DRAFT: May 2, 2008

The Impact of Health on Labour Supply over the Life Cycle

Derek Hum*
Wayne Simpson*
Eyob Fissuh*

* Department of Economics, the University of Manitoba, Winnipeg MB R3T 5V5 CANADA. Corresponding author: Wayne Simpson (simpson@ms.umanitoba.ca; (204) 474-9274). The authors would like to acknowledge financial assistance from the Prairie Centre for Excellence in Research on Immigration and Integration (now the Prairie Metropolis Centre for Research on Immigration and Integration) and the support of the University of Manitoba Research Data Centre. We are responsible for any errors, omissions and opinions.

The Impact of Health on Labour Supply over the Life Cycle

Abstract

Estimates of the impact of health on labour supply to quantify the returns to public and private investment in health vary widely for older men and are scarce for other groups. In addition, estimates using U.S. data confound differences in health insurance status that are correlated with both health and labour market outcomes. We provide new and more comprehensive estimates through an empirically tractable life cycle model of labour supply that also reflects the accumulation of health stock and estimate the model in the Canadian context of universal health insurance coverage. Our model identifies two important sources of bias in the estimation of the impact of health on labour supply: one arising from endogenous health status, and the other arising from unobserved effects or heterogeneity. Our model also identifies the econometric issues associated with the presence of nonemployment in life cycle labour supply behaviour. Using recent panel data from the Survey of Labour and Income Dynamics, we provide basic evidence on the links between health status and labour supply over the life cycle as well as econometric estimates from appropriate nonlinear models that account for the presence of unobserved (life cycle) effects and endogenous health outcomes. Our results suggest that informal and cross-sectional evidence of a large health impact on labour supply are overstated because of endogenous health outcomes and the presence of unobserved effects, particularly for the widely studied group of older men. We estimate the health impact on labour supply to be on the order of only 4-6% for all age groups when endogeneity and unobserved effects are considered.

Subject Categories: Labour Economics (LH16), Quantitative Methods (QM6)

1. Introduction

Individuals invest in health as part of their human capital. Governments also invest heavily in health care and, to a lesser extent, in the promotion of healthy lifestyles. Implicit in these investments is the idea that there exist individual and social benefits that can be identified and, ideally, quantified. One such benefit should be superior labour market outcomes, since healthier workers are expected to be more productive or able to work longer and more intensively. Thus, individuals invest in health to further accumulation of lifetime wealth through longer and more productive work careers, and this augmented productivity in turn contributes a social benefit to justify public health programs. At the same time, an illness or accident could damage health capital and impair a work career. Where such adverse events are avoidable, there is the issue of the benefits to prevention and, in legal cases, the matter of assessing foregone employment opportunities as an important component of the claim. In all of these instances, the impact of health on the supply of market labour is important and provides motivation for our research.

Substantial research already exists on the effects of health on labour supply. Currie and Madrian's (1999) extensive survey reports a wide variety of estimates in the U.S alone, using different estimators and different measures of health and work activity. They attribute much of the variation in estimates to whether health is treated as endogenous and, if so, the instruments used. They especially note that "a glaring limitation of the existing literature is the intense focus on elderly white men, to the virtual exclusion of most other groups." Although the focus on health and labour supply of older men and the attendant retirement decision has implications for public pension policy, there is a clear need for more general estimates of the impact of health on work activity across the wider age spectrum to address other important issues.

This paper starts from the perspective of labour supply over the life cycle. We argue that current empirical research on the effect of health on labour supply suffers not only from the aforementioned limitations, but also from its focus on estimates from cross-sectional data since these estimates are known to be biased in the presence of unobservable person-specific fixed effects or heterogeneity. We therefore first develop an empirically tractable life cycle model of labour supply which illustrates this bias --- a bias that can be removed with panel data and appropriate econometric procedures. We then extend this model to capture the effects of accumulation of health stock over the life cycle and to formalize the case for treating health as endogenous. The model leads us to consider several econometric issues associated with the estimation of a nonlinear life cycle labour supply model with unobserved effects and endogenous health outcomes.

One complication that arises is the strong link between health insurance and labour supply when using U.S. data, which then requires additional data and model specifications (Currie and Madrian, 3353). Our Canadian data avoids this problem because of the presence of universal health insurance.[1] Our data source is the Canadian Survey of Labour and Income Dynamics panel, which now includes measures of health status. We provide basic evidence on the links between health status and labour supply over the life cycle in Section 2. Section 3 then develops a model of life cycle labour supply and health stock accumulation to establish the foundation for our econometric analysis. Section 4 discusses estimation issues and presents our empirical strategy. Section 5 presents our results and, in particular, contrasts those obtained from cross-sectional evidence and

[1] Differences in health coverage across provinces and the scope for private health insurance coverage are severely

panel data. Section 6 provides concluding remarks.

2. Data and Descriptive Evidence

The Survey of Labour and Income Dynamics (SLID) is a continuing panel of Canadian households begun in 1993. It combines the former Labour Force Activity Survey, an intermittent series of panel surveys conducted during the 1980s, with the annual cross-sectional Survey of Consumer Finance. The SLID design is a series of overlapping 6-year panels, with a new panel enrolled every three years. Starting with the second panel, enrolled in 1996, SLID collected self-assessed health information by asking whether respondents evaluated their personal health as excellent, very good, good, fair, or poor. The potential inaccuracies and biases of self-reported health measures are well known and have been widely discussed. Nonetheless, such measures are used regularly to analyze the impact of health on labour market activity.[2] We employ the first completed panel that included the self-assessed health question, which is the second panel of respondents from 1996 to 2001.[3] This internal SLID file provides longitudinal labour market activity information such as details on hours worked, wages and pay structures throughout each survey period (calendar year) that is rarely available in conjunction with health information.

Our focus is the impact of health on labour supply, most broadly defined as total hours worked. Average total hours worked in the sample depends on positive hours worked, conditional on employment (that is, an individual working some hours), and the employment rate (whether an individual works at all). That is, suppose respondent i in a sample of size m works n_i hours, which may be zero or positive. Then total hours worked would be

$$\sum_{i=1}^{m} n_i = \sum_{i=1}^{m} n_i \pi_i = \sum_{n_i>0} n_i\big|_{\pi_i=1} + \sum_{n_i=0} n_i\big|_{\pi_i=0} \qquad [1]$$

where $\pi_i \equiv 1(n_i > 0)$ indicates employment in the sense of positive hours worked. Changes in total hours worked, which might arise from health problems, therefore depend on changes in positive

---

limited by Canada's National Health Act and are ignored in our analysis.

[2] Banks, Crossley and Goshev (2007) have recently shown, for example, that self-assessed health status is a useful predictor of future mortality and morbidity in Canada even after accounting for standard observable characteristics, including diagnosed medical conditions. For an extensive discussion of the use of self-assessed health measures in the analysis of labour market performance, see Currie and Madrian (1999, 3313-18). We would only note that, since the measurement errors between any objective measure of health and self-reported health may be related to labour market activity (i.e. respondents may justify inactivity with biased reports of personal health), the need to treat health as endogenous is likely more acute when a self-reported measure of health status is used. Our results in this section treat health status as exogenous, but we address the issue of endogeneity bias in our econometric analysis below.

[3] Access to the internal file is provided only at specified research data centres and analysis is limited to procedures that

hours worked as well as changes in employment. We begin with the simplest health effect consisting of two states, "good, very good, or excellent health" (state 1) and "fair to poor health" (state 0). To assess the impact of health on labour supply in these two states, we compare individual hours worked in state 1, n_i^1, to individual hours worked in state 0, n_i^0:

$$\sum_{i=1}^{m} n_i^1 - \sum_{i=1}^{m} n_i^0 = \sum_{i=1}^{m} n_i^1 \pi_i^1 - \sum_{i=1}^{m} n_i^0 \pi_i^0 = \sum_{i=1}^{m} [n_i^1 - n_i^0]\pi_i^1 + \sum_{i=1}^{m} n_i^0 [\pi_i^1 - \pi_i^0] \qquad [2]$$

This is a standard decomposition in which the effect of a change in health status on total hours depends on the change in positive hours worked determined by employment in state 1 plus the change in employment status weighted by positive hours worked in state 0.[4] The mean health effect correspondingly depends on the mean change in positive hours worked and the change in the employment rate, appropriately weighted. Studies which restrict analysis to employment alone, the second term on the right hand side of equation [2], consequently miss potentially important health effects on positive hours worked.[5]

How large are the health impacts on labour supply described in equation [2]? This section gives a first assessment by ignoring problems associated with measuring health status, and by solely considering age as a conditioning factor. Our "back of the envelope" estimate of these gaps can be obtained from Table 1, which reports mean total and positive hours worked and mean employment rates for all men 21-65. These men worked an average of 1,780.9 total hours in 2001, reflecting an employment rate of 87.4% and an average of 2,063.7 hours for each man working positive hours. The gap in total hours worked for all men 21-65 is 644 hours, or 34.7% of the total hours of men with good health. This consists of a gap of 171 positive hours worked, or 8.2% of positive hours worked of men with good health, and a gap of 25.9% in employment rates. From equation [2], the gap in employment rates, evaluated at mean positive hours worked for those in poor health, is 493.7

ensure the confidentiality of individual respondents according to criteria established by Statistics Canada.

