Three Essays on Empirical Labour Economics

by

Eyob Fissuh

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

Doctor of Philosophy

Department of Economics

University of Manitoba

Winnipeg

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Three Essays on Empirical Labour Economics

By

Eyob Fissuh

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Doctor of Philosophy

Dedication

I dedicate this dissertation to the memory of my late brother Amanuel and my beautiful twin daughters Lilly and Layla.

Contents

I. Acknowledgment

Thesis writing is a painstaking process which requires commitment, perseverance, and serenity, and involves many individuals. It is not possible to name all those who have helped me either directly or indirectly. However, Professor Wayne Simpson deserves the highest level of gratitude for being an excellent mentor during my graduate school training, especially when things were critical. He is an extraordinary advisor and colleague that always finds the time for listening to, the little and big, challenges and hurdles that inevitably pop up in the course of writing a PhD thesis. His technical and editorial advice was instrumental to the inception and completion of this thesis. I have learned from him invaluable lessons and insights on handling academic as well as non-academic economic research.

I would like to thank members of my thesis advisory committee, Professor Laura Brown and Professor Liqun Wang, for critically reviewing various drafts of the dissertation and their thoughtful advice in my thesis writing process. Their advice and comments were crucial to the substance and presentation of the final output. I thank Dr. Jean Kimmel for her willingness to act as my external examiner despite her other academic and professional commitments. Her indispensable review and comments have improved the presentation and content of the dissertation.

I have received financial assistance from different institutions and individuals for my PhD training. I thank the Graduate School of the University of Manitoba for granting me the University of Manitoba Graduate Fellowship (UMGF)— $70,000 over four years. I thank Professor Sean Byrne for giving me the chance to work in the fascinating world of peace and reconciliation projects in Northern Ireland—which allowed me to coauthor more than 11 peer reviewed articles.

I thank Betty McGregor and Judy Ings, staff members of the Department of Economics, for their support and encouragement. I thank my colleagues at the Research and Data Development group, Human Resources and Skills Development Canada-Labour Program, for their support and encouragement.

I thank Statistic Canada for giving me the permission to access confidential internal files of the Survey of Labour and Income Dynamics (SLID). I thank all the staff members and visitors at the University of Manitoba Research Data Centre (RDC) for their cooperation during the data analysis stage of my research.

I thank all my family members for their love and encouragement. My parents deserve my deepest level of gratitude for their love and sacrifice throughout my life. I thank my brothers and sisters: Bisrat, Hiwet, Senait, Aster and Freweini, Yosief, Melake and Abel, for their encouragement and love.

Last but not least, I would like to give my special thanks to the love of my life, the rock of our family, Bisrat Yemane, who had to bear the burden of single parenting our twin beautiful daughters Layla and Lilly during my absence from home in this long process of PhD education. Her support, patience and love enabled me to complete this thesis.

II. Abstract

This dissertation comprises three essays on empirical labor economics. The unifying theme of the thesis is the econometric methodology: using the Survey of Labor and Income Dynamics (SLID), each essay employs panel data models that control for sample selection bias and individual heterogeneity.

The first essay estimates the elasticity of labour supply of men in a life-cycle setting. This paper confirms that failure to control for the individual heterogeneity and sample selection bias produces upward biased point elasticity estimates: our model that controls for sample selection bias and individual heterogeneity produces an intertemporal labour supply substitution elasticity of 0.16 and those which do not control for these problems 0.23–0.48.

The second paper estimates the impact of health on wages in Canada using Mincer type wage offer model—that correct for sample selection and control for individual heterogeneity— and finds that the effect of health on wages is positive, as expected, but not statistically significant. We also demonstrate that treating health as an exogenous variable may give misleading results.

The third essay examines the impact of childcare costs on maternal employment in Canada. This paper extends the existing static maternal labour supply model into a life cycle model, where a typical mother is faced with a problem of maximizing life cycle utility subject to wealth and time constraints. Our study confirms that labor supply decisions of mothers are generally less responsive to childcare price: childcare price elasticity of labour supply for single mothers is -0.068 and for married mothers -0.016.

In addition to shedding light on a number of highly debated issues in labour economics, perhaps, the main contribution of this dissertation is that it demonstrates the importance of controlling for sample selection bias in empirical labour economics and that the selection process could mainly operate via time variant variables—hence the traditional fixed effects model does not suffice.

III. Introduction

This dissertation comprises three essays on empirical labor economics. The first essay examines the sensitivity of parameter estimates of the life cycle labour supply model for men to the presence of unobservable individual heterogeneity and sample selection bias. The life cycle labour supply model is derived from a standard consumer problem where a typical man aims at maximizing utility over his lifetime subject to a wealth constraint. By estimating a life cycle labor supply model we are able to distinguish between different types of wage elasticities. The focus, however, is on the intertemporal labour supply elasticity with respect to wages, keeping the marginal utility of wealth constant.

The second essay explores the impact of health on wages. In principle, there are at least two pathways by which health can affect wages. Firstly, we can treat health as a form of human capital and in its simplest form the human capital argument of health predicts that health can be treated as an investment in the production of future income. Secondly, according to the signaling hypothesis we can argue that health can be used as a signal for productivity and employers can use it as screening device. Both of these contesting hypotheses predict that healthy people will command higher wages than their unhealthy counterparts. The common element of these arguments is that health and wages are positively correlated. This essay employs appropriate econometric models from the class of sample selection panel data models to disentangle the effect of health on wages. In doing so the paper tries to control for biases arising from unobservable heterogeneity and sample selection.

The third essay examines the impact of childcare costs on maternal employment in Canada. This paper situates the maternal labour supply model in a life cycle setting where a typical mother is faced with a problem of maximizing life cycle utility subject to wealth and time constraints. We first employ standard linear panel data models to estimate the responsiveness of maternal labour supply to changes in childcare cost to compare our findings with what is documented in the literature. Then we employ panel data models which control for individual heterogeneity and the sample selection problem.

It is noteworthy that the essays in this dissertation use a similar econometric framework and all essays employ longitudinal data from the Survey of Labour and Income Dynamics (SLID). In all cases we attempt to account for unobservable individual heterogeneity and the problem of sample selection. In each essay the significance of controlling for individual heterogeneity and sample selection is emphasised, but we also consider the problem that the selection process could largely work via time variant variables. As expected, our research results suggest that accounting for individual heterogeneity using the traditional fixed effects technique may not be enough to control for the bias introduced by sample selection.

Chapter 1

1. The Sensitivity of Life Cycle Labor Supply Model Parameter Estimates to Sample Selection Correction: Panel Data Evidence from Canada

Abstract

This paper explores the application of several panel data models in estimating a life cycle labor supply model. The estimated intertemporal substitution elasticity is compared across the different panel data estimators. The results indicate that life cycle labor supply estimates are sensitive to the econometric specification of unobserved individual heterogeneity and non-random sampling. This paper confirms that failure to control for the individual heterogeneity and sample selection bias produces upward biased parameter estimates. Using the second panel of the Canadian Survey of Labor and Income Dynamics (SLID) we find that panel data models which do not control for sample selection bias give an intertemporal substitution elasticity between 0.23 and 0.48 compared to 0.16 for a model which controls for this problem. Moreover, we find insignificant wealth effects which imply that the intertemporal elasticity is a good approximation of both the own compensated and uncompensated labor supply elasticity. Our empirical results reveal that it is important to distinguish between the effect of evolutionary and parametric wage changes on labor supply decisions of individuals over their life cycle.

1.1. Introduction

Labor supply is one of the most popular areas of research in labour economics. Labour supply analysis is crucial to the assessment of a wide range of public policy issues including income support programs, child care policies, and taxation. The advent of rich microeconomic data, particularly household panel data, and the dramatic improvement in computing technology in the last four decades has spurred empirical research in labor supply. One outcome of this research agenda has been the recognition that the wage elasticity computed from static labour supply models using cross-sectional data cannot distinguish between the effects of evolutionary and profile changes in wages on labour supply. Most of these studies focus on female labour supply, however, and research on men has received very little attention.

Life cycle labor supply decision-making implies that economic agents aim at maximizing lifetime utility subject to their resource constraints. Situating the labor supply decision of an individual in a life cycle environment allows estimation of econometrically meaningful parameters (MaCurdy, 1981).[1] In particular, the proper estimation of a life cycle labor supply model permits estimation of the intertemporal elasticity of substitution, which is important for several reasons. Firstly, the elasticity estimates can be used for welfare analysis in areas such as income taxation incidence, social benefit analysis, pensions, childcare and the like.[2] Secondly, they are important links to the macroeconomic analysis of fluctuations in real business cycle theory.[3] Thirdly, the intertemporal elasticity of substitution is also important to the evaluation of government policies which aim at influencing savings (Altonji and Ham, 1990).

The life cycle labor supply model introduces person-specific effects into labor supply decisions. The econometric implication is that cross-sectional regressions are contaminated with omitted variables bias and inconsistency problems. We employ rich panel data from the Survey of Labour and Income Dynamics (SLID) to estimate life cycle labor supply model parameters for men that correct for biases arising from the presence of individual heterogeneity and sample selection. The paper also examines the sensitivity of life cycle labor supply parameter estimates to changes in econometric estimation strategy vis-à-vis heterogeneity and sample selection. To this effect we employ sample selection models that control for unobservable individual

[1] Interpretation of these estimates requires identification assumptions, including the assumption that the marginal utility of wealth is additively separable from the wage rate (Donni, 2007).
[2] Kumar (2005), for example, provides an elegant application of the effect of taxation on female labor supply in the USA.
[3] The literature identifies one of the biggest challenges of the RBC theory to be the empirically small intertemporal substitution elasticity of labor (Romer, 2006; Rebello, 2003; see Table 8 below for summary of estimates). Kimmel and Ksiener (1998) argue that the intertemporal elasticity of substitution is biased downward by data aggregation, and they advocate the use of more disaggregated data to produce the size and sign of the intertemporal elasticity of substitution needed to support RBC theory. Kimmel and Ksiener try to address both the sample selection and fixed effects in their study.

effects. To increase the reliability of our estimates we try to improve on the instrumentation of wages by introducing job tenure as an instrument in the wage offer equation. This paper contributes to the literature by estimating theory consistent intertemporal substitution elasticities for men in Canada which, to the best of our knowledge, has not been done before.[4] We hope that this paper will help to fill an important gap studying our understanding of the life cycle labor supply behavior of men in Canada. Moreover unlike most studies in the literature (e.g., MaCurdy 1981, Osberg and Phipps 1993, Reilly 1994), we do not restrict our sample to those who reported non-zero number of hours and we argue that such a restriction potentially causes sample selection bias, yielding inconsistent parameter estimates.

Our results suggest that point estimates of the intertemporal substitution elasticity with respect to annual hours worked lie in the range of 0.16 to 0.48 with the lower bound being the point estimate from the model which controls for individual heterogeneity and sample selection. We also find wealth effects, as well as the elasticity of wages with respect to the marginal utility of wealth, that are negative but statistically insignificant which implies that the intertemporal substitution elasticity is a reasonable approximation to the uncompensated and compensated labor supply elasticities. Perhaps the most important lesson of this paper is that failure to control for the individual heterogeneity and sample selection bias produces upward biased labour supply parameter estimates.

The rest of the paper is organized as follows. Section two presents the theoretical underpinning of the life cycle labor supply model. Section three develops the

[4] Hum, Simpson and Fissuh (2006) use SLID data to study the effect of health on labor supply behavior of men in Canada. They estimate logit models which are theory consistent.

econometric strategy. Section four presents the estimation results. Section five concludes.

1.2. Theoretical framework

The theoretical setting of our paper is based on the standard theory of consumer behaviour where a consumer is faced with the problem of maximizing lifetime utility subject to wealth constraint. A similar theoretical framework has been employed by Heckman (1976), Heckman and MaCurdy (1980), MaCurdy (1981) Jackubson (1988) Card (1998) French (2005), and Cahuc and Zylberberg (2004), among others. The paper does not aim at detailed derivation of the model. We only note the main results.

Assume that a consumer faces a choice between leisure and work over his lifetime. Assume also that the utility function is quasiconcave at age t which is given by $u[c_t, l_t; x_t]$, where c_t , l_t and x_t are within-period consumption, leisure time and a vector of observed time variant individual characteristics of the agent respectively. Thus lifetime utility of a typical man from time t to T is given by

$$u[u_t(c_t, l_t; x_t), u_{t+1}(c_{t+1}, l_{t+1}; x_{t+1}), \ldots, u_T(c_T, l_T; x_T)] \quad [5] \qquad [1.0]$$

The consumer is limited by the time path of his wealth constraint[6]:

$$a_t = (1+r_t)a_{t-1} + d_t + w_t h_t - p_t c_t \qquad [1.1]$$

[5] Limitation of our utility function among others includes the persistence of consumption habit, human capital formation, trade off between leisure, training and consumption (See Hotz et al 1988).
[6] This formulation ignores the role of progressive tax rates which might impact the separability of choices. See Blomquist(1985) for the effect of progressive taxes on the separability of choices and Kumar(2005) for the effect of a nonlinear budget constraint on the labor supply response parameters.

where $h_t = (l^* - l_t)$ such that l^* is total number of hours available for work; a_t is real asset at time t; a_{t-1} is real asset endowment of the consumer at time $t-1$; w_t is exogenously given within period wage level; p_t is within period price of consumption c_t; d_t is within period non-wage income[7]. We also assume that the consumer can freely borrow and save at each point in time at an interest rate r_t and the time discounting factor is ρ .

Hence the utility maximization problem of our consumer is one of

$$\underset{\{c_t, l_t, a_t\}}{Max\ U} = \sum_{t=0}^{T} u(c_t, l_t; x_t)(1+\rho)^{-t} \qquad [1.2]$$

Subject to $\sum_{t=0}^{T} \kappa_t [a_t - (1+r_t)a_{t-1} - d_t + w_t h_t + p_t c_t](1+r_t)^{-t}$

This problem could be solved by Lagrange technique. The first order conditions for maximization, normalizing the price of consumption to 1, are given by

$$u_{c_t}[c_t, l_t, x_t] - \frac{1+\rho}{1+r}\kappa_t = 0 \qquad [1.3]$$

$$u_{l_t}[c_t, l_t, x_t] - \kappa_t \frac{1+\rho}{1+r} w_t \geq 0 \qquad [1.4]$$

$$\kappa_t = (1+r_{t+1})\kappa_{t+1} \qquad [1.5]$$

where $u_{c_t}[.]$ and $u_{l_t}[.]$ are the marginal utilities with respect to within period consumption and leisure. Along the optimal path, conditions [1.3]-[1.5] unambiguously suggest that the marginal rate of substitution between consumption

[7] Strictly speaking non-wage income is $d_t + r_t a_{t-1}$.

and leisure remains the same at each time period. Note that this optimal path equilibrium is an outcome of the additive separability assumption invoked onto the utility function and is not a general result. Given the first order conditions, explicit solutions for leisure[8] and consumption demand can be written as:

$$c_t = c(w_t, (1+\rho)^t (1+r)^{-t} \kappa_t, p_t, x_t) \qquad [1.6]$$

$$h_t = h(w_t, (1+\rho)^t (1+r)^{-t} \kappa_t, p_t, x_t) \qquad [1.7]$$

$$\text{where } h_t = \begin{cases} h_t > 0 \ \text{ if } -\frac{\partial u}{\partial l_t}\Big|_{\kappa_t=\kappa_0} \leq (\frac{1+\rho}{1+r})^t w_t \\ h_t = 0 \ \text{ if } -\frac{\partial u}{\partial l_t}\Big|_{\kappa_t=\kappa_0} > (\frac{1+\rho}{1+r})^t w_t \end{cases}$$

These explicit solutions for consumption demand and labor supply show that regardless of the participation decision of our consumer over the lifetime, labor supply and consumption demand are entirely determined by the functional form of the utility functions, κ_0, taste factors, the interest rate, the discount rate ρ and contemporaneous wage rate. Most importantly the consumption demand and labor supply functions given by [1.6] and [1.7] are defined for a given marginal utility of wealth (κ_t). In the literature these demand functions are knows as κ constant demand functions or Frisch demand functions (Browning and Meghir 1991, Heckman and MaCurdy 1980).

The marginal utility of wealth (κ_t) deserves a bit of a discussion because it has a crucial repercussion on the econometric methods that should be adopted. κ_t summarizes the effect of lifetime tastes and prices and other variables included in the preference function. It is clear from equation [1.3] and [1.4] that the reservation

[8] Recall that $h_t = (l^* - l_t)$.

wage in the life cycle model depends on lifetime tastes and prices summarized by the marginal utility of wealth, κ_t. The marginal utility of wealth in turn depends on all variables present in the model. To see this, substitute the demand functions into the constraint to get an implicit expression for the optimal κ_t:

$$\sum_{t=0}^{T}\kappa_t[a_t-(1+r_t)a_{t-1}-d_t+w_t h_t(w_t,(1+\rho)^t(1+r)^{-t}\kappa_t,p_t,x_t) \\ +p_t c_t(w_t,(1+\rho)^t(1+r)^{-t}\kappa_t,p_t,x_t)](1+r_t)^{-t}=0 \qquad [1.8]$$

The κ constant demand functions help us to explore the responsiveness of labor supply to changes in wages. However, unlike the Marshallian demand or Hicksian demand elasticity functions which keep income and utility constant respectively, the elasticity computed from this function keeps marginal utility of wealth, κ_t,constant. That is to say, the elasticity measure from [1.7] is given by

$$\left.\frac{\partial h_t}{\partial w_t}\frac{w_t}{h_t}\right|_{\kappa_t=\kappa} \qquad [1.9]$$

This elasticity keeps the marginal utility of wealth constant and the labour supply elasticities computed from this technique are usually referred as Frisch elasticities. As we will see in a sequel, the Frisch elasticity should be larger than the Hicksian and Marshallian elasticities of labor supply. As mentioned above, κ_t depends on the time path of the interest rate and discounting factor and all other variables in the model. Hence, it is appropriate to use Frisch elasticity to measure the impact of wage changes through time on labor supply. However, if we want to measure the impact of wage variation across consumers on labor supply we need to estimate the effect of a change in wage profile on κ_t.

The key innovation of this Frisch labour supply model is the treatment of κ_t, which suggests the need to control for individual heterogeneity in an estimation of an empirical life cycle labour supply model. For our empirical estimation strategy we will need to manipulate expression [1.5] which is the time path of the Euler equation for the optimal path of marginal utility of wealth, κ_t. Expression [1.5] can be written in a more convenient way. Repeated iteration of the Euler equation yields

$$\ln \kappa_t = \ln \kappa_0 - \sum \ln(1 + r_t) \qquad [1.10]$$

It is evident from [1.10] that the path of the marginal utility of lifetime wealth has a fixed effects component which is specific to an individual, $\ln \kappa_0$, and a component which is path dependent representing a common effect. This way of decomposing the marginal utility of wealth proves to be empirically very useful. We will see that in detail in the next section. Note that equation [1.10] suggests that the time path of κ_t is independent of variations in wages which implies that the appropriate elasticity measure to gauge the effect of wage variation at a certain point in the lifetime is the Frisch elasticity. Expression [1.10] has an important econometric implication. Empirical labor supply models which ignore individual heterogeneity are intrinsically flawed. [9] Given that κ_t is a function of individual attributes and other taste shifters, assuming no correlation of the error term and regressors will produce biased and inconsistent estimates. In fact, our preliminary results show that the random effects estimates are downward biased and Hausman test rejects the null hypothesis of no

[9] This result implies that nonexperimental cross-sectional studies of labor supply, which constitute the lion's share of the empirical labor supply literature, are theory inconsistent and their estimates are biased and inconsistent. Experimental studies might get around this problem through proper randomization.

systematic difference between the random effects and fixed effects estimates. This study will use a variant of fixed effects model to circumvent the aforementioned problem.

Before proceeding to the empirical model it is useful to discuss some of the key implications of our theoretical model. There are some important propositions that follow from expressions [1.6]-[1.7]. Assume that leisure and consumption are normal goods and the utility function is concave. Thus, interior solution conditions [1.6]-[1.7] imply:

$$\frac{\partial c_t}{\partial \kappa} < 0 \;, \frac{\partial c_t}{\partial (\frac{1+\rho}{1+r})^t} < 0 \;, \frac{\partial h_t}{\partial \kappa} > 0, \; \frac{\partial h_t}{\partial (\frac{1+\rho}{1+r})^t} > 0, \; \frac{\partial h_t}{\partial w_t} > 0 \qquad [1.11]$$ [10]

The above propositions have important testable implications. For example in the empirical analysis we expect *a priori* the coefficient of the wage variable in the Frisch labour supply function to be positive. Expression [1.11] also implies that the coefficient of the wage variable in the equation for initial wealth should be negative because the second derivative of the utility function with respect to wealth should be negative.

[10] The proofs for these propositions are available in Heckman (1976).

Our Empirical Model

Let the utility function of our individual consumer take the following form:[11]

$$u_i(c_{it}, l_{it}) = \theta_{1it}[c_{it}]^{\Phi_1} - \theta_{2it}[h_{it}]^{\Phi_2} \qquad [1.12]$$

where $\Phi_1 > 0$ and $\Phi_2 > 0$ are common time invariant parameters and; $\theta_{1it} > 0$ and $\theta_{2it} > 0$ are time-specific age shifters. It imperative to note that θ_{1it} and θ_{2it} are functions of individual background variables such as human capital endowments (such as health[12] and education), race, gender, age, the number of children under the age of five, region of residence and the like. Given the above preference function, our consumer faces a standard problem of maximizing his lifetime utility subject to a wealth constraint. The Lagrangian function for the problem looks like:

$$\begin{aligned} L = & \sum \frac{\theta_{1it}[c_{it}]^{\Phi_1} - \theta_{2it}[h_{it}]^{\Phi_2}}{(1+\rho)^t} \\ & - \sum_{t=0}^{T} \kappa_{it}[a_{it} - (1+r_t)a_{it-1} - d_{it} - w_{it}h_{it} - p_t c_{it}](1+r_t)^{-t} \end{aligned} \qquad [1.13]$$

Assuming no corner solutions and normalizing price of consumption to 1,[13] the first order conditions of the above optimization problem are:

---

[11] MaCurdy (1981), Chamberlain (1984), Jakubson (1988) and Blundell and MaCurdy (1999) use similar empirical models. However our estimation strategy differs from those employed in these studies. We work in a situation of complete certainty, but incorporation of uncertainty into the model is straightforward and it does not change the essential results of the model.

[12] We experiment with exogenous and endogenous health but we report the results from the models which treat health as endogenous variable.

