to 5.6%. As such, homogenous transition probabilities were assumed, affording more degrees of freedom. That is, $\tau_{j|i}^{12} = \tau_{k|j}^{23} = \tau_{l|k}^{34} = \tau_{m|l}^{45} = \tau_{n|m}^{56}$. Table 11 shows the homogeneous transition probabilities.

Table 11: Homogenous transition probability for M model (n= 15 130, n=13 615 622 after weighting)

		t+1 Transition probabilities (%)	
	Category	good	poor
t	good	95.5	4.9
	poor	36.6	63.4

These transition probabilities indicate that 95.5 % of the respondents rated themselves as good health if they were in good health in the previous time period, while 4.9% of respondents who rated themselves as being in good health in the previous time period transitioned into poor health at the subsequent time period. In contrast, 36.6% of those who were in poor health status transitioned into good health status by the next time period, while 63.4% of those who were in the poor health category stayed in the poor health category through the next time period.

4.2.2. Latent Markov Model (LM)

Table 12 compares the transition probabilities between the manifest Markov and the latent Markov model.

Table 12: Comparison of manifest and latent Markov chain model with homogeneous transition probabilities, 2 classes (n= 15 130, n=13 615 622 after weighting) (%)

		Manifest Markov chain t+1		Latent Markov chain t+1	
		good	poor	good	poor
t	good	95.5	4.9	98.4	1.6
	poor	36.6	63.4	5.9	94.1

A comparison of the estimated transition probabilities in the two models reveals an overestimation of the transition from poor health to good health in the manifest Markov model. In the manifest Markov chain 36.6% of the respondents are categorized as transitioning from poor health to good health whereas in the latent Markov model this number is reduced to 5.9%. As described in the methods section, the latent Markov model explicitly models the measurement error. Measurement error arises from the fact that the outcome is measured imperfectly. That is, health is not the only factor influencing the decision on self-rated health status. Much of the research assumes that the measured version of the actual latent status is similar and therefore treats the indicator as the actual (Y*=Y). However, there are other factors that may influence how a person answers self-rated health. These factors include misinterpretation of the question, falsifying the answers, and data entry errors (Hernan, 2009). When these 'external' factors are modeled as unobserved heterogeneity, the discrepancy between the manifest and latent Markov chain arises.

4.2.3 Model Selection

Table 13 presents the likelihood ratio, chi square, BIC, AIC and the sample size adjusted BIC (SSABIC) for various extensions of the Markov model fit. For each model, both time homogeneous transition probability and heterogeneous probability were fit. A number of assumptions were imposed on all models. The first assumption is the autoregressive nature where the current state is only dependent on one previous time period. The second assumption is holding measurement error constant over time (applicable to models that involve latent variables), and the third assumption relates to the second assumption in that the measurement error was held equal across chains for models that have multiple chains to account for heterogeneity (Kaplan, 2008; Van de Pol & Langeheine, 1990).

For ease of interpretation,

Table 14 was constructed by ranking the fit criterions from smallest (1) to largest (10). For example, in the first row, the manifest Markov model with time homogenous transition probabilities had the largest L^2, X^2, and AIC values (denoted by 10^{th} ranking) and second largest BIC and SSABIC values (denoted by 9^{th} ranking). As mentioned in methods, lower values indicate a better fit. The last column was created as equal to the sum of all rankings across the fit criterions. Therefore the sum can be used to guide model selection as higher ranking denotes lower values of fit criterions.

Table 13: Fit statistics for various Markov models

Model	Transition probabilities time homog.	df	L^2	X^2	BIC	AIC	SSABIC
Manifest Markov	Yes	60	835.6	1998.4	43419.8	43396.9	43410.2
	No	52	808.1	1982.5	43455.9	43372.1	43421.0
Latent Markov	Yes	58	185.4	385.4	40826.1	40787.9	40810.2
	No	50	178.8	357.5	40885.7	40786.5	40844.4
Mixed latent Markov							
2 chain	Yes	50	80.6	105.6	40393.7	40294.6	40352.4
	No	42	72.8	102.8	40464.1	40303.9	40397.3
3 chain	Yes	40	47.78	61.4	40427.9	40252.5	40354.8
	No	-	32.4	35.9	40838.0	40327.1	40625.1
3 chain- mover stayer stayer	Yes	42	51.4	68.2	40414.8	40254.7	40348.0
	No	38	45.5	62.7	40443.7	40253.1	40364.3

Table 14: Ranking of fit criterions. 1= Lowest value, 10= Highest value

Model	Transition probabilities time homog.	df	L^2	X^2	BIC	AIC	SSABIC	Sum of Rankings
Manifest Markov	Yes	60	10	10	9	10	9	48
	No	52	9	9	10	9	10	47
Latent Markov	Yes	58	8	8	6	8	7	37
	No	50	7	7	8	7	8	37
Mixed latent Markov								
2 chain	Yes	50	6	6	1	4	2	19
	No	42	5	5	5	5	5	25
3 chain	Yes	40	3	2	3	1	3	12
	No	-	1	1	7	6	6	21
3 chain- mover stayer stayer	Yes	42	2	4	2	3	1	12
	No	38	4	3	4	2	4	17

The initial model, the manifest Markov model, fit poorly whether homogenous or heterogeneous transition probabilities were assumed. Allowing for heterogeneous transition probability results in a slightly better fit but still ranked fairly low in comparison to other models. The initial model was then extended to include change at the latent level. Under the latent Markov model, there is a decrease in all criteria. Thus, allowing for measurement error improves the fit of the single Markov chain models.

Models that allow for heterogeneity tended to fit better. Various types of mixed latent Markov models were tested:

- 2 chain: This model assumes that the population can be divided into two Markov chains. The transition probabilities are allowed to vary between chains but not within chains.
- 3 chain: Similar to the two chains, three chains assumes that the population consists of three sub-groups. Transition probabilities are not fixed for all chains but the homogenous transition probabilities within each chain was imposed.
- 3 chain- mover stayer stayer: This model has the same premise as the three chain except that the transition probabilities for the stayer chains are fixed. The transition probability of the stayer chain is set to 1 for staying in the initial category and 0 for moving into another category as shown by the two transition matrix below.

A	Good	Poor
Good	p (good|good) =1	p (poor|good) = 0
Poor	p (poor|good) =0	P (poor|poor) = 0

B	Good	Poor
Good	p (good|good) =0	p (poor|good) = 0
Poor	p (poor|good) =0	P (poor|poor) = 1

A depicts the four transition probabilities for those in constant good health and **B** depicts the four transition probabilities for those in constant poor health. The third chain is the mover chain and includes those who fluctuate between the two health states. The mover chain's transition probability is not fixed and is estimated. In terms of health trajectories, this translates to constant good health (good health status for all six time periods), constant poor health (poor health status for all six time periods) and those who fluctuate (e.g., good, poor, poor, good, poor, good).

The BIC, AIC, and SSABIC values have narrow range within the mixed latent Markov models. The BIC ranges by 444.3 (40838-40393.7), AIC ranges by 72.4 (40327.1-40252.5) and the SSABIC ranges by 277.1 (40625.1-40348.0). The two most viable models are 3 chain with time homogenous transition probability and 3 chain- mover stayer stayer with time homogenous transition probability.

With any model modification, there should be substantive interpretation of the added parameter and theoretical justification should be made. Selecting models based strictly on empirical data may lead to non-interpretable results. By taking into consideration these caveats, the choice of final model should balance interpretability, fit and parsimony. The 3 chain mixed latent Markov – mover stayer stayer is an attractive choice in terms of interpretability and theoretical justification. The proposition that the population can be separated into three groups based on their health trajectory seems warranted as there are groups who will remain in constant poor health (e.g., chronic long-term disability), constant good health and those who fluctuate (e.g., acute conditions or health shocks). The fit criterion also indicates that it is one of the better fitting models. Therefore, we conclude that a mixed latent Markov with three chains (two stayers and one mover) with homogenous transition probabilities within each chain and constant measurement error at each time period is the most appropriate model for the data.

4.2.4 Mixed Latent Markov model (MLM)

This section shows the results from the mixed latent Markov model (MLM) with the predictors and confounders added to the model. The baseline characteristics of the participants are shown by trajectory membership in Table 15. The majority (78.05%) belonged to the constant good health trajectory, followed by constant poor health trajectory (13.95%) and the remainder to fluctuating health (7.99%). As found in the previous descriptive table using manifest self-rated health, a similar socioeconomic gradient was found with the health trajectories.

Table 15: Study participant characteristics in 2002 (at baseline) by health trajectory

	Health trajectory (%)		
Characteristic	Constant good health	Fluctuating health	Constant poor health
Overall	78.05	7.99	13.95
Male	80.69	7.15	12.17
Female	75.47	8.82	15.71
Education			
< High school	57.38	21.45	21.16
High school	76.70	8.17	15.13
Non-university post-secondary	80.37	6.26	13.37
University	89.38	1.77	8.85
Employment status			
Employed	84.09	2.66	13.25
Unemployed: looking for work	66.40	10.05	23.55
Student	81.46	3.37	15.17
Out of labour force	57.09	27.88	15.02
Other	65.03	25.53	9.44
Income			
0-10000	40.44	26.36	33.20
10K-20K	66.36	15.40	18.23
20K-30K	77.97	7.49	14.55
30K-40K	82.91	6.09	11.00
40K-50K	85.13	2.96	11.91
50K-60K	88.90	3.03	8.07
60K-70K	90.15	1.81	8.04
70K+	93.53	1.49	4.98
Immigrant	72.61	7.65	19.74
Non-immigrant	79.36	8.07	12.56
Visible minority	72.84	5.81	21.35
Non visible minority	78.78	8.30	12.92
Marital status			
Married	80.97	5.65	13.38
Single	73.57	9.83	16.60
Other	65.34	20.98	13.68
Age			
25-29	92.21	1.97	5.83
30-34	87.09	2.23	10.68
35-39	85.50	4.61	9.89
40-44	75.23	7.36	17.41
45-49	74.15	9.85	16.01
50-54	67.84	13.97	18.19
55-59	63.08	17.86	19.06

Indeed, the proportion of respondents in the stable good health trajectory increased along with increasing income and education, as well as with employment status. In those with less than high school education, 57.38% were in the stable good health trajectory whereas 89.38% of those who had at least a university degree were in the stable good health trajectory. With regards to employment status, the greatest proportion of stable good health was found in those who were working (84.09%) followed by student status (81.46%). Those who were unemployed had the greatest proportion of people who were in stable poor health (23.55%). Over a quarter of those who were classified as ‘out of labour force’ and ‘other’ were in the fluctuating health category. Finally, an income gradient was evident as increasing income was associated with greater proportions of respondents in the stable good health trajectories, but decreasing proportions of both fluctuating and stable poor health trajectories

The following table (Table 16) presents the relative risk ratios (RRR) with its corresponding 95% confidence intervals from the MLM model. The RRR were obtained by exponentiation of the multinomial logit coefficients (Gould, 2000) . The RRR calculates the effect of the variable in question on the probability of belonging to the comparison group (constant good health trajectory or fluctuating health trajectory) in comparison to the reference group (poor health trajectory). In general, a RRR greater than one indicates that there is an increased risk of membership in the comparison group compared to the reference group as the predictor changes. A RRR less than one indicates the opposite in that the risk of membership in the comparison group is lessened, that is, the individual is more likely to belong to the reference group as the predictor changes. Changes in the predictor are shown by comparison of each level to the reference category. For example, for education, the reference category are those who have less than a high school education, so the changes in education are high school compared to less than high school, non-university post graduate compared to less than high school, and at least university degree compared to less than high school.