[4] A similar decomposition of the tobit model into the effect on the probability of being at the limit and the effect for those above the limit is provided by McDonald and Moffitt (1980). Our empirical results depend on the usual identifying assumption that we can create an appropriate comparison group of individuals whose labour market activity would have been the same as the group with good health had they not had good health (i.e. had they had poor health).

[5] As we discuss in section 2, this includes a large number of studies of retirement among older men and a more limited number of studies of participation rates. In this paper we focus on the employment rate (whether an individual has worked positive hours or not) rather than the participation rate (whether an individual is working or seeking employment), but the difference between employment and participation rates (the unemployment rate) will be small for longer survey periods, such as the calendar year in the SLID.

hours, or 76.7% of the total gap of 644 hours.[6] This suggests that roughly three-quarters of the impact of health on labour supply is captured by changes in employment rates.

This approach can also be used to compare middle-aged and older men in the sample. For men aged 35-50, the gap is 693.7 total hours worked, or 33.3% of the total hours of workers with good health. The gap in employment rates, evaluated at mean positive hours worked for those in poor health, is 526.2 hours or 75.8% of the gap in total hours worked. For men aged 51-65, the gap is 623.4 total hours worked, or 39.3% of the total hours of workers with good health. The gap in employment rates, evaluated at mean positive hours worked for those in poor health, is 490.5 hours or 78.7% of the gap in total hours worked. In other words, these simple calculations indicate three results for further analysis. First, there are substantial differences in labour supply for men with good and poor health, on the order of one-third of the total labour supply of men with good health. Second, the aggregate differences based on reported health state appear to be similar for middle-aged men and older men, providing no basis in the literature to focus exclusively on older men and to exclude middle-aged men. Third, about three-quarters of the difference can be attributed to differences in employment rates in both groups, implying that appropriate (nonlinear) estimation methods must be used to account for this employment decision.

Panel data offer a dynamic perspective of health and labour supply. This perspective is important insofar as apparent differences in labour supply by health status may arise for other unobserved reasons. If so, then changes in labour supply by health status may provide a more accurate assessment of the impact of health. The final column of Table 1 allows us to calculate a rough "difference in differences" (hereafter DID) estimate of the effect of health for all men aged 21 to 65, and for men 35 to 50, and 51 to 65. All men aged 21-65 with good health increased total hours worked from 1996 to 2001 by 124.6 hours, while those with poor health decreased total hours worked by 164.5 hours, yielding a DID estimate of the effect of health on hours worked of 289.1 hours per year, an estimated health impact of 15.6% of the hours worked by those in good health that is less than half the estimated gap of 644 hours (34.7%) derived from comparing total hours worked for each group. These estimates do not account for other relevant factors and, in particular, the fact that men in poor health will be older than men in good health, since health status declines with age. For men aged 35 to 50, the DID estimate of the health effect is 189.6 hours, less than

[6] The sum of the two gaps, 664.7 hours (171+493.7), is only approximate. One reason is that the gap in positive hours is not evaluated at the employment state for workers in good health. This matters in individual cases where the employment of workers in poor health rises relative to those in good health.

one-third of the estimated gap of 693.7 hours from differences in total hours worked. For men 51 to 65, the DID estimate is 173.6 hours, again less than one-third of the estimated gap of 623.4 hours from differences in the level of hours between the two groups.

An additional dynamic perspective is obtained by comparing changes in total hours worked between 1996 and 2001 with changes in health status, recognizing that the onset of poor health may have a more important impact on labour supply than ongoing health concerns. From Table 1, we can obtain rough "fixed effect" (hereafter FE) estimates of the impact of a unit health change (from better to same health, or same health to worse health) on changes in total hours worked by calculating one-half the change in hours worked for those with better health and those with worse health. The FE estimated unit health effect is 104.8 hours per year for all men (5.9% of mean hours for this group), 210.8 hours for men 35 to 50 (10.5%), and 125.8 hours for men 51 to 65 (8.6%). These estimates of the impact of health are again clearly considerably smaller (by a factor of three or four) than our original cross-sectional estimates based on comparisons of the levels of hours worked by health status and, we would note, our estimated health effects for middle-aged men remain at, or above, those for older men.

This description of the evidence ignores a number of issues, such as the effect of factors other than health on labour supply outcomes and biases arising from the measurement of health status. We address these issues in the remainder of the paper.

3. Our Model

3.1. Life Cycle Labour Supply

There is a vast literature on life cycle labour supply models whereby individuals allocate time between leisure and work to maximize discounted lifetime utility. A convenient approach for empirical analysis assumes a utility function u(·) strongly separable in consumption and leisure that is maximized over a certain lifetime of T years[7]:

$$\sum_{t=0}^{T}(1+\rho)^{-t}[u_{1t}(c_t)+u_{2t}(l_t)] \qquad [3]$$

where ρ is the rate of time preference. This function is maximized subject to a lifetime budget constraint of the form

[7] An uncertain lifetime is more realistic, particularly in the context of a discussion of health effects, although Jakubson (1988,306-7) shows that uncertainty is inconsequential in his model if forecasting errors are zero. He includes time

$$\sum_{t=1}^{T}(1+\rho)^{-t}c_t \le y + \sum_{t=1}^{T}(1+\rho)^{-t}w_t(\bar{l}-l_t) \qquad [4]$$

where ρ is also assumed to be the real rate of interest for borrowing and lending,[8] and $\bar{l}$ is the maximum leisure available per period such that $\bar{l}-l$ represents hours worked at wage rate w in each time period.

We adopt Jakubson's (1988) tractable formulation that the utility of leisure takes the specific form

$$u_{2t}(l_t) = a_t\left[{(l_t)^{\delta}}/{\delta}\right] \qquad [5]$$

where a is a person-specific taste modifier and δ reflects preference for leisure.[9] This yields a first order condition for leisure that can be solved for an interior solution of the form:

$$\log\left(\bar{l}/l_t\right) = \log\bar{l} + \left(1/_{1-\delta}\right)[\log w_t - \log a_t + \log\lambda_t] \qquad [6]$$

A corner solution, defining whether one works positive hours or consumes only leisure, is given by the usual labour supply condition that

$$\log\bar{l}/l = 0 \ if \ \ w_t < \left.\frac{\partial u_{2t}/\partial l}{\partial u_{1t}/\partial c}\right|_{l=\bar{l}} = \frac{a_t\bar{l}^{\,\delta-1}(1+\rho)^{-t}}{\lambda_t(1+\rho)^{-t}} \ or$$
$$\log w_t - \log a_t - (\delta-1)\log\bar{l}_t + \log\lambda_t \le 0 \qquad [7]$$

This condition implies that the right hand side of equation [6] is negative or zero. Thus, we can write the labour supply equation as

$$\log\left(\bar{l}/l_t\right) = \begin{cases} \log\bar{l} + \left(1/_{1-\delta}\right)[\log w_t - \log a_t + \log\lambda_t] \text{ if RHS} > 0 \\ 0 \ if \ \log\bar{l} + \left(1/_{1-\delta}\right)[\log w_t - \log a_t + \log\lambda_t] \le 0 \end{cases} \qquad [8]$$

Labour supply and labour market participation thus depend on the wage rate, tastes, and the (unobserved) marginal utility of wealth.

Jakubson parameterizes the wage rate and the taste modifier as functions of observables x_t,

---

dummies to capture nonzero-mean forecast errors in a finite panel.

[8] The rates of time preference and interest may differ but, in any case, the rate of time preference is not observed. Jakubson (1988) captures any difference in these rates by an age term in his empirical work.

[9] That is, the marginal utility of leisure is $u_l=\alpha/l^{1-\delta}$, which is larger the larger is the value of δ.

which might include human capital, household composition, and geographical variables and, for our application, health status. Thus,

$$\begin{aligned} \log w_t &= x_t\gamma_1 + \varepsilon_{1t} \\ \log a_t &= x_t\gamma_2 + \varepsilon_{2t} \end{aligned} \qquad [9]$$

The marginal utility of wealth term, $\log \lambda_t$, is represented as a person-specific fixed term *c*, an unobserved effect, so that equation [8] can be written as

$$\log\left(\bar{l} \big/ l_t \right) = \begin{cases} \beta x_t + c + \varepsilon_t & \text{if RHS} > 0 \\ 0 \quad if \quad \beta x_t + c + \varepsilon_t \leq 0 \end{cases} \qquad [10]$$
$$\beta = \gamma_1/(1-\delta) - \gamma_2, \quad \varepsilon_t = \varepsilon_{1t}/(1-\delta) - \varepsilon_{2t}$$

Cross sectional analysis of a form of equation [10] will yield biased estimates --- including biased estimates of the effects of health on labour supply --- because unobserved effects *c* are correlated with the included observables x_t through the lifetime budget constraint. Jakubson finds, for example, that cross-sectional estimates of the effect of children on the labour supply of married women are much larger than estimates from panel data that control for unobserved effects. This invites suspicion that other control variables, such as health, may be similarly affected even if health is treated as endogenous. Indeed, our simple calculations in section 2 reporting that health impacts are much smaller when using changes in health and labour supply across the panel suggest this may be the case.