[13] It is possible to relax this using the approach suggested by Ham and Reilly (2002) and Browning and Meghir (1991). Unfortunately, we do not have the information on prices of leisure and consumption to conduct these teste. In this paper we hope that this assumption is not very damaging. Moreover the implicit assumption of no liquidity constraint might cause some bias in our model. Domeij and Floden (2001) argue that failure to control for the possibility of borrowing constrains causes a downward biased labor supply estimates. More specifically they argue that the estimates will be about 50

$$\frac{\partial L}{\partial c_{it}} \Rightarrow \frac{\Phi_1 \theta_{1it} c_{it}^{\Phi_1 - 1}(t)}{(1+\rho)^t} - \kappa_{it}(1+r_t) = 0$$

$$\frac{\partial L}{\partial h_{it}} \Rightarrow \frac{\Phi_2 \theta_{2it} h_{it}^{\Phi_2 - 1}}{(1+\rho)^t} - \kappa_{it} w_{it}(1+r_t)^{-t} = 0 \qquad [1.14]$$

$$\frac{\partial L}{\partial a_{it}} \Rightarrow \quad \kappa_{it} = (1+r_{t+1})\kappa_{i(t+1)}$$

$$\ln h_{it} = \frac{1}{\Phi_2 - 1}[\ln \kappa_i - \ln \Phi_2 - \ln \theta_{2it} + t(\rho - r) + \ln w_{it}] \qquad [1.15]$$

where expression [1.15] assumes that the marginal utility of wealth is time invariant. If we assume that taste factors towards leisure are independently distributed as $\theta_{2it} = \exp(-\alpha x_{it} - f_i + e_{it})$ we can write the labor supply equation of an individual as follows

$$\ln h_{it} = \beta[\ln \kappa_i - \ln \Phi_2 + \alpha x_{it} + f_i + t(\rho - r)] + \beta \ln w_{it} + u_{it} \qquad [1.16]$$

where $\beta = 1/(\Phi_2 - 1)$ and $u_{it} = -e_{it}\beta$, which is a random error that varies over time with mean zero. Note that f_i represents unobservable variables which constitute the unmeasured component of individual preference. We can rewrite the individual κ constant supply equation as follows

$$\ln h_{it}(t) = \nabla_i + \delta t + \mu x_{it} + \beta \ln w_{it} + u_{it} \qquad [1.17]$$

where $\nabla_i = \beta[\ln \kappa_i - \ln \Phi_2 + f_i]$, $\delta = \beta(\rho - r)$ and $\mu = \beta\alpha$

downward biased. Unfortunately SLID does not have information on asset and wealth of individuals which could be used to relax this assumption. If the claim by Domeij and Floden (2001) is true this downward biases will be offset by the sample selection bias in models which do not control for sample selection bias. Comparison of our estimated β with their findings reveals that the results are similar. Domeij and Floden (2001) report the estimated β for the USA using PSID data to be between 0.24 and 0.56.

Expression [1.17] is the key equation of interest. However equation [1.17] assumes exogenously given wages which may not be plausible. Wages are a function of observable and unobservable individual characteristics such as education, experience, race, gender, ability and the like. Moreover, as it was mentioned in the introduction wages are not observed for individuals who do not work. In other words, wages will only be observed if annual number of hours worked is positive. This is a typical incidental truncation problem where wages are a function of number of hours worked. To deal with this nonrandom sampling and endogeneity of wages, we propose that the wage offer equation adopt the following structure:

$$\ln w^*_{it} = \lambda_1 x_{it} + \sigma_i + \varepsilon_{it}, \qquad i = 1,\dots,n \qquad t = t_i,\dots,T_i \quad [1.18]$$

[14]

$$s^*_{it} = \lambda_2 z_{it} + \varsigma_i + v_{it} \qquad v_{it} \mid z_i \sim N(0,1) \quad [1.19]$$

$$s_{it} = 1 \quad \text{if} \quad s^*_{it} > 0 \quad [1.20]$$

where $\ln w^*_{it}$ is a latent endogenous wage with an observable counterpart $\ln w_{it}$. s^*_{it} is latent participation decision with an observable counterpart s_{it}. Equation [1.18] is the wage offer equation of interest and equation [1.19] is a reduced form labour force participation equation. x_i and z_i contain vectors of exogenous individual characteristics such as job tenure, years of experience, years of schooling, and health. It is conceivable that most of the variables that enter the wage equation will also determine participation in the labour market. In our empirical models we will impose some fairly standard exclusion restrictions. λ_1 and λ_2 are vectors of unknown

[14] Note also that $t = t_i,\dots,T_i$ implies that the panel structure could be unbalanced. We conduct a test for attrition bias and the null hypothesis could not be rejected after taking the sample selection problem in to account.

parameters and ε_{it} and v_{it} are random error terms with $E(\varepsilon_{it} / \upsilon_{it}) \neq 0$. We assume that (ε, υ) is independent of z_i (where z_i might contain elements of x_i[15]), σ_i and ς_i are individual fixed effects which are time invariant. To anticipate our results, our empirical analysis will provide evidence of sample selection bias, as expected.[16] Thus, our employment of a sample selection model will help to predict wage offers for non-workers and possibly circumvent the problem of endogenous regressors in our labor supply equation.

Note that ∇_i is an individual-specific component which represents the unmeasured component of time invariant individual preferences. It is also very useful to remember that ∇_i contains κ_i and f_i. It follows that ∇_i is a function of the interest rate, time preference, taste shifters and all other variables included in the wealth constraint. We treat ∇_i as a fixed effect which is person-specific and the econometric estimation will explicitly account for this. It is noteworthy to emphasize that estimates from cross-sectional studies which do not control for the unobservable individual specific factors are biased and inconsistent when $E[\nabla_i \mid x] \neq 0$. Appropriate estimation of the empirical labor supply model [1.7] gives the intertemporal substitution elasticity (β). Also note that expression [1.7] suggests that age will directly enter as an argument in the labor supply model if and only if $\rho \neq r$. We include age and age squared in our empirical model. However, if we want to compare the effect of wage variation across consumers we need to specify a functional form for ∇_i.

[15] It is conceivable that most of the variables which influence participation in the labour market will also affect wages; hence we can expect that z contains most of elements of x. Ideally, we would like to have some exclusion rule here for efficiency reason.
[16] The tests for sample selection bias are contained in Tables 5 and 6 below.

As the theoretical model indicated, the marginal utility of lifetime wealth depends on future wages, initial wealth and some other characteristics. Unfortunately, theory provides no guidance regarding the exact functional form of the relationship. We assume that the fixed effects component of the labor supply model follows a standard functional form which is empirically tractable:[17]

$$\nabla_i = x_i\phi + \sum_{t=1}^{T}\gamma_t \ln w_{it} + a_{it}\zeta + b_i \qquad [1.21]$$

The concavity assumption for the utility function implies that $\gamma_t < 0$ and $\zeta < 0$. That is to say, the marginal utility of wealth diminishes with increases in wealth and wages. It is crucial to recall that the marginal utility of wealth, like permanent income in the permanent income hypothesis of consumption, remains constant throughout the life cycle of an individual (Friedmand 1957; Heckman, 1976). In this equation x_i represents the time invariant observed characteristics of the individual.[18] In this study x_i contains own education, father's education, mother's education, race and immigration status. $\phi, \gamma,$ and ζ are unknown parameters of the model. b_i is assumed to be identically and independently distributed across consumers. The fixed effects can be imputed from the estimation of equation [1.17]. However, we need to make some assumptions to proceed with the estimation of [1.21]. Empirical estimation of the time path of wages is usually not available and we need to use some projections. As in MaCurdy (1981) we propose that the lifetime path of wages is a quadratic function of age and some other age invariant characteristics.

[17] MaCurdy (1981), Blundell and MaCurdy (1999) and Cahuc and Zylberberg (2004) use a similar specification.
[18] x_i contains the time invariant component of x_{it} in the utility function [1.00].

$$w_{it} = \omega_{0i} + t\omega_{1i} + t^2\omega_{2i} + v_{1it} \quad [1.22]$$

Where $\omega_{ji} = n_i d_j$ [1.23]

n_i is a vector of time invariant exogenous determinants of lifetime wages. The time invariant variables included in our empirical model are the education level of the consumer, background variables (education level of the mother and father and race) and time dummies for each year. d_j is vector of unknown parameters and t is the age of an individual $t = 0,1,2,...,T+1$. v_{1it} is the identically and independently distributed error term

We also need to specify the equation for initial wealth. We will assume that wealth can be approximated by the time path of lifetime investment income. That is, we assume that $r_t a_{it} = i_{it}$ where i_{it} is within-period investment income of an individual. [19]As in Cahuc and Zylberberg (2004)[20] we assume that wealth is a quadratic function of age and time invariant background variables of an individual:

$$i_{it} = \alpha_{0i} + t\alpha_{1i} + t^2\alpha_{2i} + v_{2it} \quad [1.24]$$

where t represents age and

$$\alpha_{ji} = g_i p_j \quad j = 0,1,2,3,4,...,n \quad [1.25]$$

where p_j is a vector of unknown parameters. g_i is a vector of background variables which include own education level, mother's education, father's education, race, and

[19] In this study we employ investment income as property income is missing from SLID.

[20] MaCurdy (1981) and Blundell and MaCurdy (1999) also follows similar specification.

time dummies for each year in the panel. i_{it} is investment income and v_{2it} is the identically and independently distributed error term. Note that $\alpha_{0i} = a_{i0} r$, which is initial wealth.

Given equation [1.12] and equation [1.14] we can rewrite [1.11] as follows:

$$\nabla_i = x_i \phi + \sum_{t=1}^{T} \gamma_t \ln w_{it} + a_{it} \zeta + b_i$$

$$\nabla_i = x_i \phi + \omega_{0i} \bar{\gamma}_0 + \omega_{1i} \bar{\gamma}_1 + \sigma_{1i} \bar{\gamma}_1 + \sigma_{2i} \bar{\gamma}_2 + \alpha_{0i} \zeta + \eta_i \qquad [1.26]$$

where $\bar{\gamma}_j = \sum t^j \gamma_t$

In reduced form, equation [1.16] can be represented as:

$$\nabla_i = G_i \Omega + \eta_i \qquad [1.27]^{21}$$

Joint estimation of [1.17], [1.22], [1.25] and [1.26] provides all the parameters needed to gauge the labor supply response to parametric changes in wages.

We can get three types of elasticity estimates from this model.[22] The first one is intertemporal labor substitution elasticity. The intertemporal substitution elasticity is given by β. The value of β would be positive if leisure is normal good.[23] This provides a direct elasticity of substitution for hours work in any two periods. It

[21] We are implicitly assuming that the error terms across the equations are independent. If they are not independent it would be difficult to stack them in [1.17].
[22] Strictly speaking we can define five types of elasticities.
[23] That is to say, if the wage increases in a certain period this increases the price of leisure. If leisure is a normal commodity, the substitution effect will lead to more work. Note that we do not have income effects here because, if there is any effect, it should operate through the marginal utility of wealth which is constant.

calculates the response to evolutionary change in wages by keeping the marginal utility of wealth constant. Using the Slutsky equation, we can easily derive the own price and cross price elasticities. It is very useful to differentiate the intertemporal elasticity (β) the uncompensated own price effect ($\beta + \gamma_t$) and the uncompensated cross price effect (i.e. γ_t).

The compensated own price elasticity is

$$\left.\frac{\partial h_{it}}{\partial w_{it}}\right|_{u_i=u} = \left.\frac{\partial h_{it}}{\partial w_{it}}\right|_{a_0=a} - \frac{\partial h_{it}}{\partial a_{0i}} h_{it}$$

$$\frac{\partial \ln h_{it}}{\partial \ln w_{it}} = \left.\frac{\partial \ln h_{it}}{\partial \ln w_{it}}\right|_{a_0=a} - \frac{\partial \ln h_{it} w_{it}}{\partial a_{i0}} h_{it}$$

$$= \beta + \gamma_t - \zeta h_{it} w_{it}$$

And the cross compensated elasticity is

$$\left.\frac{\partial \ln h_{it}}{\partial \ln w_{jt}}\right|_{u_i=u} = \left.\frac{\partial \ln h_{it}}{\partial \ln w_{jt}}\right|_{a_0=a} - \frac{\partial \ln h_{it}}{\partial a_{i0}} h_{it} w_{jt}$$

$$= \gamma_t - \zeta h_{it} w_{it}$$

The compensated elasticities are useful to compare the behaviour of different categories of the labor force. For instance we can compare the behaviour of immigrants and natives with different wage profiles but the same life cycle utility. The parameter γ_t measures the labor supply response to parametric change in the wage profile. It is imperative we remind ourselves that compensated elasticity and intertemporal elasticities may not be the same. Despite this clear theoretical distinction it is not uncommon to see confusion in the literature in interpreting the coefficients of wages in the life cycle labor supply model. Assuming leisure is normal in all time periods $\beta > \beta + \gamma_t - \zeta h_t w_t > \beta + \gamma_t$. Moreover, while both the intertemporal elasticity and the compensated elasticity are positive the uncompensated elasticity could be either positive or negative. If the wealth effect is zero all of these elasticities reduce to β.

1.3.. Methodology: Estimation Strategy

Since we are interested in studying the sensitivity of labour supply estimates to econometric estimation we will employ different variants of nonlinear panel data regression models. In the linear panel fixed effects model estimation is made easier by time demeaning variables. In this case the constraint or person-specific effects are wiped out. Unfortunately in the nonlinear panel data case there is the notorious incidental parameter problem that will force us to estimate a huge number of constants. However, the problem is inherently statistical but not practical. We will discuss this in a sequel.

It is not the aim of the current study to provide detailed proofs of the econometric models to be employed. Since the detailed proofs of the models that are employed can be found in standard econometric text books such as Wooldridge (2002), Amemiya (1985) Greene (2003) and Baltagi (2005), we will develop the models without going into the details of proofs.

1.3.1. Tobit model

Basic Model

Consider the following estimable version of equation [1.17]. Rewrite expression [1.17] for each individual man as follows

$$\ln h^*_{it} = \mu x_{it} + \beta \ln w_{it} + \nabla_i + u_{it} \qquad [1.28]$$

where $\ln h_{it}^{*}$ is the latent hours worked, $\mu = \beta\alpha$, $u_i(t) = -e_i(t)\beta$ and $\nabla_i = \beta[\ln \kappa_i - \ln \Phi_2 + f_i]$, which is the person-specific effect from equation [1.16] and x_{it} and μ are conformable vectors. The list of wage predictors included in equation [1.18] includes years of schooling, tenure (in years), tenure squared , race, regional dummies which are supposed to capture the variations in the labor markets across provinces and price differences, and dummies for each year in the panel[24]. ∇_i contains any individual heterogeneity in preferences or skills and the like which can be taken to be constant over time t. As was clear from the theoretical discussion above, the wealth effect and non-wage income effects will only enter into the picture through this individual effect. u_{it} is the idiosyncratic error structure from equations [1.17] to [1.26] and are assumed to be identically and independently distributed with $N(0,\sigma^2)$. If we allow for the possibility of a corner solution we may rewrite [1.28] as

$$\ln h_{it} = \begin{cases} \mu x_{it} + \beta \ln w_{it} + \nabla_i + u_{it} & \text{if } \ln h_{it}^{*} > 0, \\ 0 & \text{otherwise.} \end{cases} \qquad [1.29]$$

We assume that there is random sample of N men over T periods of time which may be balanced or unbalanced. In one case where the sample is nonrandom (such as sample selection problem) our estimation strategy will have to be adjusted (Wooldridge 1995 and 2002). We will discuss more about this in the sample selection model. However it is essential to assume that $E[u_{it} \mid \nabla_i, x_{it}] = 0$. In other words, we assume that the error term is independent of the regressors and individual effects.

[24] For notational convenience we have dropped the age variable. However, expression [1.17] suggests that age will enter as an argument in the labor supply model if $\rho \neq r$. We include age and age squared in our empirical model.

Equation [1.29] defines a multivariate correlated Tobit system if we stack the equations over time. The usage of Tobit model is tantamount to assuming that the data are truncated because of corner solutions. We will see in the sample selection model that the truncation could be an outcome of participation decision in the labor market. Note that $\ln h_{it}^{*}$ is the latent variable with its observed counterpart of $\ln h_{it}$.

There are three standard ways of estimating the Tobit model [1.29] depending on the assumptions made about the individual effects and the regressors in the system. The first basic estimation technique is the pooled cross-section model. In this model we treat each cross-section as a separate observation and we estimate standard cross section Tobit model over the NT observations. In this model the individual effects will be absorbed by the common constant term. It goes without saying that this approach is not theory consistent. However, this approach will help us to assess the extent of bias introduced by ignoring the person-specific effect. The estimates from this approach will also allow us to compare our estimates with some cross-sectional studies in the literature.

The second Tobit specification is the random effects (RE) model. The random effects and fixed effects models allow for the treatment of ∇_i. In the random effects model we assume that the unobserved individual effects are uncorrelated with the regressors in the model. The advantage of random effects over the pooled cross-section panel data is that it is relatively efficient under the some assumptions. However, given the theory that we have discussed in the previous section, it is less likely that the unobserved characteristics are to be uncorrelated with the independent variables in the model. Also, as opposed to the traditional treatment of random effects model, we model the group specific constants as randomly distributed cross-sectional units. The uncorrelated random effects model would be given by the following

$$\ln h_{it} = \begin{cases} \mu x_{it} + \beta \ln w_{it} + (\nabla + e_i) + u_{it} & \text{if } y_{it}^* > 0, \\ 0 & \text{otherwise}. \end{cases} \qquad [1.30]$$

However, theory suggests the presence of person specific effect contradicts the basic assumption of the random effects model. We have seen from equation [1.8] that the individual effect is a function of all variables in the model.[25] Thus it would be implausible to assume uncorrelated random effects model. Rather we consider the "correlated random effects" model of Chamberlain (1984). Accordingly, we need to make some additional assumptions to proceed.

As in Chamberlain (1984), consider that the individual specific effects in equation [1.28] are a linear function of the observable characteristics of the individual. That is to say, assuming that ∇_i and x_{it} have finite second moment

$$\nabla_i = \hbar_1 x_{i1} + \hbar_2 x_{i2} + \lambda_3 x_{i3} + \ldots + \hbar_T x_{iT} + \varepsilon_i \qquad [1.31]$$

[25] To be more precise it will be a function of all the variables in the model and some other non-observables as the utility function only includes some key variables. We can modify this by including a variable which represents all other relevant variables as arguments in the utility function. Thus the fixed effect will be a function of all variables which are observable and non-observable. Given that our utility function has only three arguments and the budget line, the marginal utility of wealth will be a function of the variables in the maximization problem (budget set and utility function variables.

We assume that $E[\varepsilon_i \mid x_i, u_i] = 0$ [26], $u_{it} \mid x_i \sim N(0, \sigma_u^2)$ and $\varepsilon_i \mid x_i \sim N(0, \sigma_\varepsilon^2)$. As opposed to the standard random effects model, this approach does not assume anything about $E[\nabla_i \mid x_i]$ and it allows for possible correlation between ∇_i and x_i just like the standard fixed effects model. If all the $\hbar$s are zero then there is no correlation between ∇_i and x_{it}. In this study the $\hbar$s are statistically different from zero implying some form of correlation between ∇_i and x_{it}. To incorporate this information into the system we substitute [1.31] into [1.30] to get

$$\ln h_{it} = \begin{cases} \mu x_{it} + \beta \ln w_{it} + \hbar_1 x_{i1} + \hbar_2 x_{i2} + \hbar_3 x_{i3} + \ldots + \hbar_T x_{iT} + \varepsilon_i + u_{it} & \text{if } \ln h_{it}^* > 0, \\ 0 & \text{otherwise.} \end{cases} \quad [1.32]$$

The system in [1.32] has a normally distributed error component[27] and it is still a system of Tobit equations. What is evident from [1.32] is that all values of x enter the current labor supply via their correlation with ∇_i. We can write [1.32] in a reduced form as follows

$$y_{it} = \begin{cases} \pi_{t1} x_{i1} + \pi_{t2} x_{i2} + \pi_{t3} x_{i3} + \ldots + \pi_{tj} x_{iT} + \beta \ln w_{it} + a_{it} & \text{if } y_{it}^* > 0, \\ 0 & \text{otherwise.} \end{cases} \quad [1.33]$$

where $a_{it} = \varepsilon_i + u_{it}$ is normally distributed. Equation [1.33] implies the following restrictions, using matrix form with Π being the matrix of reduced form coefficients (ignoring the coefficient of wages, β, for convenience)

---

[26] This formulation is very helpful, mainly if we rule out any serial correlation in $\{u_{it}\}$ conditional on (x_i, c_i). However, given that the time period in our sample is only 6 years we may expect some form of correlation between ε_i and x_i. We also expect some form of serial correlation among the x_i.

[27] Since u_i and ε_i are normally distributed, their sum should also be normally distributed as the sum of normal distribution is a normal distribution.

$$\Pi = \mu I_T + i_T \hbar \qquad [1.34]$$

where i_T is a T-vector of 1s and I_T is T-dimensional identity matrix.

This approach of estimating the random effects model warrants discussion because of its advantage in minimizing the infamous incidental parameter problem inherent in these types of models. Even though this approach of estimation is not our preferred model it does try to control for individual heterogeneity and our preferred model resembles this approach in spirit. Heckman and MaCurdy (1980) indicate that the presence of the individual fixed effects component ∇_i in [1.30] results in a potential incidental parameter problem. In [1.30] we need to estimate $N+T$ parameters which imply that the number of parameters to be estimated increases without bound with an increase in the sample size. Given that our sample contains an unbalanced panel of more than 30,000 it is not hard to gauge the nature of the problem. In the linear panel data models we can get around this problem by employing fixed effects model where time demeaning of the variables in the model removes the fixed effects in the model. Unfortunately in nonlinear panel data models (such as ours), we are forced to estimate a huge number of constants. It needs to be emphasized that the problem is not practical, but inherently statistical. Given fixed time length, not only the estimates of ∇_i will be inconsistent and biased but also (μ, β) will also be inconsistent and biased. The problem is that (μ, β) is a function of ∇_i. The advantage of the "correlated random effects" model is that the conditional assumption about the distribution of ∇_i allows us to reduce the number of parameters to a constant number which does not grow with sample size. Equation [1.34] shows that the number of parameters in the system is $K+TK+1$. Increase in sample size gives consistent estimates of Π and hence we get consistent estimates of (μ,β) and $\hbar$ for a fixed time T. Thus we see that the Chamberlain approach allows us to get around the incidental parameter problem but at a cost of additional assumptions about the distribution of ∇_i. This model could be estimated by quasi-maximum likelihood or conditional likelihood [28].