Table 16 Relative risk ratios and 95% CI for a multinomial analysis of main SEP predictors for the health trajectories

	Constant good health (ref. constant poor health) RRR (95% CI)	Fluctuating health (ref. constant poor health) RRR (95% CI)
Income		
0-10000	Reference	Reference
10K-20K	**3.00 (2.39-3.74)**	**1.60 (1.22-2.10)**
20K-30K	**4.20 (3.36-5.24)**	**1.49 (1.13-1.98)**
30K-40K	**5.43 (4.30-6.87)**	**1.75 (1.29-2.40)**
40K-50K	**5.27 (4.10-6.78)**	1.21 (0.83-1.77)
50K-60K	**7.49 (5.56-10.08)**	**1.93 (1.24-3.03)**
60K-70K	**7.15 (4.99-10.24)**	1.08 (0.57-2.07)
70K+	**12.76 (8.44-19.30)**	0.96 (0.44-2.11)
Education		
< High school	Reference	Reference
High school	**1.78 (1.43-1.91)**	**0.67 (0.57-0.85)**
Non university post-secondary	**1.78 (1.54-2.04)**	**0.67 (0.54-0.81)**
University	**2.82 (2.35-3.39)**	**0.39 (0.29-0.54)**
Employment status		
Unemployed: looking for work	Reference	Reference
Employed	1.26 (1.00-1.63)	**0.46 (0.32-0.65)**
Student	1.29 (0.83-2.00)	1.64 (0.82-3.30)
Out of labour force	1.26 (0.97-1.63)	**7.18 (5.01-10.28)**
Other	**1.88 (1.19-2.98)**	**5.65 (3.28-9.72)**

Note: Results are adjusted for sex, age, immigration status, visible minority status, and marital status, all measured at baseline. SEP was also measured at baseline. The reference class in this multinomial estimation is the constant poor health trajectory.

Contrast 1: Constant good health trajectory to constant poor health trajectory

For income, the \$0 -10000 income category was used as the reference category. For a change from \$0-10000 to \$10K-20K, holding all other variables constant, the relative risk ratio of belonging to the constant good health trajectory versus the constant poor health trajectory is 3.00 (95% CI: 2.39-3.74). For the change from \$0-10000 to \$20K-30K, holding all other variables constant, the relative risk ratio of belonging to the constant good health trajectory versus the constant poor health trajectory is 4.2 (95% CI: 3.36-5.24). The RRR increases at each subsequent income bracket compared to the first income bracket in a dose-response pattern, indicating that the participants were more likely to belong to the constant good health trajectory rather than the constant poor health trajectory with each increase in income.

The reference category for education was 'less than high school' achievement. Those who achieved a high school diploma were 1.78 (95% CI: 1.43-1.91) more likely to be in the good health trajectory rather than the poor health trajectory. The ratio of belonging to good health trajectory increases to 2.82 (95% CI: 2.35-3.39) when comparing those who have at least a university degree compared to those who have less than high school.

The only significant point estimate for employment status was the other category which included those who were volunteering, had no main activity or had listed 'other' category. As the 95% confidence interval straddles one for the employed, student and out of labour force, the estimates are not significant.

Contrast 2: Fluctuating health trajectory to constant poor health trajectory

As evidenced by the wide confidence intervals for some estimates, caution should be taken with interpreting the fluctuating health trajectory as the cell counts for certain items were less than 30. In addition to the small cell counts, fluctuating health trajectory exhibits greater heterogeneity as it groups together those who have at least one good and at least one poor health period regardless of length of the good or poor health period. The fluctuating health trajectory also exhibited a gradient for the first three comparison categories. As expected, the income gradient for the fluctuating health trajectory was not as strong as the constant good health trajectory. For example, the change from $0-10000 to $10-20K had a RRR of 3 (95% CI: 2.39-3.79) for constant good health versus constant poor health whereas the same jump had a relative risk ratio of 1.6 (95% CI: 1.22-2.1) for fluctuating health versus constant poor health.

In comparison of fluctuating health trajectory to the constant poor health trajectory, the RRR is less than one in all cases (high school, non-university post-secondary, and university) compared to less than high school. This suggests that an individual with greater educational achievements is more likely to belong to the constant poor health trajectory.

With regards to employment status, the relative risk ratio is less than one (RRR=0.46 (95% CI: 0.32-0.65)) for the employed relative to the unemployed. This indicates that employed respondents were more likely than the unemployed to belong to the constant poor health trajectory rather than the fluctuating health trajectory. However, those who were out of the labor force were significantly more likely to be in the fluctuating health trajectory compared to those who were unemployed (RRR=7.18 (95% CI: 5.01-10.28)).

Chapter 5

DISCUSSION

In this chapter we explore the implications of the analysis presented in the previous sections. The primary findings were: 1) more than three quarter of Canadians of working age were in constant good health trajectory 2) income and education were statistically significant in distinguishing membership in the constant good versus poor health trajectory 3) employment status at baseline did not appear to play a role in determining membership in the constant good health trajectory.

5.1 Descriptive analysis of health trajectory

The majority of Canadians of working age were in the constant good health trajectory (78.05%), followed by the constant poor health trajectory (13.95%) and the fluctuating health trajectory (7.99%). These results support those of a previous international study by McDonough et al. (2010), who also found that the majority of respondents in the U.S. and Great Britain were in the constant good health trajectory (38.1% and 47.2% respectively). However, they also found a much larger proportion of respondents fit in the fluctuating health category (43.9% in the U.S. and 38.2% in Great Britain). The discrepancy in proportion of respondents in the different trajectories may be due to differences in methodology used to determine good and poor latent health class.

The databases used by McDonough and colleagues had a longer follow up periods: the U.S. database (PSID) ranged 13 years from 1990 to 2003 and the Great Britain database (BHPS) also had a follow up of 13 years 1991 to 2004. The SRH information was analyzed every two years due to the inconsistency in collection of information. In addition, the authors did not dichotomize the observed SRH responses into good and poor health, thus substantially increasing the likelihood of movement from one of the five categories to another response category.

5.2 Predictors of health trajectory

To what extent does health trajectory membership vary depending on SEP?

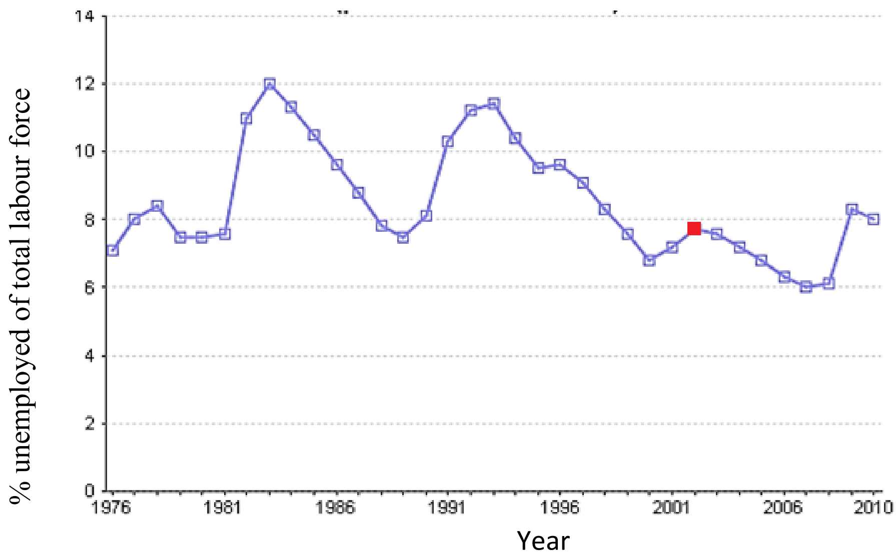
The association of three SEP indicators - income, education and employment status - with health trajectories was assessed. Our findings suggest the existence of a dose-response relationship, or gradient, between income and membership into good health trajectory rather than the poor health trajectory. Similarly, increasing levels of educational attainment translated into greater likelihood of belonging to the good health trajectory over the poor health trajectory. Finally, employment status was not found to be significant for determining membership in the constant good health trajectory versus the constant poor health trajectory. Concerning the fluctuating health versus constant poor health trajectory membership, there was also some association with higher income predicting membership in the fluctuating health trajectory, though a dose-response relationship could not be established. Being out of the labor force (vs. being unemployed) was also associated with being in the fluctuating health category rather than in the constant poor health trajectory. In contrast, higher levels of education and employment (vs. unemployment) were associated with constant poor rather than fluctuating health.

One of the strengths of this study was in the measurement of income. Indeed, by linking to income tax administrative records, the SLID allows access to income information that is both more complete and less biased than typical self-reported income information. Bias in income reporting can stem from several sources, including recall bias and social desirability, both of which are likely to be at a minimum with administrative records, as there are financial and legal penalties for misreporting one's income tax. Furthermore, administrative records more accurately and thoroughly record the multiplicity of sources of income (see Appendix A), including transfers, which may otherwise go unreported with simple income self-reports. Finally, as this information includes income tax paid and social transfers received, we can more accurately assess disposable income, which is a better predictor of the material conditions experienced by the household. In agreement with previous studies which have shown income to be a predictor for various health measures (Benzeval & Judge, 2001; Gunasekara et al., 2011; Hajat, Kaufman, Rose, Siddiqi, & Thomas, 2010; Mackenback et al., 2005), income was found here to have an independent effect on determining membership in health trajectories. This relationship is graded with higher income groups having a higher relative risk ratio of belonging to the constant good health trajectory, and is consistent with other studies that have also reported a gradient between income and health (Ecob & Davey Smith, 1999; Stronks, van de Mheen, van den Bos, & Mackenbach, 1997).

The findings of this study on education and health conform with the general literature that has shown the association between education and mortality (Davey Smith, Neaton, Wentworth, Stamler, & Stamler, 1996; Mustard, Derksen, MBerthelot, & Roos, 1998) and cardiovascular disease (Kaplan & Keil, 1993). The association between education and health is partially shaped by employment opportunities and earning potential which in turn allows for access to goods and services. Finally, education is the most exogenous variable of the three SEP measures as it is most likely to have been achieved prior to health conditions (Duncan et al., 2002).

Contrary to previous studies (Benavides et al., 2000; McLeod & Kessler, 1990; Stronks et al., 1997), employment status was not found to be a predictor of membership in health trajectories after accounting for education and income and a host of other confounders. There are several explanations for the non-significant point estimates for employment. First, relative to the U.S., from where most of the relevant studies came, the greater generosity of the Canadian government's unemployment program and associated social policies may act as a buffer against negative health consequences. That is, if loss of employment is associated with poorer health due to loss of income and thereby loss of material goods, this effect may be minimized with benefits such as employment insurance which provides up to 55% of average weekly earnings if the person has lost his/her job through no fault of his/her own (Service Canada, 2010). Access to publicly subsidized health insurance may also prove to mitigate the negative consequences of unemployment, relative to the U.S., where health insurance status strongly depends on employment status. Second, this result may be due to the measurement of employment status. The question used to capture employment status asked for a respondent's main activity at the end of the reference year. In order to capture their main activity during the entire previous period, we would have had to rely on other indicators, such as the total hours worked in the previous year. Third, although much of the literature points towards employment status having a significant effect on health, a consideration of unemployment in the current context should be given. Indeed, unemployment levels in Canada are not constant (Figure 8). Panel 4 of the survey covered 2002 to 2008, which spans a period with one of the lowest unemployment rate in Canada from 1976 to 2010. These low unemployment rates during this period may indicate that those who are unemployed are a sub-group of particularly "resistant" unemployed individuals. In addition, changes in the labor market having to do with the proliferation of non-standard work such as temporary and contract-based jobs may also limit the comparison with other studies having been done in previous periods (Quesnel-Vallee et al., 2010).