3.2. The Role of Health in Life Cycle Labour Supply

Treating health status as exogenous ignores the crucial role that health capital plays in life cycle behaviour, including labour supply --- an idea that derives from Grossman (1972). One simple way to illustrate this point in the context of a life cycle labour supply model is to specify utility in terms of healthy leisure days that are, in turn, determined by resources (time) invested in health. Let ν_t be the proportion of leisure days that are healthy and let the proportion of leisure days vary directly with the stock of health, q_t, which in turn is produced by the input of time, h_t to health-producing activities:

$$\nu_t = g_t^{\nu} q_t^{\gamma^{\nu}} \qquad [11]$$

$$q_t = g_t^{q} h_t^{\gamma^{q}} \qquad [12]$$

where g_t^{ν}, g_t^{q} are person-specific health modifiers, akin to tastes modifier a_t in equation [5], and γ^{ν}, γ^{q} are parameters representing the elasticity of healthy days with respect to the stock of health, and the stock of health with respect to health time inputs, respectively. From equations [11] and [12] we can write

$$\nu_t = g_t^{\nu} q_t^{\gamma^{\nu}} = g_t^{\nu} \left[g_t^{q} h_t^{\gamma^{q}} \right]^{\gamma^{\nu}} = g_t \left[h_t \right]^{\gamma} \qquad [13]$$

This allows us to respecify equation [5] as:

$$u_{2t}(l_t, h_t) = a_t \left[\frac{(l_t)^{\delta}}{\delta} \right] \nu_t = a_t \left[\frac{(l_t)^{\delta}}{\delta} \right] g_t h_t^{\gamma} \qquad [14]$$

We then modify equations [3] and [4] to reflect health accumulation; i.e., we maximize

$$\sum_{t=0}^{T} (1+\rho)^{-t} [u_{1t}(c_t) + u_{2t}(l_t, h_t)] \qquad [15]$$

subject to the lifetime budget constraint

$$\sum_{t=1}^{T} (1+\rho)^{-t} c_t \le y + \sum_{t=1}^{T} (1+\rho)^{-t} w_t (\bar{l} - l_t - h_t) \qquad [16]$$

such that the opportunity cost of health-improving activities is the wage rate. This formulation yields a straightforward modification of the labour supply conditions in equations [6] and [7]:

$$\begin{aligned} &\log\left(\bar{l} \big/ l_t \right) = \log \bar{l} + \left(1 \big/ 1-\delta \right) [\log w_t - \log a_t - \log g_t - \gamma \big/ \gamma^{q} \left(\log q_t - \log g_t^{q} \right) + \log \lambda_t] \\ &\log \bar{l} / l = 0 \ \mathit{if} \ \log w_t - \log a_t - \log g_t - \gamma \big/ \gamma^{q} \left(\log q_t - \log g_t^{q} \right) - (\delta - 1) \log \bar{l}_t + \log \lambda_t \le 0 \end{aligned} \qquad [17]$$

These conditions in turn generate a labour supply equation with health effects that is a significant but fairly straightforward modification of equation [8]:

$$\log\left(\bar{l} \big/ l_t \right) = \begin{cases} \log \bar{l} + \left(1 \big/ 1-\delta \right) [\log w_t - \log a_t - \log g_t - \gamma \big/ \gamma^{q} \left(\log q_t - \log g_t^{q} \right) + \log \lambda_t] \ \text{ if RHS} > 0 \\ 0 \ \mathit{if} \ \log \bar{l} + \left(1 \big/ 1-\delta \right) [\log w_t - \log a_t - \log g_t - \gamma \big/ \gamma^{q} \left(\log q_t - \log g_t^{q} \right) + \log \lambda_t] \le 0 \end{cases} \qquad [18]$$

Labour supply and labour market participation now also depend on the state of health and an additional person-specific health modifier, in addition to the usual wage rate, tastes, and the unobserved marginal utility of wealth. Moreover, the state of health at any age t is governed by a first-order condition of the form:

$$\log h_t = \left(\frac{1}{\gamma - 1} \right) [\log w_t - \log a_t - \delta \log l_t - \log g_t - \log \frac{\gamma}{\delta}] \qquad [19]$$

Since log l_t is given from [18] as

$$\log l_t = \begin{cases} \left(\frac{1}{\delta - 1}\right) [\log w_t - \log a_t - \log g_t - \gamma \log h_t + \log \lambda_t] \; if \; RHS > 0 \\ \log \bar{l} \; if \; RHS \leq 0 \end{cases} \qquad [20]$$

we can solve for a reduced form equation in either log h_t or q_t to construct an instrument:

$$\log h_t = \frac{1}{\gamma^q} \left[\log q_t - \log g_t^q\right] =$$

$$\begin{cases} \left(\frac{1}{\gamma - 1}\right)\left(\frac{1}{1 - \delta\gamma}\right)[\left(\frac{1}{1-\delta}\right)\log w_t - (1-\delta)\log a_t - \delta \log \lambda_t - (1-\delta)\log g_t \\ \qquad - \log \frac{\gamma}{\delta}] \; if \; RHS \, of \, [20] > 0 \\ \left(\frac{1}{\gamma - 1}\right)\left(\frac{1}{1 - \delta\gamma}\right)[\log w_t - \log a_t - \delta \log \bar{l} - \log g_t - \log \frac{\gamma}{\delta}] \; if \; RHS \, of \, [20] \leq 0 \end{cases} \qquad [21]$$

or

$$\log q_t = \begin{cases} \left(\frac{\gamma^q}{\gamma - 1}\right)\left(\frac{1}{1 - \delta\gamma}\right)[\left(\frac{1}{1-\delta}\right)\log w_t - (1-\delta)\log a_t - \delta \log \lambda_t - (1-\delta)\log g_t - \log \frac{\gamma}{\delta}] \\ \qquad + \log g_t^q \; if \; RHS \, of \, [20] > 0 \\ \left(\frac{\gamma^q}{\gamma - 1}\right)\left(\frac{1}{1 - \delta\gamma}\right)[\log w_t - \log a_t - \delta \log \bar{l} - \log g_t - \log \frac{\gamma}{\delta}] + \log g_t^q \; if \; RHS \, of \, [20] \leq 0 \end{cases} \qquad [22]$$

That is, the state of health is seen to depend also on labour supplied according to [19], and therefore cannot be treated as exogenous. In their extensive survey of the U.S. literature, Currie and Madrian (1999) argue that the bias from treating health as endogenous is potentially large but largely uninvestigated. Moreover, equation [22] tells us that a proper instrument for health will depend on both wages and observed socio-demographic variables that capture interpersonal differences in tastes and health.

3.3 The Role of Health Insurance

The model represented by equations [18] and [22] must also account for the effect of health insurance when health insurance coverage varies across individuals. Health insurance in the U.S. is primarily provided by employers for the working age population. Consequently, individual labour supply decisions affect health insurance coverage, which in turn affects health status through time and labour supply outcomes. The unobserved effect λ_t is then correlated not only with labour supply and health outcomes, but also with health insurance coverage. Currie and Madrian (1999, 3358-9) argue that health insurance coverage is likely to be correlated with any omitted variables in the labour supply equation, including the unobserved effect λ_t, thereby imparting bias to estimates of the effect of health on labour supply. They characterize the "potential for omitted variable bias [as] something that should be taken seriously." In these circumstances, Canada presents a potentially useful natural experiment that holds health insurance coverage constant across individuals so that the independent effect of health on labour supply can be estimated.

4. Estimation Issues

Estimation of life cycle labour supply models, such as equation [8] or [10], involves two complications: first, the presence of unobserved effects which are correlated with included regressors, and, second, a nonlinear model specification associated with a corner solution (no hours worked). When the model is linear, the appropriate estimation procedure to deal with unobserved effects is now well known and fairly straightforward when panel data is available for each individual over time. One may estimate the unobserved effects model (Hausman and Taylor, 1981) by either random or fixed effects methods. Fixed effects estimation is consistent even if the unobserved effects are correlated with regressors in the model and can be estimated by least squares estimation of either a transformed (differenced in time) model or a more computationally demanding model with dummy variables to identify each individual fixed effect. The latter approach is rarely needed because the individual fixed effects are typically nuisance parameters of little interest to researchers. Random effects can be estimated by generalized least squares but the estimates are only consistent if the unobserved effects are uncorrelated with the regressors, an hypothesis that can be tested against the fixed effect estimates.