The third specification is the fixed effects (FE) model. In this variant of the estimation strategy we can directly tackle, at least theoretically, the unobservable individual heterogeneity. The advantage of this approach over the random effects estimation is that we are not required to make orthogonality assumption about ∇_i directly. The disadvantage is that the number of parameters to be estimated increases with sample size. This approach required T to be very large to get consistent estimates of ∇_i. As was discussed above with fixed T, as is the case in our sample, ∇_i is not only inconsistent but also it contaminates (μ,β) with inconsistency and bias.[29] This might introduce the incidental parameter problem but we hope that the incidental parameter problem does not significantly affect our results. Greene (2001) argues that the extent of the bias is small when T is larger than 2. A comparison of the random effects model and the fixed effects model should provide a good sense of the problem. If the results from these two approaches do not differ significantly, it may be the case that the incidental parameter problem is not very serious. However, it is also possible that the RE and FE are similar because the incidental parameter problem is very serious but is offset by the bias arising from $E[\nabla_i \mid x_{it}] \neq 0$.

---

[28] It is also possible to employ minimum distance estimation (Wooldridge, 2004; Jakubson, 1988)

[29] See Chamberlain(1984), Wooldridge(2002) or Greene(2003) for detailed discussions and proofs.

1.3.2.. Sample Selection Model

By definition, equation [1.17] represents the life cycle labour supply decision of every man of working age whether or not he was working at the time of the survey. However, because we can only observe the number of hours worked for the employed, we select our sample on the basis of participation in the labor market.[30] To deal with this infamous problem of nonrandom sampling we propose the following panel data[31] structure for the labor supply equation:

$$\ln h^*_{it} = \mu x_{it} + c_i + u_{it}, \qquad i = 1,\ldots,n \qquad t = t_i,\ldots,T_i{}^{32} \qquad [1.35]$$

$$s^*_{it} = z_{it}\psi + \varsigma_i + v_{it} \qquad v_{it} \mid z_i \sim N(0,1) \qquad [1.36]$$

$$s_{it} = 1 \quad \text{if} \quad s^*_{it} > 0 \qquad [1.37]$$

$$y_{it} = (y^*_{it})(s_{it}) \qquad [1.38]$$

where $\ln h^*_{it}$ is latent hours worked with observable counterpart $\ln h_{it}$ and s^*_{it} is latent the participation decision with observable counterpart s_{it}. Equation [1.35] is the labor supply function which is the equation of interest and equation [1.36] is a reduced form for the propensity to participate in the labor market. x_i and z_i contain vectors of exogenous individual characteristics including age, years of schooling, health, and imputed wages. It is conceivable that most of the variables that enter the wage equation will also determine participation in the labor market. In our empirical models we will impose some standard exclusion restrictions. The vectors of unknown parameters are μ and ψ and the random error terms are u_{it} and v_{it} with

[30] This selection problem has a long history in the literature (Gronau 1974, Lewis 1974, Heckman 1978, Wales and Woodland 1980, Vella 1998). Our estimation technique should account for the non-random nature of the sample as failure to do so may yield inconsistent estimates (Heckman 1978).
[31] See Baltagi (2005), Hsiao(2003), Greene(2003), Wooldridge(2002) for a discussion of panel data modeling. See Vella (1998) for a readable survey of sample selection models.
[32] Note also that $t = t_i,\ldots,T_i$ implies that the panel structure could be unbalanced.

$E(u_{it}/\upsilon_{it}) \neq 0$. Note that, for convenience, μ now includes β, the coefficient of $\ln w_{it}$, unlike the previous section. We assume that (u,υ) is independent of z_i (where z_i might contain elements of x_i[33]), and ∇_i and ς_i are individual fixed effects which are time invariant.

There are three traditional variants of the sample selection panel data model. The first specification is the pooled sample selection model. This model is nothing but a cross sectional sample selection model. The second variant is the fixed effects model and the third specification the random effects model.

Define the following deviation forms of the variables as

$$\ddot{x}_{it} = x_{it} - \frac{\sum x_{ir}s_{ir}}{\sum s_{ir}} \text{ if } \sum s_{ir} > 0 \qquad [1.39]$$

$$\ln \ddot{h}_{it} = \ln h_{it} - \frac{\sum \ln h_{ir}s_{ir}}{\sum s_{ir}} \text{ if } \sum s_{ir} > 0 \qquad [1.40]$$

Hence the fixed effects estimator for unbalanced ($\mu_{FE}(u)$) and balanced panel ($\mu_{FE}(B)$) are as follows

$$\mu_{FE}(u) = \left(\sum_{i=1}^{N}\sum_{i=1}^{\tau} \ddot{x}_{it}'\ddot{x}_{it}s_{it}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{i=1}^{\tau} \ddot{x}_{it}'\ln \ddot{h}_{it}s_{it}\right) \qquad [1.41]$$

$$\mu_{FE}(B) = \left(\sum_{i=1}^{N}\sum_{i=1}^{\tau} \ddot{x}_{it}'\ddot{x}_{it}d_{i}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{i=1}^{\tau} \ddot{x}_{it}'\ln \ddot{h}_{it}d_{i}\right) \qquad [1.42]$$

where $d_i = \{(\Pi_{t=1}^{\tau} s_{it}) = 1\}$

[33] It is conceivable that most of the variables which influence participation in the labor market will also affect wages; hence we can expect that Z_i contains most of elements of x_i. Ideally, we would like to have some exclusion rule here for efficiency.

For the consistency of our fixed effects estimator we require that $E[\ddot{u}_{it} \mid \ddot{x}_{it}, s_{it}] = 0$. In other words, the sample selection should operate via individual fixed effects which will be removed by time demeaning the variables. Note that this assumption will break down if the selection process is operating through some time variant unobservable.

The third standard type of specification is the random effects model. To get the random effects model we follow Verbeek and Nijiman (1992). Define $\ln h_{it} = (\ln h_{i1},...,\ln h_{iT})'$, $x_{it} = (x_{i1},...,x_{iT})'$, and $u_{it} = (u_{i1},...,u_{iT})'$. Assume that all the variables in the labor force participation equation are available and define the number of units $s_{it} = 1$ as T_i and define $T_i x T_i$ matrix R_i transforming $\ln h_{it}$ into the T_i dimensional vector of observed $\ln h_i^o$. Note that matrix R_i is obtained by deleting the rows of the T dimensional identity matrix corresponding to $s_{it} = 0$. Defining the unit vector I, the variance covariance of the error term in equation [1.35] can be written as $\Omega = \sigma_c^2 ii' + \sigma_u^2 I$. Given this random error structure, the random effects estimator for the unbalanced and balanced panel case are :

$$\mu_{RE}(u) = \left(\sum x_i^{o}{}' \Omega_i^{-1} x_i^o \right)^{-1} \left(\sum x_i^{o}{}' \Omega_i^{-1} \ln h_i^o \right) \qquad [1.43]$$

$$\mu_{RE}(B) = \left(\sum x_i^{o}{}' \Omega_i^{-1} x_i^o d_i \right)^{-1} \left(\sum x_i^{o}{}' \Omega_i^{-1} \ln h_i^o d_i \right) \qquad [1.44]$$

where $\Omega_i = R_i \Omega R'_i$ and $x_i^o = R_i x_i R'_i$

For the consistency of [1.43] and [1.44], we need $E[u_{it} + \nabla_i \mid x_{it}, s_{it}] = 0$. Thus, the random effects estimator will be inconsistent if the selection is operating either through the individual or/and the idiosyncratic error.

These are the traditional panel data estimators but, as we discussed above, it is most likely that the assumptions required for consistent estimation by these models will be violated. In light of these, we follow Wooldridge (1995) for testing and correcting the sample selection problem.[34] Under some general mild conditions, this model will produce consistent estimates of μ . This will be our preferred model but, since we are interested in studying the sensitivity of labour supply estimates to econometric estimation, we will also report results from the estimation of the standard linear panel data models and variants of the Tobit panel data models.

Testing and correcting for sample selection

There have been a number of suggestions on the detection of sample selection bias in panel data models (Wooldridge 1995, Verbeek and Nijiman 1992 1996, Vella 1998, Vella and Verbeek 1994) [35]. We follow Wooldridge (1995) for testing and correcting sample selection problem in our data. The basic premise of this approach is that it parameterises the conditional expectation required for the consistency of the pooled estimator:

$$E[\nabla_i + u_{it} \mid x_i, s_{it} = 1] = E[\nabla_i \mid x_{it}, s_{it} = 1] + E[u_{it} \mid x_{it}, s_{it} = 1] = 0 \quad \forall t \qquad [1.45]$$

[34] The basic testing procedure is similar to that of Ridder (1990) Nijiman and Verbeek (1992) and Vella and Verbeek (1994), who base their approach on simple variable addition tests. This section follows Rochina-Barrachina and Dustmann (2007) in notation.

[35] Wooldridge (1995) claims that the estimation of the selection model is of second order importance as the objective is to derive a test and we are not interested in the selection equation parameters *per se*. We will test this claim concerning whether the different specifications of the selection make any difference or not.

That is to say, the approach parameterises assumption [1.45] and adds the parameters as additional regressors in the main equation. The key assumptions are:

A1: the conditional expectation of ς_i *given* $z_i = (z_{i1},...,z_{iT})$ *is linear.* $\varsigma_i = z_{i1}\kappa_1 + ... + z_{iT}\kappa_T + \partial_i$ where ∂_i is random component.

A2: the errors terms in the selection equation, $\Diamond_{it} = \varsigma_i + v_{it}$ *are independent of* $\bar{z}_i$ *and normal* $(0,\sigma\ )$, *where* $\bar{z}_i = (x_i, z_i^+)$ *with* $x_i = (x_{i1},...,x_{iT})$ *and* z_i^+ *containing the non-overlapping elements in* z_i

A3: $E[u_{it} \mid x_i, v_{it}] = E[u_{it} \mid v_{it}] = \eta_t v_{it}$, $t = 1,2,...,T$. This is an assumption about joint normal distribution of the error terms in equation [1.35] and [1.36].

A4: $E[\nabla_i \mid x_i, v_{it}] = E[\nabla_i \mid x_{i,} v_{it}] = L[\nabla_i \mid 1, x_{i,} v_{it}]$, $s_{it} = 1[x_{it}\delta + \nabla_i + a_{it} > 0]$. This assumption implies that the fixed effect is a linear function of x_i and the error term in equation [1.36]. The violation of this assumption leads to inconsistent estimates (Wooldridge 2002). This can be represented as follows

$$E[\nabla_i \mid x_i, v_{it}] = \pi x_i + \phi v_{it} \qquad [1.46]$$

Where with the help of the law of iterated expectations, we can write [1.46] as follows

$$E[\nabla_{it} \mid x_i, v_{it}] = \pi\ x_i \qquad [1.47]$$

Hence [1.35] could be rewritten as

$$E[\ln h_{it} \mid x_i, v_{it}] = \beta x_{it} + \pi x + \eta_t v_{it} \qquad [1.48]$$

Conditioned on $s_{it} = 1$ we can write [1.48] as follows

$$E[\ln h_{it} \mid x_i, v_{it}] = \beta x_{it} + \pi x + \xi_t \lambda(x_i \psi_t) \quad [1.49]$$

It is possible to obtain a consistent estimate of μ by first estimating a labor force participation equation using a probit regression s_{it} on x_i for each panel j and saving the inverse Mills ratios ($\hat{\lambda}_{it}$) for all i and t. The next crucial step is to run pooled OLS regression using the selected sample: $\ln h_{it}$, on $x_{it}, x_i, \hat{\lambda}_{it}, d_t\hat{\lambda}_{it},..., d_T\hat{\lambda}_{it}$ for all $s_{it} = 1$ where $d_t - d_T$ are the time dummies. Wooldridge (1995) shows that we can get consistent estimates of [1.49] using either OLS or the minimum distance estimator.[36] Note this approach allows for the correlation between the unobservables in the selection equation, v_{it} and the unobservables in the wage offer equation (∇_i, u_{it}) so that the selection process might operate via both the error term from the main equation u_{it} and the unobservable individual effect ∇_i. However, we need to adjust the standard errors for general hetroscedasticity and autocorrelation and for the first stage estimation.

Wooldridge (1995) adopts a variable addition test to detect the presence of a sample selection problem. Under the null hypothesis of $E[u_{it} \mid x_i, s_{it}, \nabla] = 0, \ t = 0,1,...,T$ the inverse Mills ratios for each cross section from equation [1.36] should not be significant in an equation estimated by the fixed effects method. To elaborate, in the first step estimate the inverse Mills ratio from [1.36] for each cross-section. The next step is to estimate equation [1.35] using fixed effects model on the selected sample by including the inverse Mills Ratios as additional regressors and then testing for sample selection using a t test on the inverse Mills Ratios in this fixed effects model.

[36] See Wooldridge (1995) for the detailed derivation of the model.

Wooldridge (1995a) shows that the limiting distribution of the t statistic under the assumption $E[u_{it} \mid x_i, s_{it}, \nabla] = 0$ is not affected by the estimation adopted for the participation equation [1.36]. As long as the standard errors are robust and adjusted for hetroskedasticity, we can trust the student t test.

1.4. Data

The data employed for the estimation of the life cycle labor supply models are drawn from the second panel of Survey of Labor and Income Dynamics (SLID). Our sample contains men between the ages of 21 to 65. We have data for 6 years: 1996, 1997, 1998, 1999, 2000 and 2001. Unlike most studies in the literature (e.g., MaCurdy 1981, Osberg and Phipps 1993, Reilly 1994), we do not restrict our sample to those who reported non-zero number of hours. We argue that such a restriction potentially causes sample selection bias, yielding inconsistent parameter estimates. The SLID is a continuing panel of Canadian households that began in 1993. It combines and replaces the former Labor Force Activity Survey, an intermittent series of short panel surveys conducted during the 1980s, with the Survey of Consumer Finance, an annual cross-sectional survey. It therefore provides detailed information on labour supply, wages, and demographic characteristics as well as valuable tax record information on income from various sources. The SLID design is a series of overlapping 6-year panels, with a new panel enrolled every three years.

It is common in the literature to calculate the annual number of hours worked by multiplying the number of weeks worked and the usual number of hours worked per week. Furthermore, the hourly wage is calculated by dividing reported annual earnings by estimated annual hours worked. These calculations usually introduce measurement bias and yield inconsistent parameter estimates. Any measurement error in the annual number of hours worked will carry over to the hourly wage rate (Ziliak and Kniesner 1999, Mroz, 1987) and affect our estimates (Conway and Kniesner 1999). In SLID the annual number of hours and the composite hourly wage are calculated from extensive interviews with detailed questions on each job and payment that individuals get in the survey period. Questions are posed on the number of jobs held and the hours worked pay by pay, the number of weeks worked, and the number of weeks absent from work. Respondents are encouraged by a detailed questionnaire to retrieve information or, in the case of income tax records, to permit access to their personal files through the Canada Revenue Agency to produce as reliable information as possible on hours worked and the hourly wage.

The wage predictors include years of schooling, years of experience and its square, job tenure and its square, a dummy variable for visible minority, regional dummies and time dummies. We introduce the tenure variable, defined as the number of years worked with the current employer, to improve the instrumentation of wages and the precision of the structural parameter. We impute wages using the sample selection model indicated in equation [1.9]. Other income is calculated as the difference between individual total annual wages and total household income.

The health variable is self-reported health status: respondents are asked to rate their health as excellent, very good, good, fair or poor. We also generate a dummy variable by collapsing the ordinal health measure into two categories: good health if the self assessed health is excellent, very good or good; and poor health, if the self assessed health is fair health or poor health. However, because of the established research, our health variable is instrumented. We follow the life cycle consistent health model.[37] The health predictors include the imputed wage, age and its square, time dummies and the functional limitation. The age variables are expected to proxy depreciation in the health stock and the time dummies are expected to capture technological changes and other changes in health production over time. The health equation is estimated using a fixed effects logit model,[38] but we also use the instruments without including imputed wage to explore the sensitivity of our results. The summary descriptive statistics of the key variables are provided in Table 1.

------------------Table 1 here------------------------

1.5. Estimation results

In this section we present the estimation results using the different models discussed above.[39] We first present the results from a sample of individuals who report non-zero number of hours using standard panel data models. This serves a double purpose. Firstly, we will be able to compare our results to the many studies in the literature because most of the old literature that restrict analysis to a sample of employed men. Secondly, these results can be used as a benchmark to explore the extent of bias introduced by ignoring the sample selection problem.

Table 2 presents estimation results from traditional panel data models using the selected sample only. The first column of Table 2 reports the pooled OLS results. The second and third columns report the results for the random effects (RE) and fixed

[37] Dustmann and Windmeijer (2000) derive life cycle consistent demand for health following the same approach that we adopt to derive our life cycle labor supply model.
[38] This might introduce the incidental parameter problem but we hope that the incidental parameter problem does not significantly affect our results. Greene (2003) argues that the extent of the bias is small when T is larger than 2.
[39] Most of the computation is done using LIMDEP and STATA version 10.

effects (FE) models. As Table 2 demonstrates, three of these panel data models produce remarkably different estimates. An LM test rejects the null hypothesis of variance constancy, implying that the classical regression model with a single constant is inappropriate for our data. This does not necessarily imply that the RE model is the best model to describe the generated data because there is a competing FE specification. In fact, a Hausman test with a null hypothesis of no systematic difference between the RE and the FE estimates ($H_0 : \beta_{FE} = \beta_{RE}$) is rejected in support of the FE model. Note also that the preferred specification according to our economic theory is the FE model. A comparison of the implied intertemporal labour elasticity reveals that the FE model produces the largest estimate followed by the RE model. More specifically, the implied intertemporal labor substitution for the pooled, RE and FE model are 0.23, 0.27 and 0.31 respectively. These results are very similar to those reported in the literature that are restricted to a sample of employed men. Altonji (1986) presents the survey of labor supply estimates for men in the USA and reports the intertemporal labor substitution elasticity to be between 0 and 0.35.

----------------------Table 2 here------------------------

We now turn to the standard panel data models without excluding those who reported zero hours. Table 3 summarise the results from the standard OLS, RE and FE estimates without taking into account the truncation in the data. In Column 1 and 2 we present the standard OLS and Random Effects models. The results are similar but not the same. For example, the implied intertemporal elasticity of substitution are 0.48 and 0.39 for the RE and OLS respectively. However, both the RE and pooled OLS specifications are not consistent with our theory because they ignore individual heterogeneity. Column 3 of Table 3 presents the FE estimates. A Hausman test of the correlation between individual heterogeneity and the regressors rejects the null hypothesis of $H_0 : \beta_{FE} = \beta_{RE}$. Table 3 demonstrates that the FE estimates are lower relative to the pooled OLS and the RE point estimates. For example, the coefficients of the imputed wage are 351.72, 707.98 and 591.36 for the FE, RE and pooled OLS respectively. One explanation for the lower values of the FE estimates might be that the partial elimination of the upward bias that could be represented by the fixed effect component of the selection mechanism. Of course, the traditional OLS and RE are biased because they fail to account for the individual heterogeneity if $E[\nabla_i \mid x_i] \neq 0$.

---------------------------------Table 3 here---------------------------

So far, there has not been an explicit account of those with zero hours reported in our sample. In this section we treat zero hours as an outcome of a corner solution and estimate different versions of panel data Tobit models. Table 4 presents the results from the three standard Tobit models: pooled, RE and FE models. The second column of Table 4 contains the results from the pooled Tobit model. This model ignores the panel structure and treats the data as extended cross-section. However, this form of estimation does not exploit the panel nature of the data and is inconsistent with our economic theory. Comparison of these results with the cross-sectional Tobit models employed in the literature reveal that the implied intertemporal substitution elasticity is in the range of those reported in the literature. The implied intertemporal substitution elasticity is 0.41.

Column 1 of Table 4 reports the results of the uncorrelated RE Tobit model. This model assumes that the idiosyncratic error terms are independent of the regressors, which is not consistent with our economic theory. Table 4 demonstrates that the RE model estimates are significantly different from that of the pooled OLS model. Also note that most of the estimates from the RE model are smaller compared to that of the pooled OLS. For example, the coefficient of "imputed wage" has decreased by about 40 percent, which implies a Frisch labor supply elasticity of 0.25 but statistically insignificant at less than 10 % level. The correlated RE (Chamberlain type) was also estimated but it gave results similar to that of the FE specification that is reported in Column 3. For this reason, we only report the results of the FE model.[40]

Column 3 of Table 4 presents the point estimates from FE Tobit model. The FE approach models the fixed effects directly. In this model the fixed effects are assumed to be uncorrelated with the regressors in the model. This formulation is theory consistent. Table 3 demonstrates that the FE estimates are quite different from the RE model estimates. Unfortunately, there is no statistical test available to compare the RE and FE estimates as they are based on different statistical grounds. The best test available is economic theory which is in favour of the FE model. But examination of some of the key variables would facilitate the comparison. For example, the coefficient of "Education" is positive and significant in the pooled and RE Tobit models but it changes to a statistically significant negative coefficient as we move to the FE model. The same is true with some of the marital status indicators. The implied Frisch elasticity from the FE model is calculated to be 0.34 which is about 0.07, lower than that of the pooled Tobit model estimate. This could be explained by

[40] Note that the computation of the standard errors of the marginal effects is problematic and LIMDEP calculates the marginal effects with respect to the latent number of hours and for this reason we report the parameter estimates.

the partial elimination of upward bias which could have been introduced by any form of incidental truncation.

Now we turn to the sample selection models. Before we start discussion of the results, however, we present the results of our sample selection tests. The test results for the presence of sample selection bias in our data are reported in Tables 5 and 6. Table 5 presents estimates of the coefficients of the inverse Mills ratios (λ_t) from cross-sectional selection equations. As expected, all the estimates are negative and statistically significant at less than the 1% level. We follow Wooldridge (1995, 2002) to test for the sample selection problem. The first column of Table 6 reports the results of a FE model on the selected sample and most of the coefficients of λ_t turn out to be statistically significant at less than 5 percent level of significance. A joint test of significance rejects the null hypothesis that the λ_t are jointly zero, $H_0 : \lambda_t = 0, t = 1,...,T$. These tests provide sample evidence that our panel data is contaminated with some sort of sample selection problem.[41]

Having established the prospect of sample selection problems, we next discuss the results from the sample selection models. The sample selection model of labor supply effectively separates the participation decision from the choice of hours of labor supplied, conditioning on participation. We estimate the pooled OLS and FE versions of the sample selection model. The last two columns in Table 4 report the results of pooled and FE sample selection models respectively. Comparison of the standard

[41] Since attrition bias could be a possible explanation here, we tested for it. Following Vella and Verbeek(1998) we conduct a variable addition test for attrition by including s_{t-1} as a regressor in equation [1.23] with and without πx_i. The null hypothesis of attrition bias was decisively rejected after controlling for sample selection bias.

pooled OLS with that of the FE estimates reveals that there is sharp difference on the estimated parameters. Note that the standard sample selection model is theory inconsistent and the fixed effects model is more attractive from the theoretical point of view as it controls explicitly for individual effects.

