ESSAYS ON HEALTH AND RETIREMENT IN CANADA

ESSAYS ON HEALTH AND RETIREMENT IN CANADA

By

SIMO GOSHEV, B.Econ., M.Sc., M.A.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2008) McMaster University
(Economics) Hamilton, Ontario

TITLE: Essays on Health and Retirement in Canada

AUTHOR: Simo Goshev, B.Econ. University of National and World Economy, Sofia, Bulgaria
M.Sc. McMaster University
M.A. McMaster University

SUPERVISOR: Professor Michael Veall

NUMBER OF PAGES: x, 156

Abstract

My dissertation is composed of an introductory chapter followed by three independent chapters focusing on two themes: health and retirement. The last chapter concludes.

After the introductory chapter, the second chapter investigates whether self-reported general stress is a mediator in the relationship between socio-economic status (SES) and health. I use a six-year long panel of the Canadian Survey of Labour and Income Dynamics and employ dynamic econometric modelling techniques to study men and women who are major income earners in their families. I find little evidence that general stress is a pathway from SES to health. While the results suggest a strong negative association between stress and health for both men and women, they provide little support to the hypothesis of a significant effect of income on stress, consistent with the direction of the SES-health gradient.

The third chapter studies whether self-assessed health status (SAH) contains information about future mortality and morbidity, beyond the information that is contained in standard "observable" characteristics of individuals (including pre-existing diagnosed medical conditions). Using a ten-year span of the Canadian National Population Health Survey, we find evidence that SAH does contain private information for future mortality and morbidity. Moreover, the extra information in SAH is greater at older ages. Our results suggest that a shift from defined benefit to defined contribution pension arrangements may carry with it the cost of exacerbated adverse selection in the

market for annuities, especially at older ages. That would make it more difficult for older individuals to insure longevity risk.

The fourth chapter looks at whether differences in early retirement pathways are associated with differences in post-retirement outcomes of health, stress and dwelling tenure. I use a sample of men from the Canadian Survey of Labour and Income Dynamics, years 1996 to 2004. I find that differences in pre-retirement health indicators (such as self-assessed health and disability), as pathways to early retirement, are likely to be associated with differences in post-retirement health and stress. In addition, the results suggest that "involuntary" retirees (men who may have retired because of health and/or health related conditions) are more likely to experience worse post-retirement outcomes (in terms of health and stress) than men who retire "voluntarily". Retirement circumstances are found to have no statistically significant effect on dwelling tenure.

Acknowledgement

I would like to express my gratitude to my supervisor, Professor Michael Veall, for providing me with the opportunity to work in the area of economics and for his expert guidance and invaluable assistance throughout the course of my doctoral studies. I would like to extend special thanks to my committee members Professor Thomas Crossley and Professor Byron Spencer for their guidance, assistance and support.

I wish to thank my parents for their love and unconditional support.

Preface

The second chapter of this thesis was prepared jointly with Professor Thomas Crossley and Professor James Banks and it has been submitted for journal publication. I was primarily responsible for the empirical analysis and participated in all other stages of the study.

Table of Contents

List of Tables

Chapter II

Chapter III

Chapter IV

List of Figures

Chapter III

Chapter IV

I

Introduction

This thesis is composed of relatively independent chapters and focuses on two themes: health and retirement. The next chapter evaluates the role of general stress as a mediator in the relationship between socio-economic status and health. The third chapter investigates whether older individuals possess more private information about their future mortality and morbidity than younger individuals. The fourth chapter evaluates whether differences in early retirement pathways are associated with differences in post-retirement outcomes. The fifth chapter offers a brief conclusion.

The chapter "*Stress as a Pathway in the SES-Health Relationship: Evidence from the Canadian Survey of Labour and Income Dynamics*" investigates whether self-reported general stress is a mediator in the relationship between socio-economic status (SES) and health. Ever since Grossman's economic model of demand for health (Grossman 1972), much applied research has focused on establishing a causal association between SES and health. Recent econometric literature utilizing micro-level panel data has provided some evidence of a direct association. Indirect causal relationships through mediators have also been a focus of applied work. However, the role of general stress as a pathway has been largely left out of mainstream research and thus has not been thoroughly explored.

My work moves beyond the essentially cross-sectional investigations of stress as a mediator by employing an approach similar to the study of Granger causality in

regression models. I utilize limited dependent variable panel data methods and look for an inter-temporal link from general stress to self-assessed health (SAH) and from annual personal income to general stress. My sample of analysis includes men and women who are major income earners in their families.

I find little evidence that general stress is a pathway from SES to health. While my results suggest a strong negative association between stress and health for both men and women, they provide little support to the hypothesis of a significant effect of income on stress, consistent with the direction of the SES-health gradient.

The next chapter "*Looking for Private Information in Self-Assessed Health*" studies whether self-assessed health status contains information about future mortality and morbidity, beyond the information that is contained in standard "observable" characteristics of individuals (including pre-existing diagnosed medical conditions). Our research interest is largely motivated by the transition from defined benefit to defined contribution pension arrangements and the possible alteration of the longevity risk workers are exposed to. In particular, in an attempt to avoid the negative consequences of an aging population (and hence an aging workforce) on their private pension programs, companies in many developed countries are moving away from a defined benefit (DB) and towards defined contribution (DC) pension arrangement with their employees. Under DB plans workers commit to a retirement income stream (or at least the formula associated with one) at a relatively early stage in their working lives. Under DC plans that commitment is not made until retirement, by which time adverse selection could play a significant role in the market for annuities: only those who are healthy and expect to live

longer would make such purchases. That would drive down the income stream that the annuity seller is prepared to provide and thus make it more expensive for the older population to insure longevity risk.

Our results indicate that even after controlling for pre-existing conditions, socioeconomic characteristics, and a range of risk factors, self-assessed health predicts future mortality and morbidity. In addition, we find some evidence that this effect strengthens with age. We interpret these findings as supportive of the idea that individuals have private information about their likely future health and lifespan and this information increases with age. Thus, any change in pension arrangements that effectively delays the commitment to annuitize may carry with it the cost of exacerbated adverse selection.

The fourth chapter "*Early Retirement Pathways and Post-Retirement Outcomes in Canada*" looks at whether differences in early retirement pathways are associated with differences in post-retirement outcomes. Unlike the abundance and diversity of studies on the determinants of early retirement, the research on post-retirement outcomes other than consumption has been limited. Recent studies have suggested that involuntary retirement is likely to be associated with worse post-retirement outcomes such as financial dissatisfaction (Alan et al. 2007), reduction of spending (Smith 2006), worse health and possibly a higher likelihood of a post-retirement labour force participation (Pyper and Giles 2002). I extend that research by looking at outcomes such as self-assessed health, self-reported general stress and dwelling tenure.

My results show that differences in pre-retirement health indicators (such as self-assessed health and disability), as pathways to early retirement, are likely to be associated with differences in post-retirement health and stress. In addition, "involuntary" retirees (men who may have retired because of health and/or health related conditions) are more likely to experience worse post-retirement outcomes (in terms of health and stress) than men who retire "voluntarily". Retirement circumstances are found to have no statistically significant effect on dwelling tenure.

References

Alan, S., Atalay, K., Crossley, T. (2007), The Adequacy of Retirement Savings: Subjective Survey Reports by Retired Canadians, SEDAP Research Paper No. 199

Grossman, M. (1972), On the Concept of Health Capital and the Demand for Health, *Journal of Political Economy*, vol. 80, no. 2, pp. 223-255

Pyper, W., Giles, P. (2002), Approaching Retirement, *Perspectives of Labour and Income*, vol. 3, no. 9, pp. 5-12

Smith, S. (2006), The Retirement-Consumption Puzzle and Involuntary Early Retirement: Evidence from the British Household Panel Survey, *Economic Journal*, vol. 116, pp. C130-C148

II

Stress as a Pathway in the SES-Health Relationship: Evidence from the Canadian Survey of Labour and Income Dynamics

1. Introduction

The direct relationship between socio-economic status (SES) and health has been an object of theoretical modelling and applied investigation. However, ambiguities in the nature and direction of causality still remain. That emphasizes the need for further research. One approach, commonly utilized in the literature, is to move beyond direct causation and look for mediators (or pathways) of impact. A mediator is a personal characteristic or behaviour that is causally associated with health and SES. It also functions as a transmitter of effects between SES and health.

One characteristic acknowledged frequently in the literature as a possible pathway is stress. However, it has received limited research attention. In this paper, we offer further insights into the mediatory role of self-reported general stress in the relationship between personal income (as a proxy for SES) and self-assessed health.

The theoretical foundation of our work is Grossman's standard economic model of demand for health (Grossman 1972). Grossman regards health as a stock. All inputs and behaviours are treated as investments to that stock. These investments create increments to health and since the increments are dependent on the entire history of personal characteristics, choices and behaviours, so is the current stock of health. Stress may affect the stock of health: the amount of stress in one period may have an impact on

the health stock in future periods. Moreover, it may also be a pathway from income to health. A shock in current income could potentially affect an individual's amount of stress which in turn could influence the stock of health.

To characterize the mediatory role of stress, we consider two separate inter-temporal associations: the association between stress and health (conditional on initial income and initial health), and the association between current income[1] and stress (conditional on initial health and initial stress). If our results suggest the presence of links between stress and health and between income and stress, consistent with a positive SES-health gradient, then there would be evidence to conclude that stress is likely to have a mediatory effect.

The structure of our problem suggests application of an approach similar to the study of Granger causality in regression models, although here limited dependent variable methods must be used given the nature of the data. We employ the longitudinal sample of the Canadian Survey of Labour and Income Dynamics (SLID). The SLID is a detailed survey of the labour behaviour and income of Canadians which also offers categorical indicators of self-assessed health and stress. This unique combination of economic and health-related variables makes the SLID a particularly useful data source for our analysis.

We proceed with an overview of the related literature. Section 3 offers a description of the dataset and the methodology used in the study. The estimation results are reported in Section 4 and discussed in Section 5. Section 6 concludes.

[1] Because of data limitations, we are not able to identify uniquely the permanent and transitory components of current income.

2. Literature Review

The two-way relationship between socio-economic status and health has been an area of significant empirical interest in epidemiology and health economics. A body of recent econometric literature utilizing micro-level panel data and various measures of health and SES provides some evidence of a direct association (Contoyannis et al. 2001, Contoyannis et al. 2004, Banks et al. 2007) as well as causation (Smith 1999, Mulatu et al. 2002, Adams et al. 2003, Attanasio 2003, Smith 2004 and 2005, Adda et al. 2007) between health and SES.

In addition to studying the direct relationship, there has also been research into the indirect, or mediated, relationship between health and SES. One specific mediator is stress. Various indicators of stress are found to be associated with a number of health and SES indicators. Studies link health-related lifestyles and behaviors as well as psychosocial distress to either health or SES (McEwen and Stellar 1993, Schulz et al. 1995, Kelley et al. 1997, Kaplan and Manuck 1999, Sutherland et al. 2002). Moreover, life events and various stressors are shown to be associated with health and mortality (Gardner 2004, Lantz et al. 2005). Further evidence from the labour economics literature suggests that work and personal finance-related life events are causally linked to stress-related health conditions (Browning et al. 2006).

Few papers have focused specifically on the mediatory role of stress in the SES-health relationship. Mulatu et al. (2002) explore the inter-relationships among health (proxied by presence of health conditions), SES (measured by education, family income and occupational status) and psychological distress (characterised by anxiety and self-

deprecation) by utilizing two consecutive waves of a panel including 351 men and 356 women. Controlling for initial conditions, they find that distress is a pathway from SES to health.

Grzywacz et al. (2004) study the very short-term inter-connectedness of education (used as a proxy for SES), daily stress, and physical and mental daily health by estimating a hierarchical model specification. By utilizing a dataset of approximately 1000 individuals followed over eight consecutive days and controlling for past-day's health and psychological distress in each regression, they find that individuals with higher educational attainment reported better physical health and lower distress. In addition, their results indicate that the association between stress and health is dependent on socioeconomic status.

Two studies using the Canadian National Population Health Survey (NPHS) investigate whether lower income is associated with a lower self-rated health and whether psychosocial stressors mediate part of this social gradient. Orpana and Lemyre (2004) utilize the 1994-1995 NPHS and consider exposure to recent life events and chronic stressors. They find that exposure to stressors is associated with poor self-assessed health. Across income adequacy groups, their results suggest that stress exposure accounts for 16% to 26% of the relationship between income group and poor self-rated health among men and for 6% to 15% among women.

In a later paper, Orpana et al. (2007) extend their analysis by following up the 1994-1995 NPHS individuals (who reported their health as excellent, very good, or good) over two years. Their findings show that individuals in the two lowest household income

quintiles had significantly greater odds of experiencing a decline in self-reported health than those in the highest quintile. Stressors are found to be associated with this decline in self-assessed health: they explain 16% of the relationship between the lowest income quintile and the decline in self-reported health, and 10% of the relationship between the second lowest income quintile and the decline in health.

Thus, the literature suggests that stress is likely to be a mediator in the SES-health relationship. The present work aims at expanding the understanding about the mediatory role of stress. We consider a self-reported general stress indicator and employ a large sample of major income earners obtained from two 6-year-long panels of a longitudinal dataset. We utilize a Granger causal framework of modeling. This particular setting allows us to characterize more precisely the inter-temporal health-stress and stress-income associations. As compared to earlier studies, it provides us with the opportunity to employ a large sample and to look over a longer time span; using a six-year follow-up is a significant improvement compared to almost all of the previous work concentrated on stress. We are also able to apply advanced estimation techniques designed to take account of the dynamic nature of the relationships and resolve potential unobserved heterogeneity issues. Our estimation approach involves specifications of an increasing level of complexity which allows us to assess our result sensitivity. In addition, we perform sub-sample analyses as a way to obtain a more detailed view on the gender and over-the-life-cycle characteristics of the mediatory effect of stress.

3. Data, Sample and Methodology

3.1. Survey Details

The paper uses the longitudinal component of the Canadian Survey of Labour and Income Dynamics (SLID). SLID is administered by Statistics Canada and collects data on the labour market activity, income and related socio-economic and demographic characteristics of the Canadian population. The first year of SLID annual data is 1993.

SLID's longitudinal component collects information on individuals in Canada, age 16 and over[2], by following them for a period of six consecutive years. To keep the longitudinal sample representative of the Canadian population, a new panel of respondents is introduced every 3 years. This design also ensures continuity of the SLID sample. Every panel includes about 30,000 adult individuals, members of about 15,000 households.

The sampling frame of the Canadian Labour Force Survey (LFS) is used for the SLID. Moreover, the samples for SLID are selected from the monthly LFS. The Statistics Canada guide to SLID points out that "the LFS sample is drawn from an area frame and is based on a stratified, multi-stage design that uses probability sampling. The total sample is composed of six independent samples, called rotation groups, because each month one sixth of the sample (or one rotation group) is replaced" (Statistics Canada 2004). Two rotation groups of the LFS constitute a SLID panel.

SLID data collection is by computer-assisted interviewing by telephone. At the beginning of each panel, background information about respondents is collected. The

[2] Excluding residents of Yukon, Northwest Territories and Nunavut, residents of institutions and persons living on Indian reserves. These restrictions remove less than 3% of the population.

following 6 regular interviews have a split format with labour topics covered in January and income topics covered in May. In both cases, questions refer to the previous calendar year. The income interview is scheduled in May to take advantage of the income tax filing period when respondents are expected to be more familiar with their income. Usually, over 80% of respondents agree to provide access to their administrative income tax records.

3.2. Variables

Health, stress and personal income are the variables of primary interest to us. The health status variable is derived from the responses to the question "Compared to other people [respondent]'s age, how would you describe [respondent]'s state of health? Would you say it is..." with possible answers (1) "Excellent?", (2) "Very good?", (3) "Good?", (4) "Fair?", (5) "Poor?". Hence health has five ordered categories, each corresponding to a particular answer. It is recorded for persons age 16 and over, and is self-reported.

The question: "Would you describe [respondent's] life as . . ." having possible answers (1) "Very stressful?", (2) "Somewhat stressful?", (3) "Not very stressful?", (4) "Not at all stressful?" defines the stress status variable. It has four ordered categories each corresponding to a particular answer. Stress is observed for persons age 16 and over, and is self-reported.

The personal income variable reports annual personal after-tax income, measured in Canadian dollars in a reference year. Income is continuous and is recorded for all respondents of age 16 and over.

A description and additional information about the health, stress and income variables are provided in Table 1. Tables 2 and 3 offer the transition matrices of health and stress over the two 6-year-long panels of SLID.

The set of variables used as controls consists of age, gender, mother tongue, immigrant status (immigrant if been in Canada for less than 10 years), educational attainment, disability status, family composition, annual labour force status (LFS), dwelling tenure (owner/renter), rural/urban resident, and region of residence in Canada. The choice of the controls set is standard and conforms with the literature. The complete specification of the controls and their summary statistics are presented in Table 4, Panel 1.

A potential issue with the health and stress variables is that they are self-reported. That characteristic could make them susceptible to subjective interpretation[3] (although the health question in particular attempts to limit such possibility by requiring a comparative evaluation). Unlike the Canadian National Population Health Survey, SLID does not provide an alternative measure of health or stress. Alternative measures of health such as the Health Utility Index (HUI) for example, try to capture respondent's health more precisely. This usually is achieved by asking series of questions about specific physical and functional characteristics and then aggregating the answers following a particular formula In SLID however, there are no auxiliary variables that could serve that purpose.

[3] As well as index and cut-point shift issues (Lindeboom and Van Doorslaer 2004) as explained in a later footnote.

Although self-reported health and stress might be noisier than alternative measures, they provide a subjective health or stress assessment. A number of studies (Mossey and Shapiro 1982, Okun et al. 1984, McCallum et al. 1994, Idler and Kasl 1995, Idler and Benyamini 1997, Schwarze et al. 2000, Burstrom and Fredlund 2001, Van Doorslaer and Gerdtham 2003 and Banks et al. 2007) find self-assessed health to be a predictor of future mortality and morbidity once current health conditions and risk behaviours are controlled for. These findings suggest that self-reported measures include extra information which other measures are incapable of extracting.

One particular advantage of the SLID is that it provides detailed and highly reliable income data. This enables us to look specifically at personal income and include it as a continuous variable in our analysis. We convert the personal income of every respondent to 1996 dollar value. Furthermore, we construct an income spline (on a logarithmic scale) to study income effects within income categories. That seems particularly important since the effect of income may vary over income groups.

Lastly, there are variables in the list of controls that could potentially have mediatory roles in the health-SES relationship. These are disability status and labour force participation status. For example, a low health stock could potentially lead to a disability which may affect current personal income negatively and consequently reduce the stock of health even further. Also, a change in labour force participation status (e.g. a transition from full-year employment to part-year employment) caused by a decrease in health may potentially result in lower current personal income. These issues however are left to future work.

3.3. Sample Restrictions and Sub-Samples Design

For our analysis the initial longitudinal SLID sample is subjected to a number of restrictions.

Firstly, we impose a time restriction due to the survey design. The questions on health and stress status are first asked in 1996. This automatically removes the period 1993 to 1995 from consideration. Since a panel of three years is regarded as incomplete and insufficiently long to be included in the final sample, we employ only the second and third panels of the SLID (years 1996 to 2001 and 1999 to 2004).

Secondly, by including the annual labour force status variable in the analysis, the sample is implicitly restricted to individuals in the age-range 16 to 69. The reason behind that limiting condition is SLID's design which does not provide labour force information for respondents of age 70 and over. We further restrict the sample to individuals in the age range 20 to 60. This is targeted at keeping the sample within the age range of predominantly active labour force participation, so that we can capture more precisely the effects of personal income.

Thirdly, SLID's sample is subject to unit and item non-response. While unit non-response is not a concern (since it is handled by Statistics Canada before data release), item non-response remains an issue. To resolve it, we exclude cases in which item non-response is present.

In addition, we limit our analysis to the major income earner of an economic family in the first period of each panel. With this restriction we aim to increase the

homogeneity of our sample. In particular, respondents who are not major income earners in the first period may be subjected to different health-stress-income effects.

Individuals who move within the 6-year period of each panel may do so due to reasons related to health, stress or income. To avoid potential bias caused by the inclusion of these respondents, we constrained our sample to persons with an unchanged location of residence for the duration of each panel.

We then pool the observations in the second and third SLID panels. Thus, we obtain a 6-year-long panel of 5520 respondents (3647 male and 1873 female). We perform separate analyses by gender and age-groups (Table 4, Panel 2 offers the respective sample sizes). This provides for a more comprehensive description of the mediatory effect of stress and overcomes possible issues related to reporting heterogeneity[4].

Finally, we test for non-random attrition. The test results are discussed below.

3.4. Methodology

3.4.1. Principal Model

The health-stress and stress-income relationships are modelled using the following general Granger type causal specification:

[4] The concern is that ordered responses on health (and stress) questions may differ across populations or even across subgroups of a population. Using the Canadian National Population Health Survey data Lindeboom and Van Doorslaer (2004) find evidence of index shifting (a parallel, equal distance and same direction shift in the cut-points separating the categories of a variable) and cut-point shifting (a non-parallel, unequal distance and different direction shift in the cut-points separating the categories of a variable) in self-assessed health caused by age and gender.

$$y_{it} = \rho y_{it-1} + \mathbf{x}_{it-1}\boldsymbol{\beta} + \mathbf{w}_{it}\boldsymbol{\delta} + \mathbf{z}_i\boldsymbol{\tau} + u_i + \varepsilon_{it} \quad (1)$$

$$(i=1,...,N \text{ and } t=2,...,T)$$

Here, y_{it} represents the outcome at time t; y_{it-1} is the one period lagged dependent variable. Vector $\mathbf{x}_{it-1}$ consists of one period lagged variables, such that $E(x_{it}\varepsilon_{is}) = 0$ for all i, t and s. Vector $\mathbf{w}_{it}$ includes time-variant variables with the property $E(w_{it}\varepsilon_{is}) = 0$, again for all i, t and s. Vector $\mathbf{z}_i$ contains time-invariant variables. In this specification, u_i is the unobserved heterogeneity component. The error term ε_{it} is unique for every individual and time, and is assumed to be normally distributed, uncorrelated across individuals and time, and uncorrelated with the unobserved heterogeneity parameter.

One could think of specification (1) as a way of obtaining future predicted values of the dependent variable. For example, if one is interested in a person's health, then the future predicted health status of that individual will be based on their current health status, stress status and personal income as well as on their personal socio-economic characteristics.

Specification (1) is employed to construct a health model and a stress model. However, as health and stress are categorical variables, specification (1) cannot be estimated directly. There are two approaches that can be used in that situation. The first approach (following Mundlak 1978, Heckman 1981, Chamberlain 1984 and Wooldridge 2005) treats the latent dependent variable as unobservable. Hence the modelling is based on the observed categorical variable defined over the values of the latent. Under this

approach, formulation (1) involves a categorical dependent variable on the left-hand side and a categorical lagged dependent variable on the right-hand side.

The second approach proposed by Bover and Arellano (1997) also treats the latent dependent variable as unobservable. In contrast however, Bover and Arellano estimate a dynamic model in latent form by combining its reduced form with its demeaned specification and then imposing particular parameter restrictions.

In this study we apply the first modelling approach. The reason is two-fold. Firstly, Bover and Arellano do not present econometric evidence for the applicability (and reliability) of their procedure to ordered outcomes. Secondly, their estimator is essentially a within-group estimator and hence belongs to the class of the fixed effects techniques. The utilization of a fixed-effects estimation approach in our work is impractical because of the characteristics of our regressors. Almost all right-hand side variables, including these of primary interest, are categorical with multiple levels. Hence, an application of a fixed-effects estimator would make the transitions over the multiple categories of a variable uniquely unidentifiable.

We re-specify formulation (1). In particular, y_{it} now stands for the ordered categorical outcome whereas the parameter ρ is substituted by a vector of parameters λ_j corresponding to the j categories of the one-period lagged dependent categorical variable y_{it-1}.

We use three econometric techniques to estimate the health and stress models. These are: (i) a pooled (stacked) ordinary least squares (OLS) with a cardinalized dependent variable, (ii) a pooled ordered probit (POP), and (iii) a random effects ordered

probit (REOP). The reason to employ these different types of estimators is to investigate to what extent the results in the study are sensitive to a particular technique and identify those findings that are consistent across techniques.

The three techniques of analysis involve different estimation methods that rely on different sets of assumptions. Thus, across model comparisons are possible by looking at either relative effects (e.g. OLS vs. rest) or average partial effects (e.g. POP vs. REOP). Moreover, to ensure that the partial effects obtained from the REOP are comparable with those from the POP, we make an adjustment to the REOP regression coefficients using the estimated value of the formula $\beta_a = \beta/(1+\sigma_v^2)^{0.5}$. Here, β_a is a population average estimate, β is a REOP parameter estimate and $(1+\sigma_v^2)$ is the estimated total error variance in the REOP model. This adjustment is required since the estimated error variance in the REOP is not necessarily equal to unity, as is the case for the POP model. Finally, we compute the average partial effects following Wooldridge (2005).

Specification (1) indicates that both health and stress models have identical $\mathbf{w}_{it}$ vectors. They include age, family status and educational attainment. Vector $\mathbf{z}_i$ is also identical in both models and includes gender, immigrant status, a mother tongue indicator, rural/urban residency, and the region of residence. The $\mathbf{x}_{it-1}$ vector, $\mathbf{x}_{it-1} = \left(x_{it-1}^1, x_{it-1}^2, \mathbf{x}_{it-1}^3\right)$, however is model specific. In the health model $\mathbf{x}_{it-1}$ is comprised of stress status (x_{it-1}^1), the logarithmic-scale linear spline (with 4 nodes[5]) of annual personal income (x_{it-1}^2), and annual labour force status and disability (both elements of

[5] The nodes of the income spline in terms of actual income are at 20, 40, 60 and 80 thousand dollars.

the $\mathbf{x}^3_{\mathbf{it-1}}$ sub-vector); in the stress model $\mathbf{x}_{\mathbf{it-1}}$ contains health status (x^1_{it-1}), the logarithmic-scale linear spline of annual personal income (x^2_{it-1}) and $\mathbf{x}^3_{\mathbf{it-1}}$ as defined above.

3.4.2. Individual Heterogeneity Model

Two problems could potentially arise in the estimation of (1). Firstly, the error term, composed of the individual specific effect u_i and the white noise ε_{it}, could be correlated with one or more of the right-hand side variables through the u_i component. Secondly, because of the dynamic nature of (1), the dependent variable in the initial period (the initial condition) enters the likelihood function. In case the initial condition is endogenous and the process modelled is not initially in equilibrium, that may render the estimation inconsistent (Heckman 1981).

A popular approach proposed by Mundlak (1978) and Chamberlain (1984) for handling the first potential problem is to parameterize the individual effect u_i. The second problem (given that both the initial condition may well be endogenous and the process modelled may be out of equilibrium initially) could be handled in two ways: (i) model the initial condition using right-hand side variables and possibly other exogenous variables (Heckman 1981), or (ii) model the distribution of the individual effect u_i conditional on the initial value of the dependent variable (Wooldridge 2005). The approach proposed by Wooldridge (2005) targets both potential issues and therefore is the one we apply.

We choose a parametric model for u_i of the following form

$$u_i = \gamma_0 + \gamma_1 y_{i1} + \gamma_2 q_{i1}^1 + \gamma_3 q_{i1}^2 + \mathbf{w}_i \boldsymbol{\gamma}_4 + v_i \quad (2)$$

$$(i=1,...,N)$$

where y_{i1} is the initial value of the outcome in (1) (γ_1 stands for a set of coefficients corresponding to the levels of y_{i1}), $\mathbf{w}_i$ contains the means over T-periods of the continuous variables in the $\mathbf{w}_{it}$ as well as the means over T-periods of the categorical variables in $\mathbf{w}_{it}$ and $\mathbf{x}_{it-1}^3$. Also, q_{i1}^1 is the initial value of the first variable in $\mathbf{x}_{it-1}$ (that is again, stress in the health model and health in the stress model) and q_{i1}^2 is the initial period log-scale level of annual personal income. The parameter γ_2 represents a set of coefficients corresponding to the levels of q_{i1}^1. Here, $v_i \sim N(0, \sigma_v^2)$ and is independent of all other right-hand side variables by assumption.

This particular formulation of (2) conforms with the specifications used in the literature (Chamberlain 1984, Contoyannis et al. 2004, Wooldridge 2005) except for the q_{i1}^1 and q_{i1}^2 components. The inclusion of q_{i1}^1 is based on the state-persistence of both health and stress; Tables 2 and 3 provide evidence of that. Then controlling for the initial condition of stress in the health model and the initial condition of health in the stress model would allow for correlation between the individual effect and the respective initial value. Presumably, that would eliminate the correlation between the individual component u_i and x_{it-1}^1.

Having q_{i1}^2 in (2) is a similar way of taking out the correlation between the individual component u_i and x_{it-1}^2. Theory suggests the use of permanent income instead of the initial period level of annual personal income. While permanent income is not available in the SLID, we can proxy it as the mean over the T-periods of the annual personal income that we observe. The problem with that approach is the relatively short time-span of our panel ($T = 6$ years). It is conceivable that if included, this proxy would raise concerns about multicollinearity.