The estimation procedure is more complicated when the model is nonlinear, as it must be in a labour supply model represented by equation [10]. We can rewrite [10] in standard econometric notation for individual i at time t ($i = 1, \ldots, N; t = 1, \ldots, T$) as:

$$y_{it} = \begin{cases} x_{it}\beta + c_i + \varepsilon_{it} & \text{if RHS} > 0 \\ 0 \ \ if & x_{it}\beta + c_i + \varepsilon_{it} \leq 0 \end{cases} = F(x_{it}\beta + c_i) + \varepsilon_{it} \qquad [23]$$

Function $F(\cdot)$ is a nonlinear transformation function (Davidson and MacKinnon, 2004, 453), such as a tobit specification. Random effects models can still be estimated but fixed effects models are considerably more problematic. In particular, the marginal effects we seek to estimate now depend on the fixed effect: $\frac{\partial E[y_{it}]}{\partial x_{it}} = F'(\beta x_{it} + c_i).\beta$. Thus, differencing of $E[y_t]$ no longer eliminates the fixed effect c_i, so the estimated marginal effects will be inconsistent. In the case of the conditional logit estimator, it is possible to condition the fixed effects out of the likelihood function and estimate β independently of c_i, and hence obtain consistent estimates of β, but this is not generally feasible for other discrete choice models, including other (unconditional or ordered) logit estimators, probit estimators, and the tobit estimators that are commonly used. Without a feasible and consistent fixed effect estimator, we are unable to assure that random effects estimates of β, or pooled estimates of β that ignore heterogeneity, are consistent.

4.1. Obtaining Consistent Estimates of β When Health in Exogenous

Recent literature suggests a number of methods to obtain consistent estimates of β for panel data when we cannot assume $E[c_i \mid x_{it}] = 0$ in models like equation [23]. A direct, but computationally demanding, solution to this problem is to use $N-1$ individual dummy variables to capture the fixed effects, as in Greene (2002). In addition to the computational problems associated with specifying and including $N-1$ additional regressors, we now have to estimate a nonlinear model using numerical optimization procedures (nonlinear least squares or maximum likelihood). Since we must estimate each $c_i\ (i = 1,\ldots,N)$ with at most T observations, where T is typically quite small, these estimates are not consistent, and since the estimate of β depends on the estimate of c_i, it will not be consistent either.[10] This incidental parameters problem imparts "large and persistent" bias to binary and ordered choice models in Greene's (2002) Monte Carlo experiments, although he finds the bias to be acceptably small for the tobit estimator we use in this paper.

An alternative approach is to treat equation [23] as a model of sample selection; that is, to

[10] That is, the asymptotic variance for the ML estimate of c_i is O[1/T] and, since T is fixed, the estimate is inconsistent.

estimate the model only for $y_{it} > 0$ and correct for the selection bias associated with dropping observations for which $y_{it} = 0$. For panel data, two-step estimators have been proposed that extend Heckman's (1979) approach of treating sample selection as an omitted variables problem that can be tested and corrected by residuals estimated from a sample selection model. Woolridge (1995) proposes the estimation of a pooled tobit model of equation [15][11] to yield residuals $\hat{\varepsilon}_{it}$ which are then included in a standard fixed effects model of the form:

$$\ddot{y}_{it} = \ddot{x}_{it}\beta + \rho\ddot{\varepsilon}_{it} + \omega_{it}, \quad \ddot{y}_{it} \equiv y_{it} - \sum_{r=1}^{T} s_{ir}y_{ir} \big/ T_i, \quad \ddot{x}_{it} \equiv x_{it} - \sum_{r=1}^{T} s_{ir}x_{ir} \big/ T_i,$$
$$\ddot{\varepsilon}_{it} \equiv \hat{\varepsilon}_{it} - \sum_{r=1}^{T} s_{it}\hat{\varepsilon}_{it} \big/ T_i, \quad T_i \equiv \sum_{t=1}^{T} s_{it} \qquad [24]$$

Estimation by least squares permits a t-test for selection bias ($\rho \neq 0$) with robust standard errors and provides consistent estimates of the impact of health on labour supply from $\hat{\beta}$.[12]

4.2 Obtaining Consistent Estimates of β When Health in Endogenous

A third estimation problem arises when introducing health into life cycle behaviour represented by equations [18] and [22]. Since equation [18] modifies equation [8], we rewrite it in standard econometric notation corresponding to equation [23] as:

$$y_{it} = \begin{cases} \beta x_{it} + \gamma q_{it} + c_i + \varepsilon_{it} \ \text{ if RHS} > 0 \\ 0 \ \ if \quad \beta x_{it} + \gamma q_{it} + c_i + \varepsilon_{it} \leq 0 \end{cases} = F(\beta x_{it} + \gamma q_{it} + c_i) + \varepsilon_{it} \qquad [25].$$

The additional health modifier from equation [18] is now represented by observable and unobservable personal characteristics contained in $\beta x_{it} + c_i$. Bound et al (1999) argue that an appropriate instrument to address endogeneity and measurement bias arising from the use of self-reported health address will include both exogenous observable personal characteristics and more objective health measures, implying that we should rewrite our health status equation [22] as:

$$q_{it} = w_{it}^{q}\kappa + x_{it}^{q}\pi + z_{it}^{q}\theta + \xi_{it} \qquad [26]$$

See Greene (2002).

[11] Alternatively, he proposes using cross-sectional tobit models for each time period in the panel.

[12] Vella and Verbeek (1999) extend this approach to allow both time-varying and fixed selection effects, which can be decomposed from the residuals of a random effects tobit model. In a fixed effects specification like [18], however, the fixed selection effects would disappear.

where w_{it}^{q} is the wage rate, x_{it}^{q} includes observable taste and health modifiers, and z_{it}^{q} includes objective measures of health status. We only observe self-reported health status, η_t, however, not true health status, q_{it}. Since η_{it} is a categorical variable, it will also generate a nonlinear latent variable model that we can write as:

$$\eta_{it} = G(q_{it}^{*}) = G(w_{it}^{q}\kappa + x_{it}^{q}\pi + z_{it}^{q}\theta) + \xi_{it} + \mu_{it} \qquad [27]$$

where $G(\cdot)$ is a nonlinear transformation function consistent with an ordered probit or logit specification. We can obtain consistent estimates of κ, π and θ, and hence a consistent estimate of $\hat{q}_{it} = w_{it}^{q}\hat{\kappa} + x_{it}^{q}\hat{\pi} + z_{it}^{q}\hat{\theta}$, from an ordered probit model under the assumption that the disturbance $\xi_t + \mu_t$ is normally distributed. We would then use $\hat{q}_{it}$ as an instrument for q_{it} in equation [25] to address endogeneity issues and use the estimation techniques discussed in section 4.1.

4.3 Model Specifications

We investigate the relative importance of endogeneity and unobserved effects bias in two steps. First, we estimate models of labour supply with exogenous health corresponding to equation [8], using pooled and random effects estimators that are inconsistent in the presence of unobserved effects bias and fixed effects estimators that are uniformly consistent. We then estimate labour supply models with endogenous health corresponding to equations [18] and [22] using pooled, random effects and fixed effects estimators. Our labour supply specifications are tobit models that estimate the impact of health on annual hours worked for all respondents, including those with zero hours worked. We employ two measures of health: first, our "binary" measure defined to distinguish good and poor health as in section 2 and, second, what we refer to as our "full" measure that distinguishes all five categories of health. The binary measure gives an easy direct interpretation of the impact of health and labour supply, while the full measure can provide a more detailed and nuanced interpretation. Our model specifications are presented in Table 2 and discussed briefly below.

We use direct measures of wages from SLID to construct a wage equation that depends on human capital (schooling and work experience), weekly hours, geographic location, and immigrant/visible minority status. The wage equation corrects for sample selection bias arising from the censored wages of those who did not work (or did not report a wage) during the survey period. This imputed wage is then included in labour supply equations [8] and [18] along with observed characteristics that may be associated with tastes. We do not differentiate the factors associated with tastes related to health and labour supply and specify the same set of commonly used taste modifiers in equations [8] and [18]; namely, those related to age, years of schooling, family structure (marital status and number of children in the household), ethnicity (visible minority and immigrant status), and location (urban/rural and region of residence). We also include a measure of "other income"; this is a standard explanatory variable in static labour supply models but should have reduced significance in a properly specified life cycle model.

Our health equation conforms to equation [26]. It assumes that true (latent) health status is a linear function of the wage rate (imputed), exogenous observable personal characteristics (including age, education, and immigrant status) and, following Bound et al (1999), an objective health measure that identifies the presence of functional limitations.[13]

5. Econometric Results

Section 2 presented basic results that imply that the effects of ill health on labour supply, particularly work participation, are large. We re-estimate these effects within an explicit life cycle model of labour supply to assess their robustness and, in particular, their sensitivity to potential biases arising from endogeneity and unobserved effects. We first discuss our model estimates in section 5.1 and then focus on the impact of health in section 5.2.