----------------Table 4 here--

According to Table 4 the Frisch elasticity for the FE and pooled OLS sample selection models are calculated to be 0.32 and 0.25 respectively. This could imply that the bias introduced by $E[c_i \mid x_{it}] \neq 0$ is nontrivial. We observe also a sharp difference in the parameter estimates for the marital status variables in both models. While the pooled sample selection model implies that widowed, separated and divorced men work more hours than single unmarried men, both the FE Tobit and FE sample selection models imply the opposite.

However the traditional FE sample selection model is consistent only if the selection process is not correlated with the idiosyncratic error. The hope is that the selection problem will be wiped out by time demeaning the variables during the estimation process as indicated in expression [1.36]. If the selection problem is also correlated with time variant unobservables, then we need to be concerned about the consistency of our estimates.

---Table 5 here-----------------------------

Having established the presence of the sample selection problem, the next step is to address the problem using some sort of correction mechanism. The second column of Table 6 reports the results of the Wooldridge's (1995) estimator (WE). According to the results in Table 6 we see stark differences between the FE estimator, which includes the λ_t, and that of the WE. To facilitate our comparison we compare the implied elasticities from both models. The implied intertemporal elasticity from the FE model is 0.36 and from that of WE is only 0.16. Note that the WE produces a value of β which is lower than that of the traditional FE sample selection model by 0.09. But the WE produces a value of β which is less than 50 percent of β implied by the FE on the selected sample reported in Table 6. The upward bias in the FE model on the selected sample with the λ_t as additional regressors could be a support for Wooldridge's (1995) claim that this FE model cannot be consistent under any circumstance. However, it is interesting to note the similarity between the intertemporal substitution elasticity implied by FE sample selection model (β=0.32) and that of the FE model on the selected sample (β=0.36) with the λ_t as additional regressors. This is not unexpected as both models control the selection problem partially.

---Table 6 here------------------------------

Our results demonstrate that other parameter estimates are also sensitive to the econometric modelling adopted. Consider the effect of the number of children in the family on labour supply. Contrary to what has been documented by most researchers we find a negative effect of fatherhood on labor supply in most of the models which do not account for sample selection problem. In most of the cases the parameter estimates for the number of children are negative and significant at less than the 1% level. This might be an indication of a change in the role of men in home production and is not without any empirical support. Vere (2005) finds that married men's labour supply falls with a second child in 1980s and 1990s in the USA. Vere argues that the home intensity effect dominates over the specialization effect and most of the adjustment with fatherhood lies with earnings rather than labor supply. However once we control for individual heterogeneity and correct for sample selection bias the WE model implies the opposite. This seems to agree with the majority view. For example, Lundberg and Rose (1999) report a positive effect of fatherhood on labor supply and earnings.

Another variable worth examining is health. The coefficient of health was negative in all of the models which do not correct for sample selection bias, implying that men with good health work less than men with poor health. However in the WE model the coefficient of health is positive and statistically significant at less than the 1% level. This is indeed very interesting because the effect of health was negative but insignificant in the fixed effects Tobit and other traditional panel data sample selection models at the 10% level. This further reinforces Currie and Madrian's (1999) claim that the effect of health on labor supply is sensitive to the instrumentation and estimation technique employed. The positive relationship makes more sense as Figure 1 shows that the number of hours worked is higher for healthy people than that of the respondents who reported poor health. Overall these findings demonstrate that the point estimates of a labor supply model are sensitive to the econometric model and that accounting for sample selection bias matters.

-- Figure 1 here --------------------------------------

Finally we want to compare our results with those in the literature. Table 8 summarizes some previous related studies on the responsiveness of men's labor supply. We mainly summarize the Frisch elasticity, the uncompensated wage elasticity and the wealth elasticity. Of course, there are obvious differences in the estimation strategies and data employed in these studies and any strict comparison should take this into account. Overall our results are not very different from previous studies and are in the range of the survey of American studies in Altonji (1986). The implied Frisch elasticity from the preferred model (WE) is similar to that of MaCurdy (1981) which employs similar specification framework but a different estimation strategy. The advantage of this study is that it controls for individual fixed effects and the sample selection problem, which are ignored in many studies. However, the findings in this paper differ from the sole Canadian study by Reilly (1994). Reilly reports the intertemporal substitution elasticity with respect to annual weeks worked to be 0.6 and estimates the elasticity with respect to annual hours worked to be as large as 0.9 but statistically insignificant. Our study reveals that the value of β (the intertemporal elasticity with respect to annual hours worked) is statistically different from zero in a range of 0.16 to 0.48. The differences could mainly be attributed to the sample selection problem inherent in his modeling. Note that the upper bound for our estimates comes from the models which do not account for individual heterogeneity and sample selection bias and the lower bound for our estimates comes from the WE model which controls for both individual heterogeneity and sample selection bias. Thus this study finds that failing to control for individual heterogeneity and sample selection problem produces upward biased estimates of intertemporal substitution elasticity.

Wage Shift Parameters[42]

The previous section estimated the intertemporal elasticity substitution with respect to labour supply. In this section we estimate the wage shift parameters. These estimates will help in providing additional information about the responsiveness of hours worked to wage changes. The estimation of the parameters which measure the parametric change in the wage profile is not straightforward since we need to obtain the observable counterparts by manipulating or exploiting the observable components of [2.17]. From equation [2.17] it is possible to predict c_i as follows.

$$E[\ln h_{it}] = E[\nabla_i + \delta t + \mu x_{it} + \beta \ln w_{it} + u_{it}]$$

$$E[\ln h_{it}] = E[\nabla_i] + E[\delta t] + E[\mu x_{it}] + E[\beta \ln w_{it}] \qquad [1.50]$$

$$\hat{\nabla}_i = \frac{1}{J}\sum_{j=1}^{J}(\ln h_{it}) - \hat{\delta t} - \mu x_{it} - \hat{\beta}\ln w_{it}) \qquad j = 1,\ldots,J$$

where the value of j stands for a cross-section of a panel and in our case it assumes values of 1-6 corresponding to each year of the panel. Note that, from the law of large numbers, asymptotically $E[\hat{\nabla}_i] = \nabla_i$. To get an observable counterpart of the wage growth equation, we use:

$$d_k \ln w_{ij} / k = \omega_{1i} + \omega_{2i}[2t(j) - k] + d_k v_{ij} / k \qquad [1.51]$$

[42] This section closely follows MaCurdy (1981) and Dustmann and Windmeijer (2000) in notation.

$$\frac{1}{2j-k-3}\left[\frac{d_k \ln w_{ij}}{k} - d_k \ln w_{i2}\right] = \omega_{2i} + \frac{d_k v_i(j)}{(2j-k-3)k} - \frac{d_1 v_i(2)}{2j-k-3} \quad [1.52]$$

where d is a difference operator. Note that the average value of [1.52] is ω_{2i}. Hence

$$\overline{\omega}_{2i} = \frac{1}{\tau-1}\sum\left[\frac{1}{j}\{\frac{d_{j+1}\ln w_{i(j+2)}}{j+1} - d_1 \ln w_{i2}\}\right] \quad [1.53]$$

Given [1.50], from [1.52] we find

$$\overline{\omega}_{1i} = \frac{1}{\tau-1}\sum\left[\frac{d_j \ln w_{i(j+1)}}{j} - \overline{\omega}_{2i}[2t(j+1)-j]\right] \quad [1.54]$$

$$\overline{\omega}_{0i} = \frac{1}{\tau}\sum\{\ln w_{ij} - \overline{\omega}_{1i}t(j) - \overline{\omega}_{2i}[t(j)]^2\} \quad [1.55]$$

It can easily be shown that, using the law of large numbers, $E[\overline{\omega}_{ih}] = \omega_{hi}$. We can employ the same strategy to get observable counterparts of the initial wealth by replacing $\ln w_{ij}$ by $i_i(j)$ in (1.55] as $\overline{\alpha}_{0i} = \frac{1}{\tau}\sum\{i_i(j) - \overline{\alpha}_{1i}(j) - \overline{\alpha}_{2i}[t(j)]^2\}$. Then we can estimate the following system of simultaneous equations to find the complete parameters to gauge the wage shifters:

$$\hat{\nabla}_i = X_i\phi + \overline{\omega}_{0i}\overline{\gamma}_0 + \overline{\omega}_{1i}\overline{\gamma}_1 + \overline{\omega}_{2i}\overline{\gamma}_2 + \overline{\alpha}_{0i}\overline{\zeta} + \eta_1 \quad [1.56]$$

$$\overline{\omega}_{hi} = g_i q_h + \eta_3 \qquad \text{where } h = 0,1,2 \quad [1.57]$$

$$\overline{\alpha}_{0i} = g_i p_0 + \eta_4 \quad [1.58]$$

where g_i contains exogenous variables and q_0, q_1, q_2 and p_0 are defined as in equations [1.22] and [1.25]. The structural parameters of interest are $\overline{\gamma}_0$, $\overline{\gamma}_1$, $\overline{\gamma}_2$ and $\overline{\xi}$.

The g_i vector determines the lifetime income and wage path. The variables included in the execution of the actual empirical analysis include the education level of the individual, the mother's education, the father's education, race, and time dummies for each year in the panel. This system of equations could be estimated by two stage least square (2SLS). Of course, the standard errors should be corrected for the first stage estimation. Note that the endogenous variables are $\overline{\omega}_{hi}$ and $\hat{\nabla}_i$.

Table 7 presents the estimation results. According to Table 7 almost all of the coefficients have the expected sign but not all of the coefficients were statistically significant at less than 10 percent level of significance. We report the results for two sets of intertemporal elasticity values, β=0.16 and β=0.30. Table 7 reveals that the value of γ_t decreases with an increase with a value of β used to impute the fixed effects. This is expected as a constant uncompensated elasticity requires this inverse relationship between β and γ_t. As we discuss in the theoretical section of the paper, we need to find $\beta+\gamma_t$ to find the uncompensated elasticities. Assuming that the average working time for a typical man is 40 years, dividing $\overline{\gamma}_0$ by 40 gives an average cross uncompensated elasticity of 0.00075. If we add this average cross-uncompensated elasticity to β=0.16 we get a value of 0.15925 which implies that the average own-uncompensated elasticity is approximately 0.16. Note that, assuming leisure is normal in all time periods, the theoretical prediction is that $\beta>\beta+\gamma_t-\zeta h_t w_t>\beta+\gamma_t$. Thus the own-compensated elasticity of labor supply is in fact 0.16. This can be interpreted as follows. A 100 percent increase in wages at time t will induce a 16% increase in labor supply (hours worked) in time t but will leave the number of hours worked in all other periods unaltered.

-------------------------Table 7 here --

Table 7 suggests that ζ is negative but statistically insignificant at 10 percent, which suggests that three of the labour supply elasticities discussed in this paper are almost the same. The measure of initial wealth employed in this study is derived from investment income. This has, of course, a number of problems as it does not capture all wealth. This could be a partial explanation for the insignificant coefficient. We also experimented with property income and other family income but the results were very similar to what is reported in Table 7. Unfortunately SLID does not contain information on total assets and wealths. Overall our results imply that equating intertemporal labor substitution and the uncompensated elasticity, as is done in cross-sectional studies, would not be a bad approximation. As it can be verified from Table 8 our results are very similar to those in the literature. Conway and Kniesner (1998), MaCurdy (1981) and Tries (1990) report similar results for the USA.

Adding the value of $\bar{\gamma}_0$ to β gives the effect of a parallel shift in the wage profile of an individual over the life cycle. According to the empirical results $\bar{\gamma}_0 + \beta$ (0.16-0.03) is found to be 0.13. This can be interpreted as follows. A 1 percent increase in all wages over the life cycle leads to an increase in life cycle labor supply in all ages by 0.13 percent. Similarly, we can gauge the effect of a change in the slope of the wage profile by adding $\bar{\gamma}_1$ and $\bar{\gamma}_2$ to β. Unfortunately these slope parameters are not statistically significant at less than 10 percent level in cases where β=0.16. However, the negative values of $\bar{\gamma}_1$ and $\bar{\gamma}_2$ imply a backward bending labor supply curve.

1.6. Conclusion

This paper attempts to estimate a life cycle labor supply model for men using sample selection corrected panel data models that control for individual heterogeneity. Unlike most studies in the literature (e.g., MaCurdy 1981, Osberg and Phipps 1993, Reilly 1994), we do not restrict our sample to those who reported non-zero number of hours. We argue that such a restriction potentially causes sample selection bias, yielding inconsistent parameter estimates. Our data set is the second panel of Canadian Survey of Labor and Income Dynamics (SLID) which provides annual labour force activity, demographic and income data from 1996 to 2001. The results confirm that life cycle labor supply estimates are very sensitive to the specification and the econometric methods employed. Failure to control for individual heterogeneity, as in cross-sectional models, and sample selection produces an upward biased intertemporal labor substitution elasticity. The models which do not account, and correct, for sample selection bias produce intertemporal labor substitution elasticity ranging from 0.23 to 0.48. However, using the WE estimator after correcting for sample selection and individual heterogeneity we estimate the intertemporal labor substitution elasticity to be only 0.16. Moreover, we find the wealth effects to be statistically insignificant implying that the intertemporal elasticity is a very good approximation to both the compensated and uncompensated labor supply elasticity.

The key message of this study is that situating the labor supply decision of an individual in a life cycle setting produces meaningful parameter estimates once individual heterogeneity and sample selection are controlled. For example, our empirical results imply that a 100 percent increase in wages at a typical age would induce an increase in labor supply at that particular age by 16 percent but would leave all labour supply decisions in all other ages unchanged. On the other hand, a 100 percent increase in wages over all ages (a parallel shift of the wage profile) would induce a 13% increase in labor supply at all ages.

Many of the usual caveats still apply. The model implicitly assumes that individuals make their labor supply decisions freely. Taxes and the fixed costs of work are not taken into account. Individuals are not assumed to be liquidity constrained and there is no human capital formation that affects labor supply decisions. Furthermore, the assumptions that we adopted to project lifetime wages and initial wealth should be remembered in interpreting the results. Finally, it would be useful in future research to assess the robustness of our results to changes in the specification of the utility function.

References

Abowd, J.M., and Card, D. (1989) "On the Covariance Structure of Earnings and Hours Changes," *Econometrica,* 57, pp. 1-45.

Altonji, J. G. (1986) "Intertemporal Substitution in Labor Supply: Evidence from Micro Data," *Journal of Political Economy,* 94(2), pp.176-215.

Altonji, J.G. and Ham, J. (1990) "Intertemporal Substitution, Exogeneity and Surprises: Estimating Life-Cycle Models for Canada," *Canadian Journal of Economics*, XXIII (1), pp. 1-43.

Amemiya, T. (1985) Advanced Econometrics, Harvard University Press, Cambridge, MA.

Angrist, D. J. and Evans, W.N. (1998) "Children and Their Parent's Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, 88(3), pp. 450-477.

Baltagi, B.H. (2005) Econometric Analysis of Panel Data, 3rd Ed., Chichester: John Wiley and Sons Ltd.

Baltagi, B.H. et al. (2003) "A Panel Data Study of Physicians' Labor Supply: The Case of Norway," CESifo Working Paper Series No. 895. Available at SSRN: http://ssrn.com/abstract=388623

Baltagi, B.H. and Chang, Y.J. (2000) "Simultaneous Equations with Incomplete Panels," *Econometric Theory,* 16, pp. 269-279.

Blomquist, N.S. (1985) "Labor Supply in a Two-Period Model: the Effect of a Non-linear Progressive Income Tax," *Review of Economic Studies,* 52, pp. 515-529.

Blomquist, N.S. (1983) "The Effects of Income Taxation on the Labor Supply of Married Men in Sweden," *Journal of Public Economics*, 22:169-197.

Blundell, R. and MaCurdy, T. (1999) "Labor Supply: A Review of Alternative Approaches," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3A, pp. 1559-1695, Amsterdam: Elsevier Science.

Browning, M. and Meghir, C. (1991), "The Effects of Male and Female Labour Supply on Commodity Demands," *Econometrica*, 59(4), pp. 925-951.

Cahuc, P. and Zylberberg, A. (2004) Labor Economics, MIT Press: Cambridge, USA

Chamberlain(1984) "Panel Data," in V. Griliches and M.D. Intrrilirgator(Eds), *Handbook of Econometrics,* Volume II, pp. pp.1248-1318, Amsterdam: Elsevier Science.

Conway, K.S. and Kniesner, T.J. (1992) "How Fragile are Male Labor Supply Function Estimates?," *Empirical Economics*, 17, pp. 169-182.

Currie, J. and Madrian, B. (1999) "Health, Health Insurance and the Labor Market," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3, pp. 3309-3416, Amsterdam: Elsevier Science.

Domeij, D. and Flode'n, M. (2001) "The Labor Supply Elasticity and Borrowing Constraints: Why Estimates are Biased," SSE/EFI Working Paper Series in Economics and Finance No.480, Stockholm.

Donni, O. (2007) "On the Identification of Frisch Labour Supplies," *Economics Letters*, 95, pp. 1-6.

Dustmann, C. and Windmeijer, F. (2000) "Wages and the Demand for Health - A Life Cycle Analysis," IZA Discussion Paper No.171, Institute for Fiscal Studies, Bonn, Germany.

Dustmann, C. and Rochinna-Barrachina, M.E. (2007) "Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage Equations," *Econometrics Journal*, 10, pp. 263-293.

French, E. (2005) "The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behaviour," *Review of Economic Studies,* 72 (2), pp. 395–427.

Friedman, M. (1957) Theory of the Consumption Function, Princeton University Press.

Greene, W.H. (2003) Econometric Analysis, 5th Ed., Prince Hall, USA.

Greene, W.H. (2003) "Fixed Effects and Bias Due to the Incidental Parameter Problem in the Tobit Model," Unpublished Manuscript, Stern School of Business, New York University.

Gronau, R. (1974) "Wage Comparisons-A Selectivity Bias," *Journal of Political Economy*, 82(6), pp. 1119-1143.

Ham, J.C. (1986) "Testing Whether Unemployment Represents Intertemporal Labor Supply Behavior," *Review of Economic Studies,* 53, pp. 559-78.

Ham, J.C. and Reilly, K.T. (2002) "Testing Intertemporal Substitution, Implicit Contracts, and Hours Restriction Models of the Labor Market Using Micro data," *American Economic Review*, 92(4), pp.905-927.

Hanoch, G. (1980) "Hours and Weeks in the Theory of Labor Supply," in James P. Smith(Ed), *Female Labor Supply,* pp. 119-65. Princeton, NJ: Princeton University Press.

Heckman, J. J. (1979) "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), pp. 153-161.

Heckman, J. J. and MaCurdy, T. (1980) "A Life Cycle Model of Female Labor Supply," *Review of Economic Studies*, 47, pp. 47-74.

Heckman, J.J. (1976) "A Life Cycle Model of Earnings, Learning and Consumption," *Journal of Political Economy*, 84(4), pp. Sl l - S44.

Hotz, V.J. and *et al* (1988) "Intertemporal Preferences and Labour Supply," *Econometrica*, 56, pp.335-360.

Hsiao, C. (2003) Analysis of Panel Data, 2nd ed., Cambridge: Cambridge University Press.

Hum, D., Simpson, W. and Fissuh, E. (2006) "The Impact of Health on Labor Supply in Panel Data," *Global Business and Economics Anthology*, selected papers from the 2006 Business and Economics Society International Conference, Florence Italy, July 15-19, 2006.

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(3), pp. 302-329.

Kimmel, J. and Kniesner, J. T. (1998) "New Evidence on Labor Supply: Employment versus Hours Elasticities by Sex and Marital Status," *Journal of Monetary Economics*, 42(2), pp. 289–301.

Kumar, A. (2005) "Life cycle Consistent Estimation of Effect of Taxes on Female Labor Supply in the U.S.: Evidence from Panel Data," unpublished manuscript, Federal Reserve Bank of Dallas, USA.

Kyriazidou, E. (1997) "Estimation of a Panel Data Sample Selection Model," *Econometrica*, 65(6), pp.1335-1364.

Leung S.F. and Yu, S. (1996) "On the Choice between Sample Selection and Two-Part Models," Journal *of Econometrics*, 72 (1-2), pp. 197-229.

Lewis, H. G. (1974) "Comments on Selectivity Biases in Wage Comparisons," *Journal of Political Economy*, 82, pp. 1145-57.

Lundberg, S. and Rose, E. (1999) "The Effect of Sons and Daughters on Men's Labor Supply and Wages," Working Paper, Department of Economics, University of Washington, Seattle.

MaCurdy, T., Green, D. and Paarsch, H. (1990) "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *The Journal of Human Resources*, 25(3), pp. 415-490.

MaCurdy, T. (1981) "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy,* 89:1059-85.

Mroz, T.A. (1987) "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica,* 55, pp. 765-800.

Nijiman and Verbeek, M. (1992) "Nonresponse in Panel Data: the Impact on Estimates of a Life Cycle Consumption Function," *Journal of Applied Econometrics*, vol.7, 243-57.

Osberg, L. and Phipps, S. (1993) "Labour Supply with Quantity Constraints: Estimates from a Large Sample of Canadian Workers," *Oxford Economic Papers*, 45(2), pp. 269-291.

Pencavel, J. (1986) "Labor Supply of Men: A Survey," in Ashenfelter, O.C. and Layard, R. (Eds), *Handbook of Labor Economics,* vol. 1, pp. 3-102, Amsterdam: Elsevier Science.

Reilly, K.T. (1994)"Annual Hours and Weeks in a Life cycle Labor Supply Model: Canadian evidence on Male behaviour," *Journal of Labor Economics,* 12(3), pp. 460-477.

Ridder, G. (1992) "An Empirical Evaluation of Some Models for Nonrandom Attrition in Panel Data," *Structural Change and Economic Dynamics*, 17, pp. 77-84.

Romer, D. (2006) Advanced Macroeconomics, McGraw Hill.

Vella, F. and Verbeek , M. (1999) Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias, " *Journal of Econometrics,* 90 , pp. 239-263.

Vella, F. (1998) "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, XXXIII (1), pp. 127-169.

Verbeek, M. (1990), "On the Estimation of a Fixed Effects Model with Selectivity Bias," *Economics Letters*, 34, pp. 267-70.