Finally, we combine (1) and (2) and estimate the resulting equation using the same econometric techniques as the ones employed earlier for the estimation of specification (1) alone.

3.4.3. Auxiliary Model

Under the assumption that the model of the unobserved individual effect is correctly specified, all estimators from the previous section will provide consistent coefficient estimates. There is however the possibility that the model for u_i is misspecified.

To provide an alternative view of the health-stress and stress-income inter-temporal associations, we employ a Granger type causal model with full lag structure of the following form

$$y_{iT} = \sum_{t=1}^{T-1} \rho_t y_{it} + \sum_{t=1}^{T-1} \mathbf{x}_{it}^{1,2} \boldsymbol{\beta}_t + \mathbf{x}_{iT-1}^{3} \boldsymbol{\beta}_t + \mathbf{w}_{iT} \boldsymbol{\delta} + \varepsilon_{iT} \quad (3)$$

$$(i=1,...,N).$$

In this specification, vector $\mathbf{x}_{it}^{1,2}$ has the following elements: in the health model, $\mathbf{x}_{it}^{1,2}$ includes stress status and the logarithmic-scale linear spline of annual personal income; in the stress model, $\mathbf{x}_{it}^{1,2}$ contains health status and the logarithmic-scale linear spline of annual personal income. The vector $\mathbf{x}_{iT-1}^{3}$ is composed of labour force status and disability in the period *T-1* and $\mathbf{w}_{iT}$ includes the full set of controls in period *T*.

We estimate the health and stress specifications of (3) by employing clustered (within an individual) OLS and ordered probit (OP) estimators.

By exhausting the time dimension of the data, and hence essentially estimating a cross-sectional-type model, we aim to obtain one-period-lag filtered effects. Thus, the interpretations of the estimates of (1) and (3) are substantially different. Here, we address the question whether a link between the variables of interest exists over the last two years of the panel, conditional on their entire observed history.

A potential problem that could plague this typical Granger-causal formulation is the unobserved heterogeneity which, clearly, we do not control for. Despite that limitation however, the additional perspective that (3) provides is of comparative interest. Moreover, this approach to a certain extent mimics the modelling approaches employed in other studies on stress.

4. Results

We first look at the results from the health and stress estimations that do not involve a model of the individual component. The coefficient estimates of previous period income, health and stress on current period health and stress are presented in Tables 5 to 7, separately for men and women. Table 5 shows the OLS estimates, while Tables 6 and 7 present the estimates of the effects on the latent health and stress obtained from the POP and REOP models, respectively.

For both men and women, reporting a lower level of general stress (less stress) is inter-temporally associated with lower future latent health, hence better future health; as health is coded from (1) "excellent" to (5) "poor", a negative coefficient estimate indicates higher health than the one reported by the reference group (major income earners indicating very stressful life). This finding of a clear gradient (relative to the base group) in the stress effect is consistent across econometric techniques (Tables 5 to 7, Columns 1 and 2) Furthermore, all estimates are statistically significant at conventional levels ($p<0.05$)[6]. The results also show that the effect of stress on future health is larger (approximately double) for females than for males.

We now look at the results obtained from the stress model (Tables 5 to 7, Columns 3 and 4), and more specifically to the estimates of the effect of personal income on future stress. Surprisingly, we find (consistently across techniques) that a marginal increase in income for those in the middle income range – i.e. those whose annual personal income is in the $40-60,000 range, is inter-temporally associated with an

[6] All tests for significance in this thesis are two-tailed.

increase in future stress; as stress is coded from (1) "very stressful life" to (5) "not at all stressful life", a negative coefficient estimate indicates higher stress. This effect is very similar across genders and is statistically significant ($p<0.05$). Changes in income over the rest of the income groups have no statistically significant impact on future stress. (Although some estimates are occasionally significant within a particular model, their significance is not robust across models.)

We next estimate the health and stress models separately for the under-40 and over-40 age-groups by gender. Thus, we investigate whether the effects of stress on future health (Table 8) and income on future stress (Table 9) vary with age. Each of Tables 8 and 9 reports the estimated effects obtained from the application of the three econometric techniques in a separate pane and over gender and age-group.

For the health model and male sample, our findings show that reporting lower stress levels (again relative to the base group of those reporting very stressful life) is associated with better future health in both age groups. However, the effect of stress in the older age-group is stronger (with very few exceptions) and consistently statistically significant across estimators. We observe a similar pattern in the female sample. The across gender-age-group comparisons support our earlier finding that the effects of stress on future health are larger for women.

The results from the stress model (Table 9) reveal that a unit increase in income for men in the middle income range is associated with an increase in future stress. This effect is statistically significant and consistent across techniques. It is also very similar in magnitude for both age-groups (slightly higher for the over-40 group). We see a similar

direction of the effect of income for those men who are over 40 years old and whose annual income is above $80,000; the estimate here is larger than the estimate for the $40-60,000 group.

In the female sample (Table 9, Columns 3 and 4) we do not observe a well established pattern of significance across estimators. Just as for men, the results indicate that an increase in income is associated with in increase in future stress for women in the middle income group; for women, however, most estimates are only marginally significant ($p<.10$). The magnitudes of the effects in that income range are again very similar across female age-groups, and similar also to the corresponding male age-group estimates.

We next re-estimate the models of health and stress to incorporate the formulation of the unobserved heterogeneity component. Tables 10 to 12 summarize our results. As before, Table 10 offers the OLS, while Tables 11 and 12 present the POP and REOP estimates, respectively. Once again, the models are fitted to both gender samples separately.

With the heterogeneity modelling, the results follow the same pattern as before. For genders, lower stress is inter-temporally related with better future health. There is a well established gradient in the effect of stress on health status for both men and women and the estimates are significant at conventional levels across estimation techniques (the only exception is the REOP estimate of Not at all Stressful, -0.078, which is significant at the 10% level). We also observe that the stress effect is once again higher in the female sample.

The stress model results (Tables 10 to 12, Columns 3 and 4) re-confirm our previous surprising finding that an increase in income for individuals having an annual personal income between $40,000 and $60,000 is inter-temporally associated with higher future stress. Again, the estimates are statistically significant across estimators. The magnitudes of the effects are slightly larger for women.

The gender-age-group results from the health and stress model estimation are presented in Tables 13 and 14. The organization of the tables follows the structure of Tables 8 and 9. By looking at the first two columns of Table 13, the health model estimation results for men, we see that although there is some evidence of a gradient in the effect of stress on future health in the under-40 age group, the estimates are predominantly insignificant. In the over-40 sample, by contrast, the majority of effects are significant at least at the 10% level. However, one could hardly find evidence (across techniques) of a well established gradient. In the female sample (last two columns of Table 13) we see a more familiar pattern: a well established gradient, significance of effects at conventional levels (except for the case of REOP, under-40 group) and, overall, larger stress estimates in the over-40 age group.

The results from the stress model (Table 14) reproduce broadly our previous unexpected findings for the effect of income on future stress (Table 9). Here, however, we find no statistically significant effect of income on stress in the under-40 age groups for either gender. In the sample of over 40 year old men, only a marginal change in income of the middle income group has a statistically significant effect on stress

($p<0.05$). The same is valid for the female sample with the exception that the effect is marginally significant. That estimate for women is also slightly larger in magnitude.

Compared to the estimates from the models that do not involve a model of the individual component (Tables 5 to 9), we see that all effects here (Tables 10 to 14) are smaller in magnitude (with very few exceptions). This suggests that our heterogeneity model is picking up variation, and that failing to control for the effect of the individual component may lead to inconsistent estimates.

Before turning our attention to the results from the auxiliary model, we look at the marginal effects of current stress on excellent future health (Table 15) and the marginal effects of income on the highest future stress category, Very Stressful (Table 16). Tables 15 and 16 have two parts; Part A corresponding to an estimation without, and Part B with, the model of the individual component. Each part lists the marginal effects of the POP and REOP models by gender.

The estimates offer no additional surprises: Table 15 shows that those individuals who report lower stress levels have a higher probability of reporting excellent future health than those individuals who report being in the highest stress category. The magnitudes of the effects across models within a gender group are very similar (e.g. Table 15, Part A). There is a clearly established gradient in the marginal effects for both men and women, irrespective of whether we control for the individual component or not. When we do control for it, the marginal effects are smaller. Again, the female sample estimates dominate those from the male sample for a particular estimator. Significance of the effects is only an issue in the REOP case, male sample, when we control for

heterogeneity (Table 15, Part B, Column 3). We see that in this case, only two of the three estimates are marginally significant.

The results presented in Table 16 show that a marginal increase in income of those individuals in the middle income group is associated with an increase in the probability of reporting the highest stress level in the next period. This finding is consistent across models and genders. The magnitude of this effect is very similar for both men and women and is practically unaffected by whether we control for heterogeneity or not. The significance of the estimates remains unchanged across specifications. An exception is the significance of the estimates for the female sample which appear to be sensitive to the specification (Table 16, Part A, Column 4 and Table 16, Part B, Column 4). All of the remaining income group estimates, as usual, are insignificant.

We now consider the results from our auxiliary model (Tables 17 and 18). Once again, this model is looking at whether stress is inter-temporally associated with health and whether income is inter-temporally associated with stress over the last two years of our 6-year-long sample, conditional on the last year controls and the observed history (years 1 through 4) of health, stress and income.

Tables 17 and 18 are organized in the same fashion as Tables 5 and 6. Table 17 offers the estimates from the cardinalized models of health and stress, whereas Table 18 presents the results from the OP estimator. The models are fitted separately to the male and female samples.

For both men and women, reporting the lowest level of stress (Not at all Stressful) is associated with a better future health, compared to the reference group (here again, the reference group is all individuals who indicate the highest level of stress). That effect is consistent across models (Tables 17 and 18, Columns 1 and 2) and is significant at least at the 10% level. In the female sample, there is also some evidence of a gradient in the impact of stress. In addition, the estimates for women are again greater in magnitude.

Unlike all our previous findings, the results for women do not support the hypothesis that income is inter-temporally related with stress. Over all specifications, the income estimates are insignificant (Tables 17 and 18, Column 4). By contrast, in the male sample (Tables 17 and 18, Column 3) we repeatedly find that a marginal increase in income for those in the middle income category is associated with an increase in future stress. Moreover, that effect is not only high in magnitude (compared to all our previous results) but is also highly significant ($p<0.01$)[7].

Finally, we test for non-random attrition following Verbeek and Nijman (1992). First, we limit our dynamic modelling to the first three years of our sample. We then add dummies capturing attrition over the last three time periods to our models and re-estimate them. The results do not suggest that attrition is a serious issue in our analysis. The attrition dummies are only occasionally significant and our main results are not significantly affected by their inclusion[8].

[7] The results are not significantly affected by the inclusion of the SLID sampling weights in the estimation.
[8] Full results are available from the author on request.

5. Discussion

We now look at our findings in the context of our research question of whether general stress is a pathway in the SES-health relationship. To establish the mediatory role of stress, our results have to provide evidence of an inter-temporal link from stress to health and from income to stress that is consistent with a positive SES-health gradient.

For both male and female samples and across all specifications we find a link between stress and health (Tables 5 to 7, 10 to 12, 17, and 18). Reporting lower stress level, compared to the base-group, was shown to be inter-temporally associated with higher future health (and an increasing probability of reporting excellent future health). This effect was also shown to be greater in magnitude for women. The gender-age group analyses (Tables 8, 9, 13 and 14) suggested that stress has a significant effect on the health of those men who are over the age of 40. That was not the case for women. For them, stress mattered irrespective of age and its effect was, overall, stronger for the over-40 age group.

On the income-stress side, the results from the principal specifications and full gender samples suggested consistently that an increase in income for all individuals in the middle income range (Tables 5 to 7 and 10 to 12) is correlated with an increase in future stress. The magnitudes of this effect were very similar in both gender samples. Income was not found to have a significant effect on stress in any other income category. The results from the gender-age analyses (Tables 8, 9, 13 and 14) showed that these findings were valid only for men and women in the over-40 age group. Again, no income category estimate was significant in the under-40 age groups.

Our results suggest that while the association between stress and health is relatively strong and has the expected negative direction (lower stress is inter-temporally associated with better health) for both men and women, an association between income and stress is found only for the group of men who are over the age of 40 and whose personal income is in the middle income range. Moreover, the direction of the latter effect implies that an increase in income is associated with higher future stress. That positive association is surprising and suggests a wrong direction for the SES-health gradient. Thus, we find little evidence of a mediatory role of general stress in the SES-health relationship.

Clearly, our conclusions on the mediatory role of stress are driven by the results from the income-stress models. The insignificant effects for both gender samples of major income earners and the wrong direction of the effects for the group of males over 40 and in the middle income range are surprising. A likely reason for our findings could be that general stress is dominated by socio-economic characteristics other than personal income. Thus, a relatively small change in income may have a marginal impact on stress. Also, males over 40 with middle-group income may work longer hours or take heavier workloads to earn more income. Those choices could be associated with higher future stress. These and other possible reasons for our findings are subject to future study.

6. Conclusion

In this paper we investigate whether general stress is a mediator in the SES-health relationship. We use a 6-year long panel of the Canadian Survey of Labour and Income

Dynamics and find little evidence that stress plays the role of a pathway. While our results suggest a strong negative association between stress and health for both men and women, they provide little support to the hypothesis of a significant effect of income on stress, consistent with the direction of the SES-health gradient.

Our findings are not entirely consistent with other findings in the stress literature. Unlike other studies (Mulatu et al. 2002, Grzywacz et al. 2004, Orpana and Lemyre 2004, Orpana et al. 2007), we find little evidence that stress plays the role of a mediator. One possible reason for that difference could be our sample: we are analyzing only major income earners; other studies do not impose that restriction. Also, the estimation approach employed in our work substantially differs from that utilized in other relevant research.

In the future work, it would be interesting to employ a health survey, such as the Canadian National Population Health Survey. It could help provide further insights into the mediatory role of stress and assess the generality of our findings.

Following the literature, we could broaden the analysis by looking at different stressors and investigate their particular mediatory roles. The application of dynamic techniques involving simultaneous modelling would be a natural estimation approach.

Finally, as we have previously suggested, indicators such as disability status and labour force participation could potentially have mediatory roles in the health-SES relationship. Studying these possibilities is an additional attractive avenue of future work.

References

Adams, P., Hurd, M., McFadden, D., Merrill, A., Ribeiro, T. (2003), Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status, *Journal of Econometrics*, vol. 112, no. 1, pp. 3-56

Adda, J., Banks, J. Von Gaudecker, H. (2007), The Impact of Income Shocks on Health: Evidence from Cohort Data, Institute for Fiscal Studies, IFS Working Paper No. W07/05

Attanasio, O., Emmerson, C. (2003), Mortality, Health Status and Wealth, *Journal of the European Economic Association*, vol. 1, no. 4, pp. 821-850

Banks, J., Crossley, T., Goshev, S. (2007), Looking for Private Information in Self-Assessed Health, SEDAP Working Paper No. 219

Banks, J., Marmot, M., Oldfield, Z., Smith, J. (2007), The SES Health Gradient on Both Sides of the Atlantic, Institute for Fiscal Studies, IFS Working Paper No. W07/04

Bhargava, A., Sargan, J. D. (1983), Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods, *Econometrica*, vol. 51, no. 6, pp. 1635-1659

Bover, O., Arellano, M. (1997), Estimating Limited Dependent Variable Models from Panel Data, *Investigaciones Economicas*, Fundación SEPI, vol. 21, no. 2, pp. 141-166

Browning, M., Dano, A. M., Heinesen, E. (2006), Job Displacement and Stress-Related Health Outcomes, *Health Economics*, vol. 15, no. 10, pp. 1061-1075

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?, *Epidemiology and Community Health*, vol. 55, no. 11, pp. 836-840

Chamberlain, G. (1984), Panel Data, In: Griliches, Z., Intriligator, M. D. (Eds.), Handbook of Econometrics, North Holland, Amsterdam, vol. 1, pp. 1247-1318

Contoyannis, P., Jones, A., Rice, N. (2004), The Dynamics of Health in the British Household Panel Survey, *Journal of Applied Econometrics*, vol. 19, no. 4, pp. 473-503

Contoyannis, P., Rice, N. (2001), The Impact of Health on Wages : Evidence from the British Household Panel Survey, *Empirical Economics*, vol. 26, pp. 599-622

Deaton, A. (1999), Inequalities in Income and Inequalities in Health, NBER Working Paper No. 7141

Ettner, S. (1996), New Evidence on the Relationship between Income and Health, *Journal of Health Economics*, vol. 15, pp. 67-85

Gardner, J., Oswald, A. (2004), How is Mortality Affected by Money, Marriage, and Stress?, *Journal of Health Economics*, vol. 23, pp. 1181-1207

Grossman, M. (1972), On the Concept of Health Capital and the Demand for Health, *Journal of Political Economy*, vol. 80, no. 2, pp. 223-255

Grzywacz, J., Almeida, D., Neupert, S., Ettner, S. (2004), Socioeconomic Status and Health: A Micro-level Analysis of Exposure and Vulnerability to Daily Stressors, *Journal of Health and Social Behavior*, vol. 45, pp. 1-16

Heckman, J. (1981), The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, In: Manski, C., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press: Cambridge, MA, pp. 179-195

Idler, E., Benyamini, Y. (1997), Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies, *Journal of Health and Social Behavior*, vol. 38, no. 1, pp. 21-37

Idler, E., Kasl, S. (1995), Self-Ratings of Health: Do They Also Predict Change in Functional Ability?, *Journal of Gerontology*, vol. 50, pp. S344-S353

Kaplan, J., Manuck, S. (1999), Status, Stress, and Arteriosclerosis: The Role of Environment and Individual Behaviour, In: Adler, N. et al. (Eds.), Socioeconomic Status and Health in Industrialized Nations, Annals of the New York Academy of Science, vol 896, pp. 145-161

Kelley, S., Hertzman, C., Daniels, M. (1997), Searching for the Biological Pathways between Stress and Health, *Annual Review of Public Health*, vol. 18, pp. 437-462

Lantz, P., House, J., Mero, R., Williams, D. (2005), Stress, Life Events, and Socioeconomic Disparities in Health: Results from the Americans' Changing Lives Study, *Journal of Health and Social Behavior*, vol. 3, pp. 274-288

Lindeboom, M., Van Doorslaer, E. (2004), Cut-Point Shift and Index Shift in Self-Reported Health, *Journal of Health Economics*, vol. 23, pp. 1083-1099

McCallum, J., Shadbolt, B., Wang, D. (1994), Self-Rated Health and Survival: 7 Years Follow-up Study of Australian Elderly, *American Journal of Public Health*, vol. 847, pp. 1100-1105

McEwen, B., Stellar, E. (1993), Stress and the Individual: Mechanisms Leading to Disease, *Archives of Internal Medicine*, vol. 153, pp. 2093-2101

Mossey, J., Shapiro, E. (1982), Self-Rated Health: A Predictor of Mortality among the Elderly, *American Journal of Public Health*, vol. 72, pp. 800-808

Mulatu, M. S., Schooler, C. (2002), Causal Connections between Socio-Economic Status and Health: Reciprocal Effects and Mediating Mechanisms, *Journal of Health and Social Behavior*, vol. 43, no. 1, pp. 22-41

Mundlak, Y. (1978), On the Pooling of Time Series and Cross-Section Data, *Econometrica*, vol. 46, no. 1, pp. 69-85

Okun, M., Stock, W., Haring, M., Witter, R. (1984), Health and Subjective Well-Being: A Meta-Analysis, *International Journal of Ageing and Human Development*, vol. 19, no. 2, pp. 111-132

Orpana, H., Lemyre, L. (2004), Explaining the Social Gradient in Health in Canada: Using the National Population Health Survey to Examine the Role of Stressors, *International Journal of Behavioral Medicine*, vol. 11, no. 3, pp. 143-151

Orpana, H., Lemyre, L., Kelly, S. (2007), Do Stressors Explain the Association between Income and Declines in Self-Rated Health? A Longitudinal Analysis of the National Population Health Survey, *International Journal of Behavioral Medicine*, vol. 14, no. 1, pp. 40-47

Schulz, R., O'Brien, A., Bookwala, J., Fleissner, K. (1995), Psychiatric and Physical Morbidity Effects of Dementia Caregiving: Prevalence, Correlates, and Causes, *Gerontologist*, vol. 35, no. 6, pp. 771-791

Schwarze, J., Andersen, H., Anger, S. (2000), Self-Rated Health and Changes in Self-Rated Health as Predictors of Mortality - First Evidence from German Panel Data, Humboldt Universitaet Berlin, Sonderforschungsbereich 373

Smith, J. (1999), Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status, *Journal of Economic Perspectives*, vol. 13, no. 2, pp. 145-166

Smith, J. (2004), Unravelling the SES-Health Connection, *Population and Development Review*, vol. 30, pp. 108-132

Smith, J. (2005), Consequences and Predictors of New Health Events, In: Wise, D. (Ed.), Analyses in the Economics of Aging, Chicago: University of Chicago Press, pp. 213-237

Statistics Canada (2004), SLID Public Use Microdata Documentation, Statistics Canada

Sutherland, C., Bybee, D., Sullivan, C. (2002), Beyond Bruises and Broken Bones: The Joint Effects of Stress and Injuries on Battered Women's Health, *American Journal of Community Psychology*, vol. 30, no. 5, pp. 609-636

Van Doorslaer, E., Gerdtham, Ulf-G. (2003), Does Inequality in Self-Assessed Health Predict Inequality in Survival by Income? Evidence from Swedish Data, *Social Science and Medicine*, vol. 57, pp. 1621-1629

Verbeek, M., Nijman T. (1992), Testing for Selectivity Bias in Panel Data Models, *International Economic Review*, vol. 33, no. 3, pp. 681-703

Wooldridge, J. (2005), Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity, *Journal of Applied Econometrics*, vol. 20, no. 1, pp. 39-54

Appendix

Table 1. DESCRIPTION OF THE HEALTH, STRESS AND INCOME VARIABLES

Variable	Short Variable Name	Description
Self-Assessed Health	Health	Health has five ordered categories: Excellent, Very good, Good, Fair and Poor, coded (1) to (5), where (1) is Excellent and (5) is Poor In all regressions where Health is an independent variable, Excellent Health is the reference category
Self-Reported Level of Stress in One's Life	Stress	Stress has four categories: (Life is) Very stressful, Somewhat stressful, Not very stressful and Not at all stressful, coded (1) to (4), where (1) is Very stressful and (4) is Not at all stressful. In all regressions where Stress is an independent variable, Very stressful is the reference category
Natural Logarithm of Personal Yearly Income	Income	Income is the natural logarithm of the Personal Yearly Income. Income enters in two ways in the health and stress models: as a log-scale personal yearly income spline and as a log-scale personal yearly income in base year

Table 2. TRANSITION MATRIX OF SELF-ASSESSED HEALTH (IN PERCENTAGES, OVER 6 YEARS)

	Future Health				
Current Health	Excellent	Very Good	Good	Fair	Poor
Excellent	54.3	33.6	10.5	1.3	0.3
Very Good	19.1	55.4	22.0	3.1	0.4
Good	8.5	32.1	47.4	10.2	1.8
Fair	3.4	12.9	30.6	41.6	11.4
Poor	1.4	4.2	11.5	27.1	55.8
Total	23.8	39.3	25.8	8.0	3.1

Table 3. TRANSITION MATRIX OF SELF-REPORTED LEVEL OF STRESS IN LIFE (IN PERCENTAGES, OVER 6 YEARS)

	Future Stress			
Current Stress	Very Stressful	Somewhat Stressful	Not Very Stressful	Not Stressful
Very Stressful	49.5	41.6	7.0	1.9
Somewhat Stressful	13.1	68.7	15.0	3.2
Not Very Stressful	5.0	38.7	44.2	12.1
Not Stressful	3.4	22.6	34.9	39.1
Total	16.7	54.1	21.7	7.5

Table 4. SUMMARY STATISTICS

Panel 1. Summary Statistics of the Control Variables

Variable	Variable Category	Mean	Standard Deviation
Log(Income):	Level	3.30	1.04
Education:	Less than high school [a]	0.18	0.38
	High school graduate	0.17	0.38
	Post-secondary non-university degree	0.50	0.50
	University graduate	0.15	0.35
Age		48.24	9.00
Family Composition:	Unattached individual in one person household [a]	0.15	0.36
	Unattached individual in multi-person household	0.03	0.16
	Married or common-law couple/no children	0.22	0.42
	Married or common-law couple with children [b]	0.40	0.49
	Female lone-parent with children [b]	0.05	0.21
	Male lone-parent with children [b]	0.02	0.13
	Other economic family types	0.13	0.34
Labour Force Status:	Employed all year [a]	0.74	0.44
	Unemployed all year	0.01	0.09
	Not in the labour force all year	0.13	0.34
	Employed part-year, unemployed part-year	0.05	0.21
	Employed part-year, not in labour force part-year	0.04	0.18
	Unemployed part-year, not in labour force part-year	0.01	0.08
	Employed, unemployed and not in labour force	0.02	0.13
Gender:	Male	0.67	0.47
	Female	0.33	0.47
Province/Region of Residence:	Atlantic Provinces [a]	0.08	0.27
	Quebec	0.30	0.46
	Ontario	0.36	0.48
	Prairie Provinces	0.15	0.36
	British Columbia	0.11	0.31
Size of Area of Residence:	Rural [a]	0.20	0.40
	Urban with less than 100,000 residents	0.39	0.49
	Urban with over than 100,000 residents	0.41	0.49
Disability:	Yes [a]	0.21	0.41
	No	0.79	0.41
Dwelling Tenure:	Owner [a]	0.82	0.38
	Renter	0.18	0.38

Continues on next page

(1)	(2)	(3)	(4)
Mother Tongue:	English [a]	0.58	0.49
	French	0.29	0.46
	Other	0.13	0.34
Immigrant:	No [a]	0.86	0.35
	Yes	0.14	0.35

Notes:
a. Reference category
b. All children under the age of 25

Panel 2. Sample Sizes by Gender and Age-Groups

	Ages		
Gender	All	Below 40	Over 40
Men	3647	1411	2236
Women	1873	743	1130

Table 5. OLS (CARDINAL) MODELS OF HEALTH, STRESS WITHOUT HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Health Males	Model: Health Females	Model: Stress Males	Model: Stress Females
Income:	< 20,000	-0.011	-0.019	0.007	-0.0001
		(0.014)	(0.017)	(0.013)	(0.015)
	20-40,000	-0.025***	-0.018**	0.001	-0.006
		(0.007)	(0.007)	(0.006)	(0.006)
	40-60,000	-0.015***	-0.012*	-0.013***	-0.015***
		(0.004)	(0.007)	(0.003)	(0.006)
	60-80,000	0.003	-0.038*	-0.009	0.017
		(0.007)	(0.020)	(0.006)	(0.018)
	> 80,000	-0.003	-0.003	-0.018*	0.035
		(0.011)	(0.031)	(0.010)	(0.030)
Health:	Very Good	0.440***	0.482***	-0.016	-0.075***
		(0.016)	(0.023)	(0.013)	(0.018)
	Good	0.884***	0.924***	-0.070***	-0.166***
		(0.020)	(0.029)	(0.016)	(0.021)
	Fair	1.448***	1.527***	-0.140***	-0.244***
		(0.035)	(0.048)	(0.026)	(0.033)
	Poor	2.078***	2.106***	-0.326***	-0.441***
		(0.066)	(0.067)	(0.046)	(0.052)
Stress:	Somewhat Stressful	-0.057***	-0.127***	0.432***	0.392***
		(0.018)	(0.024)	(0.018)	(0.022)
	Not very Stressful	-0.123***	-0.207***	0.916***	0.888***
		(0.021)	(0.028)	(0.022)	(0.029)
	Not at all Stressful	-0.174***	-0.303***	1.305***	1.300***
		(0.029)	(0.041)	(0.033)	(0.049)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Table 6. POOLED ORDERED PROBIT MODELS OF HEALTH AND STRESS WITHOUT HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Health Males	Model: Health Females	Model: Stress Males	Model: Stress Females
Income:	< 20,000	-0.012	-0.022	0.017	-0.002
		(0.019)	(0.023)	(0.022)	(0.025)
	20-40,000	-0.032***	-0.024**	0.001	-0.011
		(0.009)	(0.010)	(0.010)	(0.011)
	40-60,000	-0.021***	-0.019*	-0.022***	-0.026**
		(0.006)	(0.010)	(0.006)	(0.010)
	60-80,000	0.005	-0.060*	-0.016	0.032
		(0.011)	(0.033)	(0.011)	(0.031)
	> 80,000	-0.005	-0.012	-0.037**	0.055
		(0.016)	(0.054)	(0.018)	(0.050)
Health:	Very Good	0.692***	0.779***	-0.025	-0.125***
		(0.027)	(0.039)	(0.022)	(0.030)
	Good	1.272***	1.362***	-0.116***	-0.291***
		(0.032)	(0.048)	(0.025)	(0.036)
	Fair	1.944***	2.087***	-0.223***	-0.427***
		(0.051)	(0.075)	(0.043)	(0.058)
	Poor	2.769***	2.907***	-0.543***	-0.789***
		(0.100)	(0.110)	(0.083)	(0.100)
Stress:	Somewhat Stressful	-0.076***	-0.177***	0.806***	0.767***
		(0.024)	(0.033)	(0.035)	(0.046)
	Not very Stressful	-0.169***	-0.297***	1.547***	1.566***
		(0.029)	(0.040)	(0.042)	(0.058)
	Not at all Stressful	-0.238***	-0.429***	2.095***	2.155***
		(0.042)	(0.062)	(0.057)	(0.088)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Table 7. RANDOM EFFECTS ORDERED PROBIT MODELS OF HEALTH AND STRESS WITHOUT HETEROGENEITY MODELLING

		Model			
Lagged		Health		Stress	
Variable	Variable Level	Males	Females	Males	Females
Income:	< 20,000	-0.003	-0.025	0.007	0.002
		(0.020)	(0.024)	(0.021)	(0.025)
	20-40,000	-0.040***	-0.033***	-0.002	-0.004
		(0.012)	(0.013)	(0.012)	(0.014)
	40-60,000	-0.020***	-0.018	-0.030***	-0.029**
		(0.008)	(0.013)	(0.008)	(0.013)
	60-80,000	0.007	-0.078**	-0.015	0.041
		(0.014)	(0.036)	(0.014)	(0.037)
	> 80,000	-0.001	0.013	-0.027	0.066
		(0.021)	(0.056)	(0.021)	(0.056)
Health:	Very Good	0.263***	0.462***	-0.021	-0.119***
		(0.030)	(0.048)	(0.026)	(0.036)
	Good	0.484***	0.788***	-0.100***	-0.274***
		(0.041)	(0.073)	(0.030)	(0.043)
	Fair	0.810***	1.286***	-0.194***	-0.408***
		(0.061)	(0.100)	(0.047)	(0.064)
	Poor	1.293***	1.886***	-0.491***	-0.717***
		(0.094)	(0.140)	(0.082)	(0.095)
Stress:	Somewhat Stressful	-0.078***	-0.161***	0.324***	0.359***
		(0.029)	(0.037)	(0.035)	(0.047)
	Not very Stressful	-0.140***	-0.276***	0.647***	0.783***
		(0.034)	(0.045)	(0.047)	(0.068)
	Not at all Stressful	-0.184***	-0.368***	0.842***	1.090***
		(0.046)	(0.068)	(0.064)	(0.095)
	$\sqrt{1+\rho/(1-\rho)}$	1.224	1.136	1.219	1.173

Notes:
1. Coefficient estimates are population average
2. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
3. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income
4. ρ is the estimate of the intra-class correlation.