Figure 8: Unemployment rate in Canada from 1976 to 2010



Source: Statistics Canada, Labour force survey estimates (LFS), supplementary unemployment rates by sex and age group, annual (CANSIM Table 282-0086) Ottawa: Statistics Canada, 2011

5.3 Limitations

5.3.1 Bias in Observational Research

Selection bias

A major methodological concern for longitudinal surveys, especially those with heavy response burden, is attrition. Attrition refers to loss of respondents due to refusal, loss of contact, or death. If the lost cases are randomly distributed in the population, then the result is loss in power but this does not lead to bias. However, if the sample attrition is systematic and certain groups of people are more likely to drop out, then bias may be introduced. It is plausible to assume that at least some of the drop out cases in our sample are correlated with health. In fact, the descriptive analysis shows a decline in the proportion of respondents who rated themselves as being in excellent health over the six years. However, it is not possible to discern if this temporal trend reflects attrition due to health concerns or a real change in health decline.

Cross-sectional survey non-response may also lead to selection bias. In this analysis panel, cross-sectional response rates averaged 75.2% over the six years. As around a quarter of the target population refused to be in the survey and no information was collected on the non-responders, it is not possible to ascertain the difference between these two groups. However, with the use of longitudinal weights, the sample is weighted to represent the population of Canada's ten provinces at the baseline year, thus mitigating the effects of unit non-response.

Thirdly, due to analytical limitation, the sample was restricted to those who had full covariate information. In this context, systematic attrition, response rates, and sample restriction may all contribute to selection bias where those who remain in the analytic sample may not be representative of the target population. As shown in descriptive table in Table 7 from year 2002 to 2007, income rose by $4000 after taking into account family size and inflation. The highest level of education achieved was relatively stable over time with a slight decrease in the proportion of people who had less than a high school or only a high school diploma, suggesting age-based attrition by mortality. The proportion of people who were employed was stable for all waves whereas those who were unemployed dropped from 4.08% to 2.46% and the percentage of respondents in out of labour force category increased from 18.07% to 20.49%. A comparison of baseline characteristics to the characteristics of those who were in the final year of the survey show those who remained in the sample had a higher income, higher level of education and were less likely to be unemployed at baseline. In sum, as those of lower SEP were more likely to drop out of the survey, our estimates should be conservative.

Confounders

Confounders are variables that are associated with the exposure and causally associated with the outcome. Furthermore, confounders cannot be in the mediating pathway between the exposure and the outcome. The measured confounders are: age, sex, immigrant status, visible minority status and marital status (Figure 9). As a theoretical framework is imposed on the data in structural equation modeling, confounders are chosen a priori based on previous research and satisfy the three criteria described above.

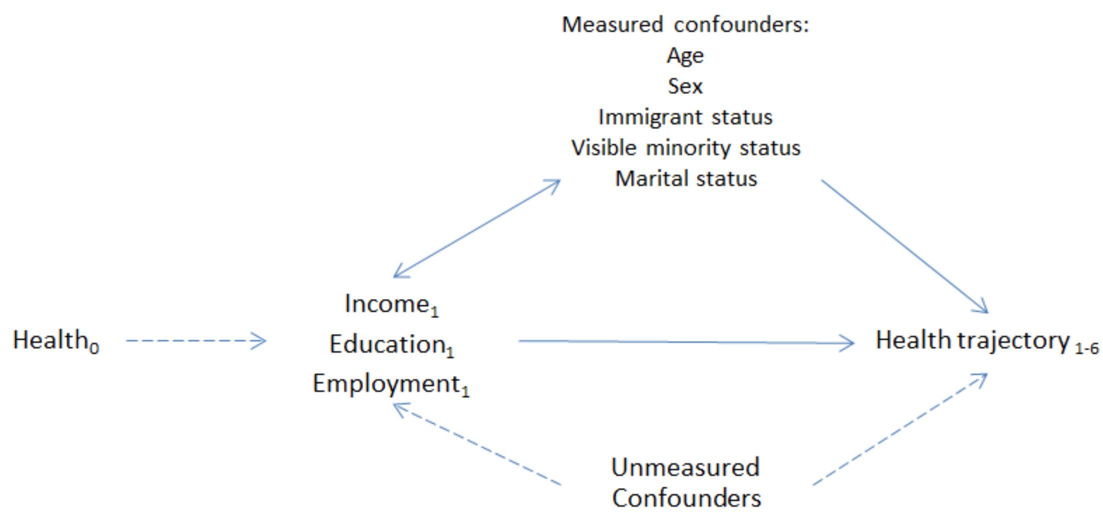


Figure 9: Graphical model of biases in socioeconomic status and health trajectory

It is important to take into account confounders as their effects are removed when estimating the association between exposure and the outcome. Unmeasured confounders pose a greater challenge especially in examining the effect of socioeconomic status on health as there are multitudes of confounders such as genetics and obesity. Taking for instance obesity, Gortmaker (1993) found that obesity both influences socioeconomic status (as measured by income, number of years of education, and household poverty) and has been found to be a risk factor for a wide variety of health outcomes (Health Canada, 2006). The inability to take into account unmeasured confounders, like obesity, may cause spurious associations to be found between SEP and health. Despite the impact that unmeasured confounders can have, all the relevant variables are seldom available in a single survey and therefore the ability to account for these confounders are limited. The effects of previous health status ($Health_0$) which includes parental SEP and childhood SEP can also impact adult SEP standings and is discussed in the next section.

Health Selection

Does being in higher SEP make you healthy, or does being healthy make you more likely to be in higher SEP?

The association between SEP and health is complex as being in low SEP (e.g., poverty) can cause poor health but being in poor health can also lead to achieving lower SEP. These two scenarios present the two competing explanations for the association between SEP and health: health selection and social causation (Figure 10).

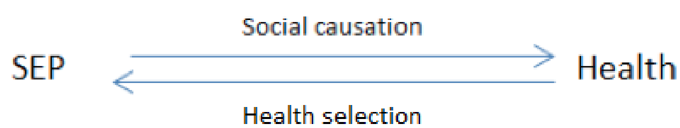


Figure 10: Bidirectional association between SEP and health

Health selection argues that characteristics associated with poor health leads to a person's failure to rise out of low SEP or to belong in low SEP. For example, analyses of the German Socioeconomic Panel (GSOEP) found a positive association between experiencing 'health shocks' and exit from labour force participation (Riphahn, 1999). *Social causation*, on the other hand, hypothesizes that it is the factors associated with a membership in certain SES that causally affect health outcomes.

Many studies have examined the validity of these two pathways using a variety of methods including instrumental variables, structural equation models, and life course theory (Haas & Fosse, 2008; Lynch, Kaplan, & Shema, 1997; Mulatu & Schooler, 2002). All these studies point towards the existence of both social causation and social selection with the strength of each direction dependent on the outcome examined and the group characteristics. McDonough and Amick (2001) argue that these two pathways are not mutually exclusive and theorized health selection as a social process. They analyzed a prospective study and hypothesized that an individual's decision to exit the labour force upon ill health is dependent on the very position that the individual occupies in the social hierarchy (e.g., level of work experience, level of education). Thus, the primacy of SEP and health continues to be debated.

One of the advantages of longitudinal data is being able to examine the temporality of the association. In this thesis, social causation was used as the primary explanation as baseline year information (SEP_1) was used to predict health trajectory membership ($Health_{1\text{-}6}$). Furthermore, from the very nature of the study design, the income information preceded the health information as the tax records refer to the previous fiscal year. Similarly, as most of our respondents had completed their education at the outset of the survey (only about 2% were reported as being still in school), education should also be seen as temporally preceding health. Thus, we did not explicitly model health selection processes. Furthermore, as the baseline year characteristics were used to predict long-term self-rated health trajectory, the causal dynamics existing between SEP and health (where retroaction loops might develop) were not explored. Finally, this baseline information did not allow the estimation of the effects of a number of material and lifestyle factors such as material deprivation and negative emotions that may have acted as mediators of this relationship (Gallo & Matthews, 2003; Lynch, Davey Smith, Kaplan, & House, 2000).

5.3.2 Methodological Considerations

Self-Rated Health

SRH was dichotomized into good and poor health. The good health category combined excellent, very good and good whereas the poor health category combined fair and poor categories. A sensitivity analysis was done to assess the effects of choosing a different threshold of good health. Good health was re-categorized as containing only the top two categories (excellent and very good) and poor health was re-categorized as containing the bottom three categories (good, fair, and poor). The results are provided in Appendix F. The trend for income remained, although it was attenuated in comparison to the original cut-off.

Model Specification Error

In structural equation modeling, the model is imposed on the data based on previous research and theories. Due to the assumption that the model has been specified correctly, it leans strongly towards confirmatory research as opposed to exploratory. Furthermore, the omission of relevant variables may lead to bias in estimates and misinterpretation (Kaplan, 2000).

Notably, different extensions of the Markov model exist that were not tested here. One variant of the Markov model is the partially latent Markov model. This model assumes multiple chains but the stayer chain does not account for measurement error. In other words, the respondents in the stayer chain are assumed to answer correctly, while those in the mover chain are corrected for measurement error (hence, partially latent). This model could be useful if there was evidence to suggest that those in the stayer chains answer with perfect reliability, an assumption we did not make here. Furthermore, the pathway chosen for this thesis was the first-order latent Markov model (Figure 11, A) where the present status is dependent only on one previous time status. Other valid models include the Socrates Markov model (Figure 11, B) (Hagenaars, 1993; Hagenaars & McCutcheon, 2002).

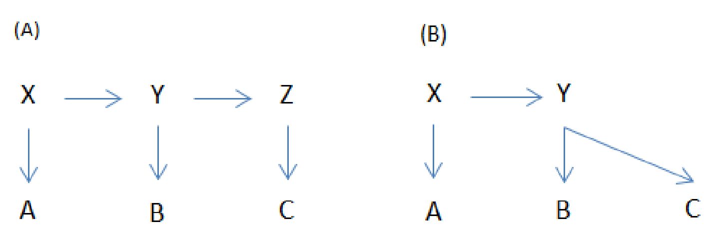


Figure 11: For a three wave longitudinal data, Latent Markov model (A) and Socratic Markov model (B). Where A B C represent manifest variables at time 1, 2, 3 respectively and X Y Z represent latent variables at time 1, 2, 3 respectively.

The ‘socratic’ latent change model (B) uses two latent variables to capture change in three manifest indicator periods. The latent variable at the second period (Y) is composed of both the second (B) and third (C) manifest variable. This model is valid when a change in health status is not expected between the second and third time period. For example, when a person suffers from a disease at time 2 and 3, it can be expected that the observed indicators capture the same information. There are a number of variations that can be modeled, but based on theory and previous research each alternative had less justification than the latent Markov model.