Vere, P. J. (2005) "Life Cycle Effects of Fertility on Parents' Labor Supply," School of Economics and Finance, University of Hong Kong. Available in <http://client.norc.org/jole/SOLEweb/Vere2005.pdf> Accessed on November 15, 2007.

Wales, T.J., Woodland, A.D. (1980) "Sample Selectivity and the Estimation of Labor Supply Functions," *International Economic Review*, 21(2), pp. 437-468.

Wooldridge, J.M. (1995), "Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions," *Journal of Econometrics*, 68, pp.115-32.

Wooldridge, J.M. (2002) Econometric Analysis of Cross-Section and Panel Data, 2nd ed., The MIT press, London, England.

Ziliak , J.P. and Kniesner, J.T. (1999)"Estimating Life Cycle Labor Supply Tax Effects," *Journal of Political Economy*, 107(2), pp. 326-359.

Ziliak, J.P. and Kniesner, J.T. (1998) “The Importance of Sample Attrition in Life Cycle Labor Supply Estimation,” *Journal of Human Resources*, 33(2), pp.507–530.

Table 1. Descriptive statists of the key variables

Variable	Observations	Mean	Std. Dev
Annual Hours worked	45455	1487.64	1055.33
Imputed log wage	45455	2.92	0.20
Health	45455	0.22	0.16
Other income	45455	38759.06	46771.90
Age	45455	43.70	10.57
Age squared (Age 2)	45455	2021.38	949.59
Education	45455	13.23	3.89
Tenure			
Experience			
Marital status variables			
Common-law=1 if in common law relationship zero otherwise.	45455	0.08	0.27
Married=1 if married and zero otherwise.	45455	0.70	0.46
Separated=1 if separated zero otherwise	45455	0.03	0.17
Divorced=1 if divorced and zero otherwise	45455	0.04	0.20
Widow=1 if widow zero otherwise.	45455	0.01	0.10
Single	45455	0.14	0.35
Visible minority	45455	0.04	0.21
Children age 0 -5 years	45455	0.24	0.57
Children age 5-15 years	45455	0.61	0.94

Table 2. Panel data estimates of men labor supply using selected sample

	Pooled		RE		FE	
Elasticity (β)	0.23		0.27		0.31	
Education	-9.34	(6.45)	-10.61	(4.71)	-47.61	(3.49)
Age	18.59	(5.01)	33.42	(6.39)	54.95	(3.06)
Age2	-0.33	(7.92)	-0.51	(8.79)	-0.82	(5.88)
Married	235.08	(18.74)	230.22	(12.98)	83.95	(2.75)
Common-law	108.48	(6.79)	112.64	(5.33)	63.44	(1.99)
Separated	65.95	(2.89)	39.50	(1.41)	-100.99	(2.49)
Divorced	55.90	(2.63)	1.88	(0.07)	-113.79	(2.62)
Widow	118.78	(2.5)	60.63	(0.98)	-43.89	(0.52)
Imputed wage	458.13	(14.62)	534.70	(13.39)	616.66	(3.34)
Health	-47.62	(1.93)	104.43	(7.87)	-99.45	(2.64)
Other income	-0.002	(29.27)	-0.004	(44.12)	-0.005	(46.41)
Children age 0 -5 years	6.21	(0.91)	1.73	(0.21)	-19.57	(1.91)
Children age 5-15 years	5.44	(1.25)	-6.72	(1.14)	-33.12	(3.90)
Constant	527.89	(6.47)	-18.24	(0.18)	221.32	0.99)
Hausman test (FE Vs RE)			312.4			

Data Source: (SLID 2001). Absolute value of t-statistics in parentheses.

Table 3. Standard panel data labor supply model estimates (full sample)

Variable	Pooled		RE		FE	
Elasticity (β)	0.39		0.48		0.25	
Education	2.62	(1.46)	1.31	(0.48)	-36.93*	(2.55)
Age	59.60*	(13.86)	56.63*	(9.92)	96.07*	(5.29)
Age2	-1.05*	(-22.63)	-1.01*	(16.60)	-1.24*	(9.15)
Married	379.58*	(24.22)	248.74*	(11.68)	58.80***	(1.78)
Common-law	219.98*	(10.73)	144.11*	(5.72)	44.71	(1.29)
Separated	144.72*	(5.04)	47.09	(1.49)	-111.42*	(2.62)
Divorced	74.63*	(2.92)	-34.80	(1.07)	-160.30*	(3.58)
Widow	218.33*	(4.44)	131.50*	(2.17)	0.05	(0.00)
Imputed wage	591.36*	(15.45)	707.98*	(15.79)	351.72***	(1.85)
Health	472.34*	(31.90)	176.13*	(14.09)	94.42*	(7.05)
Other income	-0.004*	(50.43)	-0.005*	(57.58)	-0.005*	(53.96)
Children age 0 -5 years	-79.77*	(9.07)	-38.34*	(4.19)	-23.00*	(2.12)
Children age 5-15 years	-48.32*	(8.87)	-41.95*	(6.26)	-45.56*	(5.24)
Constant	-1263.78*	(13.76)				
Hausman test (FE Vs RE)			414.94			
LM(Pooled Vs non pooled)	36204.93					

Data Source: SLID (2001) Absolute values of t-statistics in parentheses.

Table 4. Samples selection model estimates of life cycle labor supply model[a]

	Tobit Model						Sample selection model[a]	
	Random Effects		Pooled OLS		Fixed Effects		Fixed Effects	Pooled OLS
Elasticity (β)			0.41		0.34		0.32	
Constant	-1010.86	(0.0)		(12.89)				0.25
Education	7.13	(0.0)	12.12	(5.86)	-47.9	(3.04)	-47.77 (11.75)	-2.79 (1.85)
Age	52.40	(0.0)	88.82	(17.59)	160.87	97.69	56.44 (6.88)	40.69 (11.07)
Age2	-0.89	(0.0)	-1.51	(27.29)	-2.17	(13.39)	-0.84 (143.26)	-0.68 (16.49)
Married	263.91	(0.0)	447.04	(24.82)	52.47	(1.46)	83.95 (6.91)	305.07 (23.04)
Common-law	154.51	(0.0)	261.59	(11.19)	42.36	(1.13)	62.98 (1.52)	153.28 (9.10)
Separated	90.40	(0.0)	153.02	(4.65)	-197.95	(4.17)	-103.12 (2.62)	91.18 (3.88)
Divorced	26.50	(0.0)	44.76	(1.51)	-237.55	(4.70)	-115.92 (2.19)	54.99 (2.56)
Widow	118.83	(0.0)	201.23	(3.36)	-32.27	(0.33)	-43.40 (0.48)	144.08 (3.46)
Imputed wage	363.79	(0.0)	615.41	(13.75)	500.57	(2.33)	619.10 (3.04)	479.39 (15.61)
Health	-146.00	(0.0)	-247.38	(7.20)	-40.4	(0.93)	-98.92 (1.84)	-50.06 (2.27)
Other income	-0.02	(0.0)	-0.005	(50.184)	-0.005	(66.19)	-0.005 (2.34)	-0.003 (59.369)
Children age 0 -5 years	-70.67	(0.0)	-119.67	(11.99)	-49.71	(4.12)	-20.28 (1.72)	-23.60 (3.12)
Children age 5-15 years	-40.94	(0.0)	-69.35	(11.13)	-74.1	(7.46)	-33.82 (2.77)	-12.53 (2.81)
Sigma(c)		(0.0)					417.38 (51.14)	
Sigma(u)		(0.0)						
N	9947							
TN	45455							

Data Source (SLID 2001) Absolute values of t-statistics in parentheses.

Note: Health is endogenous. We estimate health equation with predictors including age, age squared, imputed wage, self reported stress level, and time dummies. The reported wage is imputed wage from sample selection corrected wage offer equation.

[a] The sample selection RE model could not converge and was not estimated.

Table 5. Estimated λ_t from the estimated cross-section selection equations

λ_t		
λ_{1996}	-0.42	(6.00)
λ_{1997}	-0.48	(5.24)
λ_{1998}	-0.44	(4.84)
λ_{1999}	-0.38	(4.22)
λ_{2000}	-0.38	(4.00)
λ_{2001}	-0.31	(3.54)

Data source: (SLID 2001). Absolute value of t-statists in parentheses.

Table 6. Fixed effects labor supply model using the selected sample only with the Inverse Mills ratios included (to test for sample selection).

Variables	Fixed Effects [b]		Wooldridge's(1995)[a]	
Constant				
Education	-111.46	(6.70)	-35.29	(10.74)
Age	11.44	(0.52)	-4.60	(0.96)
Age2	-0.78	(3.97)	-0.13	(2.31)
Children age 0 -5 years	363.77	(5.13)	8.03	(0.74)
Children age 5-15 years	-45.62	(3.80)	212.78	(8.54)
Imputed wage	732.56	(4.54)	315.65	(7.32)
Other income	-0.008	(55.83)	-0.0027	(29.45)
Health	-88.98	(2.38)	15.71	(0.25)
Married	-14.52	(0.43)	173.81	(13.03)
Common-law	73.30	(2.28)	95.56	(5.91)
Separated	-131.06	(3.21)	61.05	(2.63)
Divorced	-120.65	(2.75)	29.20	(1.35)
Widow	9.33	(0.11)	79.58	(1.61)
λ_{1996}	18236.94	(8.77)	6632.59	(18.89)
λ_{1997}	-12791.08	(5.64)	-6942.56	(13.1)
λ_{1998}	-1548.07	(0.84)	1603.74	(3.66)
λ_{1999}	-1254.27	(1.38)	-600.76	(2.92)
λ_{2000}	10414.88	(6.00)	5003.56	(10.35)
λ_{2001}	-14583.49	(7.03)	-6961.90	(16.92)
β	0.36		0.16	
N	9947			
TN	33555			
F Test for $H_0 : \lambda_i = 0$	F(6,35476)	11.8	224.14	

Data Source (SLID 2001)

Note: Health is endogenous. We estimate health equation with predictors including age, age squared, imputed wage, self reported stress level, and time dummies.

The reported wage is imputed wage using sample selection corrected wage offer equation.

Time dummies indicating each year of survey were included in the above labor supply Equation but not reported.

[a] Only the results of selected variables are reported. The interpretation of the actual size and sign of the coefficients some of the variables requires knowledge of the full set of results.

[b] We have also conducted a test following Nijiman and Verbeek (1992) by including the lagged inverse Mills's ratios only and we get similar results.

Table 7. Implied uncompensated elasticities

	$\bar{\gamma}_0$	$\bar{\gamma}_1$	$\bar{\gamma}_2$	$\bar{\xi}$
β=0.16	-0.03	-0.22	-0.94	0.00
	(2.69)	(0.98)	(0.98)	(0.00)
β=0.30	-0.015	-0.12	-0.78	0.00
	(2.09	(2.98)	(5.74)	(0.00)

Absolute value of t-statistics in parentheses

Note: we employ two stages least square instrumental variable estimation.

Table 8. Survey of selected studies labor supply of men

Author	Country of study	Substitution elasticity	Uncompensated wage elasticity	Wealth elasticity
Altonji (1986)	USA	[0.0,0.35]		
Hausman(1981)	US		[0, 0.03)	[-0.95, -1.03]
Blomquist (1983)	Sweden		0.08	[-0.03, -0.04]
Blundell and Walker (1996)	UK		0.024	-0.287
Triest (1990)	US		0.05	0
Van Soest et al(1990)	Netherlands		0.12	-0.01
MaCurdy (1981)	USA	[0.1,0.3]	[0.1, 0.3]	0
Reilly(1994)	Canada	0.6[a]		
Hum, Simpson & Fissuh (2007)	Canada	[0.05,0.26]		
Kumar(2005)	USA	[0.5,1.26]		
Ham & Reilly(2006)	USA	[0.9,1.0]		
Kniesner & Ziliak (2006)		[0.2,0.5]		
Ziliak & Kniesner (1999)	USA		[0.12, 0.15]	
Ghez & Becker (1975)	USA	[-0.068,0.44]		
Smith(1977)	USA	0.32		
Conway& Kniesner(1999)	USA			[-0.024, 0]
Kimmel and Kniesner(1998)	USA	0.39		
Kuroda and Yamamoto (2007)	Japan	[0.1,0.2]		
This study	Canada	[0.16,0.48]	[0.16,0.48]	0.00

Notes: Most of the studies adopt different estimation strategies and any comparison of the estimates must be cautious.

[c] It is some times called income elasticity. However since there is no as such income effect and if there is any it should operate via the marginal utility of wealth the designation of income effect is misleading.

[a]Elasticity is with respect to number of weeks worked.

Chapter 2.

2. The Impact of Health on Wages in Canada: Evidence from Sample Selection Corrected Panel Data Models

Abstract

This paper attempts to estimate the impact of health on wages in Canada using Mincer type wage offer models that correct for sample selection and control for individual heterogeneity. We employ the second panel of Canadian Survey of Labor and Income Dynamics (SLID) spanning from 1996 to 2001. This paper explores the application of several panel data models in estimating the impact of health on wages. The estimated health effect is compared across different panel data estimators. The results confirm that estimates from Mincer type wage offer equations are very sensitive to the econometric specification. Failure to control for individual heterogeneity and sample selection bias in most cases produces upward biased effects of health on wages. Using a model which controls for sample selection, individual heterogeneity and measurement error, this study finds that the effect of health on wages is positive, as expected, but not statistically significant. The results of this paper demonstrate the importance of controlling for sample selection bias and that the selection process could mainly operate via time variant variables and hence the traditional fixed effects model does not suffice.

2.1. Introduction

Issues related to health care provision have been part of the agenda of public policy discussion and political debate in Canada. At the heart of such policy discussions has been the relationship between health and socioeconomic status. Any health intervention policy needs to have a clear picture of the relationship between individual health and labour market outcomes such as wages, labour supply and employment (Madrian and Currie 1999). The underlying assumption behind public health investment is inherent individual and public benefits that are welfare improving. Thus an step towards a public policy intervention should be examination of the association between health status and income. There are also some cases where we need to understand the impact of health on wages. For instance, in legal disputes which involve accident and injuries causing a significant negative shock to the health capital, it is imperative to have some estimate of the valuation of such damage in terms of the lost opportunities. In these kinds of cases, among others, we need to know the impact of health on wages and other labor market outcomes.

There are at least three potential pathways where health may affect wage[43](Grossman, 1972). Firstly, like education, health may help increase productivity and hence wages. Secondly, health may be employed as a screening device by employers. Good health may be perceived by employers as a signal for productivity. In this regard individuals with a relatively large health endowment will have a higher chance of commanding a higher wage than their non-healthy counterparts. Thirdly, there may be some sort of discrimination against unhealthy individuals in the labour market. Relatively healthy individuals could be demanded, not on the grounds of productivity, but on the basis of their health endowment. The former provides the basis for the proposition that health impacts wages by affecting human capital formation and thereby influencing productivity.[44] However, there is a possibility of reverse causation. Grossman (2001) argues that if the marginal benefits of investment in health increases with wages then health should rise with wages where we have a simultaneity problem.

However, examination of the impact of health on wages is not without complications. Firstly, as it was mentioned above, health could be endogenous. Secondly, the self-

[43] Interested reader is referred to Grossman (1972a) for a formal treatment.

[44] Treating health as a stock of capital could imply that the lagged health should be related with current income. Despite the theoretical attractiveness of this line of argument the time period that health variables should be lagged is not easy to determine. In a more convincing way it would make more sense to postulate that long term health will have cumulative effect on life cycle earnings, meaning childhood investment on health may be reaped during adulthood. Unfortunately our data do not allow us to make this line of enquiry and this issue is still not under the full control of empirical research (Thomas and Strauss, 1997).

reported health employed in many studies could be plagued with measurement error and bias.[45] For example, unemployed individuals will probably report relatively lower level of health than their employed counterparts to justify their state of unemployment. Thirdly, there could be a problem of sample selection bias. Hourly wages will only be observed for those who are participating in the labour market and we can not observe wage for those who reported no hours in the survey year. Heckman (1978) shows that failure to correct for sample selection problem leads to inconsistent estimates. Lastly, there is a problem of individual heterogeneity which is not hard to imagine in the case of health. Even though health is largely endogenous it is conceivable that factors such as genetic make up and other family background are individual specific and fixed.

All the aforementioned problems call for the utilization of panel data and render cross sectional evidence unreliable. This paper employs panel data sample selection model which controls for most of the above problems in a unified framework. The current paper employs a sample selection model where the selection equation and the equation of interest have fixed effects with the fixed effects being correlated with explanatory

[45]In the simplest model of measurement error in which wages are determined by only one human capital variable that is measured with random error, the estimated attenuation bias of the wage effect of human capital is downward in proportion to the ratio of the variance of the measurement error to the variance of the measured human capital variable (Griliches, 1977). Effort to include more wage determinants that might reduce omitted variable bias also has the consequence of increasing the measurement error bias, because the added wage determinants tend to be correlated with the true human capital variables, increasing the remaining noise-to-signal ratio (Schultz, 2003). The fact that the coefficient on the variable measured with error is asymptotically biased towards zero only holds if there is only one variable measured with random error. If more than one variable is measured with error, there is very little that can be said about the direction of the bias (Maddala, 2001). It is unclear, therefore, whether estimates of the human capital returns from a wage function are improved by the inclusion of more controls, even if the controls are exogenous and correlated with wages. Thus, given the fact that we are using self assessed health and schooling to proxy health and education, we need to acknowledge that our model may suffer from measurement bias. But there could also be additional source of measurement bias associated with self-assessed health which is non-random. For example people with weak labor market attachment might underestimate their health status to justify their employment status.

variables. For testing and correcting sample selection bias this paper employs the sample selection panel data estimator suggested by Wooldridge (1995). This estimator is more flexible and requires no distributional assumptions about the behaviour of the individual fixed effects in the main equation and allows for hetroscedasticity and autocorrelation of unknown form.[46] The results of the test for sample selection problem reveal that there is sample evidence to reject the null hypothesis of no sample selection bias. The results confirm that estimates from Mincer type wage offer equations are very sensitive to the econometric specification. Furthermore, failure to control for the individual heterogeneity as in cross sectional models, and sample selection bias in most cases produces upward biased effects of health on wages. Using a model which controls for the sample selection problem and individual heterogeneity, this study finds that the effect of health on wages is positive, as expected, but not statistically significant.

2.2. Related Previous Studies

Strange as it may seem, evidence on the impact of health on wages in Canada is almost non existent. This forms the key motivation for this paper. The only study known to the author is Fissuh (2004) where he uses random effects instrumental variable estimators using a sample of employed men from SLID. Fissuh reports a positive effect of good health on wages. However, there have been a number of studies on the impact of health on wages in Europe and the U.S.A. (Pelkowski and Berger 2004, Gambin 2005, Chirikos

[46] Dustmann and Rochina-Barrachina (2007) compare Wooldridge's (1995) method with that of Kyriazidou (1997) and Rochinna-Barrachina (1999) with an application in the estimation of wage equations for Germany. Even though they report slight differences among the estimates from the different sample selection models, they conclude that it is hard to compare the estimates as the underlying assumptions could be an explanation for the differences and assert that the differences could be explained by the assumptions imposed in each of the estimators. They suggest that in any application of these estimators researchers should be careful in the interpretation of any set of estimates.

1985, Heineck 2004, Dustmann and Windmeijer 2000, Contoyannis and Rice 2003, Leung and Wong 2002, Andren and Palmer 2000, Lee 1982, Haveman *et al.* 1994, Baltagi and Hausman and Taylor 1980, Baltagi and Khant-Akom 1990). We present brief summary of these related studies below.

There are a number of studies from Germany. Jäckle (2007) examines the impact of health on wages in Germany using data from the German Socio Economic Panel (GSOEP). Jäckle employs a panel data model which corrects for sample selection and endogeneity and he reports that good health has a positive impact on wages in Germany both for men and women. Jäckle employs self-assessed health as a proxy for true health status. Jäckle uses sample selection estimators suggested by Wooldridge (1995) and Semykina and Wooldridge (2005). Dustmann and Windmeijer (2000) study the impact of wages on health demand over the life cycle where they derive the demand for health function from a life cycle framework akin to the life cycle labour supply model by MaCurdy (1980). They find a negative the intertemporal elasticity of substitution which implies that any evolutional change in wages will cause negative substitution of health time with non health time. However, they document that a permanent change in the wage profile causes a positive effect on the demand for health. Heineck (2004) estimated the relationship between height and wages in Germany using the random effects instrumental variable estimator suggested by Houseman and Taylor (1980) for panel of GSOEP data from 1991 to 2002. Heineck finds that there is no significant effect of health on wages but on hours worked. Heineck also reports an association of stature and wages for male workers from West Germany. But Heineck treats health as an exogenous variable in the wage equation.

Gambin (2005) investigates the impact of health on wages in 14 European countries by employing self- reported health as a proxy for true health. Gambin reports that health affects wages positively and the effect is greater in the men's than in the women's sample. However she reports also that the effect of acute or chronic diseases on wages seems to be stronger for women than men. Contoyannis and Rice (2003) study the relationship between health and wages in the UK. They employed random effects instrumental variable estimation for panel data models. They examine the effect of self-assessed health and psychological health on hourly wage by employing British Household Panel Survey. They report a positive effect of excellent self-assessed health on female wage rate and that reduced psychological health reduced the hourly wage rate of males. They also find that their health variables were correlated with person specific time invariant effects. Andren and Palmer (2000) study the relationship between hourly wage and annual earning with health for Sweden. They employ single equation Tobit model and they find that there is a significant effect of sickness history of workers on their earnings. However, they report that the effect of sickness history on wage rate is small.

Now we turn to the evidence from the USA.[47] Pelkowski and Berger (2004) study the impact of health on employment, wages and hours worked over the life cycle. They

[47] Madrian and Currie (1999) present an extensive survey of the link between health and labour market outcomes such as wages, number of hours worked and employment in the U.S.A.

report that a permanent health problem has a negative effect (counterproductive effect) on labour market outcomes. They use functional limitation as a proxy for health. Chirikos (1985) employs a simultaneous equation of health and wages and finds that poor health affects earnings adversely. Chirikos (1985) employs the National Longitudinal Survey from the USA. Haveman *et al.* (1994) estimate a three equation simultaneous system of health, wage and work time using a sample of adult white men from the Michigan Panel Study of Income Dynamics (PSID). They report that lagged health has a negative effect on wage, which is in line with the investment model of Grossman (1972). In addition, they also find that there is a causation running from wage to health. Lee (1982) studies the relationship between health and wage in the USA for a sample of men aged between 45 and 59. He employs a simultaneous equations approach for the estimation and documents that there is a positive relationship between wage and health, which runs in both directions. Lee employs a structural wage equation and probit health equation, because his health variable has two discrete values. However, his cross-section sample of men aged 45 to 59 casts doubt on his results, as it may be contaminated with sample selection bias.