Table 8. HEALTH MODEL: ALL ESTIMATORS, GENDER SAMPLES BY AGE-GROUPS, WITHOUT HETEROGENEITY MODELLING

		Health Model			
Lagged		Males		Females	
Variable	Variable Level	Under 40	Over 40	Under 40	Over 40
		OLS (Cardinal) Model			
Stress:	Somewhat Stressful	-0.036	-0.069***	-0.137***	-0.123***
		(0.027)	(0.023)	(0.038)	(0.030)
	Not very Stressful	-0.091***	-0.137***	-0.170***	-0.234***
		(0.033)	(0.027)	(0.043)	(0.037)
	Not at all Stressful	-0.155***	-0.176***	-0.218***	-0.340***
		(0.051)	(0.036)	(0.069)	(0.052)
		Pooled Ordered Probit Model			
Stress:	Somewhat Stressful	-0.050	-0.090***	-0.190***	-0.173***
		(0.040)	(0.031)	(0.054)	(0.042)
	Not very Stressful	-0.132***	-0.183***	-0.245***	-0.335***
		(0.049)	(0.037)	(0.063)	(0.053)
	Not at all Stressful	-0.232***	-0.233***	-0.318***	-0.482***
		(0.079)	(0.050)	(0.110)	(0.076)
		Random Effects Ordered Probit Model			
Stress:	Somewhat Stressful	-0.066*	-0.081***	-0.154***	-0.170***
		(0.046)	(0.037)	(0.058)	(0.047)
	Not very Stressful	-0.109**	-0.150***	-0.192***	-0.340***
		(0.055)	(0.044)	(0.073)	(0.058)
	Not at all Stressful	-0.217***	-0.162***	-0.207*	-0.460***
		(0.083)	(0.057)	(0.120)	(0.081)

Note: Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 9. STRESS MODEL: ALL ESTIMATORS, GENDER SAMPLES BY AGE-GROUPS, WITHOUT HETEROGENEITY MODELLING

Panel 1. OLS and Ordered Probit Estimates

Lagged Variable	Variable Level	Stress Model			
		Males		Females	
		Under 40	Over 40	Under 40	Over 40
		OLS (Cardinal) Model			
Income:	< 20,000	0.019	-0.005	-0.041**	0.009
		(0.020)	(0.016)	(0.019)	(0.018)
	20-40,000	-0.009	0.006	-0.014	-0.001
		(0.009)	(0.008)	(0.010)	(0.008)
	40-60,000	-0.011**	-0.014***	-0.017*	-0.014*
		(0.005)	(0.004)	(0.010)	(0.007)
	60-80,000	-0.020**	-0.003	-0.009	0.024
		(0.010)	(0.008)	(0.037)	(0.020)
	> 80,000	0.006	-0.028**	0.017	0.036
		(0.017)	(0.011)	(0.035)	(0.036)
		Pooled Ordered Probit Model			
Income:	< 20,000	0.043	-0.006	-0.074**	0.014
		(0.037)	(0.026)	(0.031)	(0.031)
	20-40,000	-0.017	0.010	-0.026	-0.002
		(0.015)	(0.012)	(0.017)	(0.014)
	40-60,000	-0.020**	-0.023***	-0.032*	-0.022*
		(0.009)	(0.007)	(0.018)	(0.012)
	60-80,000	-0.040**	-0.005	-0.014	0.043
		(0.020)	(0.013)	(0.073)	(0.032)
	> 80,000	0.010	-0.056***	0.033	0.054
		(0.033)	(0.021)	(0.069)	(0.057)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Panel 2. Random Effects Ordered Probit Estimates

Lagged Variable	Variable Level	Stress Model			
		Males		Females	
		Under 40	Over 40	Under 40	Over 40
		Random Effects Ordered Probit Model			
Income:	< 20,000	0.019	-0.008	-0.061	0.016
		(0.038)	(0.026)	(0.055)	(0.028)
	20-40,000	-0.018	0.008	-0.019	0.004
		(0.020)	(0.015)	(0.022)	(0.017)
	40-60,000	-0.025**	-0.032***	-0.026	-0.029**
		(0.013)	(0.010)	(0.023)	(0.016)
	60-80,000	-0.040*	-0.003	-0.001	0.055
		(0.026)	(0.017)	(0.068)	(0.044)
	> 80,000	0.012	-0.046**	0.018	0.068
		(0.040)	(0.025)	(0.130)	(0.062)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Table 10. OLS (CARDINAL) MODELS OF HEALTH, STRESS WITH HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Health Males	Model: Health Females	Model: Stress Males	Model: Stress Females
Income:	< 20,000	-0.016	-0.009	0.001	-0.008
		(0.014)	(0.016)	(0.013)	(0.014)
	20-40,000	-0.014*	0.007	0.003	-0.008
		(0.008)	(0.008)	(0.007)	(0.007)
	40-60,000	-0.016***	-0.008	-0.009**	-0.014**
		(0.004)	(0.007)	(0.004)	(0.006)
	60-80,000	0.003	-0.038*	-0.007	0.017
		(0.008)	(0.019)	(0.006)	(0.019)
	> 80,000	-0.004	0.005	-0.013	0.037
		(0.011)	(0.031)	(0.010)	(0.030)
Health:	Very Good	0.326***	0.363***	0.004	-0.062***
		(0.017)	(0.024)	(0.014)	(0.019)
	Good	0.671***	0.669***	-0.028*	-0.116***
		(0.022)	(0.031)	(0.017)	(0.023)
	Fair	1.093***	1.076***	-0.077***	-0.158***
		(0.037)	(0.053)	(0.029)	(0.037)
	Poor	1.605***	1.538***	-0.232***	-0.326***
		(0.070)	(0.076)	(0.050)	(0.060)
Stress:	Somewhat Stressful	-0.045**	-0.113***	0.351***	0.290***
		(0.019)	(0.025)	(0.018)	(0.022)
	Not very Stressful	-0.075***	-0.162***	0.717***	0.698***
		(0.023)	(0.030)	(0.023)	(0.030)
	Not at all Stressful	-0.081***	-0.220***	1.048***	1.064***
		(0.031)	(0.043)	(0.034)	(0.052)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Table 11. POOLED ORDERED PROBIT MODELS OF HEALTH AND STRESS WITH HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Health Males	Model: Health Females	Model: Stress Males	Model: Stress Females
Income:	< 20,000	-0.020	-0.011	0.008	-0.016
		(0.020)	(0.023)	(0.022)	(0.024)
	20-40,000	-0.019*	0.012	0.006	-0.016
		(0.011)	(0.012)	(0.011)	(0.012)
	40-60,000	-0.023***	-0.013	-0.016**	-0.024**
		(0.006)	(0.011)	(0.006)	(0.011)
	60-80,000	0.004	-0.062*	-0.012	0.033
		(0.011)	(0.033)	(0.011)	(0.033)
	> 80,000	-0.007	0.001	-0.028	0.061
		(0.017)	(0.055)	(0.019)	(0.051)
Health:	Very Good	0.536***	0.614***	0.006	-0.106***
		(0.027)	(0.040)	(0.023)	(0.033)
	Good	1.001***	1.034***	-0.050*	-0.208***
		(0.033)	(0.050)	(0.028)	(0.042)
	Fair	1.523***	1.543***	-0.125***	-0.290***
		(0.053)	(0.079)	(0.048)	(0.067)
	Poor	2.235***	2.241***	-0.405***	-0.605***
		(0.100)	(0.120)	(0.090)	(0.110)
Stress:	Somewhat Stressful	-0.062**	-0.163***	0.674***	0.588***
		(0.027)	(0.037)	(0.035)	(0.045)
	Not very Stressful	-0.104***	-0.241***	1.244***	1.254***
		(0.033)	(0.045)	(0.043)	(0.058)
	Not at all Stressful	-0.111**	-0.328***	1.722***	1.783***
		(0.046)	(0.067)	(0.058)	(0.090)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Table 12. RANDOM EFFECTS ORDERED PROBIT MODELS OF HEALTH AND STRESS WITH HETEROGENEITY MODELLING

		Model			
Lagged		Health		Stress	
Variable	Variable Level	Males	Females	Males	Females
Income:	< 20,000	-0.006	-0.011	-0.002	-0.010
		(0.020)	(0.026)	(0.021)	(0.026)
	20-40,000	-0.009	0.019	0.003	-0.008
		(0.013)	(0.015)	(0.013)	(0.015)
	40-60,000	-0.021***	-0.008	-0.022***	-0.025**
		(0.008)	(0.014)	(0.008)	(0.014)
	60-80,000	0.007	-0.085***	-0.009	0.043
		(0.014)	(0.038)	(0.014)	(0.037)
	> 80,000	-0.003	0.035	-0.015	0.072
		(0.021)	(0.058)	(0.021)	(0.056)
Health:	Very Good	0.147***	0.211***	0.012	-0.088***
		(0.029)	(0.041)	(0.027)	(0.039)
	Good	0.272***	0.305***	-0.029	-0.177***
		(0.037)	(0.054)	(0.032)	(0.047)
	Fair	0.469***	0.485***	-0.091**	-0.238***
		(0.055)	(0.079)	(0.050)	(0.072)
	Poor	0.830***	0.845***	-0.328***	-0.466***
		(0.091)	(0.110)	(0.087)	(0.110)
Stress:	Somewhat Stressful	-0.063**	-0.116***	0.211***	0.186***
		(0.030)	(0.040)	(0.034)	(0.046)
	Not very Stressful	-0.075**	-0.170***	0.395***	0.470***
		(0.036)	(0.050)	(0.044)	(0.062)
	Not at all Stressful	-0.078*	-0.201***	0.529***	0.696***
		(0.049)	(0.074)	(0.059)	(0.087)
	$\sqrt{1+\rho/(1-\rho)}$	1.201	1.197	1.208	1.182

Notes:
1. Coefficient estimates are population average
2. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
3. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income
4. ρ is the estimate of the intra-class correlation

Table 13. HEALTH MODEL: ALL ESTIMATORS, GENDER SAMPLES BY AGE-GROUPS, WITH HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Health Model: Males Under 40	Males Over 40	Females Under 40	Females Over 40
OLS (Cardinal) Model					
Stress:	Somewhat Stressful	-0.042	-0.048*	-0.137***	-0.092***
		(0.029)	(0.024)	(0.039)	(0.032)
	Not very Stressful	-0.051	-0.088***	-0.155***	-0.165***
		(0.036)	(0.030)	(0.045)	(0.040)
	Not at all Stressful	-0.085	-0.077**	-0.162**	-0.248***
		(0.055)	(0.038)	(0.074)	(0.055)
Pooled Ordered Probit Model					
Stress:	Somewhat Stressful	-0.061	-0.063*	-0.194***	-0.136***
		(0.043)	(0.034)	(0.057)	(0.048)
	Not very Stressful	-0.073	-0.119***	-0.227***	-0.250***
		(0.055)	(0.042)	(0.068)	(0.060)
	Not at all Stressful	-0.129	-0.101*	-0.241**	-0.374***
		(0.086)	(0.055)	(0.120)	(0.084)
Random Effects Ordered Probit Model					
Stress:	Somewhat Stressful	-0.070*	-0.060*	-0.131**	-0.102**
		(0.048)	(0.039)	(0.062)	(0.053)
	Not very Stressful	-0.053	-0.087**	-0.148**	-0.189***
		(0.059)	(0.046)	(0.079)	(0.066)
	Not at all Stressful	-0.134*	-0.053	-0.093	-0.262***
		(0.088)	(0.059)	(0.130)	(0.092)

Note: Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 14. STRESS MODEL: ALL ESTIMATORS, GENDER SAMPLES BY AGE-GROUPS, WITH HETEROGENEITY MODELLING

Panel 1. OLS and Pooled Ordered Probit Estimates

Lagged Variable	Variable Level	Stress Model			
		Males		Females	
		Under 40	Over 40	Under 40	Over 40
		OLS (Cardinal) Model			
Income:	< 20,000	0.022	-0.012	-0.034*	-0.002
		(0.022)	(0.016)	(0.018)	(0.017)
	20-40,000	0.002	0.004	-0.006	-0.010
		(0.010)	(0.009)	(0.012)	(0.009)
	40-60,000	-0.008	-0.010**	-0.007	-0.014*
		(0.006)	(0.005)	(0.010)	(0.008)
	60-80,000	-0.013	-0.003	-0.014	0.025
		(0.010)	(0.008)	(0.039)	(0.021)
	> 80,000	0.009	-0.022*	0.043	0.026
		(0.018)	(0.012)	(0.041)	(0.037)
		Pooled Ordered Probit Model			
Income:	< 20,000	0.049	-0.017	-0.061*	-0.004
		(0.041)	(0.027)	(0.032)	(0.029)
	20-40,000	0.002	0.007	-0.013	-0.018
		(0.016)	(0.014)	(0.022)	(0.016)
	40-60,000	-0.015	-0.017**	-0.014	-0.023*
		(0.010)	(0.008)	(0.020)	(0.013)
	60-80,000	-0.027	-0.005	-0.026	0.047
		(0.020)	(0.014)	(0.077)	(0.035)
	> 80,000	0.018	-0.046**	0.084	0.039
		(0.036)	(0.022)	(0.081)	(0.061)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Panel 2. Random Effects Ordered Probit Estimates

Lagged Variable	Variable Level	Stress Model			
		Males		Females	
		Under 40	Over 40	Under 40	Over 40
		Random Effects Ordered Probit Model			
Income:	< 20,000	0.023	-0.020	-0.048	0.0002
		(0.038)	(0.026)	(0.057)	(0.029)
	20-40,000	0.003	0.003	-0.004	-0.009
		(0.022)	(0.016)	(0.025)	(0.019)
	40-60,000	-0.017	-0.025***	-0.007	-0.027*
		(0.014)	(0.010)	(0.024)	(0.017)
	60-80,000	-0.027	0.0003	-0.004	0.061
		(0.026)	(0.017)	(0.069)	(0.044)
	> 80,000	0.022	-0.031	0.065	0.056
		(0.040)	(0.025)	(0.130)	(0.063)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Table 15. MARGINAL EFFECTS OF INCOME AND STRESS ON EXCELLENT HEALTH BY MODEL

A) WITHOUT HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Pooled Ordered Probit Males	Pooled Ordered Probit Females	RE Ordered Probit Males	RE Ordered Probit Females
Income:	< 20,000	0.003	0.005	0.001	0.006
		(0.005)	(0.005)	(0.006)	(0.006)
	20-40,000	0.008***	0.006**	0.010***	0.008**
		(0.002)	(0.002)	(0.004)	(0.003)
	40-60,000	0.005***	0.005*	0.006**	0.004
		(0.002)	(0.002)	(0.002)	(0.003)
	60-80,000	-0.001	0.015*	-0.002	0.020*
		(0.003)	(0.008)	(0.004)	(0.010)
	> 80,000	0.001	0.003	0.000	-0.003
		(0.004)	(0.013)	(0.006)	(0.014)
Stress:	Somewhat Stressful	0.019***	0.043***	0.021**	0.040***
		(0.007)	(0.010)	(0.010)	(0.014)
	Not very Stressful	0.044***	0.075***	0.039***	0.073***
		(0.010)	(0.015)	(0.014)	(0.021)
	Not at all Stressful	0.064***	0.113***	0.052***	0.102***
		(0.014)	(0.022)	(0.019)	(0.029)

Notes:
1. The marginal effect of a dummy variable is the change in the probability of the outcome for a discrete change of the dummy from 0 to 1
2. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
3. RE Ordered Probit stands for random effects ordered probit
4. Income is a logarithmic-scale linear spline of annual personal income and the marginal effects show the change in the probability of the outcome for a percentage change of income in the respective income category. The income intervals are closed on the left side and refer to the values of actual income.

B) WITH HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Pooled Ordered Probit Males	Pooled Ordered Probit Females	RE Ordered Probit Males	RE Ordered Probit Females
Income:	< 20,000	0.005	0.003	0.002	0.003
		(0.005)	(0.005)	(0.005)	(0.006)
	20-40,000	0.005*	-0.003	0.002	-0.005
		(0.003)	(0.003)	(0.003)	(0.004)
	40-60,000	0.006***	0.003	0.005**	0.002
		(0.002)	(0.003)	(0.002)	(0.003)
	60-80,000	-0.001	0.015*	-0.002	0.021**
		(0.003)	(0.008)	(0.004)	(0.011)
	> 80,000	0.002	0.000	0.001	-0.009
		(0.004)	(0.013)	(0.005)	(0.014)
Stress:	Somewhat Stressful	0.015**	0.038***	0.016*	0.029**
		(0.007)	(0.011)	(0.009)	(0.012)
	Not very Stressful	0.026***	0.059***	0.020*	0.043**
		(0.010)	(0.015)	(0.011)	(0.017)
	Not at all Stressful	0.028**	0.082***	0.020	0.052**
		(0.013)	(0.022)	(0.014)	(0.023)

Notes:
1. The marginal effect of a dummy variable is the change in the probability of the outcome for a discrete change of the dummy from 0 to 1
2. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
3. RE Ordered Probit stands for random effects ordered probit
4. Income is a logarithmic-scale linear spline of annual personal income and the marginal effects show the change in the probability of the outcome for a percentage change of income in the respective income category. The income intervals are closed on the left side and refer to the values of actual income.

Table 16. MARGINAL EFFECTS OF INCOME AND HEALTH ON HIGHEST STRESS LEVEL (VERY STRESSFUL)

A) WITHOUT HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Pooled Ordered Probit Males	Pooled Ordered Probit Females	RE Ordered Probit Males	RE Ordered Probit Females
Income:	< 20,000	-0.003	0.0004	-0.002	-0.0003
		(0.005)	(0.005)	(0.005)	(0.006)
	20-40,000	-0.0002	0.002	0.0003	0.001
		(0.002)	(0.002)	(0.003)	(0.003)
	40-60,000	0.004***	0.006**	0.006**	0.006*
		(0.001)	(0.003)	(0.003)	(0.004)
	60-80,000	0.003	-0.007	0.003	-0.009
		(0.002)	(0.007)	(0.003)	(0.009)
	> 80,000	0.008*	-0.012	0.006	-0.015
		(0.004)	(0.011)	(0.005)	(0.014)
Health:	Very Good	0.005	0.028***	0.004	0.027**
		(0.005)	(0.009)	(0.006)	(0.013)
	Good	0.024***	0.067***	0.021**	0.065**
		(0.007)	(0.017)	(0.010)	(0.026)
	Fair	0.048***	0.105***	0.044**	0.105***
		(0.014)	(0.025)	(0.019)	(0.038)
	Poor	0.132***	0.215***	0.126***	0.205***
		(0.031)	(0.044)	(0.044)	(0.061)

Notes:
1. The marginal effect of a dummy variable is the change in the probability of the outcome for a discrete change of the dummy from 0 to 1
2. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
3. RE Ordered Probit stands for random effects ordered probit
4. Income is a logarithmic-scale linear spline of annual personal income and the marginal effects show the change in the probability of the outcome for a percentage change of income in the respective income category. The income intervals are closed on the left side and refer to the values of actual income.

B) WITH HETEROGENEITY MODELLING

Lagged Variable	Variable Level	Model: Pooled Ordered Probit Males	Pooled Ordered Probit Females	RE Ordered Probit Males	RE Ordered Probit Females
Income:	< 20,000	-0.002	0.003	0.0004	0.002
		(0.005)	(0.005)	(0.004)	(0.006)
	20-40,000	-0.001	0.003	-0.001	0.002
		(0.002)	(0.003)	(0.003)	(0.003)
	40-60,000	0.003**	0.005**	0.005**	0.005
		(0.001)	(0.003)	(0.002)	(0.004)
	60-80,000	0.002	-0.007	0.002	-0.009
		(0.002)	(0.007)	(0.003)	(0.009)
	> 80,000	0.006	-0.013	0.003	-0.016
		(0.004)	(0.011)	(0.005)	(0.014)
Health:	Very Good	-0.001	0.023**	-0.003	0.020*
		(0.005)	(0.009)	(0.006)	(0.011)
	Good	0.010	0.046***	0.006	0.040**
		(0.006)	(0.015)	(0.007)	(0.018)
	Fair	0.026**	0.067***	0.019	0.056**
		(0.012)	(0.023)	(0.013)	(0.027)
	Poor	0.092***	0.154***	0.076**	0.119**
		(0.030)	(0.044)	(0.033)	(0.047)

Notes:
1. The marginal effect of a dummy variable is the change in the probability of the outcome for a discrete change of the dummy from 0 to 1
2. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
3. RE Ordered Probit stands for random effects ordered probit
4. Income is a logarithmic-scale linear spline of annual personal income and the marginal effects show the change in the probability of the outcome for a percentage change of income in the respective income category. The income intervals are closed on the left side and refer to the values of actual income.

Table 17. AUXILIARY MODEL: OLS (CARDINAL) MODELS OF HEALTH, STRESS

Lagged Variable	Variable Level	Model: Health Males	Model: Health Females	Model: Stress Males	Model: Stress Females
Income:	< 20,000	-0.041**	-0.007	0.016	-0.012
		(0.018)	(0.032)	(0.026)	(0.036)
	20-40,000	-0.032*	-0.003	0.030*	0.020
		(0.018)	(0.022)	(0.018)	(0.019)
	40-60,000	0.004	-0.010	-0.031***	-0.017
		(0.012)	(0.017)	(0.011)	(0.015)
	60-80,000	0.001	-0.035	-0.012	0.008
		(0.019)	(0.053)	(0.015)	(0.041)
	> 80,000	0.027	0.014	0.034	0.089
		(0.026)	(0.051)	(0.025)	(0.064)
Health:	Very Good	0.319***	0.345***	0.013	-0.037
		(0.035)	(0.053)	(0.032)	(0.044)
	Good	0.570***	0.617***	-0.013	-0.056
		(0.042)	(0.068)	(0.037)	(0.053)
	Fair	0.905***	1.075***	-0.067	-0.163*
		(0.072)	(0.120)	(0.061)	(0.086)
	Poor	1.594***	1.421***	-0.246**	-0.269**
		(0.130)	(0.160)	(0.110)	(0.120)
Stress:	Somewhat Stressful	-0.072*	-0.073	0.308***	0.255***
		(0.041)	(0.055)	(0.038)	(0.044)
	Not very Stressful	-0.049	-0.112	0.661***	0.609***
		(0.052)	(0.070)	(0.049)	(0.060)
	Not at all Stressful	-0.136**	-0.181*	0.971***	1.160***
		(0.067)	(0.100)	(0.064)	(0.091)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

Table 18. AUXILIARY MODEL: PROBIT REGRESSIONS OF HEALTH AND STRESS

Lagged Variable	Variable Level	Model: Health Males	Model: Health Females	Model: Stress Males	Model: Stress Females
Income:	< 20,000	-0.062**	-0.012	0.032	-0.023
		(0.029)	(0.049)	(0.050)	(0.071)
	20-40,000	-0.054*	-0.002	0.050	0.045
		(0.028)	(0.034)	(0.031)	(0.037)
	40-60,000	0.006	-0.018	-0.056***	-0.038
		(0.019)	(0.028)	(0.019)	(0.031)
	60-80,000	0.003	-0.046	-0.023	0.020
		(0.031)	(0.089)	(0.028)	(0.079)
	> 80,000	0.041	-0.006	0.050	0.174
		(0.044)	(0.089)	(0.045)	(0.120)
Health:	Very Good	0.573***	0.658***	0.026	-0.083
		(0.060)	(0.093)	(0.056)	(0.085)
	Good	0.942***	1.058***	-0.027	-0.121
		(0.071)	(0.110)	(0.066)	(0.100)
	Fair	1.375***	1.653***	-0.124	-0.369**
		(0.110)	(0.180)	(0.110)	(0.180)
	Poor	2.503***	2.289***	-0.465**	-0.644***
		(0.230)	(0.260)	(0.200)	(0.250)
Stress:	Somewhat Stressful	-0.099	-0.113	0.656***	0.586***
		(0.064)	(0.086)	(0.077)	(0.098)
	Not very Stressful	-0.052	-0.195*	1.243***	1.235***
		(0.080)	(0.110)	(0.093)	(0.130)
	Not at all Stressful	-0.192*	-0.281*	1.724***	2.128***
		(0.110)	(0.170)	(0.110)	(0.180)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. Income is a logarithmic-scale linear spline of annual personal income and the coefficient estimates are elasticities. The income intervals are closed on the left and refer to the values of actual income.

III

Looking for Private Information in Self-Assessed Health

1. Introduction

The goal of this paper is to investigate whether self-assessed health status (SAH) contains information about *future* mortality and morbidity, beyond the information that is contained in standard "observable" characteristics of individuals (including demographics, risk behaviors, and pre-existing diagnosed medical conditions). To the extent that SAH does have predictive power for future health shocks, we are particularly interested in how that predictive power varies with age. That is, we hope to understand how individual's uncertainty about their future health status resolves as they age, and in particular, whether people have "private information" about their future health status and whether the amount of private information changes with age.

There are a number of reasons to be interested in this question. The information content of SAH, which is easily collected and included in many surveys, is obviously a relevant issue for the great body of empirical work that uses SAH as either an explanatory variable or an outcome measure.

One particular reason to be interested in this question is because of the current trend away from defined benefit pensions and towards defined contribution pensions. Much has been made of the fact that this trend exposes workers to greater financial market risk. However, it may also alter worker's exposure to longevity risk, and this aspect of changing pension arrangements has received little, if any, attention.

Longevity risk is simply the risk that an individual may live longer than they expect. While this is, of course, a positive surprise, it can pose severe financial difficulties if the individual does not have adequate financial resources for this extra period of life. The obvious way to avoid such difficulties is to annuitize wealth. One way to think about the switch from DB to DC pensions is that individuals in DC pensions annuitize their pension wealth (if at all) at retirement. In contrast, individuals in DB pensions effectively lock into an annuity when they join the firm - typically when they are quite young.

It is well known that take-up of private annuities is surprisingly low. There are a number of reasons why this might be the case. One reason could be that annuity markets suffer from significant adverse selection. It could be that only individuals who have private knowledge that their health is good wish to purchase annuities - so the adverse selection problem is the reverse of what one has in health or life insurance.

If individuals have substantially more private information about their health/expected longevity at age 65 than at age 35, the market for annuities at age 65 will suffer from more adverse selection than the market for annuities that are locked in at age 35. Thus the DB to DC switch may make it more difficult for individuals to insure longevity risk. Brugiavini (1993) develops some of these ideas in a formal theoretical model. However, as noted above, this is an aspect of the trend to DC pensions that has not received much attention. This concern of course, rests on the presumption that individuals have more private information about their health at older ages. It is this hypothesis that we examine in this paper.