Capitalization on Chance

> *“What percentage of researchers would find themselves unable to think up a ‘theoretical justification’ for freeing a parameter? In the absence of empirical information to the contrary, I assume that the answer...is ‘near zero’”(Steiger, 1990)*

The final model, the mixed latent Markov model with three chains, was the result of fitting various modifications and extensions of the Markov model. When a number of modifications are tested on the sample, there is an increased likelihood that a modification would improve the fit of the model due to chance due to idiosyncratic characteristics of the sample (MacCallum, Roznowski, & Necowitz, 1992). Although the ‘best’ model was specified with respect to theory, parsimony, statistics, and generalizability to the population, any specification search is subject to capitalization on chance and therefore may not be generalizable to the population.

5.3.3 Time Varying Variables

Baseline covariates were used to predict membership as opposed to capturing them in their heterogeneous nature. The effect of not capturing the heterogeneity is minimized in cases where the variables are not subject to great fluctuation e.g., sex, education (as a result of age restriction), immigrant status and visible minority status. Conversely, income, employment status and cohabitation status do change over time. If the person’s baseline SEP indicators were an anomaly, they may not be reflective of the person's dynamic, 'true' SEP. In addition, there may be dose-response like relationship with repeated exposure to low indicators of SEP being more harmful to various aspects of health (Lynch et al., 1997; Power et al., 1999).

Chapter 6

CONCLUSION

Using a nationally representative longitudinal survey from Statistics Canada, this study examined the role of income, education and employment status as factors influencing membership in a health trajectory. One of the strengths of the survey used is in its measurement of income, which was derived from tax records and accounted for government transfers and deductions. The analysis captured the dynamics of SRH and took into account measurement error associated with SRH, a limitation that was often found in research that examined SEP and health longitudinally as discussed in a systematic research by Gunasekara et al. (2011).

Through model selection that balanced theory, fit, and parsimony, the population was parsed into three sub-populations. Overall, more than three quarter of Canadians were found to be in the constant good health trajectory followed by poor health trajectory and fluctuating health trajectory. Yet, in spite of access to relatively generous social programs such as having a universal health care system and benefits to mitigate the effects of inequality, we found inequality in health trajectory membership to be patterned by income and education amongst working-age Canadians. While this finding is consistent with a previous study that employed a similar analysis, our study did not find a significant difference by occupation. The discrepancy may be attributed to difference in occupation measurement. Although the potential effects of biases from observational research limits causal inference, the use of baseline SEP characteristics to predict health trajectories establishes the temporal order of events and increases the confidence in the direction of the association.

The implications of this study are manifold, but the strength of the income gradient in predicting membership in the stable good health trajectory emerged as a particularly salient finding in the current social context. Indeed, income inequality has steadily risen since the mid-1990s in Canada to levels that have not been observed since 1920s. With the effects of the economic recession in the US and EU rippling globally, coupled with unequal recovery, the hollowing of the middle class is bound to continue. As this thesis was being written, the Occupy Wall Street movement began, along with a renewed focus on the concentration of income in the top 1%.

If nothing is done to curb these trends or mitigate their effects, we are likely to see corollary increases in health inequalities in the Canadian population. Yet, as the WHO Commission on Social Determinants of Health reminds us, health inequalities that arise out of social differences are inherently unjust, because they can be mitigated with the right mix of social policies (Commission on Social Determinanats of Health, 2008).

As John Humphrey, author of the first draft of the Universal Declaration of Human Rights wrote in autobiography (1984), ‘human rights without economic and social rights have little meaning for most people’. With fiscal constraints, the health of the current economy, and increasingly more people being affected by income inequality, it will be ever more important to keep monitoring and developing interventions to curb health inequalities.

References

Adler, N., Boyce, T., Chesney, M., Cohen, S., Folkman, S., Kahn, R., & Syme, L. (1994). Socioeconomic Status and Health: The Challenge of the Gradient. *American Psychologist, 49*(1), 15-24.

Adler, N., & Ostrove, J. (1999). Socioeconomic status and health: what we know and what we don't. *Annals of New York Academy of Science* (896), 3-15.

Adler, N., & Stewart, J. (2010). *The biology of disadvantage: socioeconomic status and health* (Vol. 1186): The New York Academy of Sciences.

Alba, R. (1990). *Ethnic identity: The tranformation of white identity*. New Haven: Yale University Press.

Angel, J. L., & Angel, R. J. (2006). Minority group status and healthful aging: Social structure still matters. [Review]. *American Journal of Public Health, 96*(7), 1152-1159.

Atkinson, A., Rainwater, L., & Smeeding, T. (1995). Income distribution in OECD countries *OECD* (Vol. No 18): OECD.

Backlund, E., Sorlie, P. D., & Johnson, N. J. (1996). The shape of the relationship between income and mortality in the United States: Evidence from the National Longitudinal Mortality Study. [Journal: Article]. *Annals of Epidemiology, 6*(1), 12-20.

Bailis, D. S., Segall, A., & Chipperfield, J. G. (2003). Two views of self-rated general health status. *Social Science & Medicine, 56*(2), 203-217.

Bartley, M. (2004). *Health Inequality: An introduction to theories, concepts and methods*. Malden: Blackwell publishing Ltd.

Bartolucci, F., Pennoni, F., & Francis, B. (2007). A latent markov model for detecting patterns of criminal activity. *Journal of the Royal Statistical Society, 170*, 115-132.

Benach, J., Mutaner, C., & Santana, V. (2006). *Employment conditions knowledge network: Scoping paper: Commission on Social Determinants of Health*, World Health Organization.

Benavides, F. G., Benach, J., Diez-Roux, A. V., & Roman, C. (2000). How do types of employment relate to health indicators? Findings from the Second European Survey on Working Conditions. *Journal of Epidemiology and Community Health, 54 (7)*, 494-501.

Benyamini, Y., Blumstein, T., Lusky, A., & Modan, B. (2003). Gender differences in the self-rated health-mortality association: Is it poor self-rated health that predicts mortality or excellent self-rated health that predicts survival? *Gerontologist, 43*(3), 396-405.

Benyamini, Y., Idler, E. L., Leventhal, H., & Laventhal, E. (2000). Positive affect and function as influences on self-assessments of health: Expanding our view beyond illness and disability. *Journal of Gerontology: Psychological Sciences, 56B*(2), 107-116.

Benyamini, Y., Leventhal, E. A., & Leventhal, H. (1999). Self-assessments of health - What do people know that predicts their mortality? *Research on Aging, 21*(3), 477-500.

Benzeval, M., & Judge, K. (2001). Income and health: the time dimension. *Social Science & Medicine, 52*(9), 1371-1390.

Benzeval, M., Taylor, J., & Judge, K. (2000). Evidence on the Relationship between Low Income and Poor Health: Is the Government Doing Enough? *Fiscal Studies, 21*(3), 375-399.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? *American Economic Review, 94*(4), 991-1013.

Biemer, P. P., & Wieson, C. (2002). Measurement error evaluation of self-reported drug use: a latent class analysis of the US National Household Survey on Drug Abuse. *Journal of the Royal Statistical Society, 165*(1), 97-119.

Biemer P.P., Groves, R., Lyberg, L., Mathiowetz, N., Sudman, S. (2004). Approaches to the modeling of measurement erros. Measurement Errors in Surveys. John Wiley & Songs, Inc., Hoboken, NJ, USA

Blakely, T. A., Kennedy, B. P., Glass, R., & Kawachi, I. (2000). What is the lag time between income inequality and health status? *Journal of Epidemiology and Community Health, 54*(4), 318-319.

Bobak, M., Pikhart, H., Rose, R., Hertzman, C., & Marmot, M. (2000). Socioeconomic factors, material inequalities, and perceived control in self-rated health: cross-sectional data from seven post-communist countries. *Social Science & Medicine, 51*(9), 1343-1350.

Bosma, H., Marmot, M. G., Hemingway, H., Nicholson, A. C., Brunner, E., & Stansfeld, S. A. (1997). Low job control and risk of coronary heart disease in Whitehall II (prospective cohort) study. *British Medical Journal, 314*(7080), 558-565.

Buhmann, B., Rainwater, L., Schmaus, G., & Smeeding, T. (1988). Equivalency scales, well-being, inequality and poverty: sensitivity estimates across 10 countries using the LIS database. *Review of Income and Wealth, 34*, 115-142.

Burstrom, B., & Fredlund, P. (2001). Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes? *Journal of Epidemiology and Community Health, 55*(11), 836-840.

Cairney, J., & Arnold, R. (1998). Socioeconomic position, lifestyle and health among Canadians aged 18 to 64: A multi-condition approach. *Canadian Journal of Public Health, 89 (3)*, 208-212.

Canadian Institute for Health Information. (2008). *Reducing gaps in health: A focus on socio-economic status in urban Canada.* Canadian Institute for Health Information

Chandola, T., & Jenkinson, C. (2000). Validating self-rated health in different ethnic groups. *Ethnicity & Health, 5*(2), 151-159.

Citizenship and Immigration Canada (2009) Facts and Figures 2008: Immigration Overview: Permanent and temporary residents. Retrieved from http://www.cic. gc.ca/english/resources/statistics/facts2009/permanent/01.aspS (accessed on 3 February2011).

Collins, L., Schafer, J., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*(4), 330-351.

Collins, L., Sayer, A. (2001) *New methods for the analysis of change*. American Psychological Association.

Commission on Social Determinanats of Health (CSDH) (2008). *Closing the gap in a generation: Health equity through action on the social determinants of health*. World Health Organization

Contoyannis, P., Jones, A. M., & Rice, N. (2004). The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics, 19*(4), 473-503.

Cudeck, R., & Henly, S. J. (2003). A realistic perspective on pattern representation in growth data: Commen on Bauer and Curran. *Psychological Methods*(8), 378-383.

Davey Smith, G., Neaton, J., Wentworth, D., Stamler, R., & Stamler, J. (1996). Socioeconomic differentials in mortality risk among men screened for the multiple risk factor intervention trial: I. White Men. *American Journal of Public Health, 86*(4), 486.

de Bruin, A., Picavet, H., & Noosikov, A. (1996). *Health interview surveys: Towards international harmonization of methods and instruments.* Copenhagen: World Health Organization.

DeShon, R. (1998) A cautionary note on measurement error corrections in structural equation models. *Psychological Methods, 3(4)* 412-423

De Witte, H. (1999). Job insecurity and psychological well-being: review of the literature and exploration of some unresolved issues. *European Journal of Work and Organizational Psychology, 8*(2), 155-177.

Deeg, D. J. H., & Kriegsman, D. M. W. (2003). Concepts of Self-Rated Health: Specifying the Gender Difference in Mortality Risk. *The Gerontologist, 43*(3), 376-386.

DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general self-rated health question. *Journal of General Internal Medicine, 21*(3), 267-275.

Desesquelles, A. F., Egidi, V., & Salvatore, M. A. (2009). Why do Italian people rate their health worse than French people do? An exploration of cross-country differentials of self-rated health. *Social Science and Medicine, 68*(6), 1124-1128.

Dowd, J. B., & Zajacova, A. (2007). Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US? *International Journal of Epidemiology, 36*(6), 1214-1221.

Duncan, G. J. (1996). Income dynamics and health. *International Journal of Health Services, 26*(3), 419-444.

Duncan, G. J., Daly, M. C., McDonough, P., & Williams, D. R. (2002). *Optimal indicators of socioeconomic status for health research*: American Journal of Public Health. 92 (7) (pp 1151-1157), 2002.