2.3. Our Model

The most common economic theory informing wage determination is the human capital theory, which was pioneered by Schultz (1961), Becker (1964) and Mincer (1958 ,1974). The conventional theory of human capital views education and training as the major sources of human capital accumulation. Based on the human capital theory Mincer (1974) developed an earnings function, in which the logarithm of earnings is expressed as a linear function of the number of years of schooling completed and as a linear and quadratic function of potential experience[48]. This Mincerian earnings function has become an essential tool in research on wage earnings in developed and developing economies (Cahuc and Zylberberg 2004, Kjellstrom and Bjorklund 2002). The original Mincerian earnings function with random coefficients can be expressed as:

$$\ln w_{it} = \beta x_{it} + \varepsilon_{it} \quad , \; x = [Educ, Exper] \qquad [2.0]^{49}$$

Where $Educ =$ Years of schooling, $Exper =$ Potential years of experience, $\ln w =$ logarithm of earnings and ε_{it} is the idiosyncratic error. In this paper, the original Mincerian model [2.0] is extended in a number of directions. Firstly, we extend the Mincerian equation to allow for the impact of health on wages and other important individual factors which are expected to influence wages. Note that health could be endogenous for the reasons explained in the introduction of the current paper. For this reason, we employ some instruments to circumvent the possible endogeneity problem. We follow Dustmann and Windmeijer (2000) and Haveman *et at* (1994) in selecting the

[48] Mincer defined potential work experience as age – years of education - 6, to proxy the number of years an individual spends in the labour market, assuming they are continuously employed. The quadratic specification of the experience variable reflects the commonly observed concave pattern of age-earnings profile which is consistent with the linear human capital decay function.

[49] Equation [2.0] is a random coefficient model and implies that: i) Age-log earnings profile is U shaped and age-log earnings profile is parallel across schooling levels. Heckman, Lochner and Todd (2003) examine these implications in light of empirical evidence and report that the empirical evidence supports thee implications, even after some changes. Our preliminary data analysis also supports these implied relationships. The empirical literature shows that the logarithmic wage function is superior to the linear wage function and other functional forms and the log normal distribution is a good approximation of the empirical wage profile (Card 1999).

instruments for health.[50] Secondly, we need our model to take into account the panel nature of the data. Consider the following panel data model:

$$\ln w_{it} = \beta x_{it} + c_i + u_{it}, \qquad i = 1,\ldots,n \qquad t = t_i,\ldots,T_i \qquad [2.1]$$

Thirdly, we want our model to allow for non-random sampling. As it was mentioned in the previous section, wages are not observed for individuals who do not work. In other words, wages will only be observed if annual number of hours worked is positive. A selection problem arises if the unobservable variables which determine the decision for participation in the labour market also affect wages. In the case of health it is conceivable that some genetic factors and life situations can affect the participation decision. At the same time, these factors will also affect wage because they affect health stock. However, these two factors have different implications. The former could affect wages via the fixed effect component of the wage equation. If this was the case we could solve the problem by estimating a fixed effects model where the fixed effects will be wiped out during the time demeaning process. However, the latter affects wages via the idiosyncratic error term in the wage offer equation as they are time variant. Since health is correlated with these variables estimation of [2.1] using OLS may yield inconsistent estimates. This selection problem has been long identified in the literature (Wales and Woodland 1980, Gronau 1974, Lewis 1974, Heckman 1978, Vella 1998). To deal with this notorious

[50] Our health predictors will include age, age squared, imputed wage, mother's education, father's education, time dummies and functional limitations. Dustmann and Windmeijer (2000) present an elegant way of deriving the life cycle consistent health demand and justify for the inclusion of these predictors.

problem of nonrandom sampling we propose that the wage offer equation follows the following panel data structure[51]:

$$\ln w^*_{it} = \beta x_{it} + c_i + u_{it}, \qquad i = 1,\dots,n \qquad t = t_i,\dots,T_i{}^{52} \qquad [2.2]$$

$$s^*_{it} = \psi z_{it} + \varsigma_i + v_{it} \qquad v_{it} \mid z_i \sim N(0,1) \qquad [2.3]$$

$$s_{it} = 1 \quad \text{if} \quad s^*_{it} > 0 \qquad [2.4]$$

$$\ln w_{it} = (\ln w^*_{it})(s_{it}) \qquad [2.5]$$

Where $\ln w^*_{it}$ is a latent endogenous wage with an observable counterpart $\ln w_{it}$. s^*_{it} is latent labour force participation decision with an observable counterpart s_{it}. Equation [2.2] is the Mincerian wage function which is the equation of interest and equation [2.3] is a reduced form for the propensity to participate in the labour market. x_i and z_i contain vectors of exogenous individual characteristics such as , experience, tenure, years of schooling, health, and others. It is conceivable that most of the variables that enter the wage equation will also determine participation in the labour market. In our empirical models we will impose some exclusion principle and more specifically we will include at least one time varying variable in the selection equation that does not affect wages. β and ψ are vectors of unknown parameters and u_{it} and v_{it} are random error

[51] See Baltagi (2005), Hsiao (2003), Greene (2003), Wooldridge (2002) for discussion on panel data modeling. See Vella (1998) for a readable survey of sample selection models.

[52] Note also that $t = t_i,\dots,T_i$ implies that the panel structure could be unbalanced.

terms with $E(u_{it}/\upsilon_{it}) \neq 0$. We assume that (u,υ) is independent of z_i (where z_i might contain elements of x_i[53]), c_i and ς_i are individual fixed effects which are time invariant.

Next we present a brief discussion of the traditional RE and FE sample selection estimators to highlight their deficiencies and motivate our preferred estimator. To save on notation let $\ln w_{it} = y_{it}$. Define the following deviation forms of the variables as

$$\ddot{x}_{it} = x_{it} - \frac{\sum x_{ir} s_{ir}}{\sum s_{ir}} \text{ if } \sum s_{ir} > 0 \qquad [2.6]$$

$$\ddot{y}_{it} = y_{it} - \frac{\sum y_{ir} s_{ir}}{\sum s_{ir}} \text{ if } \sum s_{ir} > 0 \qquad [2.7]$$

Hence the fixed effects estimator for unbalanced ($\beta_{FE}(u)$) and balanced panel ($\beta_{FE}(B)$) are as follows

$$\beta_{FE}(u) = \left(\sum_{i=1}^{N} \sum_{i=1}^{\tau} \ddot{x}_{it}' \ddot{x}_{it} s_{it} \right)^{-1} \left(\sum_{i=1}^{N} \sum_{i=1}^{\tau} \ddot{x}_{it}' \ddot{y}_{it} s_{it} \right) \qquad [2.8]$$

$$\beta_{FE}(B) = \left(\sum_{i=1}^{N} \sum_{i=1}^{\tau} \ddot{x}_{it}' \ddot{x}_{it} d_i \right)^{-1} \left(\sum_{i=1}^{N} \sum_{i=1}^{\tau} \ddot{x}_{it}' \ddot{y}_{it} d_i \right) \qquad [2.9]$$

where $d_i = \{ (\Pi_{t=1}^{\tau} s_{it}) = 1 \}$

[53] It is conceivable that most of the variables which influence participation in the labour market will also affect wages hence we can expect that z contains most of elements of x. Ideally, we would like to have an exclusion rule here for the purpose of efficiency. More specifically, the Wooldridge (1995) estimator requires at least a time varying variable which affects selection but not wages. In our case, among other things, we include other income in the selection equation but not in the wage equation. The two-step estimation could be unreliable in the absence of exclusion restriction (Vella, 1998). However, Leung and Yu (1996) argue the reverse if there is a sufficient variation in one of the regressors to induce sufficient variation in the tail behaviour in the Inverse Mills ratio.

For the consistency of our fixed effects estimator we require that $E[\ddot{u}_{it} \mid \ddot{x}_{it}, s_{it}] = 0$. In other words the sample selection should operate via individual fixed effects which will be removed by time demeaning the variables. Note that this assumption will break down if the selection process is operating through some time variant unobservable variables.

To get the random effects model we follow Vebreek and Nijiman (1992). Define $y_{it} = (y_{i1},...,y_{iT})'$, $x_{it} = (x_{i1},...,x_{iT})'$, and $u_{it} = (u_{i1},...,u_{iT})'$ Assume that all the variables in the labour force participation equation are available and define the number of units $s_{it} = 1$ as T_i and define $T_i x T_i$ matrix R_i transforming y_{it} into the T_i dimensional vector of observed variables, y_i^o. Note that matrix R_i is obtained by deleting the rows of the T dimensional identity matrix corresponding to $s_{it} = 0$. Defining the unit vector I the variance covariance of the error term in equation [2.1] can be written as $\Omega = \sigma_c^2 ii' + \sigma_u^2 I$. Given this random error structure the random effects estimator for the unbalanced and balanced panel case are:

$$\beta_{RE}(u) = \left(\sum x_i^{o\prime} \Omega_i^{-1} x_i^o \right)^{-1} \left(\sum x_i^{o\prime} \Omega_i^{-1} y_i^o \right) \qquad [2.10]$$

$$\beta_{RE}(B) = \left(\sum x_i^{o\prime} \Omega_i^{-1} x_i^o d_i \right)^{-1} \left(\sum x_i^{o\prime} \Omega_i^{-1} y_i^o d_i \right) \qquad [2.11]$$

where $\Omega_i = R_i \Omega R'_i$ and $x_i^o = R_i x_i R'_i$

For the consistency of [2.10] and [2.11] we need $E[u_{it} + c_i \mid x_{it}, s_{it}] = 0$. Thus the random effects estimator will be inconsistent if the selection is operating either through the individual fixed effect or the idiosyncratic error. Our test for the presence of sample selection bias reveals that the assumption of consistency is not tenable. For this reason we report the fixed effects estimates only. Note that the beauty of the fixed effects model is that it does not require a selection equation and it does not impose any distributional assumption about the error terms. However the fixed effects model is subject to the infamous incidental parameter problem which might render the estimates biased.

These are the traditional panel data estimators which require some bold assumptions hold to obtain consistent and unbiased estimates. As discussed above, it is most likely that the assumptions required for consistent estimation of the aforementioned models will likely be violated. In light of this, Wooldridge (1995) develops an estimator for testing and correcting a selection problem.[54] Under some generally mild conditions, this model produces consistent estimates of β. In the next section, we present Wooldridge's (1995) estimator.

Wooldridge (1995) estimator

The Wooldridge (1995) estimator relies on level equations. The basic premise of this approach is that it parameterises the conditional expectation required for the consistency of the pooled estimator:

[54] The basic testing procedure is similar to that of Ridder (1990) Nijiman and Verbeek(1992) and Vella and Verbeek(1994) who base their procedure on simple variable addition test but this test provides a more general way of obtaining consistent estimates.

$$E[c_i + u_{it} \mid x_i, s_{it} = 1] = E[c_i \mid x_{it}, s_{it} = 1] + E[u_{it} \mid x_{it}, s_{it} = 1] = 0 \quad \forall t \qquad [2.12]$$

That is to say, the approach parameterises assumption [2.12] and adds the resulting variables as additional regressors in the main equation. This approach is semi parametric with regards to the wage equation as it does not require joint normality of the error terms in the selection equation and wage offer equation. Akin to the Heckman's (1978) two stage estimation, it requires marginal normal distribution of the error terms in the selection equation and conditional mean assumption of the error terms in the wage equation. However, the time dimension permits control of individual fixed effects, at a cost of invoking additional assumptions on the conditional means of the fixed effects in both equations. The approach is similar to that of Chamberlain (1984) in spirit. The assumptions are:

A1: the conditional expectation of ς_i *given* $z_i = (z_{i1}, ..., z_{iT})$ *is linear.* $\varsigma_i = z_{i1}\kappa_1 + ... + z_{iT}\kappa_T + \partial_i$ where ∂_i is random component.

A2: the errors terms in the selection equation, $\lambda_{it} = \varsigma_i + v_{it}$ *are independent of* $\bar{z}_i$ *and normal* $(0, \sigma_t)$*, where* $\bar{z}_i = (x_i, z_i^+)$ *with* $x_i = (x_{i1}, ..., x_{iT})$ *and* z_i^+ *containing the non-overlapping elements in* z_i*.*

A3: $E[u_{it} \mid x_i, v_{it}] = E[u_{it} \mid v_{it}] = \eta_t v_{it}$, $t = 1,2,...,T$. This is an assumption about joint normal distribution of the error terms in equation [2.2] and [2.3].[55]

[55] If $E[u_{it} \mid x_i, v_{it}] = E[u_{it} \mid v_{it}] = 0$ there would be no problem of sample selection.

A4: $E[c_i \mid x_i, v_{it}] = E[c_i \mid x_{i,} v_{it}] = L[c_i \mid 1, x_{i,} v_{it}]$, $s_{it} = 1[x_{it}\delta \; + c_i + a_{it} > 0]$. This assumption implies that the fixed effect is a linear function of x_i and the error term in equation [2.3]. The violation of this assumption leads to inconsistent estimates (Wooldridge 2002). This can be represented as follows:

$$E[c_i \mid x_i, v_{it}] = \pi x_i + \phi v_{it} \tag{2.13}$$

With the help of the law of iterated expectations we can write [2.13] as follows

$$E[c_{it} \mid x_i, v_{it}] = \pi \; x_i \tag{2.14}$$

Hence [2.2] could be rewritten as

$$E[\ln w_{it} \mid x_i, v_{it}] = \beta x_{it} + \pi x + \eta_t v_{it} \tag{2.15}$$

Conditioned on $s_{it} = 1$ we can write [2.15] as follows

$$E[\ln w_{it} \mid x_i, v_{it}] = \beta x_{it} + \pi x + \xi_t \lambda(x_i \psi_t) \tag{2.16}$$

Thus it is possible to get consistent estimate of β by first estimating the labour force participation equation using probit model s_{it} on x_i for each panel j and saving the Inverse Mills ratios ($\hat{\lambda}_{it}$) for all i and t. The next crucial step is to run a pooled OLS regression using the selected sample: y_{it}, on $x_{it}, x_i, \hat{\lambda}_{i1}, \hat{\lambda}_{i2}, \ldots, \hat{\lambda}_{iT}$ for all $s_{it} = 1$. Wooldridge (1995) shows that we can obtain consistent estimates of [2.16] using either OLS or minimum distance estimator. Under assumption of A1-A4, the estimator for β is consistent (Wooldridge 1995, Dustmann and Rochinna-Barrachina 2007, Vella 1998).[56] Note that this approach allows for the correlation between the unobservable in the selection equation, v_{it} and the unobservable in the wage offer equation (c_i, u_{it}). That is to say the selection process might operate via both the error term from the main equation u_{it} and the unobservable individual effect c_i. Also note that identification of β for time varying variables is possible via assumption *A3*. It is also important to mention that we need to have a time varying variable which affects the selection process but does not affect wages. This is expected to help partially circumvent the possible multicollinearity problem between the cross sectional $\hat{\lambda}_{it}$ which is left unnoticed in implementing the model but can be very damaging.

There have been a number of suggestions on the detection of sample selection bias in panel data models (Wooldridge 1995, Verbeek and Nijman 1992 1996, Vella 1998, Vella and Verbeek 1994). In this section we follow Wooldridge (1995).Wooldridge (1995) argues that for estimation purposes equation [2.3] can be either estimated by random effects or pooled cross section. The key assumption is that the selection process follows

[56] See Wooldridge (1995) for more detailed derivation of the model.

equation [2.3]. [57] Our test in this paper shows that estimated $\hat{\lambda}_{it}$ both in the RE and FE was found to be statistically significant.

Under the null hypothesis of $E[u_{it} \mid x_i, s_{it}, c] = 0,\ \ t = 0,1,...,T$ the $\hat{\lambda}_{it}$ from equation [2.3] should not be significant in an equation estimated by fixed effects method. To elaborate, let $\hat{\lambda}_{it}$ be the estimated Inverse Mills ratio from [2.3], the sample selection equation, by pooled cross section for all i over t. The next step is to test the sample selection using t test for the $\hat{\lambda}_{it}$ in the fixed effects model. Wooldridge (1995a) shows that the limiting distribution of t under the assumption $E[u_{it} \mid x_i, s_{it}, c] = 0$ is not affected significantly by the form of specification adopted for the participation equation [2.3]. As long as the standard errors are robust and adjusted for hetroskedasticity we can trust the student t test. At this stage it might be tempting to estimate [2.2] by including the $\hat{\lambda}_{it}$ as additional regressors. However, this may lead to inconsistent estimation as the root cause of the selection problem is not corrected by such procedure (Wooldridge 2002).Rather, we employ the estimator suggested by Wooldridge (1995).

2.4. Data

We employ the panel data set available from the Survey of Labour and Income Dynamics (SLID) in Canada. This rich longitudinal data is a household survey that covers five major regions of Canada: Atlantic, Prairies (Manitoba, Alberta and Saskatchewan),

[57] Wooldridge (1995) claims that the estimation of the selection model is of a second order importance as the objective is to derive a test and we are not interested in the selection equation parameters per se. We experimented with different specifications of the selection equation and the results were similar.

Ontario, British Columbia and Quebec, starting from the year 1993. The main aim of launching SLID was to provide additional dimensions to the traditional surveys on labour market activity and income. The SLID has two components where each panel contains about 15,000 households and 30,000 adults, and the panel is followed for a period of six consecutive years. The survey respondents include selected samples of adult individuals (16+). A new panel is introduced every three years, so two panels always overlap. From the year 1993 the annual Survey of Consumer Finances (SCF) was replaced with SLID so as to capture additional dimensions of the transitional dynamic of the labour market and family events in Canada.

-------------------------Table 1 here---

In the SLID, background information is collected at the beginning of each year and interviews are conducted in two subsequent stages. The first interview is conducted in January where questionnaires are administered to collect information on labour market experiences, educational activity and family relationships. The second stage is conducted in May where information on income are collected so as to take advantage of income tax as it is the time when respondents are more familiar with their tax returns. According to Statistics Canada, there has been a very high percentage of people, about 80%, willing to give permission to access their income tax files which helps avoid May income interview (Statistics Canada, 2004).

Our sample includes individuals in SLID who completed questionnaires in the six waves of year 1996 to 2001 which constitutes a sample of males between the age of 25 and 65 who gave valid responses for the variables employed in our estimation. Our sample consists of both employed (those who reported non zero wage and non zero number of hours) and non employed (with missing wage and zero number of hours) . The final sample contains 9208 individuals and a total of 38689 observations.

----------------Figure 1 here----------------------------------

Figure 1 presents smoothed age-wage profile by health status over the life cycle. Figure 1A demonstrates that there is indeed a wage gap between healthy and non healthy individuals (according to their self reported health status). It is also interesting to note that the gap does not remain constant during the life cycle. Figure 2 demonstrates that men who reported good health are relatively younger than those who reported poor health.

Table 1 presents the descriptive statistics of the key variables across the five health categories. This table demonstrates that years of schooling completed vary with health positively. What is interesting about this table is the respondents with excellent health have lower mean wage, number of hours worked , weeks worked than those who reported very good health. This could be partially explained by average age of respondents in each category. People who report excellent health are on the average younger than those who reported very good health and it is likely that younger people will have lower wages because of their experience. Another possibility could be that the perception of very good and excellent could differ among individuals. Apart from this, we observe a general positive relationship between health and most of the labor market outcomes.

--------------Figure 2 here----------------

It is common in the literature to calculate the hourly wage by dividing reported annual earnings by estimated annual number of hours worked. These calculations usually introduce measurement bias and yield inconsistent parameter estimates. Any measurement error in the annual number of hours worked will carry over to the hourly wage rate (Ziliak and Kniesner 1999, Mroz, 1987). Conway and Kniesner (1999) argue that the choice of wage measure matters. In SLID the annual number of hours and the composite hourly wage are calculated from very extensive interviews with detailed questions on each job and payment that individuals get in the survey period. Questions on the number of jobs held and the hours worked pay by pay, number of weeks worked, number of weeks absent from work and others where respondents are walked by detailed questionnaire to retrieve information and where access to the income files of respondents is obtained are supposed to produce reliable information on number of hours worked and hourly wage.

Following Mincer (1974) we include experience and experience squared in our regression analysis but use actual years of experience. Experience includes number of years work experience full year full time equivalent imputed since first starting to work[58]. Experience

[58] Note that the usual practice is to calculate experience as age minus six minus the number of years of schooling .This practice will most likely overstate the actual amount of experience because it assumes that

squared is included to test whether the effect of experience on earnings is concave. In most earnings function estimations the implied earnings-experience profile is concave (Mincer, 1974). Figure 1 tends to agree with this assertion. Moreover, Figure 1A shows that the hourly wage-age profile of the relatively healthy individuals is higher than those of the less healthy individuals at each level of age. In addition we also include tenure and tenure squared to capture any return for seniority. We also include years of schooling as a measure of education.

Our health variable is self reported health on a scale of excellent, very good, good, fair and poor. Despite the subjective nature of self reported health there have been numerous studies which document that self reported health is a good indicator of true health because it is highly correlated with objective measures of health status (Madrian and Currie 1999). The problem with self reported health is not that it is correlated with objective measures of health but that the measurement error is not randomly distributed across the sample. Our approach deals with potential sources of non-random measurement error by controlling for sample selection bias and employing some health instruments. We generate a binary health variable called "Binaryh" which assumes a value of 1 if an individual reports excellent, very good or good health status, otherwise the value is 0.[59] We have also a full scale health variable model. It is worth noting that these variables are time varying. Table 2 presents the transition of individuals across the

---

a worker is continuously employed throughout his/her life after school. However, a worker can be unemployed for a number of years after school.

[59] We also generate three dummy variables for these different categories of health status. The first dummy which was constructed is VGH standing for very good health which takes a value of one if the individual reported either excellent health or very good health status. The second one is GH standing for good health and it takes value of one if an individual reported good health status otherwise zero. The third one is PH which represents poor health status and assumes a value of one if the individual reports either poor health or fair health status, otherwise it assumes a value of zero.

different health categories from the year 1996 till 2001. The numbers in bold font are those who stayed static on their respective health category throughout. Despite the apparent dynamism in health status, the transitional matrix tends to exhibit regression to the mean in each category and some sort of state persistence.

--------------------------Table 2 here---

We also looked at the dynamics of health by age group and income quintile. Figure 5 demonstrates a clear health-income gradient. As we move from the poorest quintile (1) to the richest quintile (5) the proportion of men who report excellent health and very good health increases whereas the proportion of men who report fair and poor health decreases. We also looked at the distribution of self assessed health by income quintiles over time and the distribution is similar to the above bar graph.