Our analysis employs a ten-year span of the Canadian National Population Health Survey (NPHS). This unusual panel survey collects detailed health information from respondents every two years, and the initial sample contains a full range of ages (as opposed, for example, to the retirement and aging surveys underway in several countries, which respondents typically only enter after the age of 50). To preview our results, we find that SAH does contain private information for future mortality and morbidity. Moreover, we find some evidence that the extra information in SAH is greater at older ages.

The next section reviews the relevant literature. Section 3 provides details on the data and the estimation approach utilized. The results are presented and discussed in Section 4. Section 5 provides a concluding discussion.

2. Literature Review

The introduction of mandatory retirement savings plans and the transition from DB to DC pensions in many developed countries has led to a rapid growth in the private annuity markets in those states. Despite the growth however, those markets have continued to be "not well developed even in the most advanced OECD countries" (James and Vittas 2000). One reason for this observed underdevelopment may be the presence of adverse selection in these markets, and this possibility has been the focus of much recent research.

One approach to the study of annuity markets is to evaluate the "value per premium dollar" of annuities offered for sale (see for example Mitchell et al. 1997). Such

studies typically find values significantly below one. The insurance load in excess of reasonable administrative costs is attributed to adverse selection.

An alternative approach to test for adverse selection is to look for correlation between annuity purchases and subsequent realized risk experience. Finkelstein and Poterba (2002) observe that the UK annuitants, particularly voluntary annuitants, live longer than non-annuitants. Finkelstein and Poterba (2004) document further evidence of a systematic relationship between future mortality and annuity characteristics.

Finally, Finkelstein and Poterba (2006) construct a test for adverse selection in insurance markets that is potentially able to distinguish adverse selection from moral hazard. The test, based on observable characteristics of insurance buyers that are not used in setting insurance prices, provides evidence of the presence of adverse selection.

The only evidence on adverse selection in Canadian annuity markets that we are aware of is Milevsky (1998). Following the methodology of Mitchell et al. (1997), Milevsky calculates value per premium dollar for Canadian annuity quotes in the period 1984-1996. He focuses exclusively on 65-year old men and women and ignores the value available at other ages. Milevsky (1998) finds value per premium dollar of about 90 cents (or, equivalently, an insurance load of about 10%). The estimates vary with alternative assumptions about mortality and the term structure of interest rates. Value per dollar of premium is higher when using annuitant life tables than when using population life

tables[9]. This reflects the greater longevity of annuitants implicit in the life tables and is consistent with adverse selection.

All of these studies take the approach of inferring adverse selection from prices or quantities in annuity markets. In this paper, we follow the alternative, and complementary strategy of trying to determine directly whether individuals actually have private information about health and longevity. One reason to take this alternative approach is that it may shed light on whether adverse selection in annuity markets is "active" or "passive". Poterba (2001) points out that mortality differences between annuitants and non-annuitants might arise if there were correlations between the characteristics of annuity purchasers and longevity. Moreover, annuitant purchasers need not be aware of these correlations. For example, annuitants tend to be wealthy and have higher incomes; these factors are plausibly correlated both with annuity demand and with health and longevity. Thus while differences in the longevity of annuitants establishes that there is selection into annuitant status, it does not establish that this selection arises because of individuals acting on private information. Our approach is to look directly for private information.

The most natural way to do this would be to examine individual's responses to survey questions about their longevity expectations. Smith et al. (2001) utilize the U.S. Health and Retirement Survey (HRS) and find that longevity expectations predict mortality at the individual level. Their results also suggest that health shocks and certain

[9] An annuitant life table (an annuitant mortality table) considers specifically annuitant mortality rather than the mortality of the general population. Individual annuity life tables are used for individual annuity pricing.

health conditions negatively impact longevity expectations. Similarly using the HRS, Hurd and McGarry (2002) look at the evolution of subjective survival probabilities and their ability to predict actual mortality. They find that subjective survival probabilities do predict actual survival.

The problem with studying longevity expectations in the context of our work is that life-expectancy questions have, to date, mostly been asked in retirement surveys. These surveys only collect data from people over the age of 50. Thus these data cannot be used to compare the private information held by younger and older individuals, which is the comparison that we are most interested in.

A potential proxy measure of longevity expectations is self-assessed health (SAH). This measure is widely available and frequently employed in the economics and epidemiology literature on mortality. Therefore, to assess the amount of private information that individuals have, we look at the effect of SAH on future mortality and morbidity while controlling for a rich set of observables including pre-diagnosed health conditions and risk behaviours. The idea is to explore whether SAH contains information beyond that which would typically be available to an annuity seller.

The literature on the predictive power of SAH for future mortality and morbidity is extensive and has established that SAH is a significant predictor of future health outcomes. Early studies (Mossey and Shapiro 1982, Okun et al. 1984, McCallum 1994, Idler and Kasl 1995) find that self-rated health predicts morbidity and survival. Idler and Benyamini (1997) summarize results from U.S. and international longitudinal studies on self-assessed health as a mortality predictor. They conclude that despite the differences in

methodology and controls, self-assessed health is recognized globally as an independent predictor of mortality. Schwarze et al. (2000) confirm this finding with German data. Several recent studies looking at self-rated health, health care utilization (DeSalvo et al. 2005) and hospital episodes (Case and Paxon 2005) find that self-assessed health is a predictor of mortality and that its effect varies by gender and baseline chronic conditions.

To evaluate whether individuals possess more private information about their health at older ages, we need to look at data collected from respondents spanning the entire age range. We then have to estimate the effects of SAH on future mortality, conditional on observables and compare the information contained in the self-reported health measure across ages. Two studies, Burstrom and Fredlund (2001), and Van Doorslaer and Gerdtham (2003) using Swedish data, take a similar approach.

Burstrom and Fredlund (2001) use the annual cross-sectional Swedish Survey of Living Conditions (SSLC) for the period from 1975 to 1997, linked to Sweden's National Causes of Death Statistics (NCDS). They focus on the mortality ratios of death during the follow-up period in relation to self-reported health at the time of interview. The authors utilize a Cox proportional hazards model and find that the mortality rate ratios for persons reporting bad health compared to individuals reporting good health are high at younger ages, but that the effect declines with age.

The second study, Van Doorslaer and Gerdtham (2003), also employs pooled data from the annual SSCL for 1980 through 1986, once again linked to the NCDS. Using a Cox proportional hazards framework and a larger set of controls, Van Doorslaer and Gerdtham also find that "the effect of SAH on mortality risk declines with age".

Both these papers suggest then, that private information about future health outcomes *declines* with age. Nevertheless, these studies are based on a common Swedish data set, and it seems important to revisit this issue with other data. We do so with data from the Canadian National Population Health Survey.

3. Data and Methods

3.1. Survey Details and Sample of Analysis

The Canadian National Population Health Survey, administered by Statistics Canada, is a longitudinal health survey of the Canadian population. The three target populations of the NPHS are household residents in all Canadian provinces[10], residents expected to remain longer than six months in health care institutions, and the residents of Yukon and the Northwest Territories[11].

In all provinces except Quebec, the NPHS household component utilizes a stratified two-stage sampling design based almost entirely on the Canadian Labour Force Survey sampling design. In Quebec, the NPHS employs the design of the 1992-93 Enquête sociale et de santé. The final NPHS household sample is created by selecting households from within cluster-dwelling break-outs and then choosing a household member, 12 years old or older, as the longitudinal respondent to be followed over cycles. The survey is biennial and ongoing. The first cycle gathered data for 1994-95. The most recently released cycle, cycle five, contains data for 2002-03.

[10] Excluding populations on Indian Reserves, Canadian Forces Bases and remote areas in Quebec and Ontario.
[11] Excluding populations on Indian Reserves, Canadian Forces Bases and remote areas.

In this study we utilize the health file of the household component of NPHS. The health file contains demographic, socio-economic and comprehensive health-related information about the longitudinal respondent. Interviewing is conducted in-person and by telephone. The percentage of each method varies across cycles and provinces (Statistics Canada 1996).

There are 17,276 respondents in Cycle 1 falling to 14,532 in Cycle 3 and 12,546 in Cycle 5. Total attrition between Cycle 1 to Cycle 5 is 27.4%. The most common reason for attrition is refusal to provide information and it amounts to 61% of all attrition. In addition, however, by Cycle 5, 1279 Cycle 1 respondents are deceased. These individuals can potentially be included in our analysis when mortality is the outcome of interest. Item non-response in Cycle 5 varies from 0% to 5%.

As described in greater detail below, our empirical strategy is to model mortality between Cycles 1 and 5, and morbidity at Cycles 3 and 5, as functions of Cycle 1 information (including self-assessed health). When we model mortality our analysis sample comprises 9004 respondents (4516 male and 4488 female) aged 20 to 64 in Cycle 1. Of these 340 are deceased by Cycle 5. The differences between the numbers above (12,546 Cycle 5 respondents and 1279 deceased) and our working sample are due to the initial age restriction and item non-response in Cycle 1. When modelling morbidity, the deceased represent attrition and our sample is restricted further by item non-response in Cycle 5, which varies between 0% and 5% across items. Thus when looking at morbidity, we utilize a sample of 7439 respondents (3326 males and 4113 females).

Throughout we analyze males and females separately. This is consistent with the fact that males and females are treated differently with respect to annuity characteristics and prices in annuity markets.

We have conducted standard tests for non-random attrition; these are described below.

3.2. Variables of Interest

Our focus is on the variable self-assessed health. It has five categories: "excellent", "very good", "good", "fair" and "poor" corresponding to the answers to the question: "In general, how would you describe your health?" Table 1 presents the distribution of SAH by gender-age groups. The rates of excellent/very good health reporting steadily decrease with age for both genders. On the other hand, the rates of reporting fair/poor health exhibit a generally increasing pattern.

We consider indicators of mortality and morbidity as health outcome variables. Our analysis of mortality employs a variable that flags all deceased individuals in the period between Cycles 1 and 5. Deaths in the NPHS are confirmed against the Canadian Vital Statistic Database.

While mortality is the relevant outcome for annuities, at younger ages mortality rates are extremely low. Thus we extend our focus to indicators of morbidity. The idea is to look at aspects of morbidity that are strongly associated with mortality. Therefore, we concentrate on conditions that potentially increase the probability of death. The aspects of

morbidity we target are the presence of a "major" condition, a "medium" condition, or an "activity restriction".

An individual is identified as having a major condition if s/he is a subject to heart disease, cancer, and/or stroke. This definition is similar to that employed by Smith (1999). An individual is identified as having a medium condition if s/he has diabetes and/or hypertension. These are significant risk factors for major conditions. Activity restriction flags all respondents who because of a physical or mental condition or a health problem are limited (handicapped and/or long-term limited -- limited in the past 6 months) in the kind or amount of activity they can perform at home, school, work or other. The definitions of all indicators and their prevalence rates are provided in Tables 2 and 3.

All morbidity flags are constructed in terms of current (Cycle 3 or Cycle 5) prevalence. Since we control for Cycle 1 prevalence, we are effectively looking for changes in prevalence between Cycle 1 and Cycles 5 or 3. The questions on which these morbidity flags are based all have the following general format: "Do you have [condition] diagnosed by a health professional?"

Note that current prevalence at Cycle 5 is necessarily less than total prevalence over the entire 10-year period between Cycles 1 and 5 (and similarly for Cycle 3). The discrepancy varies by condition (see Table 3). However, we have repeated all of the analysis described below with morbidity defined as total prevalence over the relevant period, and the results were very similar to those described below.[12]

[12] Full results are available from the authors.

The set of Cycle 1 controls we employ includes flags for pre-existing health conditions including minor conditions (defined as any health condition but major or medium) in addition to major and medium conditions and activity restrictions. It also includes risk factors (body mass index and indicators of smoking and drinking) as well as a number of socio-economic and geographic characteristics including age, gender, household income, education, marital status, labour force status, mother tongue, region of residence in Canada. Summary statistics for socioeconomic control variables are provided in Table 4.

3.3. Estimation Strategy and Methodology

Our estimation strategy is as follows. First, we divide the data into age groups: 20-34, 35-49, and 50-64. Then, within each group, we estimate econometric models of the form:

$$prob(y_{t+k}^{j} = 1) = f_A(SAH_t, Z_t, y_t^1, ..., y_t^j, ...y_t^J)$$

where y_t^j is a measure of mortality or morbidity at time t; $j=1,..,J$ where J is the total number of morbidity indicators (health conditions) considered; k is a lead indicator taking the value of 3 or 5; SAH_t is self-assessed health status at time t; and Z_t is a set of observable characteristics. These last would include demographics (age and sex, marital status); socioeconomic variables (education, occupation, income groups) and risk behaviours (smoker or not).

Thus, again, we are testing whether SAH has additional predictive power for future mortality and morbidity once we control for the types of information that would

typically be observable by a seller in an annuity or insurance market: demographics, socioeconomic status, some risk behaviours and previously diagnosed conditions ($y_t^1 ... y_t^j ... y_t^J$). To determine whether private information about health accumulates with age, we compare estimates of the effect of SAH in models of this type estimated for different age groups (as indicated by the *A* (age) subscript on the function *f*).

The particular functional form we use for *f* is a logit model. From the parameter estimates, we construct two measures of the magnitude of any effect of SAH on the probability of future health outcomes. The first is the *marginal effect*. Let *q* be the probability of a future health event for individuals in our baseline SAH category and *p* be the probability of the same health event for individuals in another SAH category. The marginal effect is then the difference between the two (averaged over the relevant sample): *p* - *q*. Thus it is an *absolute* effect on risk. The second magnitude we report is the *odds ratio*. This is the ratio between the odds of a future health event for individuals in the SAH category under consideration and the odds of the same health event for individuals in our base SAH category:

$$odds\ ratio = \frac{p/1-p}{q/1-q} = \frac{p}{q}\left(\frac{1-q}{1-p}\right) = \left(relative\ risk\right)\left(\frac{1-q}{1-p}\right)$$

The odds ratio is a natural measure of effect in a logistic model. Note that for small risks, the odds ratio is approximately equal to the relative risk and thus, for small risks, the odds ratio minus one is approximately equal to the relative effect on risk (sometimes called the relative risk reduction):

$$odds\ ratio - 1 = \frac{p}{q}\left(\frac{1-q}{1-p}\right) - 1 \approx \frac{p}{q} - 1 = \frac{p-q}{q} \text{ (for small } p, q)$$

Note that, across age groups, the absolute and relative effects can move in opposite directions. For example, the absolute effect could increase with age, while the relative effect falls. This would happen if the baseline risk (q) rose faster with age than the absolute effect ($p - q$).

4. Results

We first ask whether SAH has incremental predictive power for mortality. We focus initially on the ten-year time horizon spanned by Cycles 1 and 5. Baseline risks estimated from our models are offered in Table 5. Marginal effects are presented in Table 6 for males and Table 7 for females. Marginal effects of very good or excellent SAH versus a baseline of good health are given in the first row of each table. Marginal effects of fair or poor health, again versus the baseline middle category of good health, are given in the second row. The results for the pooled sample (ages 20 to 64) are given in the first column. Table 6 indicates that, after controlling for pre-existing conditions, risk factors, and socioeconomic variables, male respondents reporting excellent/very good health in Cycle 1 are 1.5 percentage points less likely to experience death over the next 10 years, compared to males reporting good health. The corresponding odds ratio, reported in Table 8, indicates that males who report excellent or very good health are approximately one third less likely to experience death over the following 10-year period (as indicated by an odds ratio of 0.66). Both absolute and relative effects are statistically significant at

conventional levels ($p<0.05$). Men who report fair or poor health are more likely to die over the subsequent 10 years (again relative to the base group reporting good health, and controlling for initial conditions, risk factors and socioeconomic characteristics) but the effect is not statistically significant (whether measured by as a marginal effect or an odds ratio.)

Table 9 indicates that women who reported fair or poor health are approximately 65% more likely to experience death, and this effect is statistically significant at the $p < 0.1$ level. However, the corresponding marginal effect (reported in Table 7) is not statistically significant, nor is either the marginal effect or odds ratio associated with reporting very good or excellent health.

We next estimate our predictive models separately for the 20-34, 35-49 and 50-64 age groups to investigate whether the incremental predictive power of SAH varies with age. In each of Tables 6, 7, 8 and 9, results for the 20-34 age group are in the second column; results for the 35-49 age group are in the third column; and results for the 50-64 age group are in the fourth and final column. Comparisons of marginal effects for each age group are made graphically in Figures 1 and 2 (for men) and Figures 3 and 4 (for women).

For men, the marginal effect on mortality risk of reporting excellent or very good health (Table 6) is actually positive (though not statistically different from zero) for the youngest group, turns negative (but again not statistically different from zero) for the middle group and is negative and statistically different (at $p<0.01$) for the oldest group. Thus the effect noted in the pooled sample appears to be driven largely by the oldest

group. Table 6a reports tests of equality between marginal effects in different groups, and confirms that the marginal effect for the oldest group of men is statistically different from the estimated effect for the youngest ($p = 0.003$) and middle ($p = 0.021$) groups. The marginal effect of poor or fair health is marginally significant in the middle group, but not elsewhere (Table 6) and the effects for different age groups are not statistically different from each other (Table 6a).

When we present the effects in odds ratio form, in Table 8, the same finding is apparent for very good or excellent: the predictive power observed in the full sample appears to be largely driven by the oldest group. For this group, but not for the younger groups, the odds ratio is strongly statistically different from one. The effects of poor or fair health present a less interpretable pattern (as they did when presented as marginal effects). The strongest effect here is for those aged 35 to 49.

The age-group results for the female sample are in the second through fourth columns of Tables 7 (marginal effects) and 9 (odds ratios). Corresponding tests of equality of marginal effects across age groups are presented in Table 8. None of the within group-age effects (either marginal effects or odds ratios) are statistically significant, at even the $p < 0.1$ level. In part this may reflect that the baseline mortality risk is very low (about half of male risk in these age groups - see Table 5). This means that we are modelling a very rare event.

We next ask whether SAH predicts future morbidity, and particularly the emergence of conditions that are associated with mortality risk. The results follow the same pattern as for mortality. Results for males are presented in Tables 6, 6a and 8; for

females in Tables 7, 7a and 9. Marginal effects are presented in Tables 6 and 7, and Tables 6a and 7a report tests of the equality of marginal effects across age groups. Odds ratios are reported in Tables 8 and 9. Moving down each table from the mortality results, we present in turn results for major conditions (heart disease, cancer and stroke), medium conditions (diabetes and hypertension) and activity restrictions.

Beginning with the male sample, and marginal effects, we see that the effect of excellent or good health on morbidity is negative, as expected, and there is some evidence that the magnitude of these effects increases with age. The effect in the pooled (20-64) sample is statistically significant at $p < 0.01$ for medium conditions and activity restrictions, but not for major conditions.

One reason that the pooled estimate for major conditions is not statistically different from zero is that it is *positive* and statistically significant for the youngest (20-34) group. This result says that, controlling for pre-existing conditions and risk factors, a young man who reported that he was in very good or excellent health was more likely to have a major condition ten years later than a young man that reported good health. This is a surprising result, although the corresponding effect on mortality, discussed above, has the same sign (though it is not statistically different from zero). A young man who reported that his health was fair or poor was also statistically more likely to develop a major condition so there is no simple gradient here. At older ages reporting very good or excellent is associated with lower future incidence of a major condition, though the effect is never statistically significant.

For medium conditions and activity restrictions, the point estimate of the effects of reporting very good or excellent health are larger (that is, more negative) in the older age groups. However, though they are not always statistically different from zero, and, as Table 6a illustrates, the precision with which age-group-specific effects are estimated is not sufficient to allow them to be formally distinguished from each other.

As with mortality, the effects of reporting fair or poor health are less clear – very few of the estimated effects are statistically different from zero.

Turning to women, reporting very good or excellent health has a negative and statistically significant effect on the probability of having a major condition or activity restriction 10 years later. In both cases, when broken down by age, the largest and only statistically significant effect is observed in the oldest (50-64) age group. For activity restrictions and medium conditions, reporting a fair or poor health has a statistically significant effect.

The odds ratios presented in Tables 8 (for men) and 9 (for women), tell a similar story. Some of the odds ratios are extremely large, which reflects the very low baseline risk of some conditions in some age-groups (for example, major conditions among 20-34 year-olds).

We would summarize these results as follows. First, for both men and women, SAH predicts future mortality and morbidity. Second, on balance the predictive power is stronger at older ages. This is true whether we look at marginal effects or at odds ratios; this is important because the baseline risks increase with age.

We repeated the analysis just described but using a six-year (Cycle 1 to Cycle 3)

rather than ten-year time horizon. We did this for two reasons. First, it provides a general check on the robustness of our results and some sense of the time scale over which the predictive power of SAH is operative. The six-year and ten-year horizon results are compared graphically in Figures 5 and 6. A summary would be that the six-year horizon results exhibit similar patterns to the ten-year horizon results but are generally weaker. The second reason to move to a six-year horizon is that it allows us to employ the subsequent cycles to do some testing for effects of non-random attrition, following the suggestion of Verbeek and Nijman (1992). Specifically, we augment the six-year models with dummy variables capturing future attrition (attrition between Cycles 3 and 5). The results do not contain any evidence that attrition is a serious problem in our analysis. The attrition dummies are only occasionally significant and our main results do not change significantly with their inclusion.[13]

5. Discussion

In this paper we investigate whether self-assessed health status contains information about future mortality and morbidity, beyond the information that is contained in commonly observable characteristics of individuals. Using a ten-year span of the Canadian National Population Health Survey, we find that even after controlling for pre-existing conditions, socioeconomic characteristics, and a range of risk factors, self-assessed health predicts future mortality and morbidity. Moreover, we find some evidence that this effect strengthens with age. We interpret these findings as supportive

[13] Full results are available from authors on request.

of the idea that individuals have private information about their likely future health and lifespan. This in turn suggests that the apparent adverse selection in annuity markets could be at least in part "active". Individuals do seem to be aware of private information that might inform their demand for annuity products. Moreover, we find some evidence that the predictive power of SAH strengthens with age. As Brugiavini (1993) has suggested, this means that any change in pension arrangements that effectively delays the commitment to annuitize may carry with it the cost of exacerbated adverse selection.

There are a number of important ways that this research could be extended. First, our reading of the age patterns in the predictive power of SAH in Canadian data differs from results obtained by Burstrom and Fredlund (2001) and Van Doorslaer and Gerdtham (2003) with Swedish data. It is difficult to determine whether the contrast reflects a true difference in the underlying populations, or differences in the way SAH is measured across the two surveys, or some other aspect of the data and modelling. Further results from additional data sets would help to resolve the generality of these findings.

Second the NPHS could be further exploited to look at the co-evolution of SAH and diagnosed conditions through life. In particular, we are interested in understanding what events trigger revisions of SAH.

Finally, we have reported the surprising finding that at young ages, excellent/very good SAH, conditional on observables, leads to an increased risk of mortality/morbidity in the male sample. If this result is robust, it might reflect misperceptions leading to underinvestment in health or greater engagement in risky activities. This also warrants further investigation.

References

Adams, P., Hurd, M., McFadden, D., Merrill, A., Ribeiro, T. (2003), Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status, *Journal of Econometrics*, vol. 112, no. 1, pp. 3-56

Banks, J., Blundell, R. (2005), Private Pension Arrangements and Retirement in Britain, *Fiscal Studies*, vol. 26, no. 1, pp. 35-53

Brugiavini, A. (1993), Uncertainty Resolution and the Timing of Annuity Purchases, *Journal of Public Economics*, vol. 50, pp. 31-62

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?, *Epidemiology and Community Health*, vol. 55, no. 11, pp. 836-840

Case, A., Paxon, C. (2005), Sex Differences in Morbidity and Mortality, *Demography*, vol. 42, no. 2, pp. 189-214

Connelly, J., Philbrick, J., Smith, R. Kaiser, D., Wymer, A. (1989), Health Perceptions of Primary Care Patients and the Influence on Health Care Utilisation, *Medical Care*, vol. 27, pp. S99-S109.

Deaton, A. (1999), Inequalities in Income and Inequalities in Health, NBER Working Paper No. 7141

DeSalvo, K. et al. (2005), Predicting Mortality and Healthcare Utilization with a Single Question, *Health Services Research*, vol. 40, no. 4, pp. 1234-1246

Finkelstein, A., Poterba, J. (2002), Selection Effects in the United Kingdom Individual Annuities Market, *The Economic Journal*, vol. 112, pp. 28-42

Finkelstein, A., Poterba, J. (2004), Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market, *The Journal of Political Economy[Chicago]*, vol. 112, no. 1, pp. 183-208

Finkelstein, A., Poterba, J. (2006), Testing for Adverse Selection with "Unused Observables", NBER Working Paper No. 12112

Hurd, M., McGarry, K. (2002), The Predictive Validity of Subjective Probabilities of Survival, *Economic Journal*, vol. 112, no. 482, pp. 966-985

Idler, E., Benyamini, Y. (1997), Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies, *Journal of Health and Social Behavior*, vol. 38, no. 1, pp. 21-37

Idler, E., Kasl, S. (1995), Self-Ratings of Health: Do They Also Predict Change in Functional Ability?, *Journal of Gerontology*, vol. 50, no. 6, pp. S344-S353

James, E., Vittas, D. (2000), Annuity Markets in Comparative Perspective: Do Consumers Get Their Money's Worth?, The World Bank Policy Research Working Paper Series, Working Paper No. 2493

McCallum, J., Shadbolt, B., Wang, D. (1994), Self-Rated Health and Survival: 7 Years Follow-up Study of Australian Elderly, *American Journal of Public Health*, vol. 84, no. 7, pp. 1100-1105

Milevsky, M. (1998), Optimal Asset Allocation Toward the End of the Lifecycle: To Annuitize or Not to Annuitize?, *The Journal of Risk and Insurance*, vol. 65, no. 3, pp. 401-426

Mitchell, O., Poterba, J., Warshawsky, M., (1997), New Evidence on Money's Worth of Individual Annuities, NBER Working Paper No. 6002

Mitchell, O., Poterba, J., Warshawsky, M., Brown, J. (1999), New Evidence on Money's Worth of Individual Annuities, *American Economic Review*, vol. 89, pp. 1299-1318

Mossey, J., Shapiro, E. (1982), Self-Rated Health: A Predictor of Mortality among the Elderly, *American Journal of Public Health,* vol. 72, pp. 800-808

Okun, M., Stock, W., Haring, M., Witter, R. (1984), Health and Subjective Well-Being: A Meta-Analysis, *International Journal of Ageing and Human Development*, vol. 19, no. 2, pp. 111-132

Pollock, R., Suyderhoud, J. (1992), An Empirical Window on Rational Expectations Formation, *Review of Economics and Statistics*, vol. 74, no. 2, pp. 320-324

Poterba, J. (2001), Annuity Markets and Retirement Security, *Fiscal Studies*, vol. 22, no. 3, pp. 249-270

Salas, C. (2002), On the Empirical Association between Poor Health and Low Socioeconomic Status at Old Age, *Health Economics*, vol. 11, no. 3, pp. 207-220

Schwarze, J, Andersen, H., Anger, S. (2000), Self-Rated and Changes in Self-Rated Health as Predictors of Mortality - First Evidence from German Panel Data, Humboldt Universitaet Berlin, Sonderforschungsbereich 373

Smith, J. (1999), Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status, *Journal of Economic Perspectives*, vol. 13, no. 2, pp. 145-166

Smith, V., Taylor, D., Sloan, F. (2001), Longevity Expectations and Death: Can People Predict Their Own Demise?, *American Economic Review*, vol. 91, no. 4, pp. 1126-1134

Statistics Canada (1996), NPHS Public Use Microdata Documentation, Statistics Canada

Van Doorslaer, E., Gerdtham, Ulf-G. (2003), Does Inequality in Self-Assessed Health Predict Inequality in Survival by Income? Evidence from Swedish Data, *Social Science and Medicine*, vol. 57, pp. 1621-1629

Van Doorslaer, E., Koolman, X. (2004), Explaining the Differences in Income-Related Health Inequalities across European Countries, *Health Economics*, vol. 13, no. 7, pp. 609-628

Van Doorslaer, E., Wagstaff, A., Bleichrodt, H. et al. (1997), Income-Related Inequalities in Health: Some International Comparisons, *Journal of Health Economics*, vol. 16, pp. 93-112

Van Ourti, T. (2003), Socio-Economic Inequality in Ill-Health Amongst the Elderly. Should One Use Current or Permanent Income?, *Journal of Health Economics*, vol. 22, no. 2, pp. 219-241

Verbeek, M., Nijman T. (1992), Testing for Selectivity Bias in Panel Data Models, *International Economic Review*, vol. 33, no. 3, pp. 681-703

Walliser, J. (2000), Adverse Selection in the Annuities Market and the Impact of Privatizing Social Security, *Scandinavian Journal of Economics*, vol. 102, no. 3, pp. 373-394

Warshawsky, M. (2001), The Market for Individual Life Annuities and the Reform of Social Security: An Update and Further Analysis, *Benefits Quarterly [Brookfield]*, vol. 17, no. 4, pp. 24-43