Dunn, J. (2009). Canada in context: Health inequalities research in Canada: From data liberation to research proliferation to...? *Health Reports (Statistics Canada), 20*(4).

Ecob, R., & Davey Smith, G. (1999). Income and health: What is the nature of the relationship?. *Social Science and Medicine, 48*(5), 693-705.

Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement, 61*(5), 713-740.

Engel, U., & Reinecke, J. (1996). *Analysis of Change: Advanced Techniques in Panel Data Analysis*. Berlin; New York: Walter de Gruyter & Co.

Eriksson, I., Unden AL., Elofsson, S. (2001) Self-rated health. Comparisons between three different measures. Results from a population study. *International Journal of Epidemiology*. 30 (2): 326-33

Etile, F., & Milcent, C. (2006). Income-related reporting heterogeneity in self-assessed health: Evidence from France. *Health Economics, 15*(9), 965-981.

Ettner, S. (1996). New evidence on the relationship between income and health. *Journal of Health Economics, 15*(1), 67-85.

Feinstien, J. (1993). The relationship between socioeconomic status and health: a review of the literature. *Milbank Quarterly, 71*(2), 279-322.

Ferrie, J. E., Martikainen, P., Shipley, M. J., & Marmot, M. G. (2005). Self-reported economic difficulties and coronary events in men: Evidence from the Whitehall II study. [Journal: Article]. *International Journal of Epidemiology, 34*(3), 640-648.

Finkel, S. (1995) *Causal analysis with panel data.* Series: Quantitative Applications in the Social Sciences. SAGE University Paper. London.

Fischer, J. A. V., & Sousa-Poza, A. (2009). Does job satisfaction improve the health of workers? New evidence using panel data and objective measures of health. *Health Economics, 18*(1), 71-89.

Fuchs, V. R. (1979). Economics, health, and post-industrial society. *Milbank Memorial Fund Quarterly - Health & Society, 57*(2), 153-182.

Fujiwara, T., & Kawachi, I. (2009). Is education causally related to better health? A twin fixed-effect study in the USA. *International Journal of Epidemiology, 38*(5), 1310-1322.

Gallo, L., & Matthews, K. (2003). Understanding the association between socioeconomic status and physical health: Do negative emotions play a role? *Psychological Bulletin, 129*(1), 10-51.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Davey Smith, G. (2006). Indicators of socioeconomic position (part 1). *Journal of Epidemiology and Community Health, 60*(1), 7-12.

Gould, W. (2000). Interpreting logistic regression in all its forms. In J. Newton (Ed.), *Stata Technical Bulletin* (Vol. STB-53).

Gortmaker, S., Must, A., Perrin, J., Sobol, A., & Dietz, W. (1993). Social and economic consequences of overweight in adolescence and young adulthood. *New England Journal of Medicine, 329*(14), 1008-1012.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics, 19*(3), 403-420.

Gunasekara, F. I., Carter, K., & Blakely, T. (2011). Change in income and change in self-rated health: Systematic review of studies using repeated measures to control for confounding bias. *Social Science & Medicine, 72*(2), 193-201.

Haas, S., & Fosse, N. E. (2008). Health and the educational attainment of adolescents: evidence from the NLSY97. *Journal of Health and Social Behavior, 49*(2), 178-192.

Hagenaars, J. (1993). *Loglinear models with latent variables*: Sage.

Hagenaars, J., & McCutcheon, A. (2002). *Applied Latent Class Analysis*: Cambridge University Press.

Hagenaars, J. A. (1992). Exemplifying longitudinal loglinear analysis with latent variables. . In P. Van der Heijden, W. Jansen, B. Francis & G. Seeber (Eds.), *Statistical Modelling* (pp. 105-120). Amsterdam: Elsevier Science Publishers.

Hajat, A., Kaufman, J. S., Rose, K. M., Siddiqi, A., & Thomas, J. C. (2010). Do the wealthy have a health advantage? Cardiovascular disease risk factors and wealth. *Social Science & Medicine, 71*(11), 1935-1942.

Hamilton, J. (1989). A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica, 57*, 357-384.

Hansen, M. S., Fink, P., Frydenberg, M., & Oxhoj, M. L. (2002). Use of health services, mental illness, and self-rated disability and health in medical inpatients. *Psychosomatic Medicine, 64*(4), 668-675.

Harper, S., King, N., Meersman, S., Reichman, M., Breen, N., & Lynch, J. (2010). Implicit value judgments in the measurement of health inequalities. *The Milbank Quarterly, 88*(1), 429.

Health Canada. (2006). It's your health: Obesity: Health Canada.

Hernan, M., & Cole, S. (2009). Invited commentary: causal diagrams and measurement bias. *American Journal of Epidemiology, 170*, 959-962.

Huisman, M., Van lenthe, F., & Mackenbach, J. (2007). The predictive ability of self-assessed health for mortality in different educational groups. *International Journal of Epidemiology, 36*(6), 1207-1213.

Humphrey, J. (1984) *Human rights & the United Nations: A great adventure*. Dobbs Ferry, NY. Transnational Publishers.

Humphries, K., & van Doorslaer, E. (2000). Income related health inequality in Canada. *Social Science & Medicine, 50*, 663-671.

Idler, E., & Benyamini, Y. (1997). Self-Rated Health and Mortality: A Review of Twenty Serven Community Studies. *Journal of Health and Social Behavior, 38*(1), 21-37.

Idler, E. L. (2003). Discussion: Gender differences in self-rated health, in mortality, and in the relationship between the two. *Gerontologist, 43*(3), 372-375.

Idler, E. L., & Kasl, S. (1995). Self-ratings of health: do they also predict change in functional ability? *Journal of Gerontology: Psychological Sciences, 50*(6), S344-353.

Jurges, H. (2007). True health vs response styles: Exploring cross-country differences in self-reported health. *Health Economics, 16*(2), 163-178.

Jylha, M. (1994). Self-rated health revisited: Exploring survey interview episodes with elderly respondents. *Social Science & Medicine, 39*(7), 983-990.

Jylha, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science & Medicine, 69*(3), 307-316.

Jylha, M., Guralnik, J. M., Ferrucci, L., Jokela, J., & Heikkinen, E. (1998). Is self-rated health comparable across cultures and genders? *Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 53*(3), S144-S152.

Kalleberg, A. (2009). Precarious work, insecure workers: Employment relations in transition. *American Sociological Review, 74*, 1-22.

Kaplan, D. (2000). *Structural Equation Modeling*: Sage Publications.

Kaplan, D. (2008). An Overview of Markov Chain Methods for the Study of Stage-Sequential Developmental Processes. *Developmental Psychology, 44*(2), 457-467.

Kaplan, G. A., & Keil, J. E. (1993). Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation, 88*, 1973-1998.

Kawachi, I., Adler, N., & Dow, W. (2010). Money, schooling, and health: Mechanisms and causal evidence. In N. Adler & J. Stewart (Eds.), *The Biology of Disadvantage: Socioeconomic Status and Health* (pp. 56): Annals of the New York Academy of Sciences.

Kawada, T. (2003). Self rated health and life prognosis. *Archives of Medical Research, 34*(4), 343.

Kim, S., Egerter, S., Cubbin, C., Takahashi, E. R., & Braveman, P. (2007). Potential implications of missing income data in population-based surveys: An example from a postpartum survey in California. *Public Health Reports, 122*(6), 753-763.

Kleingman, A. (1989). *The illness narratives: suffering, healting, and the human condition*: Basic Books.

Krause, N. M., & Jay, G. M. (1994). What do global self-rated health items measure. *Medical Care, 32*(9), 930-942.

Krieger, n., Williams, D. R., & Moss, N. (1997). Measuring social class in US Public Health Research: concepts, methodologies, and guidelines. *annual reviews of public health, 18*, 341-378.

Kunz-Ebrecht, S. R., Kirschbaum, C., & Steptoe, A. (2004). Work stress, socioeconomic status and neuroendocrine activation over the working day. *Social Science and Medicine, 58*(8), 1523-1530.

Langeheine, R., Stern, E., & van de Pol, F. (1994). State mastery learning: dynamic models for longitudinal data. *Applied Psychological Measurement, 18*, 277.

Lanjouw, P., & Ravallion, M. (1995). Poverty and Household Size. *The Economic Journal, 105*(433), 1415-1434.

Lee, E. S., & Forthofer, R. (2006). *Analyzing complex survey data* (Vol. Series/Number 07-071). London: SAGE publications.

Levesque, I., & Franklin, S. (2000). Longitudinal and cross-sectional weighting of the survey of labour and income dynamics 1997 reference year. Ottawa: Statistics Canada.

Lindeboom, M., & Van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. [Review]. *Journal of Health Economics, 23*(6), 1083-1099.

Link, B., & Phelan, J. (1995). Social Conditions as Fundamental Causes of Disease. *Journal of Health and Social Behavior, 35*, 80-94.

Litva, A., & Eyles, J. (1994). Health or healthy: Why people are not sick in a Southern Ontarian town. *Social Science & Medicine, 39*(8), 1083-1091.

Lynch, J., & Kaplan, G. (2000). *Socioeconomic Position*. Oxford: Oxford University Press.

Lynch, J., Kaplan, G., & Shema, S. (1997). Cumulative impact of sustained economic hardship on physical, cognitive, psychological, and social functioning. *New England Journal of Medicine, 337*(26), 1889-1895.

Lynch, J., Davey Smith, G., Kaplan, G. A., & House, J. S. (2000). Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. *British Medical Journal, 320*(7243), 1200-1204.

MacCallum, R. C., Roznowski, M., & Necowitz, L. (1992). Model modifications in covariance structural analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*(3), 490-504.

Macdonald, J., Raphael, D., Labonte, R., Colman, R., Torgerson, R., & Hayward, K. (2009). Income and health in Canada: Canadian researchers' conceptualizations make policy change unlikely. *International Journal of Health Services, 39*(3), 525-543.

Macintyre, S., Maciver, S., & Sooman, A. (1993). Area, class and health- Should we be focusing on places or people. *Journal of Social Policy, 22*, 213-234.

Mackenback, J. P., Martikainen, P., Looman, C., Dalstra, J., Kunst, A., & Lahelma, E. (2005). The shape of the relationship between income and self-assessed health: an international study *International Journal of Epidemiology, 34*(2), 286-293.

Magidson, J., Vermunt, J., & Tran, B. (2007). Using a mixture latent markov model to analyze longitudinal U.S. employment data involving measurement error: Statistical Innovations

Manderbacka, K. (1998). Examining what self-rated health question is understood to mean by respondents. *Scandinavian Journal of Social Medicine, 26*(2), 145-153.

Manor, O., Matthews, S., & Power, C. (2000). Dichotomous or categorical response? Analysing self-rated health and lifetime social class. *International Journal of Epidemiology, 29*(1), 149-157.

Markov, A. (1907). *Extension of the limit theorems of probability theory to a sum of variables connected in a chain* (Vol. No. 9): Physio-mathematical College XXII.

Marmot, M., Shipley, M. J., & Rose, G. (1984). Inequalities in death-specific explanations of a general pattern. *Lancet, 5*(1), 1003-1006.

Marmot, M. G., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Davey Smith, G.(1991). Health inequalities among British civil servants: the Whitehall II study. *The Lancet, 337*(8754), 1387-1393.