--------------------Figure 3 here----------------

Another way of looking at the distribution of wages by health categories is to compare the cumulative distribution of wages by health category. Figure 4 shows that the distributions of wages for those who reported excellent health , very good health and good health stochastically dominates the distribution of wages for those who reported poor health and/or fair health. In fact Figure 5 shows that the wage distribution of those who reported good health stochastically dominates those who reported poor or fair health. However this first order stochastic dominance disappears after controlling for education and experience (age). Both Figure 6 and Figure 7 show that there is no clear stochastic dominance after controlling for education and age. This seems to substantiate the findings in the paper that after controlling for sample selection and endogeneity of health the effect of health on wages is statistically not significant.

-----------------------------------Figure 4 and Figure 5 here----------------------------------

We have included regional dummies in our wage offer equations following Contoyannis and Rice (2003). The justification for the inclusion of these regional dummies is that they would capture regional price variations. Since the wage equation is supposed to estimate the real wage, this practice is legitimate. However, we further hope in this paper that these regional dummies may also capture regional peculiarities such as cost of living, unemployment and other region specific factors. We follow Statistics Canada's classification of the region of residence of the individual respondents. In this classification, we have Ontario region, prairies, and British Columbia Region and Atlantic region. We also include provincial minimum wage (Minimum Wage) to capture some supply and demand interactions in the labor market. We hope that this variable captures some component of regional differences in unemployment and cost of living. Dummy variable indicators for visible minority, for number of children below the age of five years and for number of children between the age of 5 and 15 were also included in the model. These variables are included in our wage equation to capture the fact that people with more children tend to have less time available for labour market activities. Hence we expect this variable to capture productivity, distribution of market time across the household and experience effect which is not captured with other variables in our model.

--------------------------------------- Figure 6 and Figure 7 here--------------------------------

2.5. Estimation Results

This section presents the estimation results. Before presenting the IV and sample selection model results, we report the results of the pooled OLS and the traditional panel data estimators: FE and RE models. Table 3 reports the results from a pooled OLS both with endogenous and exogenous health variable. The pooled OLS model assumes that the regressors in the model are not correlated with individual heterogeneity. For example, if there is some individual heterogeneity on health endowment and this is associated with participation in the labor market positively then OLS estimates are upward biased. The estimates seem to produce the common standard results reported in many applications. [60]When we look at the coefficient of health variables we find that the coefficients are positive and statistically significant at less than 5% level, suggesting that wage rate is positively associated with health. What is more interesting is that the positive association between health and wages is robust to the exogeneity assumption about health. However, the coefficient of health is slightly lower when it is treated as an endogenous outcome as opposed to exogenously given variable.

--------------------------------Table 3 here--

The general result of a positive impact of health on wages is in line with our expectation. However, since in pooled OLS model it is assumed that the individual effects are not

[60] See Medrian and Curie (1999) for an extensive survey on USA.

correlated with the explanatory variables we may suspect that the estimates could be biased and inconsistent. In fact, an LM test has decisively rejected the null hypothesis of non correlation of the error terms with the explanatory variables at less than 1 % level of significance for all the models reported. Hausman type tests were also conducted to compare the fixed effects and random effects models and in all cases the test decisively rejected the null hypotheses of non correlation of the error terms with the explanatory variables (consistency of RE model) at less than 1 % level of significance. Due to the fact that in all cases the results rejected the RE specification in favour of Fixed Effects, we report the FE estimates only.

------------------------Table 4 here---

To investigate how our results change with time demeaning of our variables we present the results from the within estimators in Table 4. As was mentioned above the within estimator produces unbiased and consistent estimates, no matter if the variables are correlated with the error terms or not. However if there is a time varying variable which is driving both the wage and selection equation the estimates from this model could be further upward biased. Looking at Table 4 we find that the coefficients for the health variables are slightly smaller than the pooled OLS estimates with standard error almost remaining the same. If we look at the other variables also we observe that there are differences in the coefficients and their standard errors. These small differences may suggest that the assumption of no correlation between the individual effect and the explanatory variables is not very damaging. Note that just like the above pooled OLS the sample includes only selected sample of those who reported positive number of hours. We next discuss the results from the sample selection models.

Before presenting the results from models which correct for sample selection problem its useful and more appropriate to conduct a diagnostic check. To establish whether there is sample selection bias in our sample or not we formally test for the presence of sample selection problem in our data using the method suggested by Wooldridge (1995). According to Wooldridge (1995) the test should be conducted as follows. We first estimate the λ_t from the selection equation estimated for each cross section. In the next step, we estimate a fixed effects model using the selected sample only but including the λ_t as additional regressors. The first two columns of Table 5 report the results of a FE model, which includes λ_t as part of its regressors, when health is treated as exogenous variable.

-----------------------Table 5 here---

As suspected, most of the λ_t turn out to be statistically significant at less than 5 percent level of significance and a joint test of significance rejects the null hypothesis that jointly λ_t are zero, $H_0 : \lambda_t = 0$. These tests provide sample evidence that our panel data is contaminated with some sort of sample selection bias.[61] As expected also, most of the λ_t are negative. The interpretation is that higher level of participation (estimated from the logistic function) are related with higher wages. It is helpful to remember that λ_t is nothing but inverse of the probability of employment. The same result has been reported in the literature using similar modelling procedure for Germany (Dustmann and Rochinna-Barrachina 2007, Jäckle, 2007).

Having established the presence of a sample selection problem in our data we next discuss the results from the sample selection models discussed above. However we can not trust the results from the standard fixed effects sample selection models. The upward bias from the presence of a common time varying variable which affects the structural equation and the selection equation could be exacerbated with the employment of FE model. For this reason we estimate the Wooldridge (1995) model because it can assist in controlling individual heterogeneity and selection problem in a unified framework. The results from this model specification are reported in Table 6. The first two columns of Table 6 present the results of cases where health is treated as an exogenous variable. In the case of binary health the effect of health is relatively smaller than what has been reported in Table 3 and Table 4. In the case of the full health scale variable not only the size of the coefficient of health is smaller than what has been reported by the previous

[61] There might also be attrition. In this paper we hope that by controlling sample selection problem we are capturing the attrition bias if there is any. However, if the mechanism which drives attrition is not the same as the sample selection mechanism our results would be prone to attrition bias.

models but also it is statistically different from zero[62]. Column 3 of Table 6 reports a case where binary health is endogenous and the coefficient of health is not statistically

------------------------Table 6 here---

significantly different from zero. This result may imply that accounting for the individual heterogeneity using the FE technique is not enough to control for the upward bias introduced by sample selection bias.[63] Controlling for endogeneity reduces the estimates and controlling for selection reduces the magnitude of the effects further. This is an indication of the importance of controlling for sample selection bias and that the selection process is mainly operating via time variant variables and hence the FE model does not suffice. This is potentially the most important result of this paper.

To determine if there is reverse causality problem we have estimated instrumental variable 2SLS models. We estimate both the RE and FE models of panel data

------------------------Table 7 here---

[62] We could try with four dummies to capture the non linearity. However, we would run into an interpretation problem.

[63] We also tried to estimate the fixed effects sample selection model. However, there were computational problems. The sample selection model is not globally concave and hence the iteration procedure in the second stage of the estimation broke down in all the specifications that were attempted as the Hessian matrix was singular and indefinite. This is not uncommon in panel data sample selection estimations. Also note that the fixed effects model is subject to the infamous incidental parameter problem. Given that the estimation procedure involves two stages, the estimates from this model are also inefficient. For this reason, we employed 30 bootstrap replications to improve on the results. However, the estimation was terminated during the initial stage of the bootstrapping hence the models could not be estimated. For technical details of estimating a sample selection model with fixed effects see LIMDEP (2003) manual.

instrumental variable models. In the RE model we assume all the exogenous variables and instruments in the system to be uncorrelated to the error terms. In the FE version we allow for the possible correlation of individual heterogeneity and the regressors.[64] Table 7 reports the results from the IV (instrumental variable) random effects and fixed effects models. According to the results in Table 7 the instrumental variable models reveal that both the binary health and the full scale health variables are significantly different from zero at less than 5 percent level of significance except the full health scale variable in the fixed effect model. When we compare the results of IV2SLS (instrumental variable two stage least square) with that of the previous models we find that the health coefficient from the IV2SLS gives higher coefficients. This is not unexpected if there is any measurement error.[65] This upward bias could be explained by the presence of some time varying variable which affects both the wage offer equation and selection process. However, it is worth remembering that the standard errors from this estimation need to be corrected. We use Baltagi-Chang (2000) method to estimate the standard errors.[66]

Examination of the implied effects of the other variable in our model shows that our results are similar to what has been reported in the literature. In all specifications, we find a concave relationship between logwages and experience. Our results suggest that an

[64] The set of additional instruments in addition to the variables in the wage offer equation include age squared, father's education, mother's education and functional limitation.

[65] In the presence of a classical measurement error (the measurement error is not correlated with the true health measure) the coefficient of health variable is expected to be biased towards zero. We plan to extend the current research in to with errors in variables model.

[66] We use STATA 10 to conduct the computation.

additional year of experience increases wages by about 6-7%, at the mean level of years of experience. We also document a concave relationship between the number of years in a specific job and wages. When we look at the coefficient of education, for instance, we get very similar results as in the literature for developed countries. The estimated rate of return to education is about 4-5 percent which is within the range of the reports for developed countries (Psacharopoulos and Patrinos, 2002). Men from the visible minority group earn lower wage than men from the non visible minority group, ceteris paribus and provincial minimum wage is positively associated with wage.

--------------------------------Table 7 here--

2.6. Conclusion

This paper examines the effect of health on wages in a manner which accounts for a number of possible problems; the problem of unobservable individual fixed effects such as genetic endowment; measurement error from the employment of self assessed health; and sample selection problem associated with endogenous choice of participation into the labour market which might render estimates inconsistent. We utilize the estimator suggested by Wooldridge (1995) and some IV panel data estimators to take care of the aforementioned problems. The fact that there are unobservable components of health panel data is useful to control for individual heterogeneity. The results indicate that sample selection bias causes upward biased estimates of health effects on wages. After controlling for endogeneity of health and sample selection bias our results reveal that the effect of health on wages is, as expected, positive but not statistically significantly different from zero. The key message of this paper is that accounting for sample selection bias and endogenity of health is crucial in any attempt to uncover the health wage nexus.

The conventional earnings function should take into account the sample selection bias in order to obtain consistent and unbiased estimates of the impact of health and other human capital on wages.

There are a number of caveats to our model. We need to remember that the current paper aims at exploring the effect of contemporaneous health on wages. However, treating health as a stock of capital could imply that the lagged health should be related with current income. Despite the theoretical attractiveness of this line of argument the time period that health variables should be lagged is not easy to determine. In a more convincing way it would make more sense to postulate that long term health will have a cumulative effect on life cycle earnings, meaning childhood investment on health may be reaped during adulthood. Unfortunately our data do not allow us to make this line of enquiry and this issue is still not under full control of empirical research (Thomas and Strauss, 1997). We hope that the IV estimation employed in this paper can partially cure the endogenity introduced by omitting lagged health measures in our model. However, it is left for further research to account for all possible sources of endogeneity, errors in variables and dynamics of health in a unified framework. We have also abstracted from the impact of health on labor supply and hence on wages. We hope that the effect of health on labor supply is partially captured by the effect of health on the selection equations via the Inverse Mills Ratio. Lastly, our measure of health is self reported health and it is important to experiment with objective health measures to examine the robustness of our results.

References

Abowd, J. M., and Card, D. (1989) "On the Covariance Structure of Earnings and Hours Changes", *Econornetrica,* 57, pp. 41 1-45.

Andren, D. and Palmer, E. (2000) "The effect of sickness on Earnings," http://www.handels.gu.se/epc/data/html/pages/PDF/gunwpe0045.pdf, unpublished manuscript, Uppsala University, retrieved on Feb 10, 2004.

Baltagi, B. (2005) Econometric Analysis of Panel Data, 3rd ed., John Wiley and Sons, Ltd.

Baltagi, B. and Khant-Akom, S. (1990) "On Efficient Estimation with Panel Data: An Empirical Comparison of Instrumental Variables Estimators," *Journal of Applied Econometrics*, 5, pp. 401-406.

Baltagi, B.H. and Chang, Y.J. (2000) "Simultaneous Equations with Incomplete Panels," *Econometric Theory,* **16**, pp 269-279.

Becker, G.S. (1964) Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, New York: National Bureau of Economic Research.

Cahuc, P. and Zylberberg, A. (2004) Labour Economics, MIT, USA.

Chamberlain, G. (1984) "Panel Data", in V. Griliches and M.D. Intrrilirgator (Eds), *Handbook of Econometrics*, Volume II, pp. 1248-1318, Amsterdam: Elsevier Science.

Contoyannis, P. and Rice, N. (2003) "The Impact of Health on Wages: Evidence from the British Household Panel Survey," *Empirical Economics*, 26, pp. 599-622.

Currie, J. and Madrian, B. (1999) "Health, Health Insurance and the Labor Market," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3, pp. 3309-3416, Amsterdam: Elsevier Science.

Dustmann, C. and Rochinna-Barrachina, M.E. (2007) "Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage Equations," *Econometrics Journal*, 10, pp. 263-293.

Dustmann, C. and Windmeijer, F. (2000) "Wages and the Demand for Health - A Life Cycle Analysis," IZA Discussion Paper No.171, Institute for Fiscal Studies, Bonn.

Fissuh, E. (2005) On the Efficient Estimation with Panel Data: with an Application on Health and Wages in Canada", unpublished manuscript, Department of Economics, University of Manitoba, Canada.

Greene, W.H. (2003) Econometric Analysis, 5th Ed., Prince Hall, USA.

Griliches, Z. (1977) "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, 45, pp. 1-22.

Gronau, R. (1974) "Wage Comparisons-A Selectivity Bias," *The Journal of Political Economy*, 82(6), pp. 1119-1143.

Grossman, M. (1972) "On the Concept of Health Capital and Demand for Health," *Journal of Political Economy*, 80(2), pp. 223-255.

Halvorsen, R. and Palmquist, R. (1980) "The Interpretation of Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 70(3), pp. 474-475.

Hausman, J.J. and Taylor, W. (1981) "Panel Data and Unobservable Individual Effects, E*conometrica*, 49, pp. 1377-1398.

Hausman, J. J. (1978) "Specification Tests in Econometrics," E*conometrica*, 46, pp. 1251-1271.

Heckman, J. and MaCurdy, T. (1980) "A Life Cycle Model of Female Labor Supply," *Review of Economic Studies*, 47, pp. 47-74.

Haveman *et al* (1994) "Market Work, Wages and Men's Health," *Journal of Health Economics*, 13:163-182.

Heineck, G. (2004) "Up in the Skies? The relationship between Body Height and Earnings in Germany," unpublished manuscript, Department of Economics, University of Munich.

Hsiao, C. (2003) Analysis of Panel Data, 2nd ed., Cambridge University Press, UK.

Hum, D. Simpson, W. and Fissuh, E. (2006) "The Impact of Health on Labour Supply in Panel Data," *Global Business and Economics Anthology*, selected papers from the 2006 Business and Economics Society International Conference, Florence Italy, July 15-19, 2006.

Jäckle, R. (2007) "Health and Wages: Panel Data Estimates Considering Selection and Endogeneity," working paper, Ifo Institute for Economic Research, (February 27).

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(3), pp. 302-329.

Kjellstrom, C. and Bjorklund, A. (2002) "Estimating the Return to Investment in Education: How Useful is the Standard Mincer Equation?," *Economics of Education Review*, 21, pp. 195-210.

Kyriazidou , E. (1997) "Estimation of a Panel Data Sample Selection Model," *Econometrica*, 65(6), pp. 1335-1364.

Lee, L.F. (1982) "Health and Wages: A Simultaneous Equation Model with Multiple Discrete Indicators," *International Economic Review*, 23(1), pp. 199-221.

Leung S.F., Yu S. (1996) "On the Choice between Sample Selection and Two-Part Models," Journal *of Econometrics*, 72 (1-2), pp. 197-229.

Leung, S. F. and Wong, C. T. (2002) "Health Status and Labour Supply," Unpublished Manuscript, Honk Kong University of Science and Technology.

Maddala, G.S. (2001) Introduction to Econometrics, 3rd ed., England: John Wiley and Sons Inc.

Mincer, J. (1974) Schooling, Experience and Earnings, New York: National Bureau of Economic Research.

Mroz, T.A. (1987) "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," *Econometrica,* 55, pp. 765-800.

Psacharopoulos, G. and Patrinos, H.A. (2002) "Returns to Investment in Education: A Further Update," World Bank Policy Research Working Paper 2881, Washington D.C.

Schultz, T. W. (1961) "Investment in Human Capital," *American Economic Review*, 51, pp. 1-17.

Statistics Canada (2004) Survey of Labour and Income Dynamics, [http://www.statcan.ca/start.html]: Retrieved on March 28th 2006.

Thomas, D., Strauss, J. (1997) "Health and Wages: Evidence on Men and Women in Urban Brazil," *Journal of Econometrics*, 77, pp. 159-185.

Vella, F. and Verbeek , M. (1999) Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias, " *Journal of Econometrics,* 90 , pp. 239-263.

Vella, F. (1998) "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, XXXIII (1), pp. 127-169.

Wales, T.J., Woodland, A.D. (1980) "Sample Selectivity and the Estimation of Labor Supply Functions," *International Economic Review*, 21(2), pp. 437-468.

Walker, I., and Thompson, A. (1996) "Disability, Wages and Labour Force Participation: Evidence from UK Panel Data," Department of Economics, Keele University Working Paper no.96/14.

Wooldridge, J.M. (1995), "Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions," *Journal of Econometrics*, 68, pp. 115-32.

Wooldridge, J.M. (2002) Econometric Analysis of Cross Section and Panel Data, 2nd ed., London, England: The MIT Press

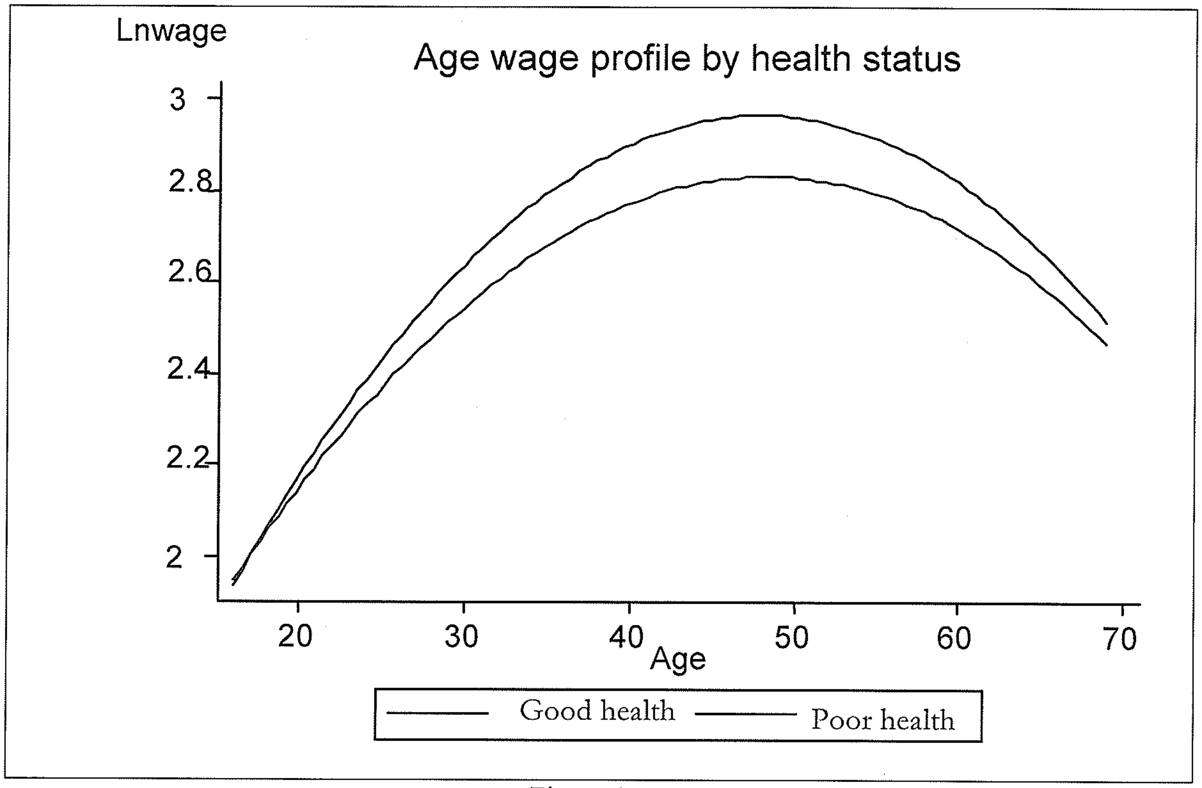


Figure 1

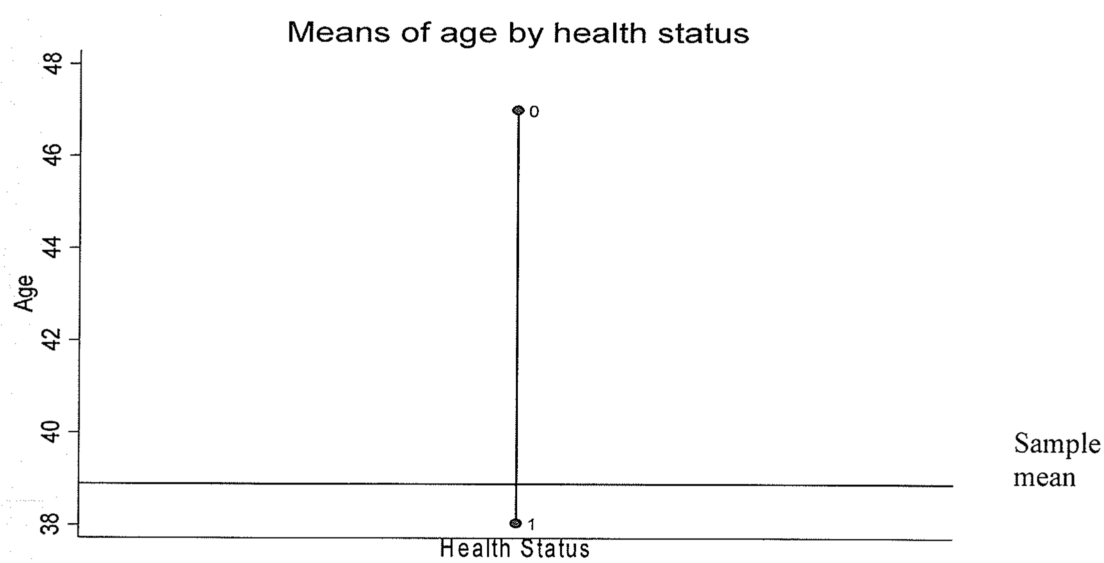
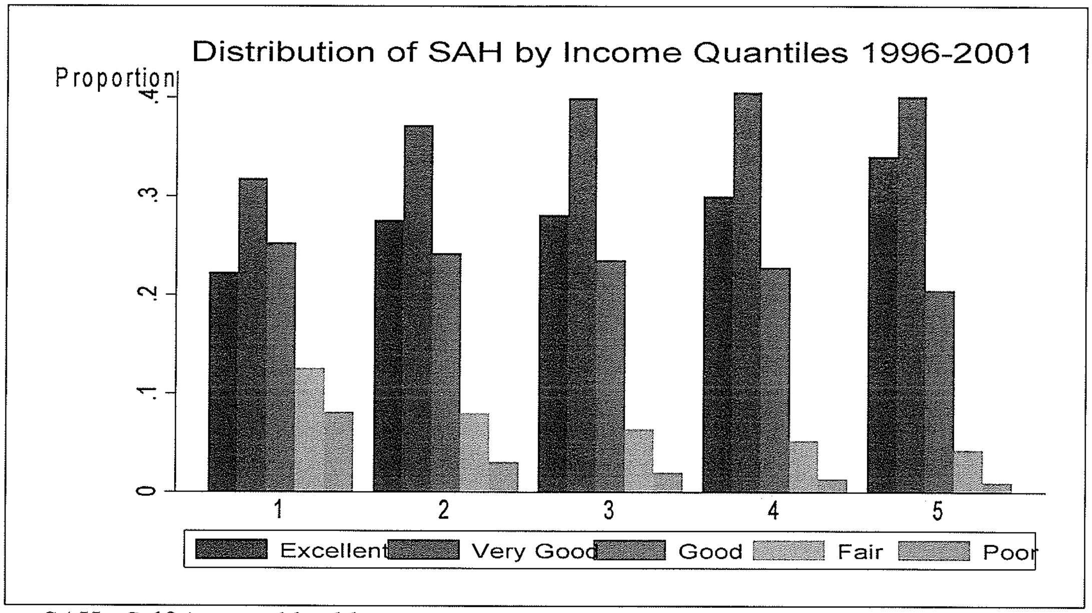