5.1 Model Estimates

Table 3 compares pooled OLS, random effects (RE) and fixed effects (FE) estimates of hours worked for the full sample (men 21-65) when health is treated as exogenous, and alternatively, when health is treated as endogenous. We present these results for the binary health scale used in section 2 and for the full five-point health scale. For the binary health scale when health is endogenous, the instrumental variable for health is the predicted latent scale for the binary probit model as in equation [27]. While the results for the full health scale are more difficult to interpret, they capture all the self-assessed health information available in the data and are included to show that the important patterns in the results for the binary health scale also apply when the full health scale is

[13] The question in SLID is a summary flag in response to a series of questions which ask whether the respondent has difficulty doing specified activities of daily living (walking, seeing, lifting, etc.) and whether the respondent has a physical or mental condition or health problem which reduces the amount of activity the respondent can do in different types of situations. For further information on the questions posed, visit http://www.statcan.ca/english/Dli/Metadata/slid/2001/75f0002mie2002002.pdf

used. For the full health scale when health is exogenous, we report a single result for the scale measure as well as a set of results for four categories of health—excellent, very good, fair and poor, treating good health as the base category. For the full health scale when health is endogenous, the instrumental variable for health is the predicted latent scale from the ordered probit model as in equation [27]. Since the binary health scale is easier to interpret, we restrict our discussion of the results to this measure unless otherwise stated.

The effects of wages and income on labour supply are similar to those in the literature. For example, for the models with health measured on a binary scale, the intertemporal wage elasticity estimate for the RE model is 0.07 when health is exogenous and 0.03 when health is endogenous. The intertemporal wage elasticity is larger, 0.21, for the FE model whether health is exogenous or endogenous, but this estimate is only slightly larger than recent results for the U.S. Ziliak and Kniesner (1999), for example, report compensated wage elasticity estimates that range from 0.13 to 0.18 for models with additive separability between consumption and leisure (as in our model), and compensated wage elasticity estimates around 0.3 when intratemporal strong separability is relaxed. The income elasticity is also small --- with estimates of 0.10 for the RE model whether health is exogenous or endogenous. Moreover, including other income should reduce significance in a properly specified life cycle model and, as expected, estimates for the FE model that control for unobserved heterogeneity imply a smaller elasticity of 0.04 for both exogenous and endogenous health with t-values less than half the size of those provided by the RE model.

Age and its square are significant and the result is robust across models in Table 3. The results imply the familiar inverted-U pattern; that is, labour supply increases with age, but at a decreasing rate, after accounting for wages and income among other factors. On the other hand, the effects of marital status and children decline sharply as one moves from the pooled and RE models to the FE model, with the effect of children becoming insignificant for all FE models.

The effect of health also differs across models, declining sharply from the pooled and RE estimates to the FE estimate for exogenous and endogenous health, and for both the binary and full health measures. The impact of health in the FE model remains significant but considerably smaller, consistent with the descriptive evidence in section 2. For the binary health measure, the effect of treating health as endogenous, which could not be captured in section 2, is also a clear reduction in the estimated health impact.[14] There is no simple basis for comparison for the full health measure

[14] We know of no test for weak instruments for our nonlinear model. Our estimates of health equations [13] and [14],

since the qualitative health scale for exogenous health and the estimated latent variable scale for endogenous health are not comparable. In subsequent discussion, we focus on the results for the binary health scale across models to facilitate comparisons of the impact of health. In addition to the health impacts for the full sample, we also partition by age by estimating labour supply models for various age groups—25-34, 35-50 and 51-65.

5.2 Health Impact on Labour Supply

As mentioned earlier, little attention has been paid to the effects of health for young and middle-aged men. Given the importance of age in the labour supply models for the full sample—after allowing for wages, income and other factors—we investigate whether the health impacts obtained for the full sample are predominately due to elderly men and, if not, how these impacts vary by age group within the life cycle model.

Table 4 presents the marginal effect of health on labour supply in terms of hours worked and as a percentage of total hours worked for four age groups of men: all ages 21-65 (corresponding to the results in Table 3 discussed in section 5.1), young men 25-34, middle-aged men 35-50, and older men 51-65. The latter age group corresponds to the category of older men that has been the focus in the literature. We again estimate pooled, RE and FE models to examine the influence of unobserved effects for exogenous and endogenous health. We again focus on the binary health measure (good vs. poor health) to simplify comparisons, but we include estimates for the full health scale in Table 3 to illustrate comparable patterns.

Our base case is the pooled OLS estimate for exogenous health for all men 21-65; this corresponds most closely to an estimate that would result from cross-sectional data. The estimated impact of good health is a labour supply increase of 783 hours per year, or 43.5% of total hours worked by all men. The RE estimate for exogenous health for all men 21-65 is slightly lower at 675 hours or 37.5%. It is interesting to note that these estimates are similar to our estimate of the health impact in section 2 (35%) when the only control variable is age. The pooled and RE estimates follow a similar pattern throughout Table 4, with the RE estimates slightly lower, and we focus on the efficient RE estimates in subsequent discussion.

The first distinctive results arise from the FE estimates which, unlike the RE or pooled estimates, are consistent when unobserved effects are correlated with included regressors. The

---

shown in Appendix A, show significant correlations of health with all included instruments—age, immigrant status, estimated health impact from the FE estimate with exogenous binary health is only 119 hours, or 6.6% of hours worked. That is, accounting for the unobserved person-specific effects that arise in the life cycle model reduces the estimated health impact on labour supply by a factor of 6 compared to the RE estimate[15]

Some of this apparent bias in estimation of the health impact arises from endogeneity. Using an instrument for the binary health scale, as specified in equation [27], the RE estimator yields a health impact of 417 hours, or 23.2%. This is about 60% of the impact estimate when health in treated as exogenous in the RE model. Still, the FE estimate is only 91 hours, or 5.1%, which reduces the health impact estimate further by a factor of more than 4.

The effects for the different age groups are also shown in Table 4. We summarize these results as follows. First, health impacts are smaller when health is treated as endogenous, and smaller still when the FE model is estimated. For young men 21-34, the impact of good health declines from 589 hours (31.4%) for the RE model when health is treated as exogenous to 395 hours (21.1%) when an instrument for health is used and declines further to 80 hours (4.2%) for the FE model with the same health instrument. For middle-aged men 35-50, the marginal health effect declines from 495 hours (24.9%) for the RE model when health is treated as exogenous to 312 hours (15.7%) when an instrument for health is used and to 86 hours (4.3%) with the FE model with the same health instrument. For older men 51-65, the impact of good health declines from 839 hours (61.3%) for the RE model when health is treated as exogenous to 494 hours (36%) when an instrument for health is used and to 83 hours (6%) with the FE model with the same health instrument. We would note that the largest effect on the estimated health instrument of accounting for the potential bias arising from unobserved heterogeneity occurs for older men, who have been the focus of nearly all the literature to date.

Our second summary result is that the estimated health impacts are significant for all age groups and show only a modest increase with age.[16] The FE estimates with an instrument for health are 4.2% for men 25-34, 4.3% for men 35-50, and 6.0% for men 51-65. This implies that the effects

education, and the presence of a functional disability disability—for both the binary and full health scales.

[15] As we discuss in section 4, there is no test for the consistency of the RE model in the nonlinear case, but the substantial differences in the wage and health coefficients between the RE and FE models suggest that the RE model is inconsistent. When we estimate linear models for only those with positive hours worked, we find Hausman test values around 50 for all four health measures (exogenous and endogenous binary and full health) which rejects the hypothesis that the RE estimates are the same as the (consistent) FE estimates.

[16] We focus on the results for the binary health scale. The FE result for young men is insignificant when the full health scale is used and health is treated as exogenous, but not when health is treated as endogenous.

of good health on male labour supply, although small, are widespread throughout the population and increase only modestly in older men.

It is difficult to compare our results to existing literature because of the diversity of health measures and the differences in specifications and estimation methods. Currie and Madrian (1999, Table 3, 3328-30), however, suggest that the impact is smaller when health is treated as endogenous in U.S. studies of hours worked. Our results corroborate this finding but we would add that treating health as endogenous substantially overstates health impacts because of bias arising from unobserved effects. Taking account of unobserved effects with panel data and a fixed effect estimator further reduces the impact of good health on labour supply by a factor of four or more.

6. Conclusion

We present a variety of estimates of the impact of health on labour supply. The easiest estimates to interpret are ones that treat health as a binary variable (i.e., good vs. poor health) although the same patterns are apparent when the full five-point health scale is used. We introduce two innovations, one theoretical and one empirical. First, we develop a life cycle model of labour supply and health that directs us to consider the bias that may arise from unobserved effects. Studies of other topics find bias from unobserved effects to be important; the results for health are no exception. Second, we use panel data from the Canadian Survey of Labour and Income Dynamics which allow us to account for unobserved effects, and to exploit an environment in which the confounding links between labour supply, health insurance and health status are neutralized by Canada's system of universal health insurance.

Our initial estimates of the impact of good health, which treat health as exogenous in cross-sectional (or pooled panel) data, suggest that the impact of health is large --- about 35% in section 2 using only age as a control and 43% in section 5 when additional controls are added and OLS regression is used. This estimate declines to 23%, however, when health is treated as endogenous and a random effects estimator is used, and declines more sharply to 5.1% when a fixed effects estimator is used to control for potential bias arising from unobserved person-specific effects. This pattern is consistently observed for the full health scale and for various age groups. It is also consistent with earlier results obtained examining the employment decision (Hum, Simpson and Fissuh Ghebretsadik, 2006). Our results suggest that estimates of health impacts that ignore unobserved effects should be treated with caution. They also invite investigation of the unobserved effects that account for much of the health impact observed in cross-sectional studies.