McCutcheon, A. (1996). Models for the analysis of categorically-scored panel data. In U. Engel & J. Reinecke (Eds.), *Analysis of Change: Advanced Techniques in Panel Data Analysis* (pp. 15-32). Berlin; New York: Walter de Gruyter & Co.

McCutcheon, A. (2002). Basic concepts and procedures in single- and multiple- group latent class analysis. In J. A. Hagenaars & A. McCutcheon (Eds.), *Applied Latent Class Analysis*. Cambridge: Cambridge University Press.

McDonough, P., & Amick, Bejamin III (2001). The social context of health selection: A longitudinal study of health and employment. *Social Science & Medicine*, 53, 135-145.

McDonough, P., Worts, D., & Sacker, A. (2010). Socioeconomic inequalities in health dynamics: A comparison of Britain and the United States. *Social Science & Medicine, 70*, 251-260.

McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiological Review, 87*, 873-904.

McEwen, B. S., & Gianaros, P. J. (2010). Central role of the brain in stress and adaptation: Links to socioeconomic status, health, and disease. *Annals of the New York Academy of Sciences, 1186*(1), 190-222.

McFadden, E., Luben, R., Bingham, S., Wareham, N., Kinmonth, A. L., & Khaw, K. T. (2009). Does the association between self-rated health and mortality vary by social class? *Social Science and Medicine, 68*(2), 275-280.

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York: John Wiley & Sons.

McLeod, J. D., & Kessler, R. C. (1990). Socioeconomic status differences in vulnerability to undesirable life events. *Journal of Health and Social Behaviour, 31*, 162-172.

Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M., & Urponen, H. (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology, 50*(5), 517-528.

Mirowsky, J., & Ross, C. (2003). *Education, social status, and health*. Hawthorne: Water de Gruyter.

Mulatu, M., & Schooler, C. (2002). Causal connections between socio-economic status and health: reciprocal effects and mediating mechanisms. *Journal of Health and Social Behavior, Journal of health and social behavior*(43).

Murray, C., Tandon, A., Salomon, J. A., Mathers, C., & Sadana, R. (2001). Cross-population comparability of evidence for health policy. *GPE Discussion Paper No. 46.* Geneva: WHO/EIP.

Mustard, C. A., Derksen, S., MBerthelot, J.-m., & Roos, L. (1998). Age-specific education and income gradients in morbidity and mortality in a Canadian province. *Social Science & Medicine, 45*(3), 383-397.

Myung, I. J. (2002). Tutorial on maximum liklihood estimation. *Journal of Mathematical Psychology, 47*(1), 90-100.

Newbold, K. B. (2005). Self-rated health within the Canadian immigrant population: risk and the healthy immigrant effect. *Social Science & Medicine, 60*(6), 1359-1370.

Phinney, J. (1996). When we talk about American ethnic groups, what do we mean? *American Psychologist, 51*(9), 918-927.

Phipps, S. (2003). The impact of poverty on health: A scan of research literature. Ottawa: Canadian institute for health information.

Popay, J., Bennett, S., Thomas, C., Williams, G., Gatrell, A., & Bostock, L. (2003). Beyond 'beer, fags, eeg and chips'? Exploring lay understandings of social inequalities in health. *Sociology of Health & Illness, 25*(1), 0141-9889.

Power, C., Manor, O., & Matthews, S. (1999). The duration and timing of exposure: Effects of socioeconomic environment on adult health. *American Journal of Public Health, 89*(7), 1059-1065.

Quesnel-Vallee, A., DeHaney, S., & Ciampi, A. (2010). Temporary work and depressive symptoms: A propensity score analysis. *Social Science & Medicine, 70*(12), 1982-1987.

Quinlan, M., Mayhew, C., & Bohle, P. (2001). The global expansion of precarious employment, work disorganization, and consequences for occupational health: Placing the debate in a comparative historical context. *International Journal of Health Services, 31*(3), 507-536.

Rabiner, L. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77*(2), 257.

Raphael, D. (2004). *Social determinants of health: Canadian Perspectives*. Toronto: Canadian Scholars' Press Inc.

Rindskopf, D. (2003). Mixture or homogenous? Comment on Bauer and Curran. *Psychological Methods*(8), 364-368.

Riphahn, R. (1999). Income and employment effects of health shocks: A test case for the German welfare state. *Journal of Population Economics, 12*(3), 363-389.

Rodin, J., & McAvay, G. (1992). Determinants of change in perceived health in a longitudinal study of older adults. *Journals of Gerontology, 47*(6), P373-P384.

Ross, C. E., & Wu, C.-l. (1995). The links between education and health. *American Sociological Review, 60*(5), 719-745.

Rust, K., & Rao, J. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*(5), 283-310.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods, 7*(2), 147-177.

Sen, A. (1993). Positional Objectivity. *Philosophy & Public Affairs, 22*(2), 126-145.

Setia, M.S., Abrahamowicz, M., Tousignant, P., Lynch, J., & Quesnel-Vallée, A. (2011). Self-rated health in Canadian immigrants: a longitudinal analysis of the Longitudinal Survey of Immigrants to Canada. *Health & Place*. 18(2) 658-670

Service Canada. (2010). Employment Insurance Regular Benefits Retrieved March 25, 2012, from http://www.servicecanada.gc.ca/eng/ei/types/regular.shtml

Singh-Manoux, A., Dugravot, A., Shipley, M. J., Ferrie, J. E., Martikainen, P., Goldberg, M., & Zins, M. (2007). The association between self-rated health and mortality in different socioeconomic groups in the GAZEL cohort study. *International Journal of Epidemiology, 36*(6), 1222-1228.

Sparks, K., Cooper, C., Fried, Y., & Shirom, A. (1997). The effects of hours of work on health: A meta-analytic review. *Journal of Occupational and Organizational Psychology, 70*(4), 391-408.

Statistics Canada. (2009). Education indicators in Canada: An international perspective: Statistics Canada.

Statistics Canada. (2011a). Consumer Price Index, historical summary Retrieved April 28, 2010, from http://www40.statcan.ca/l01/cst01/ECON46A-eng.htm

Statistics Canada. (2011b). Government transfers and income tax Retrieved December 10, 2011, from http://www40.statcan.ca/l01/cst01/famil88a-eng.htm

Statistics Canada. (2011c). Labour force characteristics by immigrant status, by detailed age group. *Employment and unemployment* Retrieved December 4, 2011, from http://www40.statcan.gc.ca/l01/cst01/labor91a-eng.htm

Steiger, J. (1990). Structural Model Evaluation and Modification: An Interval Estimation Approach. *Multivariate Behavioral Research, 25*(2), 173-180.

Strawbridge, W. J., & Wallhagen, M. R. (1999). Self-rated health and mortality over three decades. *Research on Aging, 21*, 402-416.

Stronks, K., van de Mheen, H., van den Bos, J., & Mackenbach, J. P. (1997). The interrelationship between income, health and employment status. *International Journal of Epidemiology, 26*(3), 592-600.

Sue, S. (1991). Ethnicity and culture in psychological research and practice. In J. Goodchilds (Ed.), *Psychological perspectives on human diversity in America* (pp. 51-85). Washington: American Psychological Association.

Thomas, D. (1994). Incomes, expenditures, and health outcomes: Evidence on intrahousehold resource allocation. In L. Haddad, J. Hoddinott & H. Alderman (Eds.), *Intrahousehold resource allocation in developing countries: Models, methods and policy*. Baltimore: John Hopkins University Press.

Turrell, G. (2000). Income non-reporting: Implications for health inequalities research. *Journal of Epidemiology and Community Health, 54*(3), 207-214.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Van de Pol, F., & Langeheine, R. (1990). *Mixed markov latent class models*. Oxford: Blackwell.

van Doorslaer, E., & Gerdtham, U. G. (2003). Does inequality in self-assessed health predict inequality in survival by income? - Evidence from Swedish data. *Social Science & Medicine, 57*(9), 1621-1629.

Vermunt, J., Tran, B., & Magidson, J. (2008). *Handbook of longitudinal research: Design, measurement and analysis*. Burlington, MA: Elsevier.

Vermunt, J. K., Langeheine, R., & Bockenholt, U. (1999). Discrete-time discrete-state latent Markov models with time-constant and time-varying covariates. *Journal of Educational and Behavioral Statistics, 24*(2), 179-207.

Virtanen, P., Vahtera, J., Kivimaki, M., Pentti, J., & Ferrie, J. (2002). Employment security and health. *Journal of Epidemiology and Community Health, 56*(8), 569-574.

Wagstaff, A., & van Doorslaer, E. (2000). Income inequality and health: What does the literature tell us? *Annual Review of Public Health, 21*, 543-567.

Wannamethee, G., & Shaper, A. (1991). Self-assessment of health status and mortality in middle-aged British men. *International Journal of Epidemiology, 20*(1), 239-245.

Wiggins, L. (1973). *Panel analysis*. Amsterdam: Elsevier.

Wilcox, V., Kasl, S., & Idler, E. L. (1996). Self-rated health and physical disability in elderly survivors of a major medical event. *Journal of Gerontology: Social Sciences, 17*(2), 185-205.

Wilkinson, R. (1996). *Unhealthy Societies: The Afflictions of Inequality*. London: Routledge.

Wolinsky, F., Culler, D., Callahan, C., & Johnson, R. (1994). Hospital resource consumption among older adults: A prospective analysis of episodes, length of stay, and charges over aseven-year period. *Journal of Gerontology: Social Sciences, 49*, S240-S252.

World Health Organization. (1948). Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19 June - 22 July 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

World Health Organization. (2011). Global Atlas of the Health Workforce Retrieved March 20, 2012, from http://apps.who.int/globalatlas/default.asp

Zenk, S. N., Schulz, A. J., Israel, B. A., James, S. A., Bao, S., & Wilson, M. L. (2005). Neighborhood racial composition, neighborhood poverty, and the spatial accessibility of supermarkets in metropolitan Detroit. *American Journal of Public Health, 95*(4), 660-667.

Zola, I. K. (1966). Culture and symptoms: An analysis of patients presenting complaints. *American Sociological Review, 3*, 615-630.

Appendix A
Definition of after-tax income

Source: Statistics Canada, Guideline for cross-sectional public-use microdata file: Survey of Labour and Income Dynamics (SLID) Reference year 2002.

Market income
  Earnings
    Wages, salaries and commission
    Self-employment income
  Investment income
  Retirement pensions
  Other income
(plus) Government transfers
  Child tax benefits
  Old Age Security and Guaranteed Income Supplement/Spouse's Allowance
  Canada Pension Plan/Quebec Pension Plan benefits
  Employment Insurance benefits
  Social Assistance
  Workers' compensation benefits
  GST/HST Credit
  Provincial/territorial tax credits
  Other government transfers
(equals) Total Income
(minus) Income tax
(equals) After-tax Income

- **Market income**

Market income is the sum of earnings (from employment and net self-employment), investment income, (private) retirement income, and the items under "Other income". It is equivalent to total income minus government transfers. It is also called income before taxes and transfers.

- **Earnings**

This includes earnings from both paid employment (wages and salaries) and self-employment.

- **Wages, salaries and commission**

These are gross earnings from all jobs held as an employee, before payroll deductions such as income taxes, employment insurance contributions or pension plan contributions, etc. Wages and salaries include the earnings of owners of incorporated businesses, although some amounts may instead be reported as investment income. Commission income received by salespersons as well as occasional earnings for baby-sitting, for delivering papers, for cleaning, etc. are included. Overtime pay is included.