Appendix

Table 1. DISTRIBUTION OF SELF-ASSESSED HEALTH (SAH) BY GENDER AND AGE-GROUPS

Sample	SAH	Ages			
		All	20 to 34	35 to 49	50 to 64
Males	Excellent/Very Good	0.70	0.77	0.70	0.60
	Good	0.24	0.20	0.25	0.27
	Fair/Poor	0.06	0.03	0.05	0.13
Total Sample		4516	1677	1733	1106
Females	Excellent/Very Good	0.67	0.73	0.68	0.55
	Good	0.24	0.22	0.24	0.29
	Fair/Poor	0.09	0.05	0.08	0.16
Total Sample		4488	1544	1655	1289

Table 2. DESCRIPTION OF FUTURE HEALTH CONDITIONS

Health Condition	Prevalence of Condition	Description
Deceased	Over the past 10 years	1: Individual is deceased within 10 years after the year of initial observation 0: Otherwise
Major Condition	Current	1: Individual has a Major Condition (heart disease, cancer, stroke) 10 years after the year of initial observation 0: Otherwise
	Over the past 10 years	1: Individual has experienced a Major Condition over the 10 years after the year of initial observation 0: Otherwise
Medium Condition	Current	1: Individual has a Medium Condition (diabetes, hypertension) 10 years after the year of initial observation 0: Otherwise
	Over the past 10 years	1: Individual has experienced a Medium Condition over the 10 years after the year of initial observation 0: Otherwise
Restricted (long-term)/ Restricted (LT)	Current	1: Individual has long-term disabilities or handicap 10 years after the year of initial observation 0: Otherwise
	Over the past 10 years	1: Individual has long-term disabilities or handicap over the 10 years after the year of initial observation 0: Otherwise
Minor Condition	Current	1: Individual has a Minor Condition (all but major and medium) 10 years after the year of initial observation 0: Otherwise
	Over the past 10 years	1: Individual has experienced a Minor Condition over the 10 years after the year of initial observation 0: Otherwise
Restricted (short-term)	Current	1: Because of a physical or mental condition or a health problem the individual is limited in the kind or amount of activity they can perform at home, school, work or other (for a period less than 6 months) 10 years after the year of initial observation 0: Otherwise
	Over the past 10 years	1: Because of a physical or mental condition or a health problem the individual has been limited in the kind or amount of activity they can perform at home, school, work or other (for a period less than 6 months) over the 10 years after the year of initial observation 0: Otherwise

Table 3. PREVALENCE RATES OF HEALTH CONDITIONS

		Sample			
		Male		Female	
Condition	Prevalence of Condition	Cycle 1	Cycle 5	Cycle 1	Cycle 5
Deceased	Over the past 10 years		0.05		0.03
Major Condition	Current	0.02	0.07	0.03	0.05
	Over the past 10 years		0.10		0.10
Medium Condition	Current	0.07	0.15	0.07	0.17
	Over the past 10 years		0.19		0.19
Restricted (LT)	Current	0.12	0.15	0.11	0.13
	Over the past 10 years		0.26		0.25
Minor Condition	Current	0.44	0.56	0.50	0.66
	Over the past 10 years		0.76		0.82
Restricted (short-term)	Current	0.11	0.16	0.14	0.18
	Over the past 10 years		0.29		0.33

Notes:
1. Current indicates current prevalence of a condition
2. Over the past 10 years spans the period from Cycle 1 to Cycle 5 and indicated prevalence over those 10 years. The condition could also currently exist

Table 4. SUMMARY STATISTICS – SOCIO-ECONOMIC VARIABLES

Variable	Mean	Standard Deviation
Mother Tongue: English [a]	0.58	0.49
Mother Tongue: French	0.27	0.44
Mother Tongue: Other	0.15	0.36
Immigrant	0.19	0.39
Age	39.61	11.63
Household Income: below $30,000 [a]	0.27	0.44
Household Income: $30,000-$49,000	0.30	0.46
Household Income: $50,000-$79,000	0.28	0.45
Household Income: $80,000 or over	0.15	0.36
Lower than Secondary School Education [a]	0.17	0.38
Secondary School Graduate	0.17	0.37
Post-secondary Certificate	0.27	0.45
College or University Education	0.39	0.49
Married/Common Law	0.72	0.45
Male	0.51	0.50
Smoker	0.33	0.47
Drinker	0.84	0.37
Body Mass Index	24.58	4.30
Full-time Employee	0.64	0.48
Part-time Employee	0.10	0.30
Unemployed	0.05	0.21
Self-employed	0.11	0.32
Other [a]	0.10	0.30
Residence: Atlantic Provinces [a]	0.08	0.27
Residence: Quebec	0.26	0.44
Residence: Ontario	0.37	0.48
Residence: Prairies	0.16	0.37
Residence: British Columbia	0.13	0.33

Notes:
a. Reference category

Table 5. BASELINE RISKS BY GENDER, AGE-GROUPS AND HEALTH CONDITIONS, LOGIT MODEL

Condition	Sample	Ages			
		All	20 to 34	35 to 49	50 to 64
Deceased	Male	0.044	0.013	0.022	0.118
	Female	0.024	0.006	0.017	0.064
Major Condition	Male	0.084	0.015	0.068	0.204
	Female	0.060	0.022	0.043	0.136
Medium Condition	Male	0.177	0.048	0.179	0.371
	Female	0.210	0.066	0.189	0.449
Restricted (LT)	Male	0.169	0.112	0.177	0.242
	Female	0.162	0.096	0.186	0.224

Notes:
1. Baseline risk is the probability that a person reporting good SAH experiences a particular health condition. Risks are estimated based on a logit specification.

Table 6. MARGINAL EFFECTS OF SELF-ASSESSED HEALTH (SAH) MALES, BY AGE-GROUPS AND HEALTH CONDITIONS, (LOGIT MODELS)

		Males of Age			
Condition	SAH	All	20 to 34	35 to 49	50 to 64
Deceased	Excellent/ Very Good	-0.015**	0.009	-0.005	-0.060***
		(0.007)	(0.007)	(0.009)	(0.022)
	Fair/Poor	0.010	0.048	0.042*	-0.003
		(0.009)	(0.048)	(0.023)	(0.026)
Major Condition	Excellent/ Very Good	-0.0002	0.031***	0.010	-0.040
		(0.010)	(0.010)	(0.015)	(0.033)
	Fair/Poor	0.024	0.424***	0.008	0.021
		(0.017)	(0.160)	(0.027)	(0.042)
Medium Condition	Excellent/ Very Good	-0.049***	-0.018	-0.060***	-0.059
		(0.014)	(0.015)	(0.023)	(0.037)
	Fair/Poor	-0.023	-0.017	-0.033	-0.032
		(0.019)	(0.017)	(0.038)	(0.047)
Restricted (LT)	Excellent/ Very Good	-0.048***	-0.026	-0.039	-0.079**
		(0.015)	(0.021)	(0.024)	(0.035)
	Fair/Poor	0.040*	0.036	0.102**	0.036
		(0.024)	(0.041)	(0.052)	(0.045)

Notes:
1. The marginal effect of a dummy variable is the change in the probability of the outcome for a discrete change of the dummy from 0 to 1.
2. Effects are relative to the base category, which is "good" self-assessed health
3. Standard errors are in parentheses
4. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 6A. TESTS OF EQUALITY OF MARGINAL EFFECTS ACROSS PAIRS OF AGE-GROUPS

P-VALUES (5%)

Males, by Health Condition
Marginal Effects of SAH

		Age-group	
Condition	Age-group	20 to 34	35 to 49
Marginal Effect of Excellent/Very Good (versus Good)			
Deceased	35 to 49	0.215	
	50 to 64	0.003	0.021
Major Condition	35 to 49	0.245	
	50 to 64	0.040	0.168
Medium Condition	35 to 49	0.125	
	50 to 64	0.304	0.980
Restricted (LT)	35 to 49	0.693	
	50 to 64	0.194	0.341
Marginal Effect of Fair/Poor (versus Good)			
Deceased	35 to 49	0.906	
	50 to 64	0.351	0.199
Major Condition	35 to 49	0.010	
	50 to 64	0.015	0.785
Medium Condition	35 to 49	0.701	
	50 to 64	0.766	0.985
Restricted (LT)	35 to 49	0.320	
	50 to 64	0.992	0.334

Table 7. MARGINAL EFFECTS OF SELF-ASSESSED HEALTH (SAH) FEMALES, BY AGE-GROUPS AND HEALTH CONDITIONS, (LOGIT MODELS)

		Females of Age			
Condition	SAH	All	20 to 34	35 to 49	50 to 64
Deceased	Excellent/ Very Good	0.006	0.001	0.009	0.008
		(0.006)	(0.005)	(0.008)	(0.018)
	Fair/Poor	0.012	0.004	0.010	0.020
		(0.008)	(0.015)	(0.012)	(0.020)
Major Condition	Excellent/ Very Good	-0.030***	-0.007	-0.022*	-0.067***
		(0.009)	(0.010)	(0.013)	(0.025)
	Fair/Poor	0.001	0.021	-0.006	0.013
		(0.009)	(0.021)	(0.014)	(0.026)
Medium Condition	Excellent/ Very Good	0.011	0.018	-0.022	0.059*
		(0.012)	(0.013)	(0.021)	(0.031)
	Fair/Poor	0.043*	-0.006	0.042	0.113**
		(0.022)	(0.034)	(0.038)	(0.050)
Restricted (LT)	Excellent/ Very Good	-0.039***	-0.030	-0.035	-0.071**
		(0.013)	(0.018)	(0.023)	(0.030)
	Fair/Poor	0.090***	0.050	0.175***	0.026
		(0.021)	(0.033)	(0.045)	(0.034)

Notes:
1. The marginal effect of a dummy variable is the change in the probability of the outcome for a discrete change of the dummy from 0 to 1.
2. Effects are relative to the base category, which is "good" self-assessed health
3. Standard errors are in parentheses
4. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 7A. TESTS OF EQUALITY OF MARGINAL EFFECTS ACROSS PAIRS OF AGE-GROUPS

P-VALUES (5%)

Females, by Health Condition
Marginal Effects of SAH

		Age-group	
Condition	Age-group	20 to 34	35 to 49
Marginal Effect of Excellent/Very Good (versus Good)			
Deceased	35 to 49	0.372	
	50 to 64	0.728	0.933
Major Condition	35 to 49	0.346	
	50 to 64	0.026	0.114
Medium Condition	35 to 49	0.099	
	50 to 64	0.225	0.030
Restricted (LT)	35 to 49	0.761	
	50 to 64	0.202	0.345
Marginal Effect of Fair/Poor (versus Good)			
Deceased	35 to 49	0.726	
	50 to 64	0.521	0.690
Major Condition	35 to 49	0.291	
	50 to 64	0.813	0.526
Medium Condition	35 to 49	0.349	
	50 to 64	0.049	0.256
Restricted (LT)	35 to 49	0.025	
	50 to 64	0.623	0.008

Table 8. ODDS-RATIOS FOR SELF-ASSESSED HEALTH FOR MALES, BY AGE-GROUPS AND HEALTH CONDITIONS, (LOGIT MODEL)

		Males of Age			
Condition	SAH	All	20 to 34	35 to 49	50 to 64
Deceased	Excellent/ Very Good	0.66** (0.45 - 0.95)	2.44 (0.53 - 11.22)	0.79 (0.34 - 1.83)	0.51*** (0.30 - 0.84)
	Fair/Poor	1.30 (0.83 - 2.04)	6.62* (0.83 - 52.65)	3.88*** (1.49 - 10.04)	0.97 (0.55 - 1.72)
Major Condition	Excellent/ Very Good	0.10 (0.71 - 1.40)	27.85*** (2.94 - 263.92)	1.25 (0.67 - 2.33)	0.74 (0.46 - 1.20)
	Fair/Poor	1.47 (0.91 - 2.36)	346.40*** (17.8 - 6746.1)	1.16 (0.43 - 3.12)	1.17 (0.64 - 2.14)
Medium Condition	Excellent/ Very Good	0.61*** (0.47 - 0.79)	0.60 (0.28 - 1.28)	0.58*** (0.40 - 0.84)	0.71 (0.47 - 1.08)
	Fair/Poor	0.78 (0.50 - 1.20)	0.55 (0.13 - 2.34)	0.71 (0.30 - 1.68)	0.83 (0.46 - 1.46)
Restricted (LT)	Excellent/ Very Good	0.66*** (0.52 - 0.84)	0.74 (0.46 - 1.17)	0.72* (0.49 - 1.05)	0.58** (0.37 - 0.92)
	Fair/Poor	1.39* (0.97 - 1.99)	1.49 (0.68 - 3.24)	2.11** (1.12 - 3.97)	1.28 (0.72 - 2.27)

Notes:
1. 95% confidence interval is reported in parentheses
2. Effects are relative to the base category, which is "good" self-assessed health
3. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 9. ODDS-RATIOS FOR SELF-ASSESSED HEALTH FOR FEMALES, BY AGE-GROUPS AND HEALTH CONDITIONS, (LOGIT MODEL)

		Females of Age			
Condition	SAH	All	20 to 34	35 to 49	50 to 64
Deceased	Excellent/ Very Good	1.33 (0.81 - 2.20)	1.20 (0.19 - 7.73)	1.79 (0.74 - 4.34)	1.15 (0.60 - 2.20)
	Fair/Poor	1.65* (0.95 - 2.84)	1.64 (0.06 - 45.471)	1.79 (0.55 - 5.83)	1.41 (0.73 - 2.71)
Major Condition	Excellent/ Very Good	0.53*** (0.38 - 0.76)	0.71 (0.28 - 1.82)	0.55* (0.29 - 1.06)	0.51*** (0.30 - 0.84)
	Fair/Poor	1.03 (0.70 - 1.51)	2.32 (0.67 - 8.05)	0.84 (0.36 - 1.96)	1.14 (0.69 - 1.87)
Medium Condition	Excellent/ Very Good	1.12 (0.88 - 1.42)	1.46 (0.83 - 2.57)	0.81 (0.55 - 1.19)	1.45* (0.99 - 2.13)
	Fair/Poor	1.47** (1.02 - 2.13)	0.89 (0.22 - 3.63)	1.45 (0.79 - 2.67)	1.93** (1.11 - 3.32)
Restricted (LT)	Excellent/ Very Good	0.70*** (0.56 - 0.88)	0.70 (0.45 - 1.11)	0.74 (0.51 - 1.07)	0.60** (0.40 - 0.91)
	Fair/Poor	2.05*** (1.54 - 2.73)	1.80* (0.94 - 3.44)	3.28*** (2.03 - 5.31)	1.21 (0.76 - 1.93)

Notes:
1. 95% confidence interval is reported in parentheses
2. Effects are relative to the base category, which is "good" self-assessed health
3. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Figure 1. Male Sample, Marginal Effects of Excellent/Very Good SAH by Age Groups and Health Conditions

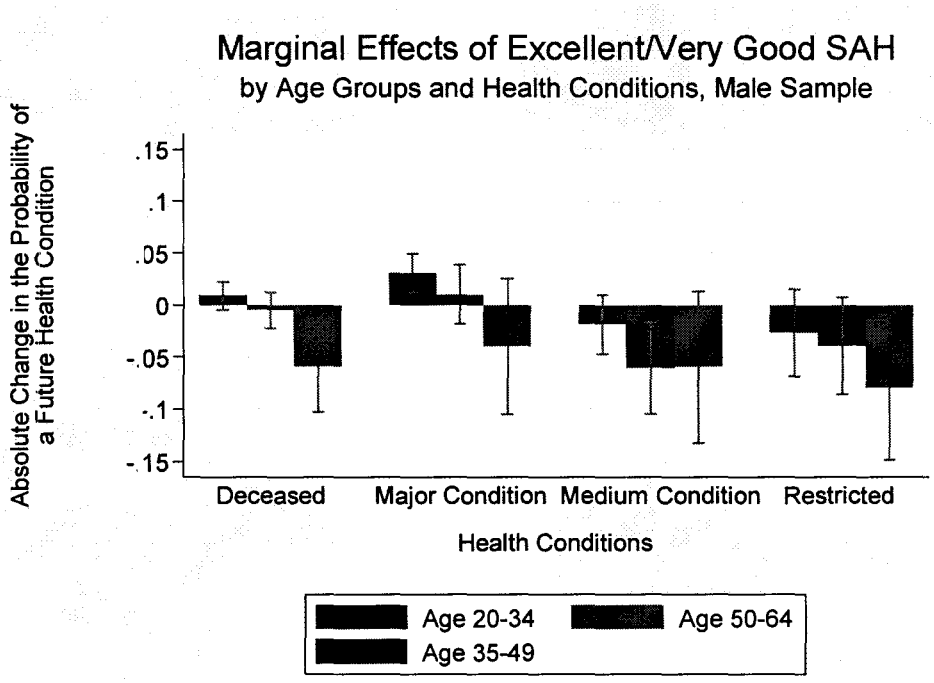


Figure 2. Male Sample, Marginal Effects of Fair/Poor SAH by Age Groups and Health Conditions

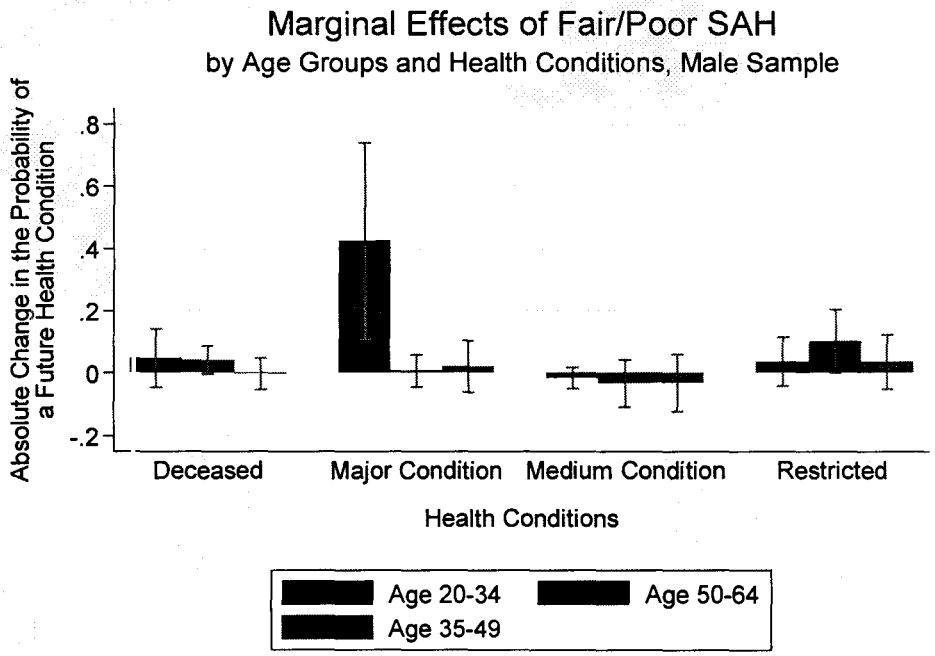


Figure 3. Female Sample, Marginal Effects of Excellent/ Very Good SAH by Age Groups and Health Conditions

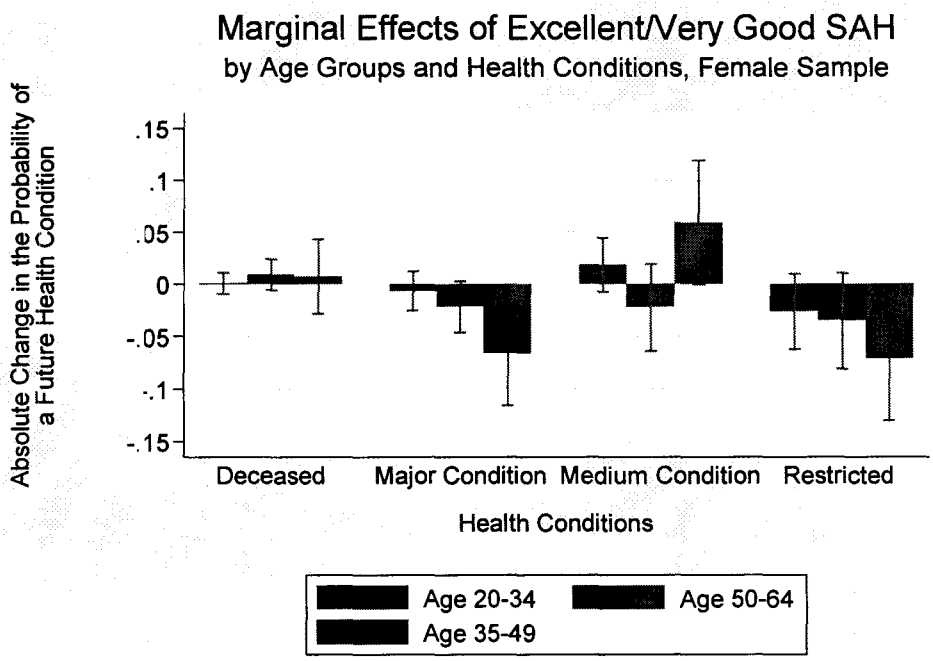


Figure 4. Female Sample, Marginal Effects of Fair/Poor SAH by Age Groups and Health Conditions

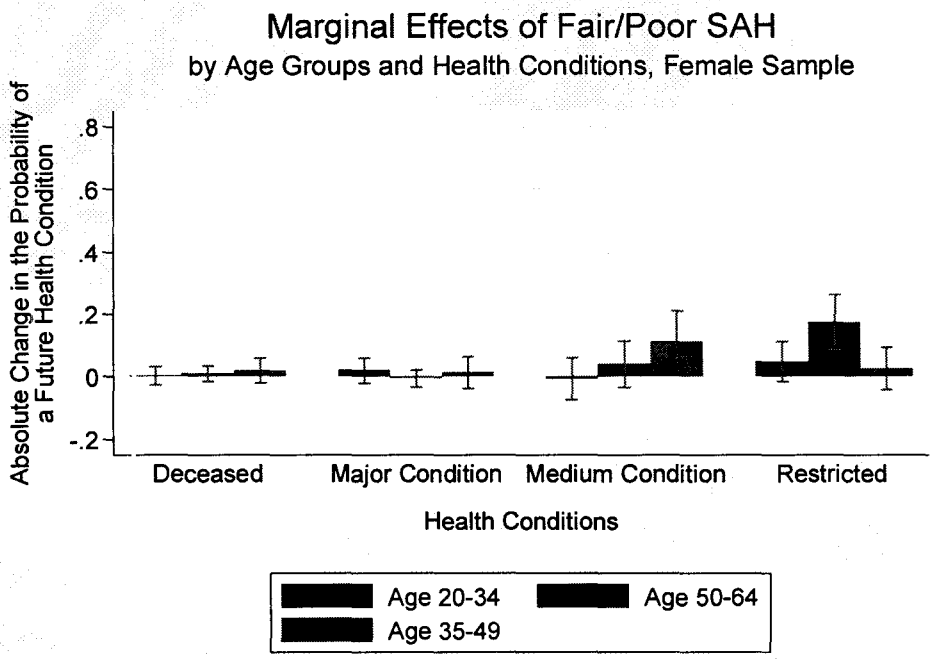


Figure 5. Male Sample, Marginal Effects of Excellent/Very Good SAH by Age Groups, Health Conditions and Horizon

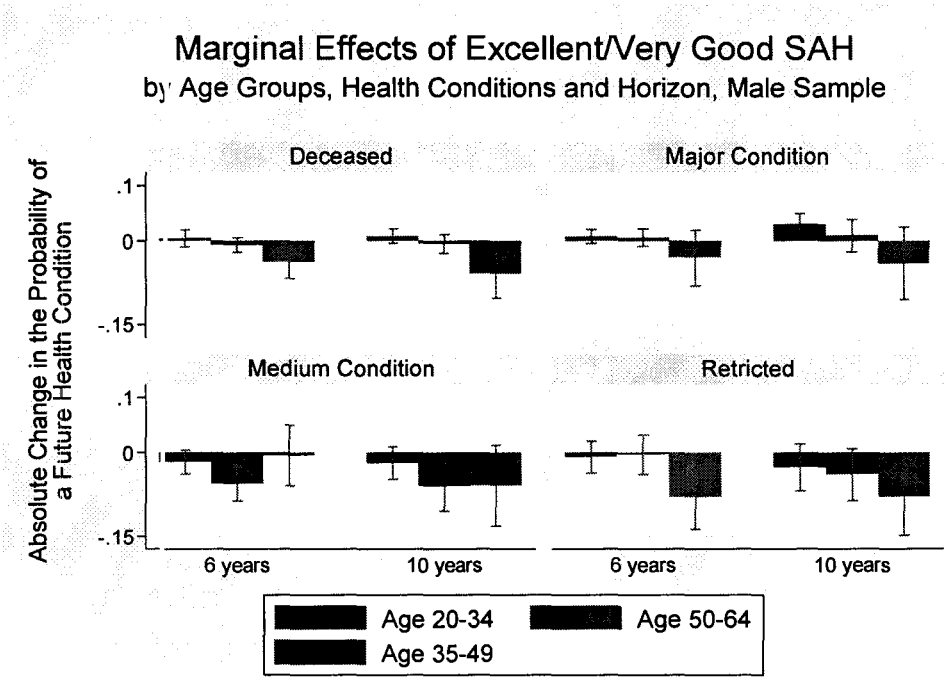
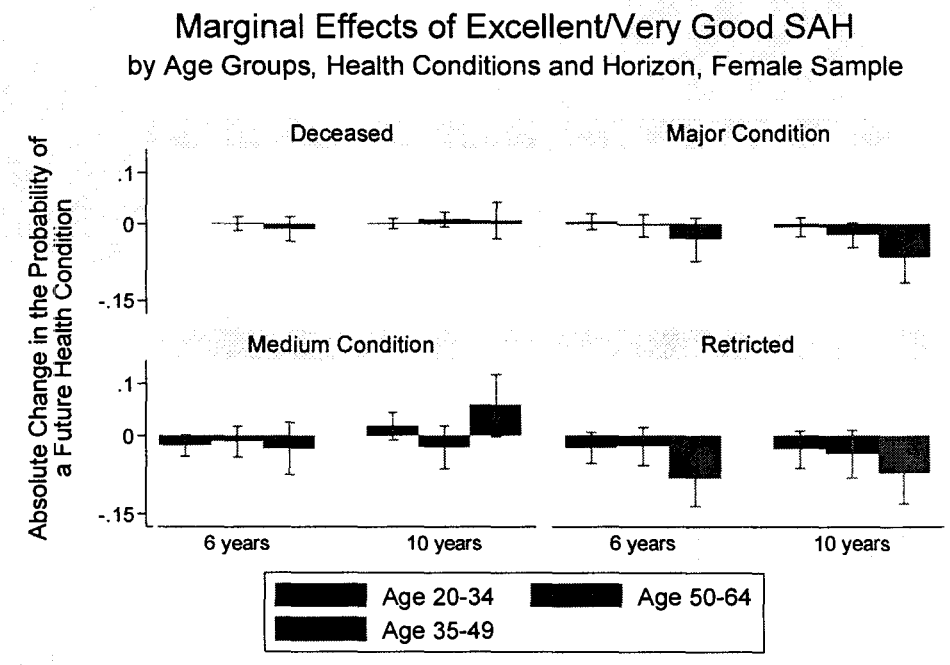


Figure 6. Female Sample, Marginal Effects of Excellent/Very Good SAH by Age Groups, Health Conditions and Horizon



IV

Early Retirement Pathways and Post-Retirement Outcomes in Canada

1. Introduction

The objective of this study is to investigate whether post-retirement observable outcomes of early retirees, persons retiring before the traditional age of 65, differ over the range of pre-retirement observable circumstances. To the extent that differences exist, we would like to understand how post-retirement characteristics are associated with a particular pathway to retirement, and more importantly how these characteristics depend on whether individuals are "pushed" into retirement or not.

There are several reasons for interest in this subject. One is related to the well known feature of the simple life-cycle model which treats retirement as an exogenous event. The event of early retirement however is inconsistent with that theoretical conceptualization and may be the result of an endogenous decision. In addition, the recent global trend toward abolishing mandatory retirement practically eliminates retirement exogeneity, giving individuals more opportunity to choose the timing of their retirement. Hence, the post-retirement outcomes they experience are presumably associated with the timing and circumstances of retirement.

Consider a male in good health and working full-time who retires before the age of 65. We would expect that this person differs in some characteristics, for example income, education, health and financial assets, from a person of the same age who does

not retire. Moreover, if this individual retires "voluntarily", meaning that he is not forced to retire, we would expect that he is well prepared (including financially) for retirement.

By comparison, a male of the same age who also retires before the age of 65 but instead is in bad health, or has been experiencing a long term disability immediately before retirement, also would differ in some characteristics from a non-retiree (and a "voluntary" retiree). However, it is more likely that this individual was forced into retirement, presumably by his health circumstances (this could be an example of "involuntary" retirement). That kind of transition to retirement may carry a likelihood of non-preparedness to face the challenges of the post-retirement period. Thus, despite their similarity in taking early retirement, these two retirees exhibit different trajectories of (or pathways to) retirement, possibly leading them to different post-retirement outcomes.