We also show that the health impact is significant for younger and middle-aged men and increases only modestly with age. Our fixed-effect estimate of the health impact on hours worked, treating health as endogenous, is 4.2% for men 21-34, 4.3% for men 35-50, and 6% for older men 51-65. The largest change in the estimated health impact from treating health and endogenous and accounting for unobserved effects in our sample occurs for older men, the focus of the literature to date. Since we include those not working in our estimates of labour supply responsed, this result includes the effects of health on early retirement in the latter group.

References

Banks, J., T. Crossley, and S. Goshev (2007) "Looking for Private Information in Self-Assessed Health," Research Institute for Quantitative Studies in Economics and Population Report No. 423, McMaster University, August

Bound, J., M. Schoenbaum, T. Stinebrickner and T. Waidmann (1999) "The dynamic effects of health on the labor force transitions of older workers," *Labour Economics* 6, 179-202

Currie, J., and B. Madrian (1999) "Health, Health Insurance and the Labor Market," *Handbook of Labor Economics, Volume 3*, O. Ashenfelter and D. Card (eds.), Elsevier Science

Davidson, R. and J. MacKinnon (2004) *Econometric Theory and Methods*, Oxford University Press

Greene, W. (2002) "The Bias of the Fixed Effects Estimator in Nonlinear Models," working paper, Department of Economics, Stern School of Business, New York University, October; available at http://pages.stern.nyu.edu/~wgreene/

Grossman, M (1972) "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy* 80, 223-55

Hausman, J. and W. Taylor (1981) "Panel Data and Unobservable Individual Effects," *Econometrica* 49, 1377-98

Heckman, J. (1979) "Sample Selection Bias as a Specification Error." *Econometrica* 47(1): 153-61

Hum, D., W. Simpson, and E. Fissuh Ghebretsadik (2006) "The Impact of Health on Labour Supply in Panel Data" in *Global Business and Economics Anthology Volume I*, December, 185-196

Jakubson, G. (1988) "The sensitivity of labor-supply parameter estimates to unobserved individual effects: fixed- and random-effects estimates in a nonlinear model using panel data," *Journal of Labor Economics* 6, 302-329

McDonald, J., and R. Moffitt (1980) "The Uses of Tobit Analysis," *The Review of Economics and Statistics* 62, 318-321

Vella, F. and M. Verbeek (1999) "Two-step estimation of panel data models with censored endogenous variables and selection bias," *Journal of Econometrics* 90, 239-263

Woolridge, J. (1995) "Selection corrections for panel data models under conditional mean independence assumption," *Journal of Econometrics* 68, 115-132

Ziliak, J., and T. Kniesner (1999) "Estimating life cycle labor supply tax effects," *Journal of Political Economy* 107, 326-359.

Ziliak, J., and T. Kniesner (2005) "The effect of income taxation on consumption and labor supply," *Journal of Labor Economics* 23, 769-796.

Tables

Table 1. Labour Supply By Health Status, 2001

	Total Hours	Positive Hours	Employment Rate	Log Odds of Empt	Change in Hours
All Men21-65	1780.9	2063.7	87.43%		91.3
Good Health	1853.6	2077.9	90.17%	2.22	124.6
Poor Health	1209.6	1907.0	64.28%	0.59	-164.5
Better Health					50.2
Same Health					114.8
Worse Health					-159.4
Men 35-50	2011.7	2180.4	92.86%		39.3
Good Health	2082.4	2193.5	95.36%		58.6
Poor Health	1388.7	2020.8	69.32%		-131.0
Better Health					233.8
Same Health					49.4
Worse Health					-187.7
Men 51-65	1466.9	2033.2	74.51%		-279.8
Good Health	1586.3	2061.2	79.24%		-246.0
Poor Health	962.9	1857.8	52.845		-419.6
Better Health					-218.2
Same Health					-257.9
Worse Health					-469.7

Source: Survey of Labour and Income Dynamics (SLID) Internal Longitudinal File (1996-2001)

Table 2. Model Specifications and Brief Definitions of Variables

Variables:	Variable Definitions (as above if blank)
Wage Equation:	Dep Var: composite hourly wage for all paid-worker jobs
Education	Total years of schooling completed
Experience	Years of full-time equivalent work experience
Experience squared	
Weekly hours	Average hours worked per week during reference year
Immigrant	Immigrant to Canada (base=Native Born)
Visible Minority	Member of a visible minority (base=not a visible minority)
Region	Prairies, Atlantic, Quebec, Ontario (base=British Columbia)
Wage selection eqn:	Dep Var: employed (reporting a wage) in reference year
Education	
Age, age squared	Age in years
Kids	Number of children at home
Married	Married or living common law
Other income	Family income other than own earnings
Region	
City	Resident of urban area (base=rural resident)
Immigrant	
Health (instrument) eqn	Dep Var: (1) 1=excellent, good or very good; 0=fair or poor health (2) 5=excellent, 4=very good, 3=good, 2=fair, 1=poor
Age, age squared	
Education of mother	
Education	
Disabled	=1 if reports functional disability; =0 otherwise
Workl	=1 if functional disability is limiting at work; =0 otherwise
Labour supply eqn:	Dep Var: (1) 1=positive annual hours; 0=no hours in reference yr (2) annual hours worked
Wage	Imputed wage predicted from wage equation
Age, age squared	
Health	(1) self-reported health or (2) health predicted from health eqn
Other income	
Education	
Kids, Married	
Region	
Visible Minority	
Immigrant	

Table 3. Pooled, Random Effect and Fixed Effect Tobit Estimates of Labour Supply for Men 21-65 Years of Age with Exogenous and Endogenous Health Status (Dependent variable is annual hours worked; t-values are in parentheses)

	Binary Health[a]						Full Health Scale[b]								
	Exogenous Health			Endogenous Health			Exogenous Health						Endogenous Health		
	Pooled	RE	FE	Pooled	RE	FE	Pooled		RE		FE		Pooled	Re	FE
Constant	-2344.32	-2344.14		-2627.58	-2627.57		-2632.79	-1704.22	-2632.79				-1702.36	-1702.4	
	(-26.45)	(-26.00)		(-29.59)	(-28.72)		(-29.06)	(19.18)	(-28.12)				(-19.68)	(-19.67)	
Log wage[c]	131.05	131.72	382.91	43.66	43.96	374.85	134.08	122.35	134.37	123.14	384.00	383.64	80.36	80.65	381.66
	(16.08)	(14.72)	(10.44)	(5.22)	(4.57)	(10.22)	(16.41)	(15.16)	(14.49)	13.76	(10.47)	(10.5)	(9.78)	(8.55)	(10.41)
Other income[d]	-2.90	-9.91	-3.38	-2.67	-9.80	-3.36	-3.02	-2.82	-11.24	-13.08	-3.39	-3.35	-2.67	-9.82	-3.36
	(-17.08)	(-57.92)	(-19.52)	(-15.79)	(-55.69)	(-19.40)	(-17.72)	(-16.73)	(-53.73)	(-64.45)	(-19.57)	(-19.3)	(-15.82)	(-55.89)	(-19.40)
Age	174.46	164.38	197.70	181.21	169.86	196.33	181.57	177.81	170.01	89.96	199.37	197.56	193.59	182.24	198.59
	(42.94)	(39.82)	(23.04)	(44.86)	(40.79)	(22.89)	(44.53)	(44.11)	(40.72)	(92.13)	(23.23)	(23.1)	(47.73)	(43.18)	(23.16)
Age squared	-229.05	-240.32	-247.69	-225.64	-235.36	-242.50	-236.48	-231.63	-246.01	-157.17	-248.93	-246.8	-240.20	-249.94	-245.16
	(-50.88)	(-52.25)	(-25.67)	(-50.43)	(-51.39)	(-25.10)	(-52.39)	(-51.87)	(-53.00)	(-104.3)	(-25.80)	(-25.6)	(-53.67)	(-53.92)	(-25.41)
Married	326.83	327.09	45.90	313.14	313.21	45.05	326.56	314.09	326.64	355.23	46.53	45.41	313.24	313.32	45.05
	(27.18)	(24.53)	(2.26)	(26.19)	(23.15)	(2.22)	(27.08)	(26.33)	(24.43)	(26.46)	(2.29)	(2.24)	(26.20)	(23.17)	(2.22)
Health	813.44	814.18	119.22	504.46	505.62	91.29	225.64		227.07		30.83		564.12	565.13	102.74
	(46.86)	(45.58)	(8.22)	(53.11)	(51.14)	(10.59)	(43.63)		(38.38)		(6.85)		(53.03)	(50.98)	(10.58)
Excellent								160.21		120.75		12.32			
								(11.44)		(6.07)		(1.14)			
Very Good								114.79		85.97		-8.15			
								(8.91)		(4.43)		(-0.92)			
Fair								-444.38		-474.07		-74.73			
								(-20.96)		(-18.41)		(-4.78)			
Poor								-1561.30		-1580.91		-403.24			
								(-43.83)		(-37.30)		(-12.83)			
Kids	21.40	21.82	4.32	13.43	13.68	3.74	17.67	20.52	17.90	-13.05	4.60	4.27	13.43	13.68	3.74
	(2.25)	(1.75)	(0.45)	(1.42)	(1.10)	(0.39)	(1.85)	(2.18)	(1.44)	(-1.05)	(0.48	(0.45)	(1.42)	(1.10)	(0.39)
Immigrant	70.83	70.94		84.55	84.60		76.94	77.16	76.99	66.14			90.28	90.33	
	(4.40)	(3.93)	513.19	(5.28	(4.83)		(4.76)	(4.84)	(4.25)	(3.67)			(5.64)	(5.16)	
σ_ε	973.40	977.21	(257.06)	966.56	968.74	512.86	946.42	963.06	978.7	967.6	485	512.4	966.66	968.85	512.86
	(250.91)	(130.41)		(3.85)	(7.40)	(1.99)	(243.3)	(251.19)	(128.95)	131.7	(125)	(257)	(3.85)	(7.40)	(257.06)
σ_c		99.89			98.20				99.24	97.8				98.22	
		(5.64)			(18.44)				(5.30)	(5.40)				(18.45)	
ln L	-286362	-53972	-260949	-286061	-53889	-260930	-286515	-285861.8	-53551	-53827	-60809	-260893	-286066	-53933	-260,930
Obs.								38,975							