- **Self-employment income**

This is net self-employment income, i.e. after deduction of expenses. Negative amounts (losses) are accepted. It includes income received from self-employment on own account, in partnership in an unincorporated business, or in independent professional practice. Income from roomers and boarders (excluding that received from relatives) is included. Note that because of the various inclusions, receipt of self-employment income does not necessarily mean the person held a job.
Self-employment income is subdivided into farm self-employment income and non-farm self-employment income. Farm self-employment income is reported by individuals who operate their own or a rented farm, either on own account or in partnership. Included are money receipts from the sale of farm products as well as related supplementary and assistance payments from governments. Income in kind is excluded.

- **Investment income**

This includes interest received on bonds, deposits and savings certificates from Canadian or foreign sources, dividends received from Canadian and foreign corporate stocks, cash dividends received from insurance policies, net rental income from real estate and farms, interest received on loans and mortgages, regular income from an estate or trust fund and other investment income. Realized capital gains from the sale of assets are excluded. Negative amounts are accepted.

- **Retirement pensions**

This is retirement pensions from all private sources, primarily employer pension plans. Amounts may be received in various forms such as annuities, superannuation or RRIFs (Registered Retirement Income Funds). Withdrawals from RRSPs (Registered Retirement Savings Plans) are not included in retirement pensions. However, they are taken into account as necessary for the estimation of certain government transfers and taxes. For data obtained from administrative records, income withdrawn from RRSPs before the age of 65 is treated as RRSP withdrawals, and income withdrawn from RRSPs at ages 65 or older is treated as retirement pensions. Retirement pensions may also be called pension income.

- **Other income**

This sub-total includes all items of market income not included elsewhere. Among them are support payments received (also called alimony and child support). The coverage of other items depends at least to some extent on the method of income data collection, whether from administrative income tax records or by interview. Those items which are included on line 130 of the T1 tax return are well covered. These include, but are not restricted to, retirement allowances (severance pay/termination benefits), scholarships, lump-sum payments from pensions and deferred profit-sharing plans received when leaving a plan, the taxable amount of death benefits other than those from CPP or QPP, and supplementary unemployment benefits not included in wages and salaries.

- **Government transfers**

Government transfers include all direct payments from federal, provincial and municipal governments to individuals or families. See the table "Classification of Income Sources" for a list of the government transfers identified separately in the latest reference year. It should be noted that many features of the tax system also carry out social policy functions but are not government transfers per se. The tax system uses deductions and non-refundable tax credits, for example, to reduce the amount of tax payable, without providing a direct income.

- **Child tax benefits**

Federal child tax benefits began in 1993 and replaced both the federal Family Allowances and the Child Tax Credit. Several provincial and territorial programs have since been introduced, in addition to Quebec family allowances which already existed before 1993. To be eligible, a person must have the primary responsibility for the care and upbringing of one or more children under the age of 18. Most benefits are calculated by setting a maximum amount per family or per child and reducing that total by a certain percentage of the family's net income. The programs which were explicitly accounted for in the data for 2002 were: the federal basic benefit and National Child Benefit Supplement (together called the Canada Child Tax Benefit), the Newfoundland and Labrador Child Benefit, the Nova Scotia Child Benefit, the New Brunswick Child Tax Benefit, the New Brunswick Earned Income Supplement, the Quebec Allocation familiale, the Quebec Allocation à la naissance, the Ontario Child Care Supplement for Working Families, the Saskatchewan Child Benefit, the Alberta Family Employment Tax Credit, the BC Family Bonus, and the BC Earned Income Benefit.

- **Old Age Security (OAS) benefits**

The Old Age Security (OAS) pension is targeted to Canadian residents aged 65 and over. OAS recipients who have little or no other income may also receive the federal

Guaranteed Income Supplement (GIS); and their spouses, if aged 60 to 64 (and not yet eligible for OAS and GIS themselves), receive the Spouse's Allowance.

- **Canada Pension Plan (CPP) and Quebec Pension Plan (QPP) benefits**

The CPP and QPP are compulsory contributory social insurance programs that provide a source of retirement income and protect workers and their families against loss of income due to disability or death.

- **Employment Insurance benefits**

Employment Insurance is a federal program which includes the following types of benefits: regular unemployment benefits, sickness benefits, maternity and parental benefits, and benefits for persons taking approved training courses or participating in job creation or job-sharing projects. To qualify, the claimant must have ceased receiving employment income and have worked a minimum number of weeks or hours of insurable employment over the preceding period.

- **Social assistance**

Social assistance covers many provincial and municipal income supplements to individuals and families. It is usually provided only after all other possible sources of support have been exhausted.

- **Workers' compensation benefits**

Workers' compensation is provided to protect all full-time and part-time employees from loss of salary due to work accidents or occupational diseases and help them to pay their medical expenses and other costs.

- **Goods and Services Tax/Harmonized Sales Tax Credit**

This credit was introduced in conjunction with the Goods and Services Tax in 1990. It is intended to offset the GST/HST for lower-income families and individuals. In Nova Scotia, New Brunswick and Newfoundland and Labrador, its name was changed to the Harmonized Sales Tax Credit in April 1997 when the administration of the tax was combined with provincial sales tax. Included is the federal Relief for Heating Expenses paid in 2001.

- **Provincial/territorial tax credits**

Included here are refundable tax credits other than those for children (included with child tax benefits). Some are designed to help low-income individuals and families to pay property taxes, education taxes, rent and living expenses, and so on. Provincial sales tax credits such as the Quebec Sales Tax Credit and the Newfoundland and Labrador HST

Credit are included. The Quebec abatement, although refundable, is not included here but rather with income taxes.

- **Other government transfers**

This includes government transfers not included elsewhere, mainly any other non-taxable transfers. In SLID, these amounts are included with “Other income”. This is partly because the coverage of any transfers not taxed through the income tax system is low. In survey interviews, there may be under-reporting of these transfers, which are mainly collected using an open question. Nonetheless, the types of transfers which have come under this heading include: training program payments not reported elsewhere, the Veteran's pension, pensions to the blind and the disabled, regular payments from provincial automobile insurance plans (excluding lump-sum payments), and benefits for fishing industry employees.

- **Total income**

Total income refers to income from all sources including government transfers before deduction of federal and provincial income taxes. It may also be called income before tax (but after transfers). All sources of income are identified as belonging to either market income or government transfers.

- **Income tax**

Income tax is the sum of federal and provincial income taxes payable (accrued) for the taxation year. Income taxes include taxes on income, capital gains and RRSP withdrawals, after taking into account exemptions, deductions, non-refundable tax credits, and the refundable Quebec abatement. In the Survey of Labour and Income Dynamics, the data are either taken directly from administrative records or estimated based on aggregate data from administrative records, as this yields better results than the amounts reported by interview.

- **After-tax income**

After-tax income is total income, which includes government transfers as defined here, less income tax. It may also be called income after tax.

Appendix B
SRH rating by proxy status

From 2002 to 2004, the January survey was used to tabulate the following tables. From 2004 onwards, the May survey was merged with the January survey.

Y2002

Pearson chi² (df,4) = 12.2, p=0.016

	SRH status					
	Excellent	Very good	Good	Fair	Poor	*Total*
In person	1202	1448	1038	251	95	11008
Proxy	3271	3913	2680	830	314	4034
Total	4473	1448	1038	251	95	15042

Y2003

Pearson chi² (df,4) = 3.85, p=0.426

	SRH status					
	Excellent	Very good	Good	Fair	Poor	*Total*
In person	2393	3410	2159	655	275	8892
Proxy	1152	1633	998	292	109	4184
Total	3545	5043	3157	947	384	13076

Y2004

Pearson chi² (df,4) = 2.15, p=0.707

	SRH status					
	Excellent	Very good	Good	Fair	Poor	*Total*
In person	2254	3228	2158	674	255	8564
Proxy	996	1497	980	284	111	3868
Total	3250	4725	3133	958	366	12432

Y2005

Pearson chi² (df,4) = 4.52, p=0.340

	SRH status					
	Excellent	Very good	Good	Fair	Poor	*Total*
In person	1943	3085	2224	633	291	8176
Proxy	857	1434	1000	264	111	3666
Total	2800	4519	3224	897	402	11842

Y2006

Pearson chi² (df,4) = 7.53, p=0.110

	SRH status					
	Excellent	Very good	Good	Fair	Poor	*Total*
In person	1736	3064	2126	657	253	7836
Proxy	676	1288	963	276	120	3323
Total	2412	4352	3089	933	373	11159

Y2007

Pearson chi² (df,4) = 0.66, p=0.956

	SRH status					
	Excellent	Very good	Good	Fair	Poor	*Total*
In person	16655	3010	2145	674	277	7761
Proxy	709	1297	929	278	112	3325
Total	2364	4307	3074	952	389	11086

Appendix C
Measurement component of the latent Markov model

What is measurement error?
How can a latent variable account for measurement error?
How 'true' is the true score?

- In a classical test theory, using method i, an observation y_i is composed of a true value F and an error term $\mathcal{E}_i$. F is not dependent on the method used as it represents true value:

$$y_i = F + \mathcal{E}_i \qquad (1)$$

- As the scale of the latent variable and the observed variable may differ (e.g., the response categories of the observed variable may differ from the latent categories), scaling constants are added:

$$y_i = a_i + b_i\, F + \mathcal{E}_i \qquad (2)$$

- A term (D_i) that represents a constant influence on the observed value throughout the repeated measurements is added. Examples of a stable disturbance on the observed variable include personal characteristics of the respondent:

$$y_i = a_i + b_i\, F + D_i + \mathcal{E}_i \qquad (3)$$

- The stable disturbance, D_i , can be decomposed into two terms:
  - M_i Variation due to methods
  - U_i Variation due to interaction between trait and method, 'unique component'

  In other words, if the method stays the same, then $M_i = 0$. If the question is answered under the same condition in all waves, then $U_i = 0$.

$$y_i = a_i + b_i\, F + M_i + U_i + \mathcal{E}_i \qquad (4)$$

  In this thesis, the method variable is equal to zero as the same question was used for all waves.

- Simplified by separating the components that are not random, the equations are:

$$y_i = T_i + \mathcal{E}_i \qquad (5)$$

$$T_i = a_i + b_i\,F + M_i + U_i \qquad (6)$$

If T_i is substituted back into equation 5, it can be seen that it is the same as equation 4. Reliability can now be defined as to the degree that T_i and y_i agrees with each other. That is, proportion of observed score variance that is attributable to true score variance:

Reliability = variance in T_i / variance T_i+error variance
= $variance\ (T_i)/\ variance\ (y_i)$

- In more conventional terms:

$$y_i = h_i T_i + \mathcal{E}_i \qquad (7)$$

$$T_i = b_i\,F + M_i + U_i \quad (8)$$

Where h_i is the coefficient linking the latent variable and the observed variable.