To study the effects of early retirement pathways on post-retirement outcomes, we employ a sample of men who retire within the three panels of the Canadian Survey of Labour and Income Dynamics (SLID). The SLID is a unique source of high quality socio-economic and demographic longitudinal data which provides a unique perspective on the labour behaviours and income of Canadians.

We focus on three indicators of post-retirement outcomes: self-reported health status, self-reported stress level and dwelling tenure. The first two indicators are regarded as proxies for subjective well-being, while the third is intended to capture financial strain. Further details on these variables will be presented below.

The paper proceeds with a review of the related retirement literature. Section 3 provides details on the SLID and our sample of analysis. Section 4 offers the econometric

methodology. Section 5 reports and Section 6 discusses the results. Section 7 concludes.

2. Literature Review

The fraction of males leaving work before the traditional pension age has been increasing for decades since the 1970s and only recently has it shown signs of reversal. This downward tendency in the labour force participation rate and the recovery observed over the past 10 years has been common across OECD countries. The latest decreases in early exits are generally attributed to financial need, better health, improved health care, institutional changes (such as abolishing mandatory retirement and/or early retirement incentives) or even changes that re-emphasize work as a means of providing a sense of purpose (Bailey 2005, Townson 2006). As well, the decrease in the exit from the labour force among older married men has been attributed partially to the increases in their spouses' labour participation (Schirle 2007).

Studies agree that health is among the most common reasons for early retirement. McGarry (2004) finds that subjective health affects retirement expectations significantly and that changes in the retirement expectations are influenced much more by changes in health than by changes in income or wealth. Health and health related conditions such as various forms of depression and disability were also found to be positively associated with early retirement (Tompa 1999, Blekesaune and Solem 2005, Karpansalo et al. 2005, Møller Danø et al. 2005).

Several papers (Blekesaune and Solem 2005, Siegrist et al. 2007, Mein and Ellison 2006, Lund and Villadsen 2005, Soidre 2005, Dorn and Sousa-Poza 2005) point

to occupational and work conditions as particularly important reasons for an early exit from the labour force. A negative attitude towards work motivates individuals to retire, while an improvement in working conditions and the psychosocial quality of work are strong incentives to stay at work. In addition, the interaction between individual attributes and organizational practices is shown to affect retirement outcomes (Vickerstaff 2006). Thus, the employer is found to have a role in determining the manner and timing of individual retirement.

Wealth, income and institutional incentives are generally recognized as factors that affect early retirement, although there is some disagreement on the importance that institutionally devised incentives have for early exit. Banks and Smith (2006) find that people at both ends of the wealth distribution are less likely to be working even in their 50's. They also suggest that retirement decisions in the UK are strongly influenced by institutional factors. Schils (2004) shows that the retirement behaviour of older workers is dependent largely on retirement policies and regime types. In the UK, where a market-oriented welfare system is in place, the rate of early exits is lower compared to the early exits in Germany and the Netherlands. (The welfare states of the latter two countries are characterized by more generous and universal social security arrangements.)

Dorn and Sousa-Poza (2005) study early retirement in Switzerland and conclude that the wage rate is an important determinant. Their results indicate that both high and low wages reduce the probability of an early exit and also point out the significance of social security coverage. This is in line with the findings of Larsen and Pedersen (2005) who suggest that early retirement programs have a positive effect on early exit in

Denmark. A Canadian perspective on the significance of income and private pensions is provided by Tompa (1999). He shows that many who receive early retirement benefits under the Canadian Pension Plan (CPP) are people with higher life-time earnings. In addition, those who receive private pensions are more likely to make an early exit to publicly funded retired-worker benefits.

Canadian studies provide conflicting evidence of the impact of institutions (and institutional change) on the decision to retire early. Baker and Benjamin (1999) consider early retirement behaviour of older males (men between the ages of 60 to 64) and find that even though public pension receipt significantly increased after the introduction of early retirement provisions to the Canadian Pension Plan (CPP) and the Quebec Pension Plan (QPP), little immediate effect on labour market behaviour was observed[14]. In addition, Compton (2002) finds no evidence that expected CPP/QPP benefits or current wage levels influence the retirement decision. (She also finds no evidence that wealth affects the retirement decision.) By contrast, using a unique administrative data set, the Longitudinal Worker File, combined with data from the Canadian Census, the Labour Market Activity Survey and the SLID, Baker et al. (2003) find that the Canadian Income Security system has substantial effects on the retirement decisions of both males and females.

Other determinants of early retirement identified in the literature include education, industry, occupation (Dorn and Sousa-Poza 2005) as well as individual and partner's employment status (Tompa 1999, Dorn and Sousa-Poza 2005, Schellenberg et

[14] The early retirement provision was introduced to the QPP almost three years before the matching reform of the CPP was enacted.

al. 2006, Schirle 2007), career instability (Stone et al. 2006) and well-being (Szinovacz 2006).

Unlike the abundance and diversity of studies on the determinants (implicitly – the pathways) of early retirement, the research on post-retirement outcomes has primarily focused on consumption (Hamermesh 1984, Banks et al. 1998, Bernheim et al. 2001, Hurd and Rohwedder 2003, Miniaci et al. 2003, Aguiar and Hurst 2004, Haider and Stephens 2007). The common finding is that average consumption falls significantly after retirement. (Hurd and Rohwedder 2006, Ameriks et al. 2007 and Alan et al. 2007 suggest that often this drop is fully anticipated by the households.) However, the related question of how the pathway to retirement and the retirement circumstances are associated with the individual post-retirement outcomes other than consumption has received surprisingly little attention.

Smith (2006) shows that among UK households of men who retired involuntarily, spending falls in retirement. Using Canadian data (cycles 5, 9 and 16 of the General Social Survey and the 1975 Retirement Survey), Alan et al. (2007) find that involuntary retirement and specifically involuntary retirement associated with poor health is significantly correlated with financial dissatisfaction in the post-retirement period.

Van Solinge (2007) and Rennemark and Berggren (2006) also suggest that individuals retiring involuntarily, including those who retire for health reasons, differ from voluntary retirees and working individuals of the same age in observable characteristics, and also experience a decrease in perceived health.

A Canadian study (Pyper and Giles 2002) of the labour force behaviour of workers in their 50's and 60's shows that those who ended career jobs voluntarily are less likely to work again during the following two years than those who experienced an involuntary job separation.

Thus, the pathways to retirement appear to affect post-retirement outcomes. In particular, the type of retirement (voluntary/involuntary) seems to be associated directly with post-retirement well-being.

An interesting question that has received limited attention is whether individuals are willing to substitute ownership for rentership in the event of an "involuntary" early retirement. A number of studies (Skinner 1996, Ostrovsky 2002, Venti and Wise 2004) have found that housing wealth is not treated as any other financial asset in later life; households are generally more reluctant to consume it. Browning and Lusardi (1996) and Browning and Crossley (2001) provide the rationale behind this phenomenon: due to precautionary motives agents may not regard all their sources of wealth as substitutes. Thus, housing wealth may be viewed as a financial asset to be consumed after retirement and/or in extraordinary circumstances (e.g. unexpected health shocks). Hence, it is interesting to know whether "involuntary" early retirement could be a trigger for tapping into this source of wealth by substituting ownership for rentership.

Our focus here is to investigate whether and how the pathway to retirement is related to individual self-perceived health, self-reported stress and dwelling tenure in retirement. The use of a high-quality longitudinal dataset allows us to identify individuals who transitioned to retirement and to know their pre- and post-retirement characteristics.

Hence, our analysis relies on observables rather than recall and avoids possible bias arising from post-retirement rationalisations, which is a primary concern in studies based on surveys of retirees. On the other hand, our study is constrained by the 6-year length of the SLID panel and the relatively small number of individuals making the transition to retirement.

3. Survey Data and Sample of Analysis

3.1. Survey Details

The Canadian Survey of Labour and Income Dynamics is administered by Statistics Canada and collects data on the labour market activity, income and related socio-economic and demographic characteristics of the Canadian population. The first year of SLID annual data is 1993.

SLID's longitudinal component collects information on individuals in Canada, age 16 and over[15], by following them for a period of six consecutive years. To keep the longitudinal sample representative of the Canadian population, a new panel of respondents is introduced every 3 years. This design also ensures continuity of the SLID sample. Every panel includes about 30,000 adult individuals, members of about 15,000 households.

The sampling frame of the Canadian Labour Force Survey (LFS) is used for the SLID. Moreover, the samples for SLID are selected from the monthly LFS. The Statistics Canada guide to SLID points out that "the LFS sample is drawn from an area frame and

[15] Excluding residents of Yukon, Northwest Territories and Nunavut, residents of institutions and persons living on Indian reserves. These restrictions remove less than 3% of the population.

is based on a stratified, multi-stage design that uses probability sampling. The total sample is composed of six independent samples, called rotation groups, because each month one sixth of the sample (or one rotation group) is replaced" (Statistics Canada 2004). Two rotation groups of the LFS constitute a SLID panel.

SLID data collection is by computer-assisted interviewing by telephone. At the beginning of each panel, background information about respondents is collected. The following 6 regular interviews have a split format with labour topics covered in January and income topics covered in May. In both cases, questions refer to the previous calendar year. The income interview is scheduled in May to take advantage of the income tax filing period when respondents are expected to be more familiar with their income records. Usually, over 80% of respondents agree to provide access to their administrative income tax records

3.2. Sample Restrictions

For our analysis the initial longitudinal SLID sample is subjected to a number of restrictions.

Firstly, we impose a time restriction due to the survey design. The questions on health status, stress level and major activity (which are key variables in our analysis) were first asked in 1996. This automatically removes the period 1993 to 1995 from consideration. Thus, we employ years 1996 to 1998 of Panel 1 and the following two complete panels (years 1996 to 2001 and 1999 to 2004).

Further, we restrict our analysis to males to conform with the bulk of the existing literature on retirement.

We limit our study to men who are in the age range 40 to 58 in the initial year of every SLID panel, who retire in any of the following years and stay retired for at least one year after retirement. Retirement under the age of 40 is very rare. The upper age limit is imposed to constrain the age of retirement to 64 in the final year of our panels (6 years later). Naturally, we also exclude all men who retire in the first or last year of every SLID panel.

Details on the primary definitions of retirement and the sample sizes associated with them are provided in Tables 1 and 2. Table 3 offers a description of all retirement definitions considered in the study. The definitions of retirement are discussed below.

3.3. Variables

3.3.1. Indicators of Retirement

Our retirement indicators are derived from the variables *end-of-year major activity* and *annual labour force status*.

The major activity is identified through responses to the question "I'd like to ask you a few questions about your main activity at the end of [year]. Was your main activity . . ." with possible answers (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 or disabled?", (8) "Doing volunteer work?" and (9) "Other?".

Based on major activity, a person is defined as having retired during the survey period if:

(1) *He was self-reported as "Retired" in any two consecutive years between years 2 and 5 of a panel and was not self-reported as retired in the year preceding the first year of self-reported retirement*[16](Definition A), or

(2) *He was self-reported as not at school, not working at a job/business, not looking for work and not having a long-term illness two consecutive years between years 2 and 5 of a panel and was self-reported as having a long-term illness, working at a job/business or looking for work in the year preceding this two-year period* (Definition B).

There are pronounced differences between definitions A and B. The first defines as retired only those men who self-reported themselves as retired. By contrast, Definition B extends the range of activities that constitute retirement and also limits the states from which the transition to retirement could take place; Definition B focuses on withdrawal from the labour force. Table 4 offers the percentage of retirees in the SLID sample (Panel A) and the percentage distribution of retirees by age-group (Panel B), all based on Definition B[17].

The annual labour force status is constructed based on a vector of monthly labour force status indicators and its categories are: (1) employed all year, (2) unemployed all year, (3) not in the labour force all year, (4) employed part-year, unemployed part-year,

[16] In the first SLID panel, a person is defined as having retired if he was self-reported as "Retired" in the last two years of Panel 1 (years 5 and 6) and was not self-reported as retired in year 4. The same rule for timing, regarding Panel 1, is applied to all subsequent definitions of retirement.

[17] Summary statistics for the rest of retirement definitions are not presented because of disclosure limitations. Definition B statistics offer a close approximation.

(5) employed part-year, not in the labour force part-year, (6) unemployed part-year, not in the labour force part-year, and (7) employed, unemployed and not in the labour force during the year. Labour force status is recorded for respondents age 16 to 69.

Using the annual labour force status, a respondent is defined as having retired if:

(1) *He was not in the labour force for two consecutive years between years 2 and 5 of a panel and was in the labour force for at least part of the year preceding this two-year period* (Definition C), or

(2) *He was not in the labour force and was not disabled for two consecutive years between years 2 and 5 of a panel and was either disabled and not in the labour force or was in the labour force for at least part of the year preceding the two-year period of not being in the labour force* (Definition D).

The more general Definition C defines retirees based strictly on withdrawal from the labour force. A potential problem with this indicator is that not all men who are not in the labour force for two consecutive years are retired. Long-term disability could influence labour force participation and although it could potentially lead to retirement, it is not a certain outcome, especially for younger men. Definition D accounts for a possible effect of disability by excluding the temporarily disabled individuals who are not in the labour force from the retiree group.

The reason we consider both the end-of-year major activity indicator[18] and the annual labour force status of respondents is that each provides information unmatched by

[18] From 1996 to 1999 the major activity indicator refers to the major activity during the reference year. From 1999 and onwards, this indicator refers to the end-of-the-reference-year major activity. We conducted tests for the change in definition and found that it had no statistically significant effect on our estimates.

the other. While annual labour force status captures the labour behaviour throughout a year, it is less informative about non-labour behaviours (e.g., someone could report himself as not being in the labour force for an entire year but this does not necessarily mean that he is retired). On the other hand, the major activity indicator is more informative about the non-labour behaviours but it refers only to activities at the end of a specific year. (Thus, if someone retires during a year but his major activity at the end of that year is different, he may be misclassified as non-retired.)

We recognize that our retirement definitions are flagging different sub-populations of retirees. It is possible that men identified as retired by the major activity variable could report an annual labour force status that is inconsistent with the pattern characteristic for fully retired individuals. Thus, definitions A and B capture both full and partial retirement. In contrast, definitions C and D identify only full retirees.

To be defined as a retiree, all four definitions require that the major activity or labour force status of an individual remains unchanged over two consecutive years. The main reason for selecting this particular post-retirement period is the limited time-dimension of our panels as well as the very small number of men making a transition in a particular year. To obtain a sample size sufficient for analysis, we pool retirees within a panel (e.g. within one complete, 6-year long SLID panel we pool the individuals who retire in years 2 to 5 and thus roughly quadruple the number of retirees) and also across panels. Table 2 shows the sample sizes that correspond to definitions A to D. Imposing a longer post-retirement period would result in a sample size too small for analysis. On the

other hand, a post-retirement period of only one year would increase the heterogeneity in our sample and would likely reduce the reliability of our results.

Also, due to the nature of our data, some retirees have a longer observed post-retirement period than others (e.g., men who retire in year 2 of a complete panel are observed for three more years than men who retire in year 5) while others have a longer pre-retirement period. For consistency, we restrict the observed retirement period to two years and the pre-retirement period to one for all retirees. This translates into strict and well defined horizons. (Other definitions relax this restriction somewhat, as described below.)

3.3.2. Alternative Indicators of Retirement

Table 3 offers a description of the full range of retirement definitions used in this study. The table has two panels: Panel 1 provides a symbolic description whereas Panel 2 offers a verbal description of the retirement definitions. The indicators are divided into three groups: (i) indicators based on major activity (definitions 1 to 8); (ii) indicators based on annual labour force status (definitions 9 to 11), one of which is combined with disability (definition 11); and (iii) indicators based on the union of major activity and labour force status (definitions 12 and 13)[19]. Within each group we distinguish two types, differing in terms of the observed pre-retirement history: the first considers only one year of pre-retirement history (odd numbered definitions from 1 to 11 and D12); the second takes into account the entire available and observed pre-retirement period (even

[19] Definitions A to D are listed in Table 3 as D1, D7, D9 and D11 respectively.

numbered definitions from 2 to 10 and D13). We use those definitions to study the sensitivity of our results to the retirement specification.

3.3.3. Pathways to Retirement

The set of pre-retirement characteristics used to identify a pathway to retirement includes *end-of-year major activity*, *annual labour force status*, *self-assessed health* and *disability status*.

To create the pathway variables for use in model estimation, we collapse the categories of the end-of-year major activity variable from nine down to five by combining going to school, keeping house, caring for other family members, and doing volunteer work with "other". We also recategorize the annual labour force status indicator based on the patterns of labour force participation. The transformed variable distinguishes among employed all year, unemployed all year, individuals who are not in the labour force all year and those who are unemployed or not in the labour force part-year. With these variables we intend to capture the character of respondent's self-reported labour and non-labour activities as well as his labour force (or job) attachment which, as the body of retirement literature suggests, may be associated with retirement and hence could affect the post-retirement outcomes.

It is natural to include self-assessed health status in the set of pre-retirement characteristics since health is one of the major determinants of retirement. Self-assessed health (SAH) has five ordered categories which are the answers to the question "Compared to other people [respondent]'s age, how would you describe [respondent]'s

state of health? Would you say it is..." with possible answers: (1) "Excellent", (2) "Very good", (3) "Good", (4) "Fair", or (5) "Poor". In our analysis we sometimes use all five categories but collapse them into two (by combining excellent, very good and good as "good", and fair and poor as "bad") when the cell sizes are small.

Our disability status variable is a flag indicating whether the person suffered from any long-term physical condition, mental condition or health problem in the pre-retirement year. The inclusion of this indicator in the set of characteristics is based on substantive evidence in the literature of its significance for retirement.

The post-retirement outcomes we consider are SAH, self-reported stress and dwelling tenure. In our post-retirement models, we utilize the five-category SAH. The stress variable is an answer to the question "Would you describe [respondent's] life as . . ." with possible answers: (1) "Very stressful?", (2) "Somewhat stressful?", (3) "Not very stressful?", (4) "Not at all stressful?". Finally, the dwelling tenure is a dummy taking the value of unity if the respondent (or a member of his family) owns his current residence and zero otherwise. Table 5 presents the transition matrices of these three outcomes; dwelling tenure exhibits a strong state persistence while health and stress seem to be less state-dependent.

We also consider a number of additional variables (the controls) which, as we hypothesize, have important effects on post-retirement outcomes. Those are age, educational attainment, investment income (including interest from bank accounts and other deposits, interest from loans and mortgages, regular income from trust funds or estates, net dividends and other investment income), family composition, marital status,

an indicator of self-employment, immigrant status (immigrant if been in Canada for less than 10 years) and an indicator for whether the person is a member of an employer provided pension plan.

Table 6, Panels A and B present the pre- and post-retirement summary statistics of the sample of retirees (under Definition B). We can see that there are no unusual dynamics in our sample. However, several things are worth noting. Firstly, while the fraction of men who are not in the labour force increases substantially after retirement (as compared to before retirement), its value remains below unity. In contrast, the major activity indicator shows that there are no men who are in the labour force in the year after retirement. This difference is expected and is caused by the characteristics of Definition B presented earlier. Secondly, the fraction of married men increases in the year after retirement and so does that of disabled men. Finally, we see a substantial increase in mean investment income, both personal and family, in the year after retirement. While it is difficult to establish the exact cause for this, one possibility is that some retirees may receive lump sums upon retirement or may "cash in" retirement assets such as Registered Retirement Savings Plans.

4. Methodology

We adopt the following estimation approach. First, for each of the 13 definitions of retirement we identify all those in the three panels of SLID who retired. We then pool the retirees from all panels to obtain our sample for analysis. Finally, for each of the three

post-retirement outcomes (health status, stress level and dwelling tenure) we estimate a model of the form:

$$y_{jR+1} = f(y_{jR-1}, P_{jR-1}, W_{jR-1}) \quad (1)$$

where y_{jR+1} is the post-retirement outcome of retiree j and y_{jR-1} is the pre-retirement value of the same characteristic for the same individual. The vector P_{jR-1} includes the set of pre-retirement characteristics used to identify a pathway to retirement (the pathway variables) and W_{jR-1} contains the set of pre-retirement year controls[20]; R denotes the year of retirement.

A few points deserve particular attention. Firstly, only one of the two pathway variables, major activity and annual labour force status, is included in the model. Thus, two separate versions of every model are estimated[21]. Secondly, the stress and dwelling specifications include an additional regressor, vector Z_{jR+1}. It contains the post-retirement annual labour force status of a retiree when a definition of retirement is based on major activity; when the definition is based on annual labour force status, Z_{jR+1} includes the post-retirement major activity of a retiree. The reason for having $R+1$ labour and non-labour activity indicators in these models is that they could have an effect on the $R+1$ outcomes and hence should be accounted for. Finally, the set of controls, which is common to all models, includes age, educational attainment, family investment income,

[20] In final model estimation we use standardized investment income. Also, we exclude the indicators of marital status and pension plan membership from the set of controls since their effects are insignificant once family composition and labour force status are controlled for.

[21] Exceptions to this rule are models based on definitions 12 and 13 where both variables are included in the model.

family composition, an indicator of self-employment, and an indicator of immigrant status.

The econometric form of f in specification (1) changes depending on whether y_{jR+1} is binary or has multiple categories. In the former case f has a probit form, and in the latter f has an ordered probit form.

Our objective is to characterize how retirees differ among themselves in the post-retirement period conditional on their observed pre-retirement characteristics (and implicitly on the type/pathway of retirement they experience).

We estimate specification (1) separately for health status, stress level and dwelling tenure. Other studies (e.g., Smith 2006 and Alan et al. 2007) have suggested that involuntary retirement is likely to be associated with worse post-retirement outcomes. To test for that we define involuntary retirees as those who are unemployed or not in the labour force for at least some period during the year preceding retirement, who report fair or poor health, and/or have a long-term illness or disability. We would expect individuals employed all year, in good or better self-assessed health and having no disability to be the voluntary retirees.

We hypothesize that men who have been forced to retire would be less likely than voluntary retirees to report high self-assessed health, more likely to report higher levels of stress and just as unwilling to substitute ownership for rentership.

5. Results

5.1. Models Utilizing Indicators of Retirement Based on Major Activity

We begin by examining the results of the models using retirement definitions A and B. Tables 7 to 9 present the estimates of the health, stress and dwelling tenure models. The first two columns of every table report the model estimates under Definition A; the second two columns show the estimates under Definition B. All three models are estimated twice: once with pre-retirement major activity as part of the set of regressors, Columns 1 and 3, and once with pre-retirement labour force status, Columns 2 and 4.

The results for the health models[22] (Table 7, Columns 1 to 4) show the presence of a strong and statistically significant ($p<0.05$) gradient in health: men are more likely to experience worse health after retirement the worse their state of pre-retirement health. Moreover, disability in the pre-retirement period is associated with worse post-retirement health, irrespective of the retirement definition. Thus, one who retires because of bad health or disability is likely to have worse post-retirement health than one who does not possess those particular pre-retirement characteristics.

Conditioning on pre-retirement health, men who live in an owned (either by them or a member of their family) residence are more likely to report better health in post-retirement, compared to renters. The effect is statistically significant under Definition A (Columns 1 and 2) but it is of somewhat lower magnitude and less significant ($p<0.10$) under Definition B (Columns 3 and 4). Neither major activity nor annual labour force status are found to have a significant effect on post-retirement health at group levels.

[22] Since health status is coded from (1) to (5) where (1) is excellent and (5) is poor health, a positive effect in the health models indicates lower health compared to the reference group.

We next consider the stress model estimates[23] (Table 8). As anticipated, reporting lower pre-retirement stress is associated with lower post-retirement stress, especially for the lower stress categories. Men who define themselves as not very stressed or not stressed in pre-retirement are more likely to report lower stress than those who define themselves as very stressed in pre-retirement. Controlling for pre-retirement stress, men in poor health are likely to indicate significantly higher stress than men in fair or better health. (This result is statistically significant only under Definition A.) Disability in pre-retirement is found to be associated with higher stress in post-retirement. The statistical significance of this result holds under both retirement definitions (Table 8, Panel 1, Columns 1 to 4).

Retirees who own their current residence (or live in a residence owned by a member of their family) in pre-retirement are more likely to report lower levels of stress in post-retirement, compared to renters. Just as before, we find that major activity and annual labour force status have no statistically significant effects on post-retirement stress at group levels (controlling for post-retirement labour status).

Finally, we review the estimates of the dwelling tenure regressions (Table 9). We find a strong state-dependence in dwelling ownership across retirement definitions. This result is in line with the literature. On the other hand, our results show no evidence of an association between health status (tested as a group) and dwelling ownership in post-retirement. Also, we see no statistically significant association between disability status and ownership.

[23] Since stress status is coded from (1) to (4) where (1) is very stressed and (4) is not stressed, a positive effect in the stress models indicates higher stress compared to the reference group.

Although the labour and non-labour indicators (major activity and annual labour force status) are found to be insignificant predictors of post-retirement dwelling ownership (tested as groups), our findings indicate that those who are looking for work in the pre-retirement period are less likely to be owners in post-retirement compared to men who are working at a job or business (Table 9, Columns 1 and 3). This effect is significant for both retirement definitions.

5.2. Models Utilizing Indicators of Retirement Based on Labour Force Status

We now consider the results of the models using retirement definitions C and D[24]. Tables 10 to 12 present the estimates of the health, stress and dwelling tenure models. The first two columns of Tables 10 and 11 report the model estimates under Definition C; the second two columns show the estimates under Definition D. As before, all three models are estimated twice: once with pre-retirement major activity as part of the set of regressors, Columns 1 and 3, and once with labour force status, Columns 2 and 4. Table 12 reports estimates only under Definition D; the dwelling tenure model under Definition C cannot be estimated because of insufficient variability in the dependent variable.

As before, the health models estimates (Table 10, Columns 1 to 4) indicate the presence of a strong and statistically significant gradient in health. In addition, disability in the pre-retirement period is found to be associated with worse post-retirement health

[24] The models using retirement definitions C and D are estimated for smaller sample sizes than the models of self-reported retirees. Table 2 provides details.

under Definition C but, unlike the previous estimates, under Definition D disability is insignificant. (A discussion of these findings is provided below.)

Conditioning on pre-retirement health, pre-retirement annual labour force status is found to have a statistically significant effect on post-retirement health with significance holding across both retirement definitions. The estimates under Definition C (Table 10, Column 2) suggest that men who are unemployed all year and men who are unemployed part-year or not in the labour force part-year in the year before retirement are more likely to report better health after retirement than men who are employed all year. Both estimates are statistically significant ($p<0.01$). Also, the effect of unemployment remains statistically significant under Definition D (Table 10, Column 4). While this is an unexpected finding, results from previous studies provide a possible explanation; further discussion follows.

We next consider the stress model estimates (Table 11). There is once again a well-established pattern of state-dependence in stress. Its statistical significance generally holds across the two retirement definitions. However, contrary to the earlier results, disability and dwelling tenure are not found to be associated with stress. Major activity and annual labour force status at group levels are also found to be insignificant for stress. However, men reporting themselves as looking for work in the period before retirement are likely to report lower stress than men whose major activity is working at a job or business (Table 10, Column 3). These findings are discussed below.

In line with the conclusions in the literature, our results of the dwelling tenure estimations (Table 12) show that dwelling ownership after retirement is associated only

with dwelling ownership before retirement. Retirement circumstances appear to have no statistically significant effect on post-retirement dwelling tenure.

5.3. Models Utilizing the Alternative Indicators of Retirement

Figures 1 to 7 present graphically the estimates based on all 13 definitions of retirement. Figures 1 to 4 offer coefficient estimates from the health models; Figures 5 to 7 show estimates from the stress models. Asterisks indicate statistical significance ($p<0.05$).

Figure 1 presents the effects of the various pre-retirement major activities (compared to the baseline) on post-retirement health. The effects are rather scattered and generally statistically insignificant across retirement definitions. In contrast, the coefficient estimates of the disability dummy are relatively similar in magnitude and preserve their statistical significance across 12 of 13 retirement definitions (Figure 2).

Figure 3 shows that the estimates of the labour force status categories (compared to the baseline) appear to be statistically significant only for models in which retirement is defined by annual labour force status (i.e., definitions 9, 10 and 11). Their magnitudes and directions are in line with our findings from previous sections.

Higher educational attainment, as compared to less than high school, is found to be associated with better health in retirement (Figure 4). The significance of the effects varies across retirement definitions. Also, there is no evidence of a well established gradient.

Figures 5 to 7 offer no surprises. The estimates of the categories of major activity (Figure 5) are generally insignificant (compared to baseline); the same holds for the estimates of annual labour status (Figure 6). As seen before, disability loses its significance once retirement definitions based on labour force status are employed (Figure 7). However, one result deserves attention: men who are not looking for work, not retired and not long-term ill are found to be more likely to report higher stress in post-retirement than men who work at a job or business (Figure 5). This effect is fairly consistent across retirement definitions (whenever estimated) and statistically significant for almost all definitions based on annual labour force status[25].

Finally, the effects of the other pre-retirement characteristics are not offered here since their magnitudes and significance do not vary substantially across definitions. Similarly, we do not present the estimates of the dwelling tenure models since they provide no further insights into the relationship between the post-retirement dwelling tenure and pathway to retirement.