Data: SLID longitudinal file (1996-2001)

Notes: [a] Binary health measure takes on a value of one if health is reported to be excellent, very good or good and takes on a value of zero if health is reported to be fair or poor

[b] Full health scale takes on a value of 5 if health is reported to be excellent, 4 if it is very good, 3 if it is good, 2 if it is fair and 1 if health is reported to be poor. We report results for both the scale measure and the separate categories of health (with good health as the base category); the latter measures are less restrictive but the results across models are more difficult to summarize.

[c] Log wage is imputed from a model of log wage offers which corrects for selection bias arising from non-participation and censored wages. The wage equation includes years of schooling, actual full-time equivalent work experience and its square, annual hours worked, region, and foreign/native born status. (See Table 2.)

[d] Other family income in $,000

See Table 2 for a definition of variables; t-statistics are in parentheses

Table 4. Summary of health effects, including marginal effects as % of hours worked, in labour supply equations for men by age group

	All Ages 21-65			Ages 21-34			Ages 35-50			Ages 51-65		
	Pooled	RE	FE	Pooled	RE	FE	Pooled	RE	FE	Pooled	RE	FE
Exogenous Health:												
Binary health measure	782.61	675.38	119.22	592.22	589.48	87.80	727.15	495.22	113.69	847.71	839.21	107.27
	(46.86	(45.58)	(8.22)	(14.04)	(19.88)	(2.42)	(38.42)	(4.78)	(5.62)	(24.61)	(23.37)	(4.04)
% marginal effect	43.5%	37.5%	6.6%	31.6%	31.4%	4.7%	36.5%	24.9%	5.7%	61.9%	61.3%	7.8%
Full Health Scale	217.00	185.39	30.83	121.25	121.25	3.83	248.25	162.37	41.23	218.21	215.76	31.20
	(43.63)	(38.38)	(6.85)	(11.62)	(14.37)	(0.40)	(34.57)	(40.84)	(5.50)	(24.83)	(23.05)	(3.99)
% marginal effect	12.1%	10.3%	1.7%	6.5%	6.5%	0.2%	12.5%	8.2%	2.1%	15.9%	15.7%	2.3%
Endogeneous Health[a]												
Binary health measure	485.84	416.88	91.29	396.54	395.22	79.68	470.27	311.90	85.80	498.02	493.67	82.84
	(53.11)	(51.14)	(10.59)	(18.73)	(28.15)	(3.72)	(41.99)	(56.58)	(7.38)	(27.52)	(25.84)	(4.94)
% marginal effect	27.0%	23.2%	5.1%	21.1%	21.1%	4.2%	23.6%	15.7%	4.3%	36.4%	36.0%	6.0%
Full Health Scale	543.29	465.47	102.73	441.95	440.45	89.60	526.15	347.07	96.59	557.06	552.91	93.25
% marginal effect	(53.03)	(50.98)	(10.58)	(18.65)	(27.99)	(3.72)	(41.95)	(56.64)	(7.38)	(27.57)	(25.87)	(4.94)
	30.2%	25.9%	5.7%	23.5%	23.5%	4.8%	26.4%	17.4%	4.9%	40.7%	40.4%	6.8%

Source: Table 3 for All Ages 21-65 and Appendix B (Tables B1, B2, and B3) for other age groups; t-values for the corresponding regression coefficient are in parentheses. Marginal effects (change in hours worked for a unit change in the health measure) are computed by the LIMDEP software and expressed as a percentage of average hours worked for the sample.

Appendix A

Table A1. Health Instrument Model Results

Ordered probit regression

Number of obs	= 38975
LR chi2(6)	= 9108.82
Prob > chi2	= 0.0000
Pseudo R2	= 0.0858

Log likelihood = -48502.212

Full health	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	-.0397469	.0042731	-9.30	0.000	-.048122	-.0313717
age2	.030302	.0047402	6.39	0.000	.0210114	.0395927
educ	.0358286	.0015	23.89	0.000	.0328887	.0387685
immigrant	-.0684743	.0175142	-3.91	0.000	-.1028015	-.034147
disable	-1.285468	.0163601	-78.57	0.000	-1.317533	-1.253402
ln(wage)	.0349446	.0092813	3.77	0.000	.0167535	.0531357
/cut1	-3.107225	.0947896			-3.293009	-2.921441
/cut2	-2.253014	.0939253			-2.437104	-2.068924
/cut3	-1.224	.0936184			-1.407489	-1.040511
/cut4	-.0947762	.0934429			-.277921	.0883685

Probit regression

Number of obs	= 38975
LR chi2(6)	= 6762.92
Prob > chi2	= 0.0000
Pseudo R2	= 0.2575

Log likelihood = -9749.2213

Binary health	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
age	-.0197611	.0079648	-2.48	0.013	-.0353717	-.0041505
age2	.0047614	.0085837	0.55	0.579	-.0120625	.0215852
educ	.0382859	.002785	13.75	0.000	.0328274	.0437444
imigrant	-.0651579	.0311482	-2.09	0.036	-.1262073	-.0041085
disable	-1.447354	.0209712	-69.02	0.000	-1.488456	-1.406251
ln(wage)	.1146258	.0183119	6.26	0.000	.0787351	.1505164
_cons	1.854163	.1791565	10.35	0.000	1.503023	2.205304

Data Source: SLID longitudinal file (1996-2001)

Appendix B

Table B1. Labour supply model of men aged 21-34

	Full Health			Binary health		
	Pooled	RE	FE	Pooled	Re	FE
Exogenous Health						
Imputed wage	88.94	88.94	286.13	88.86	90.31	284.34
	(5.73)	(6.88)	(4.86)	(5.75)	(6.71)	(4.83)
Other income	-3.5	-3.5	-3.41	-3.3	-7.11	-3.43
	(-10.22)	(-13.74)	(-8.17)	(-9.68)	(-20.10	(-8.22)
Age	68.12	68.12	231.3	69.66	68.26	229.88
	(0.89)	(0.66)	(3.86)	(0.91)	(0.67)	(3.84)
Age2	-98.65	-98.65	-345.19	-105.01	-105.5	-342.39
	(-0.76)	(-0.57)	(-3.44)	(-0.81)	(-0.61)	(-3.42)
Married	285.29	285.29	62.72	285.61	287	62.63
	(12.72)	(11.9)	(1.97)	(12.79)	(11.4)	(1.97)
Health	123.04	123.04	3.83	600.84	601.6	87.8
	(11.62)	(14.37)	(0.4)	(14.04)	(19.88)	(2.42)
Kids	83.27	83.27	-4.88	85.29	85.9	-4.9
	(6.05	(5.57)	(-0.32)	(6.21)	(5.61)	(-0.32)
Immigrant	-97.95	-97.95		-103.32	-103.49	
	(-2.95)	(-3.03)		(-3.12)	(-3.10)	
Endogenous Health						
Imputed wage	49.58	49.68	282.18	19.64	19.74	276.24
	(3.19)	(3.48)	(4.8)	(1.24)	(1.36)	(4.69)
Other income	-3.04	-7.13	-3.39	-3.04	-7.11	-3.39
	(-8.96)	(-32.85)	(-8.13)	(-8.95)	(-32.73)	(-8.13)
Age	118.74	118.31	233.4	109.24	108.8	231.49
	(1.56)	(1.18)	(3.9)	(1.44)	(1.08)	(3.87)
Age2	-165.37	-165.45	-344.57	-154.4	-154.49	-342.34
	(-1.29)	(-0.98)	(-3.44)	(-1.21)	(-0.91)	(-3.42)
Married	272.83	272.97	62.85	272.88	273.02	62.85
	(12.29)	(11.01)	(1.98)	(12.3)	(11.01)	(1.98)
Health	448.11	448.36	89.6	402.07	402.31	79.68
	(18.65)	(27.99)	(3.72)	(18.73)	(28.15)	(3.72)
Kids	81.81	81.89	-4.81	81.6	81.67	-4.81
	(6.01)	(5.46)	(-0.32)	(5.99)	(5.45)	(0.32)
Immigrant	-99.29	-99.31		-103.76	-103.78	
	(-3.02)	(-3.04)		(-3.16)	(-3.18)	