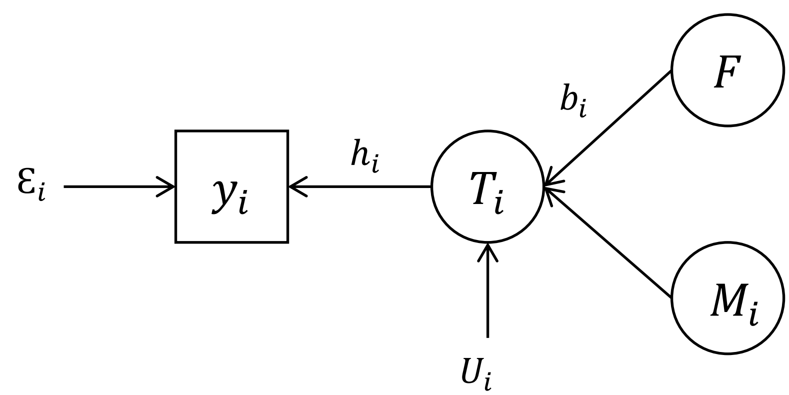


Figure 12: Path diagram for one observed indicator (yi) for one latent variable (Ti)
Source: Biemer, Groves, Lyberg, Mathiowetz, and Sudman (2004)

In summary, the observed term (y_i) is composed of a random error ($\mathcal{E}_i$) and the latent value (T_i). The latent value is a composition of the true value (F), method used (M_i) and unique component (U_i). The advantage of latent variables, then, is in accounting for the variance in the observed variable (y_i) which is equal to $h_i$when latent variable and observed variables are standardized (Finkel, 1995).

- In a repeated measure scenario with six time points:

$$y\ (t_i) = h(t_i)\ T(t_i) + \mathcal{E}(t_i) \qquad (9)$$

$$T\ (t_i) = \tau^{t-1,t} T(t_{i-1}) + U(t_i) \qquad (10)$$

As the same question was asked for all six years, the term to represent variation due to method is not needed M_i. The *i* term now refers to time period rather than method used. Equation 9 refers to the measurement part of the model, whereas equation 10 refers to the structural part.

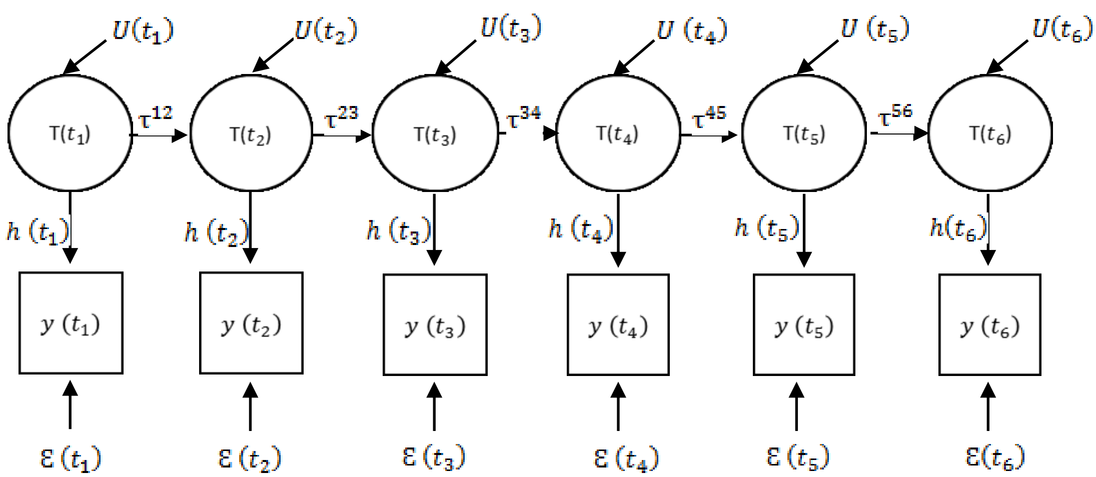


Figure 13: Path diagram for one observed indicator and one latent variable for six time periods

The true value, then, refers to a value at the latent level that was obtained by correcting for the unreliability of the question at the observed level. The latent variable still contains error due to its unique variance (U_i). In sum, then, although a latent variable approach does capture some aspect of error, we must however recognize that no amount of methodological sophistication can compensate for poor data.

As the outcome is a dichotomous, equation 9 and 10 have to be reparameterized in order to meet one of the basic assumption $E\ (\mathcal{E}_i|i) = 0$ (mean of each individual's hypothetical error distribution is equal to zero). Assuming that there is both false positive and false negative, the error term can only take three values of 0,-1,1 (see scenarios 1-4) which violates this basic assumption.

Scenario 1	$y_i = 1$ and $F = 1$, then $\mathcal{E}_i = 0$	
Scenario 2	$y_i = 1$ and $F = 0$, then $\mathcal{E}_i = 1$	False positive
Scenario 3	$y_i = 0$ and $F = 1$, then $\mathcal{E}_i = -1$	False negative
Scenario 4	$y_i = 0$ and $F = 0$, then $\mathcal{E}_i = 0$	

The same concept of measurement error still pertains with reparameterization (DeShon, 1998; Finkel, 1995).

Sources:

- For an overview of measurement error in structural modeling, refer to Chapter 24 & 28 of Biemer, Groves, Lyberg, Mathiowetz, and Sudman (2004) 'Measurement Errors in Surveys'.
- For notations referring specifically to reliability in categorical latent variables refer to Chapter 9 of Collins & Sayer (2001) 'New Methods for the Analysis of Change'
- For interpretation of measurement error and examples, refer to DeShon (1998) 'A cautionary note on measurement error corrections in Structural Equation Models' p.418-420 and Finkel (1995) 'Causal Analysis with Panel Data' p.45-58.
- For an explanation of measurement error using directed acyclic graphs (DAGs), refer to Hernan, M., Cole, S.R. (2009) Invited commentary: Causal diagrams and measurement bias. American Journal of Epidemiology. 170(8) 959-62

Appendix D
Specifications of models

Model	transition probabilities time homog.	iterations	starts	final stages starts	minimum convergence
Manifest Markov	Yes	300	1000	50	0.100
	No	300	1000	50	0.100
Latent markov	Yes	300	1000	50	0.100
	No	300	1000	50	0.100
Mixed latent markov					
2 chain	Yes	100	500	30	0.100
	No	200	800	30	0.100
2 chain- mover stayer	Yes	200	800	50	0.100
	No	200	1500	50	0.100
3 chain	Yes	200	2000	100	0.100
	No	200	2000	100	0.100
3 chain- mover stayer stayer	Yes	200	2000	100	0.100
	No	200	4000	100	0.100

Starts: Number of initial stages starts
Final stage starts: Number of final stage starts
Iterations: Number of initial stages iterations

Appendix E
SRH validation with functional limitation

Y2002
Not released from Statistics Canada due to confidentiality concerns.

Y2003
In the main document.

Y2004

	Functional limitation % (n)		
SRH	Yes	No	Total
Excellent	6.3 (205)	93.7 (3050)	3255
Very good	13.6 (643)	86.4 (4092)	4735
Good	30.4 (953)	69.6 (2183)	3136
Fair	72.7 (701)	27.3 (263)	964
Poor	95.6 (350)	4.4 (16)	366
Total	2852	9604	12456

Y2005

	Functional limitation % (n)		
SRH	Yes	No	Total
Excellent	6.0 (168)	94.0 (2631)	2799
Very good	12.8 (579)	87.2 (3941)	4520
Good	28.7 (926)	71.25 (2295)	3221
Fair	73.1 (657)	26.9 (242)	899
Poor	96.3 (388)	3.7 (15)	403
Total	2718	9124	11842

Y2006
Not released from Statistics Canada due to confidentiality concerns.

Y2007
Not released from Statistics Canada due to confidentiality concerns.

Appendix F
Data Dictionary

Variable	Variable code in SLIDRet	Question/ description	Response categories	Modified
SRH	crhlt26	In general, how would you describe your state of health? Would you say it is . . .	1 Excellent 2 Very good 3 Good 4 Fair 5 Poor	0 Good (1,2,3) 1 Poor (4,5) In Mplus: 1 Good 2 Poor
Income	atinc27	During [reference year], what was your income from the following sources? All items on the after tax income were asked separately.	Open ended	Divided by family equivalence.
Family equivalence	eq2sc27	Calculated as sum of weight for each person in the family	Open ended	n/a
Education	hlev2g18	Highest level achieved	1 no schooling 2 elementary 3 some secondary 4 secondary school graduation 5 other beyond high school 6 some trade school 7 some community college 8 some university 9 diploma or certificate of community college 10 bachelor degree 11 Master's /MD/PhD	1 <high school graduation 2 graduated high school 3 non university postsecondary certificate 4 university degree or certificate
Employment	mjact26	I'd like to ask you a few questions about your main activity at the end of [reference year]. Was your main activity . . .	1 Working at a job or business 2 Looking for work 3 Going to school 4 Keeping house 5 Caring for other family members including young children 6 Retired 7 Long term illness	1 Employed (1) 2 Unemployed (2) 3 Student (3) 4 Not in labour force (4, 5, 6, 7) 5 Other (8, 9, 90)

			or disability 8 Doing volunteer work 9 No main activity 90 Other	
Visible Minority	Vismn15	Flag to indicate whether a person belong to a visible minority group	1 Yes 2 No	Recoded to 0,1
Immigrant	Immst15	Are you now, or have you ever been, a landed immigrant?	1 Yes 2 No	Recoded to 0,1
Marital status	Marsq26	What is [respondent name]'s marital status? Is [he/she]:	1 married 2 living common-law 3 widowed 4 separated 5 divorced 6 single, never married	Dummy variables created: Married Common law Separated Divorced Widowed Single
Sex	Sex99	Enter [respondent name]'s sex. If necessary, ask: (Is [respondent name] male or female?)	1 Male 2 Female	Recoded to 0,1
Age	Age26	What is [respondent name]'s age?	Open ended	Dummy variables for 5 year interval groups
Jan Proxy	prox1f26	Flag indicating if phase 1 interview (labour) information for a person was provided by a proxy.		
May Proxy	prox2f26	Flag indicating if phase 2 interview (income) information for a person was provided by a proxy. Not applicable after reference year 2004		

Appendix G
Sensitivity Analysis

Sensitivity analysis defining good health to include very good and good and defining poor health to include good, fair and poor. Relative risk ratios and 95% CI for a multinomial analysis of main SEP predictors for the health trajectories

	Constant good health (ref. constant poor health) RRR (95% CI)	Fluctuating health (ref. constant poor health) RRR (95% CI)
Income		
0-10000	Reference	Reference
10K-20K	**2.04 (1.49-2.79)**	0.97 (0.78-1.20)
20K-30K	**1.51 (1.11-2.06)**	0.82 (0.67-1.01)
30K-40K	**1.66 (1.22-2.26)**	0.82 (0.67-1.01)
40K-50K	**2.08 (1.50-2.88)**	0.99 (0.79-1.24)
50K-60K	**1.76 (1.24-2.49)**	0.86 (0.67-1.09)
60K-70K	**2.17 (1.45-3.26)**	0.84 (0.63-1.13)
70K-80K	1.11 (0.66-1.87)	**0.64 (0.45-0.92)**
Education		
< High school	Reference	Reference
High school	0.09 (0.07-0.12)	0.15 (0.12-0.19)
Non university post-secondary	0.04 (0.04-0.06)	0.08 (0.06-0.10)
University	0.008 (0.006-0.01)	0.03 (0.02-0.03)
Employment status		
Unemployed: looking for work	Reference	Reference
Employed	**0.34 (0.25-0.45)**	**0.72 (0.56-0.92)**
Student	1.26 (0.74-2.14)	1.19 (0.80-1.76)
Out of labour force	**1.79 (1.30-2.46)**	**0.66 (0.50-0.87)**
Other	1.32 (0.81-2.17)	0.81 (0.53-1.25)