6. Discussion

In this study we consider four primary retirement indicators: definitions A and B based on major activity, and definitions C and D based on annual labour force status. Each group of definitions treats retirement differently. Definitions A and B define retirement as self-reported while C and D derive retirement from the pattern of observed labour force participation. We account for these differences in our discussion.

[25] The likely reason for this finding is that the "other activity" category is dominated by men who report themselves as "Caring for other family members (including young children)".

We first ask whether different retirement pathways are associated with different post-retirement outcomes of health and stress for self-reported retirees. Our findings indicate that pre-retirement health (self-reported health and disability) is likely to be associated with post-retirement health outcomes (Table 7). Men who report bad health and men who are disabled are more likely to report worse health in the year after retirement, compared to the respective baselines. Also, pre-retirement stress levels and health related indicators (e.g. SAH and disability) are likely to be related to post-retirement stress outcomes (Table 8). In both models, we find no evidence that labour force participation and non-labour activities are associated with post-retirement health or stress. This finding is certainly unexpected. However, it suggests that although these two indicators may have an impact on the decision to retire (as indicated in the literature), they are unlikely to have long-term effects on health and stress. Some possible reasons for this may be the social safety net and the universal health care coverage available in Canada. On an individual level, it could be that people are able to predict (and thus prepare for) or even create their retirement circumstances with respect to labour force participation and, especially, early retirement. Clearly, if this holds true it would eliminate any negative effects on health and stress in post-retirement.

We next ask whether different retirement pathways are associated with different post-retirement outcomes of health and stress for individuals with "derived" retirement. Health related indicators (SAH and disability) are once again found to be associated with post-retirement health (Table 10). However, in certain cases (Definition D), disability is not a significant predictor. This variation in significance is most likely a result of the

differences between definitions C and D. As explained above, Definition C is more general and thus identifies as retired all men who are not in the labour force, irrespective of whether they are disabled or not in the post-retirement period. By contrast, Definition D does not carry disability into retirement (but also substantially reduces the number of retirees in the sample). Thus, part of the disability effect estimated under Definition C is attributable to men who are disabled in pre-retirement and continue to be disabled in post-retirement.

The health model estimates (Table 10) suggest that, conditioning on pre-retirement health, men who were unemployed in the entire pre-retirement period are likely to report better health than retirees who were employed all year. A possible explanation is that unemployment (involving job search and perhaps a stigma) tends to yield lower self-assessed health than retirement (no job search and no stigma). A thorough analysis of these and related issues is a subject of future work.

The stress model results (Table 11) indicate that pre-retirement stress is likely to be associated with stress in the post-retirement period. On the other hand, SAH and disability are generally insignificant. The estimate of looking for work as a major activity under Definition D indicates that men in that category are likely to report lower stress than men who are working at a job or business. This result is in line with our findings for health and is subject to the same economic interpretation.

We now look at the results of the dwelling tenure models. We show that, irrespective of the retirement definition, retirement circumstances have no effect on dwelling ownership in post-retirement. That result is consistent with the findings in the

literature. A possible explanation of the lack of effect would be that individuals will not easily substitute ownership for rentership. Other possible responses that individuals could take in the event of an early retirement, such as downsizing, participation in inverse mortgage programs and even relocation, are subjects to future study.

Finally, we consider the pathway variables whose effects are significant across most retirement definitions. Along with the pre-retirement dependent variables, disability is clearly the one that has the most stable effect and is generally statistically significant.

In conclusion, it is natural to speculate that adverse experiences followed by next-period retirement may be indicative of an "involuntary" retirement. In that context, our results could be interpreted in two ways. Firstly, they suggest that differences in pre-retirement health indicators (such as self-assessed health and disability) are likely to be associated with differences in post-retirement health and stress. Secondly, "involuntary" retirees (i.e., men who may have retired because of health and/or health related conditions) are more likely to experience worse post-retirement outcomes (in terms of health and stress) than "voluntary" retirees. In this respect, our findings are fully consistent with those in the early retirement literature.

Our results on dwelling tenure are also in line with the literature: it appears that retirement circumstances have little effect on dwelling ownership.

7. Conclusion

In this paper we investigate whether differences in retirement pathways are associated with differences in post-retirement outcomes of health, stress and dwelling

tenure. We use a sample of men from the Canadian Survey of Labour and Income Dynamics, for the years 1996 to 2004. We find that differences in pre-retirement health indicators (such as self-assessed health and disability), as pathways to early retirement, are likely to be associated with differences in post-retirement health and stress. In addition, our results suggest that "involuntary" retirees (i.e., men who may have retired because of health and/or health related conditions) are more likely to experience worse post-retirement outcomes (in terms of health and stress) than "voluntary" retirees. Retirement circumstances are found to have no statistically significant effect on dwelling ownership.

Our findings are entirely consistent with others reported in the literature. However, due to data restrictions, our study is far from complete. One possible extension would be to employ a more comprehensive longitudinal dataset and create a more detailed description of early retirement patterns and post-retirement outcomes. In particular, the goal would be to construct more precise pathways of early retirement (also involving consumption/expenditure) and use a larger set of post-retirement indicators (again including consumption/expenditure) for improving the comparative analysis.

An additional avenue of research would be to look at the particular actions, if any, that individuals take with respect to their dwelling in the event of an early retirement. Is it the case that people downsize, participate in inverse mortgage programs or simply relocate, while continuing to own?

Another possible direction of research would be to extend the early retirement analysis to women. Studies indicate that the labour force participation of women is unlike

that of men. Hence, studying their pre- and post-retirement characteristics and associations is a necessary step towards achieving a complete understanding of early retirement patterns and post-retirement experiences.

Finally, investigating how the individual characteristics of one family member influence the post-retirement outcomes of another family member (e.g. how the labour force status of the wife affects the post-retirement outcomes of the husband) is another attractive avenue for future research on early retirement.

References

Aguiar, M., Hurst, E. (2004), Consumption vs. Expenditure, NBER Working Paper No. 10307

Alan, S., Atalay, K., Crossley, T. (2007), The Adequacy of Retirement Savings: Subjective Survey Reports by Retired Canadians, SEDAP Research Paper No. 199

Ameriks, J., Caplin, A., Leahy, J. (2007), Retirement Consumption: Insights from a Survey, *Review of Economics and Statistics*, vol. 89, no. 2, pp. 256-274

Bailey, R. (2005), They're Making Adults Younger, *Rough Notes [Indianapolis]*, vol. 148, no. 10, pp. 52-54

Baker, M., Benjamin, D. (1999), Early Retirement Provisions and the Labor Force Behavior of Older Men: Evidence from Canada, *Journal of Labor Economics*, vol. 17, no. 4, pp. 724-756

Baker, M., Gruber, J., Milligan, K. (2003), The Retirement Incentive Effects of Canada's Income Security Programs, *The Canadian Journal of Economics*, vol. 36, no. 2, pp. 261-290

Banks, J., Blundell, R., Tanner, S. (1998), Is There a Retirement-Savings Puzzle?, *American Economic Review*, vol. 88, no. 4, pp. 769-788

Banks, J., Smith, S. (2006), Retirement in the UK, *Oxford Review of Economic Policy*, vol. 22, no. 1, pp. 40-56

Bernheim, D., Skinner, J., Weinberg, S. (2001), What Accounts for the Variation in Retirement Wealth among US Households?, *American Economic Review*, vol. 91, no. 4, pp. 832-857

Blekesaune, M., Solem, P. E. (2005), Working Conditions and Early Retirement: A Prospective Study of Retirement Behavior, *Research on Aging*, vol. 27, no. 1, pp. 3-30

Breslaw, J., Stelcner, M (1987), The Effect of Health on the Labor Force Behavior of Elderly Men in Canada, *Journal of Human Resources*, vol. 22, no. 4, pp. 490-517

Browning, M., Crossley, T. (2001), The Life Cycle Model of Consumption and Saving, *Journal of Economic Perspectives*, vol. 15, no. 3, pp. 3-22

Browning, M., Lusardi, A. (1996), Household Saving: Micro Theories and Macro Facts, *Journal of Economic Literature*, vol. 34, no. 4 pp. 1797-1855

Chan, S., Stevens, A., H. (2004), Do Changes in Pension Incentives Affect Retirement? A Longitudinal Study of Subjective Retirement Expectations, *Journal of Public Economics*, vol. 88, no. 7-8, pp. 1307-1333

Compton, J. (2002), Determinants of Retirement: Does Money Really Matter?, Working Papers-Department of Finance Canada 2001-02, Department of Finance Canada

Costa, D. (1996), Health and Labor Force Participation of Older Men, 1900-1991, *Journal of Economic History*, vol. 56, no. 1, pp. 62-89

Denton, F., Spencer, B. (1997), Demographic Trends, Labour Force Participation, and Long-term Growth, IESOP: Independence and Economic Security of the Older Population, Research Paper No. 23

Dorn, D., Sousa-Poza, A. (2005), The Determinants of Early Retirement in Switzerland, *Swiss Journal of Economics and Statistics*, vol. 141, no. 2, pp. 247-283

Fenn, P., Vlachonikolis, I. (1986), Male Labour Force Participation Following Illness or Injury, *Economica, New Series*, vol. 53, no. 211, pp. 379-391

Haider, S., Stephens, M. (2007), Is There a Retirement-Consumption Puzzle? Evidence Using Subjective Retirement Expectations, *The Review of Economics and Statistics*, vol. 89, no. 2, pp. 247-264

Hamermesh, D. (1984), Life-Cycle Effects on Consumption and Retirement, *Journal of Labor Economics*, vol. 2, pp. 353-370

Harkness, J. (1993), Labour Force Participation by Disables Males in Canada, *Canadian Journal of Economics*, vol. 26, no. 4, pp. 878-889

Hartig T., Fransson U. (2006), Housing Tenure and Early Retirement for Health Reasons in Sweden, *Scandinavian Journal of Public Health*, vol. 35, no. 5, pp. 472-479

Hurd, M., Rohwedder, S. (2003), The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending and Retirement, NBER Working Paper No. 9586

Hurd, M., Rohwedder, S. (2006), Some Answers to the Retirement-Consumption Puzzle, NBER Working Paper No. 12057

Karpansalo, M., Kauhanen, J., Lakka, T., Manninen, P., Kaplan, G., Salonen, J. (2005), Depression and Early Retirement: Prospective Population Based Study in Middle Aged Men, *Journal of Epidemiology and Community Health*, vol. 59, no. 1, pp. 70-74

Kuhn, P., Robb, A. (1997), Unemployment, Skill and Labor Supply: Evidence from Canadian Microdata, 1971-1991, In: Abbott, M. et al. (Eds.), Transition and Structural Change in the North American Labor Markets, Kingston: IRC Press, pp. 41-85

Larsen, M., Pedersen, P. (2005), Pathways to Early Retirement in Denmark, 1984-2000, Institute for the Study of Labor, Discussion Paper No. 1575

Lund, T., Villadsen, E. (2005), Who Retires Early and Why? Determinants of Early Retirement Pension among Danish Employees 57-62 Years, *European Journal of Ageing*, vol. 2, no. 4, pp. 275-280

McDonald, L., Donohue, P., Moore, B. (1997-1998), The Economic Casualties of Retiring to Caregive, IESOP: Independence and Economic Security of the Older Population, Research Paper No. 28

McGarry, K. (2004), Health and Retirement: Do Changes in Health Affect Retirement Expectations?, *The Journal of Human Resources*, vol. 39, no. 3, pp. 624-648

Mein, G., Ellison, G. (2006), The Impact of Early Retirement on Perceptions of Life at Work and at Home: Qualitative Analyses of British Civil Servants Participating in the Whitehall II Retirement Study, *International Journal of Aging and Human Development*, vol. 63, no. 3, pp. 187-216

Miniaci, R., Monfardini, C., Weber, G. (2003), Is There a Retirement-Consumption Puzzle in Italy?, Institute for Fiscal Studies Working Paper No. 03/14

Møller Danø, A., Ejrnæs, M., Husted, L. (2005), Do Single Women Value Early Retirement More than Single Men?, *Labour Economics*, vol. 12, no. 1, pp. 47-71

Ostrovsky, Y. (2002), A Longitudinal Study of the Residential Mobility of the Elderly in Canada, SEDAP Research Paper No. 78

Parsons, D. (1982), The Male Labour Force Participation Decision: Health, Reported Health, and Economic Incentives, *Economica*, New Series, vol. 49, no. 193, pp. 81-91

Pyper, W., Giles, P. (2002), Approaching Retirement, *Perspectives of Labour and Income*, vol. 3, no. 9, pp. 5-12

Rennemark, M., Berggren, T. (2006), Relationships between Work-Status and Leisure Lifestyle at the Age of 60 Years Old, *European Journal of Ageing*, vol. 3, no. 2, pp. 82-88

Schellenberg, G., Turcotte, M., Ram, B. (2006), The Changing Characteristics of Older Couples and Joint Retirement in Canada, In: Stone, L. (Ed.), New Frontiers of Research on Retirement, Statistics Canada, Catalogue no. 75-511-XPE, pp. 199-218

Schils, T. (2004), Early Retirement Behaviour of Senior Workers: A Comparative Analysis of Germany, Britain and the Netherlands Using Panel Data, Paper Presented at the 1st International Conference for Panel Data Users, Zurich

Schirle, T. (2007), Why Have the Labour Force Participation Rates of Older Men Increased Since the Mid-1990s?, Department of Economics Working Paper Series, 2007-03 EC, Wilfrid Laurier University

Siegrist, J., Wahrendorf, M., Von Dem Knesebeck, O., Jürges, H., Börsch-Supan, A. (2007), Quality of Work, Well-Being, and Intended Early Retirement of Older Employees -- Baseline Results from the SHARE Study, *European Journal of Public Health*, vol. 17, no. 1, pp. 62-68

Skinner, J. (1996), Is Housing Wealth a Sideshow?, In: Wise, D. (Ed.), Advances in the Economics of Aging, Chicago: University of Chicago Press

Smith, S. (2006), The Retirement-Consumption Puzzle and Involuntary Early Retirement: Evidence from the British Household Panel Survey, *Economic Journal*, vol. 116, pp. C130-C148

Soidre T. (2005), Retirement-Age Preferences of Women and Men Aged 55-64 Years in Sweden, *Ageing and Society*, vol. 25, pp. 943-963

Statistics Canada (1994), Wave 2 – Survey of Labour and Income Dynamics, Microdata User's Guide

Statistics Canada (2004), Survey of Labour and Income Dynamics, User's Guide, Statistics Canada Web-Site

Stone, L., Nouroz, H., Genest, A., Deschenes, N. (2006), A New Perspective on Retirement Processes: Trajectories of Transitions from Work to Retirement, In: Stone, L., Nouroz, H. (Eds.), New Frontiers of Research on Retirement: Technical Annex, Statistics Canada, Catalogue no. 75-512-XPE, pp. 7-56

Szinovacz, M. (2006), Families and Retirement, In: Stone, L. (Ed.), New Frontiers of Research on Retirement, Statistics Canada, Catalogue no. 75-511-XPE, pp. 165-198

Tompa, E. (1999), Transitions to Retirement: Determinants of Age of Social Security Take Up, SEDAP Research Paper No. 6

Townson, M. (2006), Growing Older, Working Longer: The New Face of Retirement, Canadian Centre for Policy Alternatives

Van Solinge, H. (2007), Health Change in Retirement: A Longitudinal Study among Older Workers in the Netherlands, *Research on Aging*, vol. 29, no. 3, pp. 225-256

Venti, S., Wise, D. (2004), Aging, Housing and Equity: Another Look, In: Wise, D. (Ed.), Perspectives in the Economics of Aging, Chicago: University of Chicago Press

Vickerstaff, S. (2006), I'd Rather Keep Running to the End and Then Jump Off the Cliff, Retirement Decisions: Who Decides?, *Journal of Social Policy*, vol. 35, no. 3, pp. 455-472

Wise, D. (1997), Retirement Against the Demographic Trend: More Older People Living Longer, Working Less, and Saving Less, *Demography*, vol. 34, no. 1, pp. 83-95

Appendix

Table 1. DESCRIPTION OF THE PRIMARY RETIREMENT DEFINITIONS

Definition of Retirement Based on	Definition	Description
Major Activity	A	A person is **defined to have retired** if: 1) He was self-reported as retired in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as not retired in the year preceding the period of self-reported retirement.
	B	A person is **defined to have retired** if: 1) He was self-reported as not at school, not working at a job/business, not looking for work or not having a long-term illness in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as working at a job/business or looking for work or was self-reported as having a long-term illness, in the year preceding the period described in 1).
Annual Labour Force Status	C	A person is **defined to have retired** if: 1) He was not in the labour force for two consecutive years between years 2 and 5 of a panel, 2) He was in the labour force at least part of the year preceding the period described in 1).
	D	A person is **defined to have retired** if: 1) He was not in the labour force and was not disabled for any two consecutive years between years 2 and 5 of a panel, 2) He was either disabled and not in the labour force or was in the labour force at least part of the year preceding the period described in 1).

Table 2. SAMPLE DROP-OUT, COMPLETE CASES AND SAMPLE SIZES BY RETIREMENT DEFINITION

	Sample Characteristics		
Definition of Retirement	Drop-out	Complete Cases	Sample Sizes
Definition A	5.7%	76.6%	422/419/426
Definition B	5.0%	77.7%	463/459/466
Definition C	3.1%	58.6%	398/389/n.a.
Definition D	0.0%	59.5%	209/204/207

Notes:
1. Drop-out shows the percentage of respondents lost while confirming retirement status. These individuals have the pre-retirement characteristics of retirees but their next period observables are missing. Since they cannot be correctly classified they are eliminated from the analysis.
2. Complete cases indicates the percentage of retirees with no missing values for the variables other than those needed to define retirement.
3. Sample sizes indicates the number of individuals employed in model estimation. The first number reports the sample size used for estimating the health model, the second – the sample size for the stress model, and the third – the sample size for the dwelling model.

Table 3. DESCRIPTION OF THE FULL RANGE OF RETIREMENT DEFINITIONS

Panel 1. Symbolic Description

Definition of Retirement Based on	Definition	Description	
		Group At Risk	Retired If
Major Activity	D1 (=A)	$(W, L, S, H, C, I, V, O)_{\{t-1\}}$	$(R)_{\{t, t+1\}}$
	D2	$(W, L, S, H, C, I, V, O)_{\{t-A \sim t-1\}}$	$(R)_{\{t, t+1\}}$
	D3	$(W, L, I)_{\{t-1\}}$	$(R)_{\{t, t+1\}}$
	D4	$(W, L, I,)_{\{t-A \sim t-1\}}$	$(R)_{\{t, t+1\}}$
	D5	$(W, L)_{\{t-1\}}$	$(H, C, R, I, V, O)_{\{t, t+1\}}$
	D6	$(W, L)_{\{t-A \sim t-1\}}$	$(H, C, R, I, V, O)_{\{t, t+1\}}$
	D7 (=B)	$(W, L, I)_{\{t-1\}}$	$(H, C, R, V, O)_{\{t, t+1\}}$
	D8	$(W, L, I,)_{\{t-A \sim t-1\}}$	$(H, C, R, V, O)_{\{t, t+1\}}$
Annual Labour Force Status	D9 (=C)	$(LF, LFPY)_{\{t-1\}}$	$(NLF)_{\{t, t+1\}}$
	D10	$(LF, LFPY)_{\{t-A \sim t-1\}}$	$(NLF)_{\{t, t+1\}}$
	D11 (=D)	$(LF, LFPY, NLF \& D)_{\{t-1\}}$	$(NLF \& ND)_{\{t, t+1\}}$
Major Activity and Annual Labour Force Status	D12	$(W, L, I, LF, LFPY)_{\{t-1\}}$	$(H, C, R, V, O, NLF)_{\{t, t+1\}}$
	D13	$(W, L, I, LF, LFPY)_{\{t-A \sim t-1\}}$	$(H, C, R, V, O, NLF)_{\{t, t+1\}}$

Notes:

1. Major Activity is coded as follows: working at a job or business (W), looking for work (L), going to school (S), keeping house (H), caring for other family members (C), retired (R), long-term illness or disabled (I), doing volunteer work (V), and other (O)
2. Annual Labour Force Status is grouped as follows: "LF" includes men who were in the labour force all year, "LFPY" captures all men who were in the labour force only part-year, and "NLF" are all men who were not in the labour force a l year
3. "D" indicates the presence of a disability; "ND" stands for no disability
4. The meanings of the symbols used in the table are as follows: "t" is the period of retirement, "A" varies from 1 to 4 depending on the number of pre-retirement periods in which we observe a particular respondent, "," is the conjunction "or" unless when used in the timing index where it means "and", and "~" is the conjunction "to".

Panel 2. Verbal Description

Definition of Retirement Based on	Definition	Description
Major Activity	D1 (=A)	A person is **defined to have retired** if: 1) He was self-reported as retired in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as not retired in the year preceding the period of self-reported retirement.
	D2	A person is **defined to have retired** if: 1) He was self-reported as retired in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as not retired in all years (available in the SLID) preceding the period of self-reported retirement.
	D3	A person is **defined to have retired** if: 1) He was self-reported as retired in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as working at a job/business or looking for work or was self-reported as having a long-term illness in the year preceding the period described in 1).
	D4	A person is **defined to have retired** if: 1) He was self-reported as retired in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as working at a job/business or looking for work or was self-reported as having a long-term illness, in all years (available in the SLID) preceding the period described in 1).
	D5	A person is **defined to have retired** if: 1) He was self-reported as not at school, not working at a job/business or not looking for work in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as working at a job/business or looking for work in the year preceding the period described in 1).

Continues on next page

(1)	(2)	(3)
	D6	A person is **defined to have retired** if: 1) He was self-reported as not at school, not working at a job/business or not looking for work in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as working at a job/business or looking for work in all years (available in the SLID) preceding the period described in 1).
	D7 (=B)	A person is **defined to have retired** if: 1) He was self-reported as not at school, not working at a job/business, not looking for work or not having a long-term illness in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as working at a job/business or looking for work or was self-reported as having a long-term illness in the year preceding the period described in 1).
	D8	A person is **defined to have retired** if: 1) He was self-reported as not at school, not working at a job/business, not looking for work or not having a long-term illness in any two consecutive years between years 2 and 5 of a panel, 2) He was self-reported as working at a job/business or looking for work or was self-reported as having a long-term illness in all years (available in the SLID) preceding the period described in 1).
Annual Labour Force Status	D9 (=C)	A person is **defined to have retired** if: 1) He was not in the labour force for two consecutive years between years 2 and 5 of a panel, 2) He was in the labour force at least part of the year preceding the period described in 1).
	D10	A person is **defined to have retired** if: 1) He was not in the labour force in any two consecutive years between years 2 and 5 of a panel, 2) He was in the labour force at least part year in all years (available in the SLID) preceding the period described in 1).

Continues on next page

(1)	(2)	(3)
	D11 (=D)	A person is **defined to have retired** if: 1) He was not in the labour force and was not disabled for two consecutive years between years 2 and 5 of a panel, 2) He was either disabled and not in the labour force or was in the labour force at least part of the year preceding the period described in 1).
Major Activity and Annual Labour Force Status	D12	A person is **defined to have retired** if: 1) He was either self-reported as not at school, not working at a job/business, not looking for work or not having a long-term illness or was self-reported as not in the labour force in any two consecutive years between years 2 and 5 of a panel, 2) He was in the labour force at least part year or was self-reported as having a long-term illness, working at a job/business or looking for work in the year preceding the period described in 1).
	D13	A person is **defined to have retired** if: 1) He was either self-reported as not at school, not working at a job/business, not looking for work and not having a long-term illness or was self-reported as not in the labour force in any two consecutive years between years 2 and 5 of a panel, 2) He was in the labour force at least part year or was self-reported as having a long-term illness, working at a job/business or looking for work in all years (available in the SLID) preceding the period described in 1).

Table 4. RETIREMENT DESCRIPTIVES

A) RETIREMENT WITHIN THE SIX-YEAR SLID SAMPLE (Percentages)

Status	Percentage
Retired	3.8
Non-retired	96.2

Notes:
1. Periods 1 and 6 are excluded since Period 1 is the base year and Period 6 retirees could not be followed up (as period 6 is the last period in the panel)
2. Retirement defined according to Definition B

B) RETIREMENT BY AGE GROUPS (Percentages)

Age Group (in period 1)	Retired
40 to 49	12.4
50 to 54	38.9
55 to 58	48.7

Notes:
1. Periods 1 and 6 are excluded since Period 1 is the base year and Period 6 retirees could not be followed up (as period 6 is the last period in the panel)
2. Retirement defined according to Definition B

Table 5. TRANSITIONS IN HEALTH, STRESS AND DWELLING TENURE (RETIREMENT DEFINITION B)

A) SELF-ASSESSED HEALTH STATUS (Percentages)

Health in Year before Retirement	Health in Year after Retirement: Excellent/Very Good/Good	Health in Year after Retirement: Fair/Poor	All
Excellent/Very Good/Good	88.2	11.8	83.9
Fair/Poor	43.5	56.5	16.1
All	81.0	19.0	

B) SELF-REPORTED LEVEL OF STRESS (Percentages)

Stress in Year before Retirement	Stress in Year after Retirement: Very Stressed/ Somewhat Stressed	Stress in Year after Retirement: Not Very Stressed/ Not Stressed	All
Very Stressed/Somewhat Stressed	48.3	51.7	59.3
Not Very Stressed/Not Stressed	18.3	81.7	40.7
All	36.1	63.9	

B) DWELLING TENURE
(Percentages)

Dwelling Tenure in Year before Retirement	Dwelling Tenure in Year after Retirement		
	Owner	Renter	All
Owner	97.5	2.5	86.3
Renter	17.7	82.3	13.7
All	86.5	13.5	

Table 6. SUMMARY STATISTICS FOR THOSE WHO RETIRED (RETIREMENT DEFINITION B)

A) YEAR BEFORE RETIREMENT

Characteristic	Variable Category	Mean	Standard Deviation
Self-Assessed Health:	Excellent [a]	0.23	0.42
	Very good	0.36	0.48
	Good	0.25	0.43
	Fair	0.09	0.28
	Poor	0.08	0.26
Self-Reported Stress	Very stressed [a]	0.17	0.37
	Somewhat stressed	0.42	0.49
	Not very stressed	0.27	0.44
	Not stressed	0.14	0.34
Dwelling Tenure:	Renter [a]	0.14	0.34
	Owner	0.86	0.34
Labour Force Status (LFS):	Employed all year [a]	0.54	0.50
	Unemployed all year	0.07	0.25
	Not in the labour force all year	0.15	0.36
	Unemployed or not in the labour force part-year	0.24	0.43
Major Activity	Working at a job or business [a]	0.73	0.44
	Looking for work	0.10	0.29
	Long-term illness	0.17	0.38
	Retired	0.00	0.00
	Other	0.00	0.00
Education:	Less than high school (LTHS) [a]	0.26	0.44
	High school graduate	0.17	0.38
	Post-secondary non-university certificate	0.38	0.49
	University graduate	0.18	0.39
Age		55.33	3.89
Marital Status:	Married/Common law [a]	0.82	0.38
	Single	0.18	0.38
Disability:	No [a]	0.70	0.46
	Yes	0.30	0.46

Continues on next page

(1)	(2)	(3)	(4)
Self-employed:	No [a]	0.85	0.36
	Yes	0.15	0.36
Immigrant:	No [a]	0.84	0.36
	Yes	0.16	0.36
Pension Plan with Current Employer:	No [a]	0.12	0.32
	Yes	0.53	0.50
	Do not know/Not applicable	0.35	0.48
Family Composition:	Unattached individual in one person household [a]	0.13	0.33
	Married or common-law couple, no children	0.45	0.50
	Married or common-law couple with children (all children under age 25)	0.24	0.42
	Other	0.19	0.39
Investment Income:	Personal	2279.63	9217.31
	Family	3353.56	10381.48

Notes:
a. Reference category

B) YEAR AFTER RETIREMENT

Characteristic	Variable Category	Mean	Standard Deviation
Self-Assessed Health:	Excellent	0.20	0.40
	Very good	0.36	0.48
	Good	0.25	0.43
	Fair	0.11	0.32
	Poor	0.08	0.27
Self-Reported Stress	Very stressed	0.07	0.26
	Somewhat stressed	0.29	0.45
	Not very stressed	0.39	0.49
	Not stressed	0.25	0.43
Dwelling Tenure:	Renter	0.13	0.34
	Owner	0.87	0.34
Labour Force Status:	Employed all year	0.11	0.32
	Unemployed all year	0.06	0.23
	Not in the labour force all year	0.75	0.44
	Unemployed or not in the labour force part-year	0.08	0.28
Major Activity	Working at a job or business	0.00	0.00
	Looking for work	0.00	0.00
	Long-term illness	0.00	0.00
	Retired	0.89	0.31
	Other	0.11	0.31
Education:	Less than high school	0.26	0.44
	High school graduate	0.17	0.38
	Post-secondary non-university certificate	0.38	0.49
	University graduate	0.18	0.39
Age		57.33	3.89
Marital Status:	Married/Common law	0.84	0.37
	Single	0.16	0.37
Disability:	No	0.64	0.48
	Yes	0.36	0.48