Data Source: SLID longitudinal file (1996-2001). t-statistics in parentheses

Dependent variable: total annual hours worked

Table B2. Labour supply model of men aged 35-50

	Full Health			Binary health		
	Pooled	RE	FE	Pooled	Re	FE
Exogenous health						
Imputed wage	167.88	167.81	533.44	159	159.16	534.54
	(16.61)	(16.08)	(10.23)	(15.8)	(15.07)	(10.25)
Other income	-1.22	-3.75	-2.69	-1.05	-6.48	-2.67
	(-5.48)	(-14.33)	(-10.90)	(-4.73)	(-25.85)	(-10.81)
Age	8.27	4.35	19.8	-1.03	-4.1	17.6
	(0.30)	(0.13)	(0.76)	(-0.04)	(-0.12)	-0.67
Age2	-16.71	-18.41	-15.64	-5.94	-7.21	-13.99
	(-0.51)	(-0.47)	(-0.51)	(-0.18)	(-0.19)	(-0.46)
Married	311.71	311.67	83.22	303.06	303.16	82.69
	(22.00)	(26.26)	(2.92)	(21.49)	(25.04)	(2.9)
Health	220.33	220.1	31.2	855.65	855.79	107.27
	(34.57)	(40.84)	(5.50)	(38.42)	(4.78)	(5.62)
Kids	29.57	29.57	16.07	35.9	35.98	15.83
	(2.54)	(2.43)	(1.28)	(3.10)	(2.84)	(1.26)
Immigrant	-23.01	-23.02		-27.71	-27.69	
	(-1.10)	(-1.06)		(-1.34)	(-1.23)	
Endogenous Health						
Imputed wage	112.12	112.06	531.81	75.62	75.55	525.64
	(11.06)	(10.50)	(10.20)	(7.32)	(6.95)	(10.08)
Other Income	-0.79	-2.92	-2.66	-0.79	-3.13	-2.66
	(-3.59)	(-10.93)	(-10.79)	(-3.59)	(-11.80)	(-10.79)
Age	4.10	0.85	17.70	-8.28	-11.85	15.64
	(0.15)	(0.03)	(0.68)	(-0.3)	(-0.36)	(0.60)
Age squared	-2.21	-3.62	-11.01	12.43	10.88	-8.59
	(-0.07)	(-0.09)	(-0.36)	(0.38)	(0.28)	(-0.28)
Married	289.05	289.01	80.39	289.02	288.98	80.39
	(20.59)	(24.02)	(2.82)	(20.59)	(23.98)	(2.82)
Health	562.08	562.20	93.25	502.51	502.46	82.84
	(41.95)	(56.64)	(7.38)	(41.99)	(56.58)	(7.38)
Kids	18.82	18.81	14.65	19.11	19.10	14.65
	(1.63)	(1.48)	(1.17)	(1.66)	(1.52)	(1.17)
Immigrant	-15.53	-15.54		-21.53	-21.54	
	-0.75)	(-0.74)		(-1.04)	(-1.02)	

Data Source: SLID longitudinal file (1996-2001). t-statistics in parentheses

Dependent variable: total annual hours worked

Table B3. Labour supply model of men aged 51-65

	Full Health			Binary health		
	Pooled	RE	FE	Pooled	Re	FE
Exogenous health						
Imputed wage	110.58	110.54	352.49	118.7	118.84	344.73
	(5.3)	(5.33)	(3.16)	(5.7)	(5.78)	(3.09)
Other income	-5.98	-10.01	-5.11	-5.99	-7.74	-5.11
	(-15.33)	(-24.31)	(-14.81)	(-15.33)	(-35.83)	(-14.78)
Age	665.24	654.29	1010.87	684.2	674.21	1004.63
	(6.97)	(6.61)	(13.22)	(7.16)	(6.71)	(13.14)
Age2	-703.12	-712.07	-989.23	-719.72	-729.85	-984.48
	(-8.49)	(-8.30)	(-14.69)	(-8.68)	(-8.38)	(-14.61)
Married	441.3	441.28	65.22	460.55	460.64	60.67
	(12.41)	(13.43)	(0.97)	(12.94)	(14.68)	(0.91)
Health	314.46	314.77	41.3	921.74	922.05	113.86
	(24.83)	(23.05)	-3.99)	-24.61)	(23.37)	(4.04)
Kids	-69.25	-69.25	-144.36	-38.44	-38.42	-153.59
	(-0.55)	(-0.35)	(-1.15)	(-0.30)	(-0.18)	(-1.22)
Immigrant	305.38	305.44		296.93	297.06	
	(8.3)	(8.67)		(8.07)	(8.51)	
Endogenous Health						
Imputed wage	8.95	9.06	343.51	51.73	51.5	349.97
	(0.42)	(0.43)	(3.08)	(2.45)	(2.46)	(3.14)
Other Income	-5.75	-9.15	-5.08	-5.76	-11.52	-5.08
	(-14.84)	(-22.42)	(-14.72)	(-14.86)	(-28.22)	(-14.72)
Age	659.09	648.78	1004.12	674.06	662.89	1006.27
	(6.95)	(6.5)	(13.14)	(7.1)	(6.66)	(13.16)
Age squared	-685.54	-695.23	-980.12	-703	-709.62	-982.64
	(-8.32)	(-8.04)	(-14.56)	(-8.53)	(-8.22)	(-14.60)
Married	445.81	445.84	59.62	446.2	446.05	59.62
	(12.6)	(13.62)	(0.89)	(12.61)	(13.66)	(0.89)
Health	594.88	595.51	85.93	665.59	665.84	96.73
	(27.57)	(25.87)	(4.94)	(27.52)	(25.84)	(4.94)
Kids	-66.69	-66.67	-144.18	-68.4	-68.4	-144.18
	(-0.53)	(-0.32)	(-1.14)	(-0.54)	(-0.34)	(-1.14)
Immigrant	331.41	331.49		337.96	337.93	
	(9.05)	(9.25)		(9.23)	(9.40)	

Data Source: SLID longitudinal file (1996-2001). t-statistics in parentheses

Dependent variable: total annual hours worked

Table B4. Means values of the variables in the study by age groups

	Age group 21-34			Age group 35-50			Age group 51-65			All sample		
Variable	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N
Age	29.62	2.19	8486	42.20	4.48	19554	57.21	4.38	10935	43.72	10.56	38975
Kids	0.56	0.80	8486	0.24	0.56	19554	0.01	0.10	10935	0.24	0.58	38975
Annual hours worked	1877	820	8486	1991	836	19554	1370	1140	10935	1799	973	38975
Age2	8.82	1.30	8486	18.01	3.81	19554	32.92	5.06	10935	20.23	9.50	38975
Other income	16.80	26.18	8486	15.92	27.37	19554	24.93	33.98	10935	18.74	29.76	38975
Married	0.51	0.50	8486	0.74	0.44	19554	0.81	0.39	10935	0.70	0.46	38975
Immigrant	0.09	0.28	8486	0.09	0.29	19554	0.16	0.37	10935	0.11	0.32	38975
Full Health scale	4.04	0.92	8486	3.83	0.97	19554	3.51	1.11	10935	3.79	1.02	38975
Binary health scale	0.94	0.23	8486	0.91	0.28	19554	0.82	0.39	10935	0.89	0.31	38975
Full health scale (imputed)	-0.52	0.44	8486	-0.77	0.48	19554	-1.12	0.60	10935	-0.81	0.55	38975
Binary health (imputed)	1.80	0.51	8486	1.58	0.54	19554	1.13	0.70	10935	1.51	0.63	38975
Imputed wage	1.08	0.64	8486	1.27	0.61	19554	1.31	0.65	10935	1.24	0.63	38975

Data Source: SLID longitudinal file (1996-2001)

Note: Age2=Age squared /100
Other Income=(Total family income-Total individual wages)/1000
Immigrant=1 if immigrant, 0 otherwise.
Full health scale= self reported health
Binary health=1 if self reported health is excellent, very good or good and 0 otherwise.
Imputed wage=logarithm of wages imputed from the wage offer equation.
Married=1 if marital status is married, 0 other wise.
Annual hours worked= Total hours worked
Full health scale (imputed) = imputed full health scale based on ordered probit regression in Appendix A
Binary health (imputed) = imputed binary health scale based on binary probit regression in Appendix A