Continues on next page

(1)	(2)	(3)	(4)
Self-employed:	No	0.87	0.33
	Yes	0.13	0.33
Immigrant:	No	0.84	0.37
	Yes	0.16	0.37
Pension Plan with Current Employer:	No	0.09	0.28
	Yes	0.03	0.18
	Do not know/Not applicable	0.88	0.32
Family Composition:	Unattached individual in one person household	0.15	0.35
	Married or common-law couple, no children	0.54	0.50
	Married or common-law couple with children (all children under age 25)	0.17	0.37
	Other	0.15	0.36
Investment Income:	Personal	3215.49	21748.22
	Family	4387.25	22462.95

Table 7. HEALTH MODEL ESTIMATES; SELF-REPORTED RETIREMENT

Pre-retirement Characteristic	Category	Retirement Definition: Definition A (1)	Definition A (2)	Definition B (1)	Definition B (2)
Health:	Very Good	0.562***	0.566***	0.597***	0.605***
		(0.156)	(0.156)	(0.148)	(0.148)
	Good	1.340***	1.346***	1.212***	1.230***
		(0.171)	(0.170)	(0.159)	(0.159)
	Fair	1.651***	1.670***	1.527***	1.538***
		(0.236)	(0.231)	(0.220)	(0.215)
	Poor	2.873***	2.910***	2.346***	2.375***
		(0.308)	(0.296)	(0.284)	(0.276)
Major Activity:	Looking for Work	0.063		0.113	
		(0.229)		(0.189)	
	Other Non-Labour Activity	-0.048			
		(0.227)			
	Long-Term Illness	0.163		0.188	
		(0.213)		(0.193)	
LFS:	Unemployed All Year		0.058		0.147
			(0.269)		(0.232)
	Not in the Labour Force All Year		0.105		0.211
			(0.176)		(0.187)
	Unemployed or Not in the Labour Force Part-Year		-0.052		-0.051
			(0.133)		(0.125)
Dwelling Tenure:	Owner	-0.442**	-0.420**	-0.329*	-0.333**
		(0.188)	(0.185)	(0.172)	(0.169)
Disability:	Yes	0.443***	0.441***	0.556***	0.540***
		(0.147)	(0.145)	(0.140)	(0.140)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. The health model is estimated using an ordered probit specification

Table 8. STRESS MODEL ESTIMATES; SELF-REPORTED RETIREMENT

Pre-retirement Characteristic	Category	Retirement Definition: Definition A (1)	Definition A (2)	Definition B (1)	Definition B (2)
Stress:	Somewhat Stressed	0.004	0.032	0.106	0.123
		(0.169)	(0.169)	(0.158)	(0.158)
	Not Very Stressed	0.381**	0.411**	0.548***	0.554***
		(0.187)	(0.191)	(0.174)	(0.177)
	Not Stressed	0.854***	0.910***	1.098***	1.137***
		(0.217)	(0.223)	(0.211)	(0.218)
Health:	Fair (versus Excellent)	-0.047	-0.089	-0.104	
		(0.235)	(0.230)	(0.221)	
	Poor (versus Excellent)	-0.696**	-0.762***	-0.294	
		(0.295)	(0.284)	(0.278)	
Major Activity:	Looking for Work	-0.329		-0.371*	
		(0.238)		(0.197)	
	Other Non-Labour Activity	-0.387*			
		(0.230)			
	Long-Term Illness	-0.240		-0.308	
		(0.217)		(0.198)	
LFS:	Unemployed All Year		-0.550*		-0.330
			(0.283)		(0.253)
	Not in the Labour Force All Year		-0.209		-0.306
			(0.180)		(0.191)
	Unemployed or Not in the Labour Force Part-Year		-0.035		-0.089
			(0.143)		(0.134)
Dwelling Tenure:	Owner	0.427**	0.421**	0.294	0.300*
		(0.191)	(0.190)	(0.179)	(0.174)
Disability:	Yes	-0.328**	-0.328**	-0.413***	-0.429***
		(0.147)	(0.146)	(0.141)	(0.141)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. The stress model is estimated using an ordered probit specification

Table 9. DWELLING TENURE MODEL ESTIMATES; SELF-REPORTED RETIREMENT

Pre-retirement Characteristic	Category	Retirement Definition: Definition A (1)	Definition A (2)	Definition B (1)	Definition B (2)
Dwelling Tenure:	Owner	2.755***	2.831***	2.711***	2.920***
		(0.323)	(0.334)	(0.302)	(0.322)
Health:	Very Good	0.347	0.205	0.393	0.275
		(0.399)	(0.406)	(0.396)	(0.408)
	Good	0.492	0.192	0.217	0.005
		(0.416)	(0.407)	(0.379)	(0.387)
	Fair	0.029	-0.273	-0.389	-0.582
		(0.489)	(0.488)	(0.444)	(0.452)
	Poor	0.128	-0.190	-0.132	-0.303
		(0.572)	(0.562)	(0.554)	(0.544)
Major Activity:	Looking for Work	-0.905**		-0.678*	
		(0.420)		(0.347)	
	Other Non-Labour Activity	-0.374			
		(0.463)			
	Long-Term Illness	-0.473		-0.128	
		(0.403)		(0.363)	
LFS:	Unemployed All Year		-0.511		-0.611
			(0.500)		(0.443)
	Not in the Labour Force All Year		-0.141		-0.078
			(0.372)		(0.378)
	Unemployed or Not in the Labour Force Part-Year		0.316		0.337
			(0.382)		(0.358)
Disability:	Yes	-0.274	-0.290	-0.315	-0.214
		(0.330)	(0.325)	(0.310)	(0.321)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. The dwelling tenure model is estimated using a probit specification

Table 10. HEALTH MODEL ESTIMATES; RETIREMENT BASED ON ANNUAL LABOUR FORCE STATUS

Pre-retirement Characteristic	Category	Retirement Definition: Definition C (1)	Definition C (2)	Definition D (1)	Definition D (2)
Health:	Very Good	0.672***	0.698***	0.795***	0.786***
		(0.178)	(0.178)	(0.216)	(0.215)
	Good	0.966***	0.983***	1.315***	1.320***
		(0.189)	(0.188)	(0.238)	(0.235)
	Fair	1.655***	1.749***	1.212***	1.155***
		(0.235)	(0.227)	(0.425)	(0.398)
	Poor	2.508***	2.564***	2.112***	1.908***
		(0.272)	(0.263)	(0.533)	(0.520)
Major Activity:	Looking for Work	-0.442*		-0.459	
		(0.249)		(0.477)	
	Other Non-Labour Activity	-0.054		-0.506	
		(0.231)		(0.394)	
	Retired	-0.282**		-0.010	
		(0.142)		(0.199)	
	Long-Term Illness	-0.010		0.042	
		(0.207)		(0.511)	
LFS:	Unemployed All Year		-0.687***		-0.993**
			(0.193)		(0.387)
	Not in the Labour Force All Year				0.305
					(0.432)
	Unemployed or Not in the Labour Force Part-Year		-0.428***		-0.416
			(0.155)		(0.300)
Dwelling Tenure:	Owner	0.028	-0.002	-0.028	0.060
		(0.175)	(0.176)	(0.290)	(0.295)
Disability:	Yes	0.369**	0.392***	0.236	-0.311
		(0.147)	(0.146)	(0.202)	(0.322)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. The health model is estimated using an ordered probit specification

Table 11. STRESS MODEL ESTIMATES; RETIREMENT BASED ON ANNUAL LABOUR FORCE STATUS

		Retirement Definition			
Pre-retirement		Definition C		Definition D	
Characteristic	Category	(1)	(2)	(1)	(2)
Stress:	Somewhat Stressed	0.551***	0.599***	0.344	0.380
		(0.198)	(0.192)	(0.393)	(0.380)
	Not Very Stressed	0.817***	0.876***	0.650*	0.715*
		(0.207)	(0.202)	(0.379)	(0.368)
	Not Stressed	1.229***	1.238***	1.211***	1.368***
		(0.229)	(0.224)	(0.402)	(0.390)
Health:	Fair (versus Excellent)	-0.138	-0.359	-0.764	-0.580
		(0.242)	(0.225)	(0.484)	(0.408)
	Poor (versus Excellent)	-0.235	-0.464*	0.212	0.420
		(0.271)	(0.253)	(0.551)	(0.516)
Major Activity:	Looking for Work	-0.304		1.453**	
		(0.278)		(0.694)	
	Other Non-Labour Activity	-0.745***		-0.219	
		(0.249)		(0.471)	
	Retired	-0.338**		-0.231	
		(0.150)		(0.209)	
	Long-Term Illness	-0.200		0.213	
		(0.229)		(0.661)	
LFS:	Unemployed All Year		-0.261		-0.087
			(0.198)		(0.393)
	Not in the Labour Force All Year				-0.624
					(0.438)
	Unemployed or Not in the Labour Force Part-Year		0.133		-0.328
			(0.155)		(0.310)
Dwelling Tenure:	Owner	-0.129	-0.177	-0.396	-0.351
		(0.185)	(0.181)	(0.318)	(0.304)
Disability:	Yes	-0.162	-0.247*	-0.036	0.063
		(0.152)	(0.149)	(0.208)	(0.340)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. The stress model is estimated using an ordered probit specification

Table 12. DWELLING TENURE MODEL ESTIMATES; RETIREMENT DEFINITION D

Characteristic	Category	Dwelling Model (1)	(2)
Dwelling Tenure:	Owner	6.091***	5.108***
		(1.759)	(1.139)
Health:	Fair (versus Excellent)	0.858	-0.838
		(1.724)	(1.282)
	Poor (versus Excellent)	1.003	-1.675
		(1.773)	(1.444)
Major Activity:	Looking for Work	-2.005	
		(1.504)	
	Other Non-Labour Activity	-1.668	
		(1.603)	
	Retired	-0.732	
		(0.897)	
	Long-Term Illness	-2.019	
		(2.220)	
LFS:	Unemployed All Year		-0.245
			(0.877)
	Not in the Labour Force All Year		0.728
			(1.000)
	Unemployed or Not in the Labour Force Part-Year		0.935
			(0.844)
Disability:	Yes	-0.826	-0.350
		(1.045)	(0.898)

Notes:
1. Standard errors in parentheses; significance levels: *** p<0.01, ** p<0.05, * p<0.1
2. The dwelling ownership model is estimated using a probit specification

Figure 1. Effects of Major Activity on Latent Health

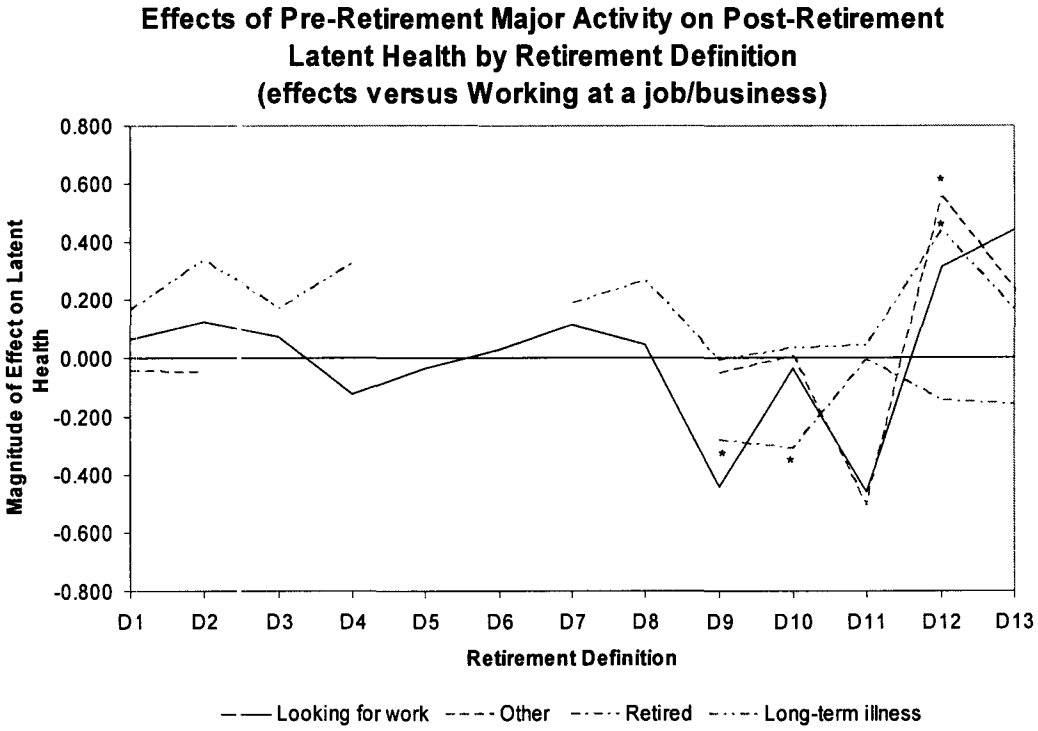


Figure 2. Effects of Dwelling Tenure and Disability on Latent Health

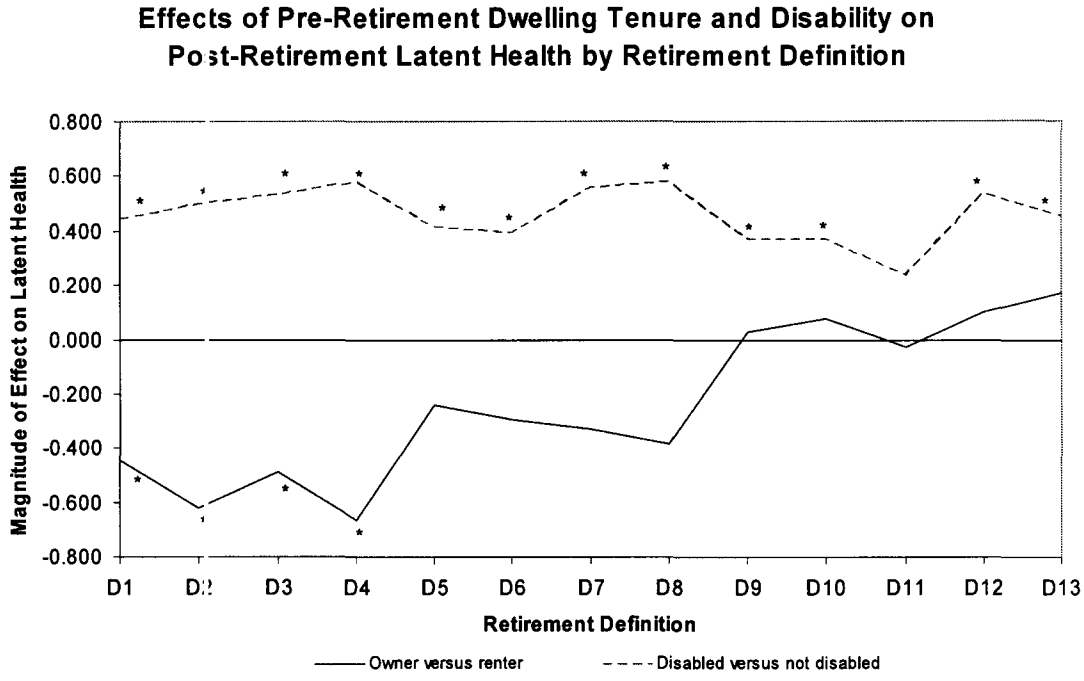


Figure 3. Effects of Annual Labour Force Status on Latent Health

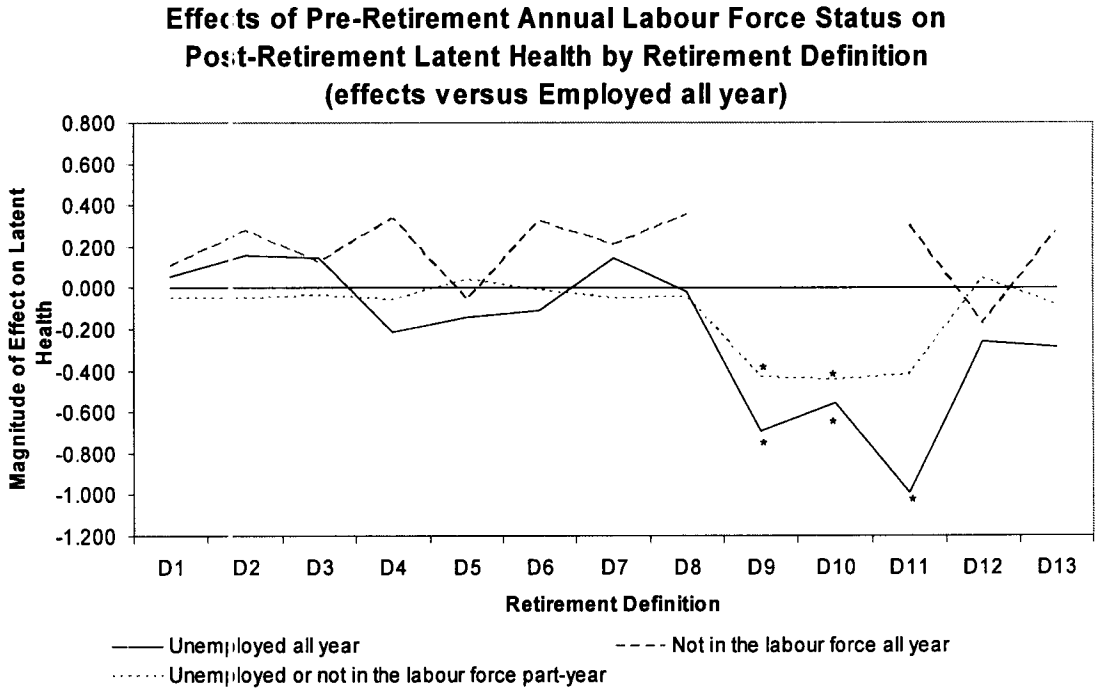


Figure 4. Effects of Educational Attainment on Latent Health

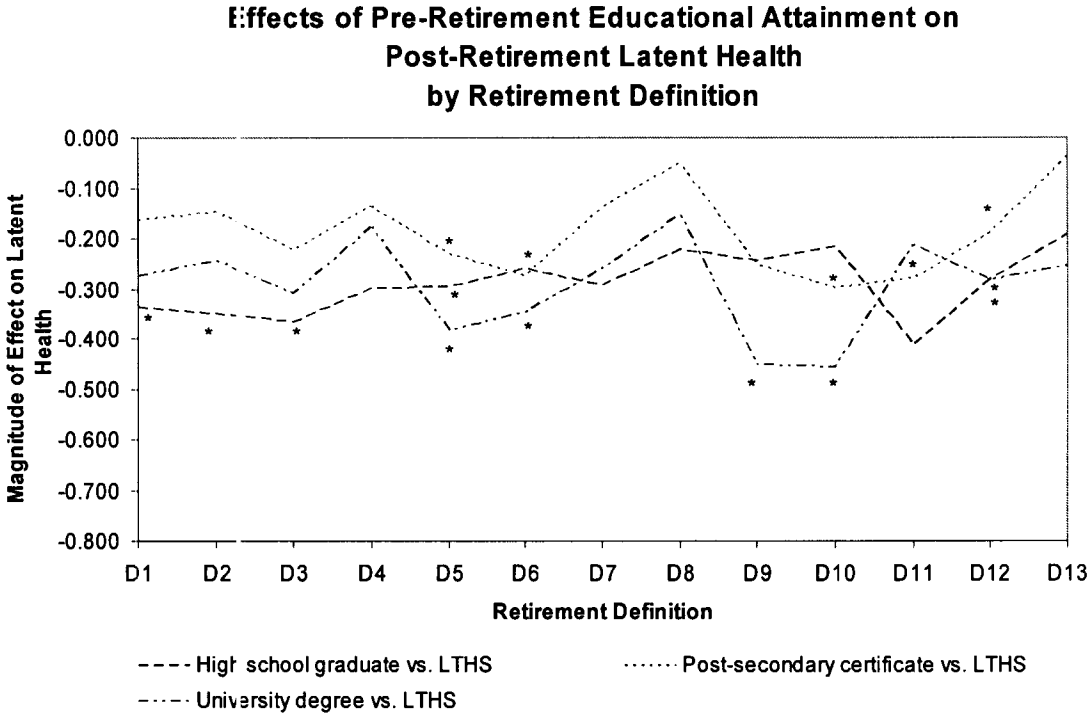


Figure 5. Effects of Major Activity on Latent Stress

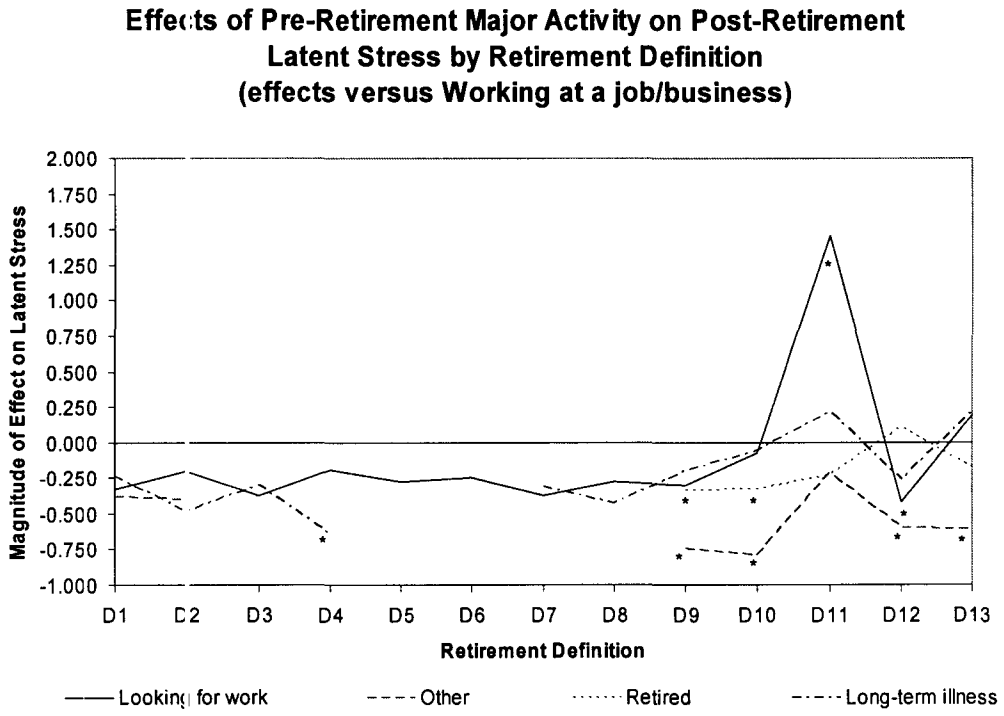


Figure 6. Effects of Annual Labour Force Status on Latent Stress

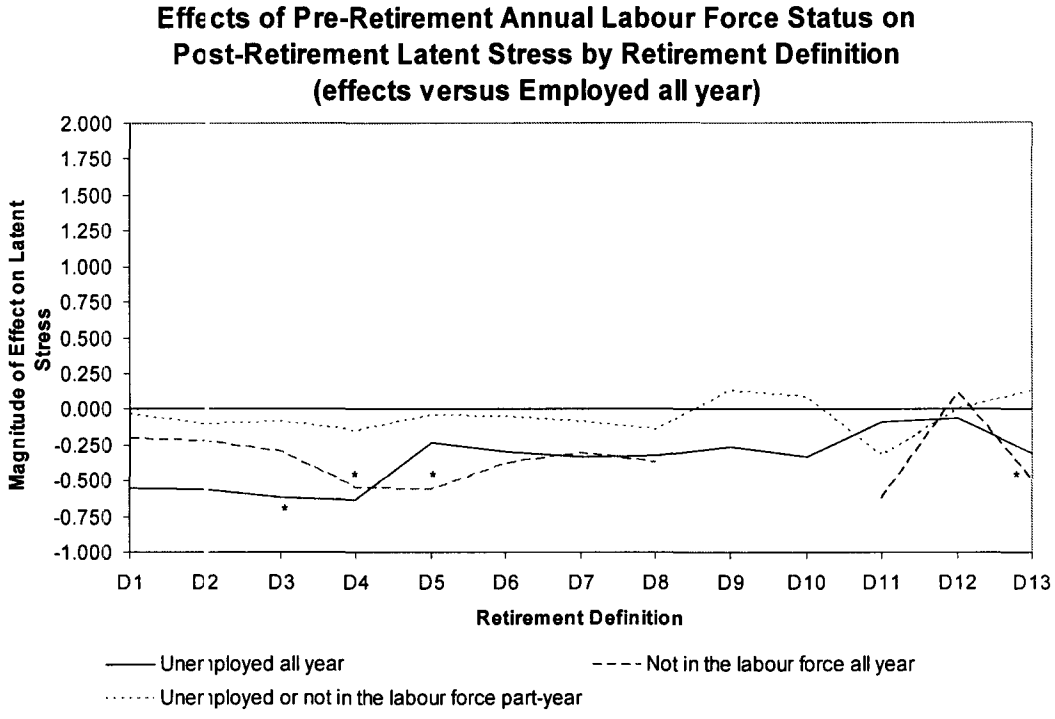
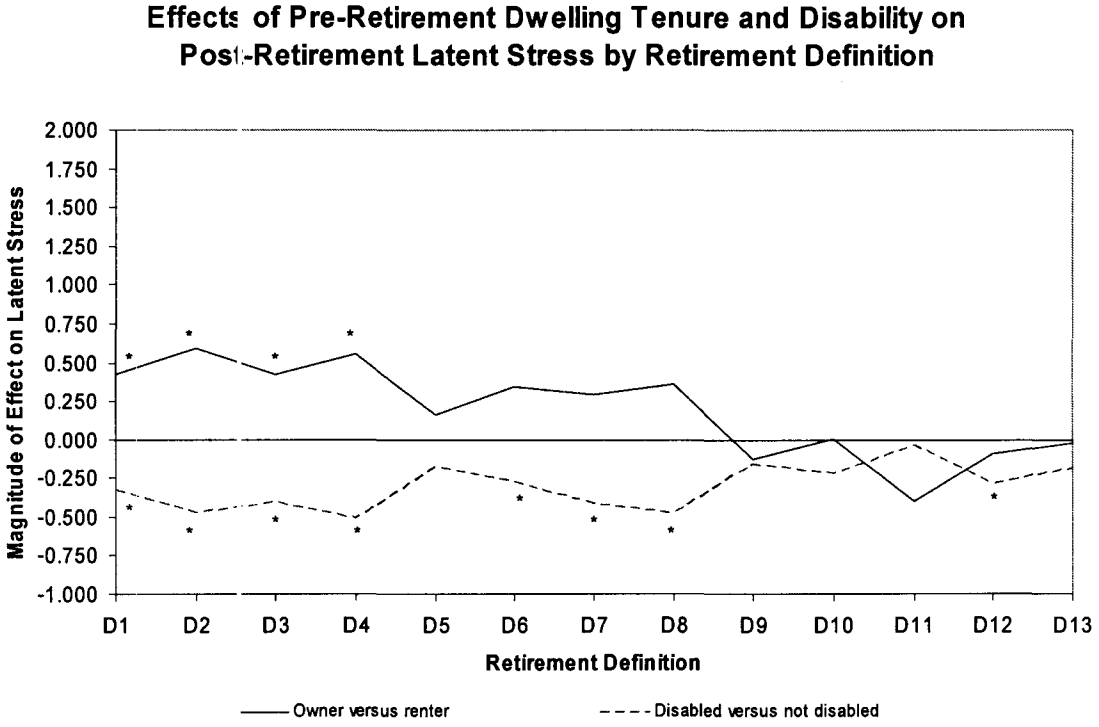


Figure 7. Effects of Dwelling Tenure and Disability on Latent Stress



V

Conclusion

This thesis is composed of an introduction and three independent chapters. The first chapter after the introduction investigates whether self-reported general stress is a mediator in the relationship between socio-economic status (SES) and health. The second studies whether self-assessed health status contains information about future mortality and morbidity. The third explores whether differences in early retirement pathways are associated with differences in post-retirement outcomes.

The first chapter after the introduction contributes to the literature on the pathways in the SES-health relationship. I find little evidence that general stress is a pathway from SES to health. While my results suggest a strong negative association between stress and health for both men and women, they provide little support to the hypothesis of a significant effect of income on stress, consistent with the direction of the SES-health gradient.

The second chapter after the introduction contributes to the literature on the predictive power of SAH for future mortality and morbidity. It also adds to the literature on adverse selection in the market for annuities. Our findings show that SAH does contain private information for future mortality and morbidity, information that increases with age. Hence, a shift from defined benefit to defined contribution pension arrangements may exacerbate the adverse selection in the market of annuities, especially

at older ages. That would make it more difficult for older individuals to insure longevity risk.

The last chapter of the thesis before this conclusion makes a contribution to the literature on early retirement. It shows that differences in pre-retirement health indicators (such as self-assessed health and disability) are likely to be associated with differences in post-retirement health and stress. These findings imply that "involuntary" retirees (men who may have retired because of health and/or health related conditions) are more likely to experience worse post-retirement outcomes (in terms of health and stress) than men who retire "voluntarily". In addition, retirement circumstances are found to have no statistically significant effect on dwelling ownership. That suggests that observable early retirement circumstances do not make men more likely to substitute dwelling ownership for rentership.