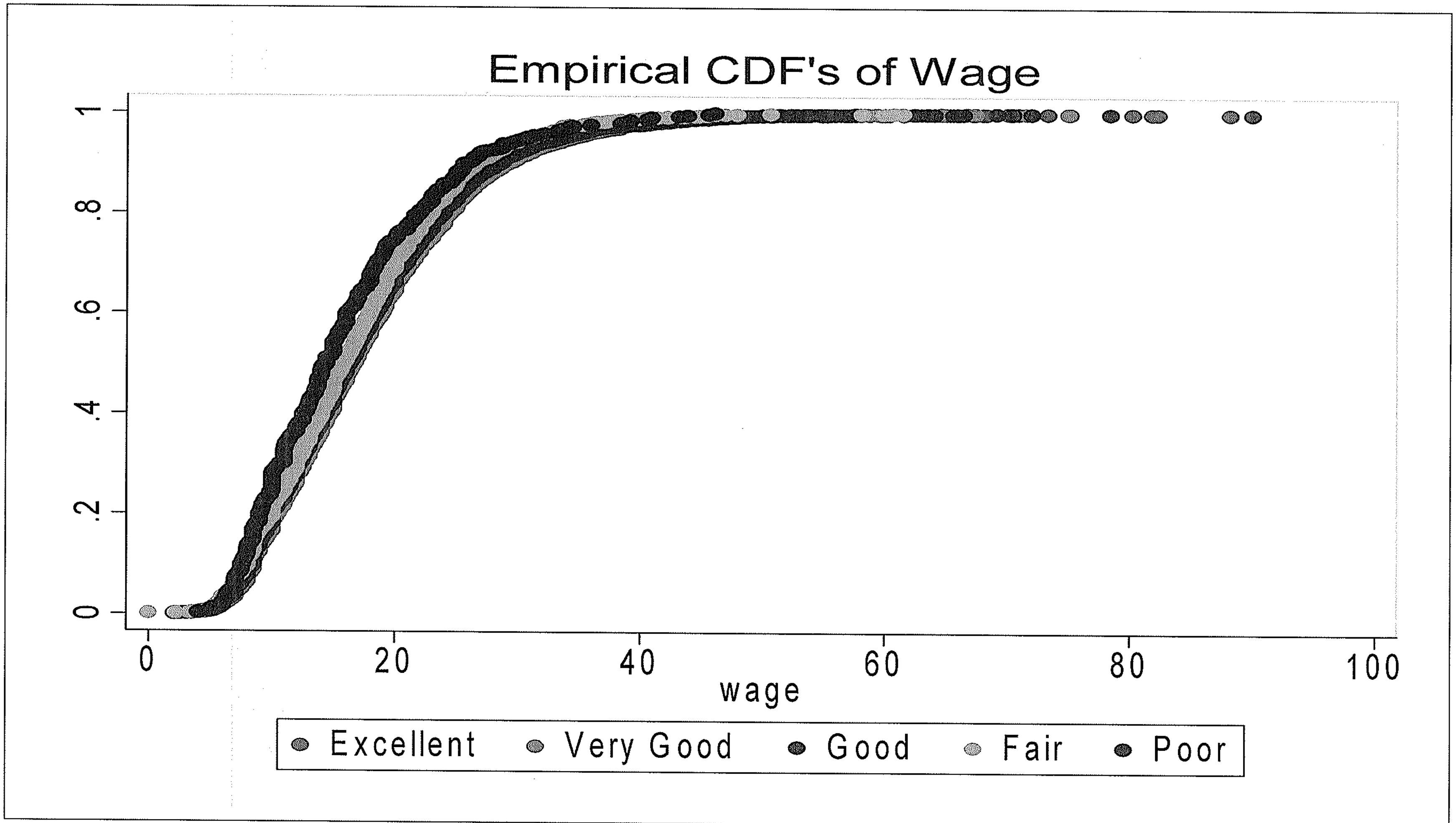
Figure 2

Note the vertical and horizontal lines represent the mean levels of the variables.



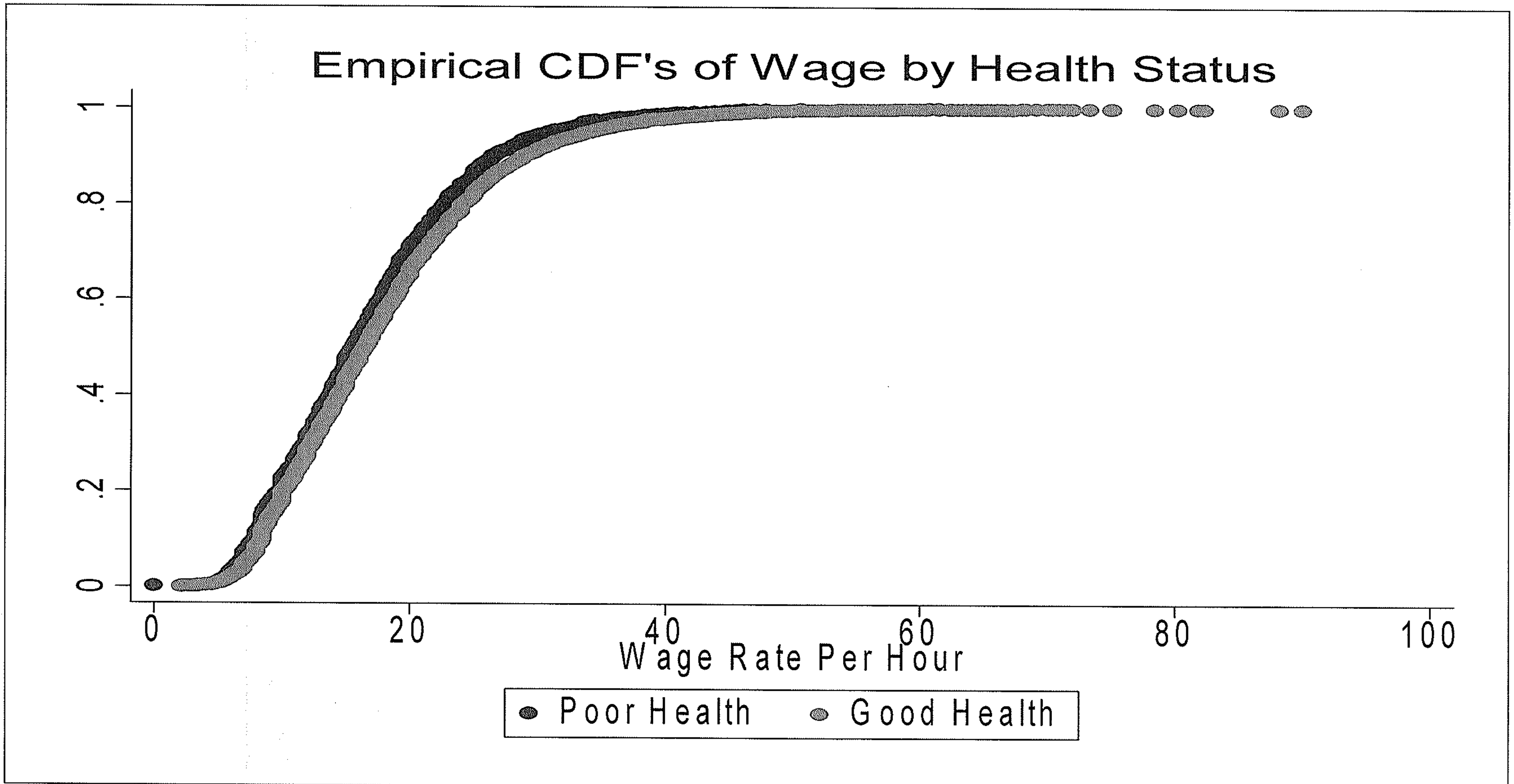
Note: SAH= Self Assessed health

Figure 3.



Data Source: SLID (2001)

Figure 4



Data Source: SLID(2001)

Figure 5

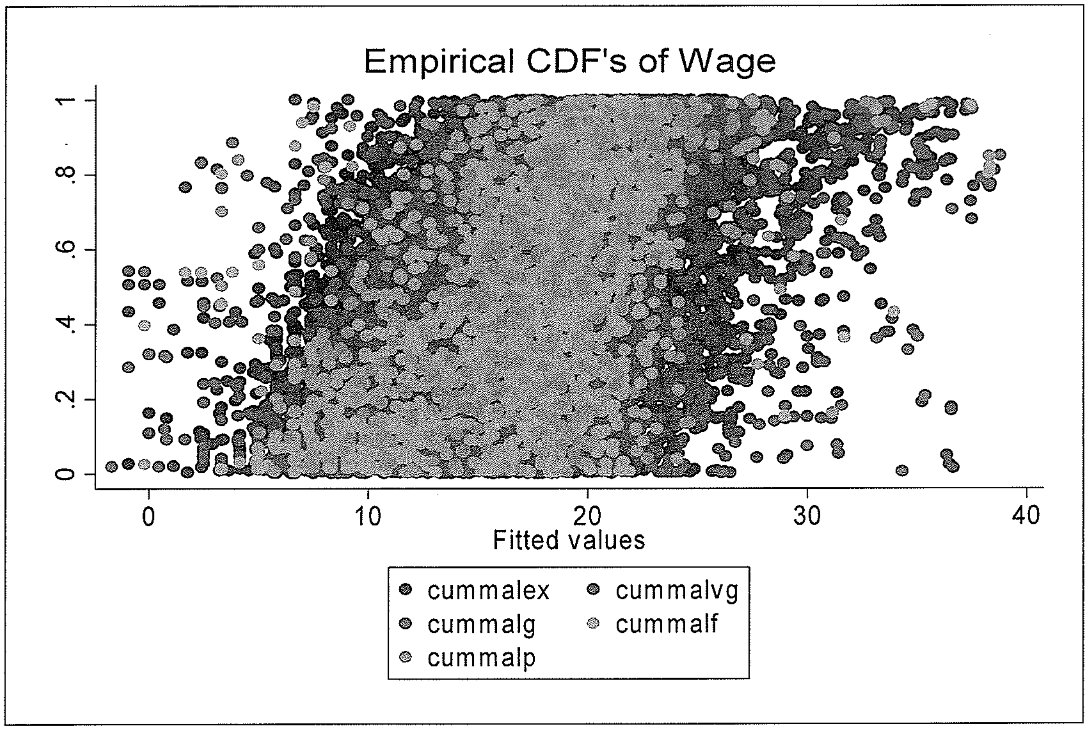


Figure 6

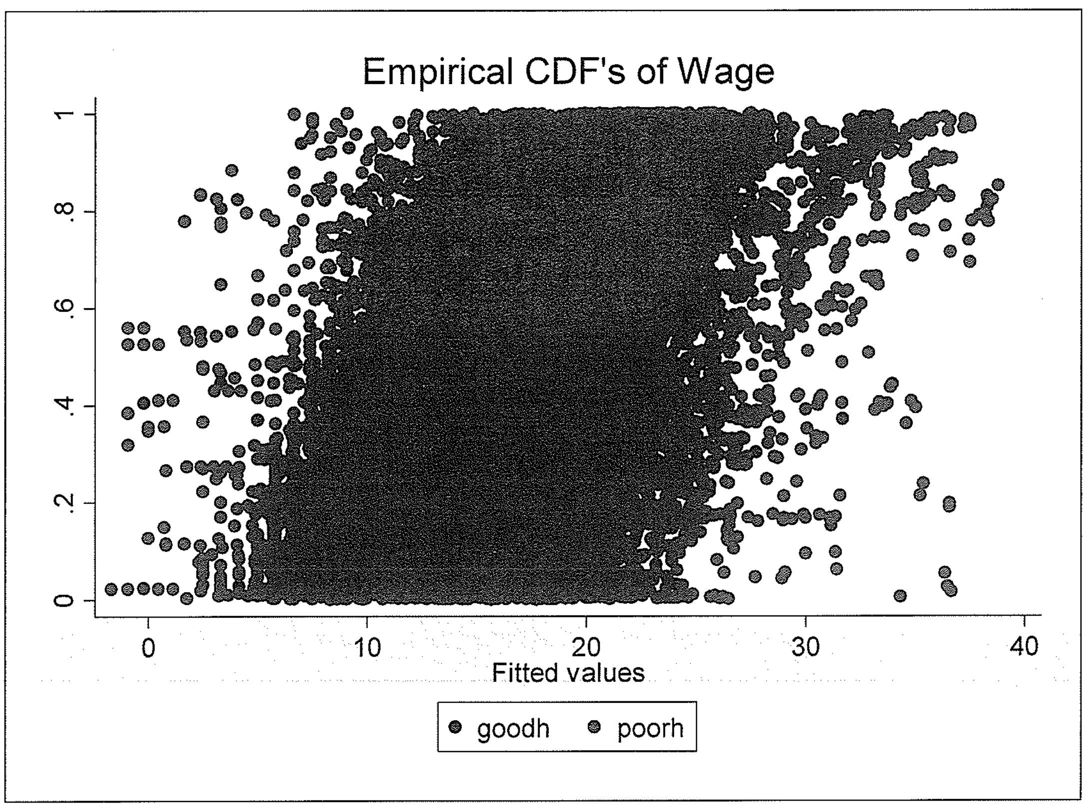


Figure 7

Table 1. Descriptive Statistics of some variables across health categories

	lnwage	Education	Tenure	Experience	Number of weeks worked	Annual Hours worked	Age
Excellent Health							
Observations	15167	18587	15569	16956	19713	19713	19713
Mean	2.64	13.47	17.94	13.81	40.42	1352.22	34.32
Std. Dev.	0.54	3.53	12.44	12.81	19.70	1038.51	13.33
Very good health							
Observations	18066	22155	20304	20082	23364	23364	23364
Mean	2.72	13.23	21.16	18.03	42.70	1461.73	38.82
Std. Dev.	0.51	3.60	12.24	12.81	18.51	1031.68	12.87
Good health							
Observations	9898	12884	12127	11581	13713	13713	13713
Mean	2.72	12.54	23.92	20.46	40.36	1374.03	42.06
Std. Dev.	0.50	3.74	12.37	12.98	20.41	1064.30	12.61
Fair health							
Observations	2298	3943	3688	3592	4204	4204	46.163
Mean	2.69	11.51	28.12	22.57	30.53	986.84	46.16
Std. Dev.	0.49	3.93	12.85	14.01	24.69	1075.88	12.71
Poor Health							
Observations	455	1617	1498	1462	1751	1751	1751
Mean	2.64	10.60	30.98	22.24	13.09	384.23	49.03
Std. Dev.	0.48	4.36	12.02	13.85	21.50	789.10	11.46

Data source: SLID (2004)

Table 2. Transitional probabilities (transitional matrix)

		Health Status (t+1)				
		Excellent	Very good	Good	Fair	Poor
Health Status(t)	Excellent	55.8	31.99	10.15	1.7	0.36
	Very good	23.6	50.99	21.36	3.4	0.61
	Good	11.86	32.86	41.8	11	2.33
	fair	5.2	15.56	31.08	36	12.41
	Poor	2.5	6.35	14.2	27	49.89

Source: SLID (2001)

Table 3. Pooled OLS results

	Exogenous health		Endogenous health	
	Full health scale	Binary Health	Full health scale	Binary Health
Constant	0.916	0.916	1.055	0.923
	(25.68)	(25.68)	(29.81)	(23.25)
Education	0.044	0.042	0.042	0.042
	(48.44)	(47.56)	(44.80)	(44.64)
Tenure	0.006	0.008	0.008	0.008
	(1.95)	(2.59)	(2.29)	(2.30)
Tenure squared	-0.048	-0.050	-0.048	-0.048
	(6.70)	(7.13)	(6.58)	(6.56)
Experience	0.041	0.038	0.040	0.040
	(12.81)	(12.08)	(12.00)	(12.02)
Experience squared	-0.031	-0.028	-0.030	-0.030
	(4.37)	(3.95)	(4.08)	(4.10)
Married	0.134	0.140	0.144	0.144
	(15.17)	(16.12)	(15.65)	(15.68)
Health	0.022	0.074	0.018	0.019
	(6.33)	(6.09)	(4.72)	(5.17)
Children age 0-5	0.025	0.024	0.024	0.024
	(3.96)	(3.76)	(3.57)	(3.55)
Children age 5-15	0.007	0.007	0.006	0.006
	(1.74)	(1.89)	(1.52)	1.51)
Regional minimum wage	0.104	0.104	0.100	0.100
	(21.31)	(21.31)	(19.28)	(19.27)
R^{-2} (Adjusted R^2)	0.42	0.42	0.42	0.42
Observations	38689	38689	38689	38689

Data Source: (SLID 2001)
Note: Absolute values of t-statistics in parentheses.
Tenure squared and experience squared are multiplied by 100.

Table 4. Fixed effects models

	Exogenous health		Endogenous health	
	Full health scale	Binary Health	Full health scale	Binary Health
Education	0.042	0.042	0.042	0.042
	(46.30)	(46.70)	(43.70)	(43.70)
Tenure	0.010	0.009	0.008	0.008
	(2.89)	(2.80)	(2.35)	(2.35)
Tenure squared	-0.053	-0.052	-0.050	-0.050
	(7.42)	(7.36)	(6.70)	(6.70)
Experience	0.037	0.037	0.040	0.040
	(11.78)	(11.75)	(11.81)	(11.81)
Experience squared	-0.027	-0.026	-0.030	-0.030
	(3.76)	(3.70)	(3.96)	(3.96)
Married	0.139	0.139	0.142	0.142
	(15.80)	(15.85)	(15.22)	(15.22)
Health	0.019	0.072	0.018	0.018
	(5.64)	(5.87)	(4.68)	(4.683
Children age 0-5	0.022	0.022	0.022	0.022
	(3.38)	(3.41)	(3.26)	(3.26)
Children age 5-15	0.007	0.007	0.006	0.006
	(1.69)	(1.69)	(1.38)	(1.38)
Regional Minimum Wage	0.100	0.101	0.097	0.097
	(20.22)	(20.42)	(18.42)	(18.42)
Observations	47646	47646	47646	47646
Groups	11338	11338	11338	11338

Data Source: (SLID 2001). Absolute values of t-statistics in parentheses.

5. Fixed effects model with selection corrections

	Binary health		Full Health scale	
Education	-0.009	(0.91)	0.002	(0.17)
Experience	0.051	(4.41)	0.053	(4.55)
Experience squared	-0.001	(4.43)	-0.001	(7.43)
Kids(age 0-5)	0.019	(1.7)	-0.002	(0.15)
Kids(age 5-15)	-0.005	(0.52)	-0.003	(0.37)
Health	-0.614	(0.91)	0.003	(0.59)
Tenure	0.043	(3.87)	0.047	(4.17)
Tenure square	0.130	(7.53)	0.136	(7.9)
Minimum wage	-0.012	(0.67)	-0.014	(0.73)
λ_{1996}	-1.429	(8.29)	0.082	(1.65)
λ_{1997}	-1.705	(10.08)	-0.209	(4.78)
λ_{1998}	-1.812	(10.84)	-0.328	(8.06)
λ_{1999}	-2.021	(12.19)	-0.548	(13.58)
λ_{2000}	-2.250	(13.59)	-0.787	(18.04)
λ_{2001}	-0.906	(5.52)	0.531	(11.09)
Constant	-0.213	(0.13)	0.865	(0.52)
Sigma_u	1.753		1.832	
Sigma_e	0.551		0.552	
Rho	0.910		0.917	
Observations	47646			
Groups	11338			

Data Source: SLID (2001). Absolute value of t-statistics in parentheses.
Note: Five regional dummies were included in the models but not reported.

Table 6. Wooldridge (1995) model estimates

	Wooldridge(1995) models				
	Exogenous health		Endogenous health		
	1	2	3	4	Fixed Effect
Education	0.006	0.056	0.053	0.053	0.056
	(0.49)	(26.83)	(18.48)	(18.03)	(26.83)
Experience	0.042	0.069	0.076	0.077	0.069
	(5.66)	(22.82)	(16.65)	(15.99)	(22.82)
Experience squared	0.000	-0.001	-0.001	-0.001	-0.001
	(0.12)	(16.63)	(11.3)	(10.9)	(16.63)
Tenure years	0.022	0.021	0.016	0.016	0.021
	(8.86)	(8.47)	(4.26)	(4.00)	(8.47)
Tenure square	-0.043	-0.042	-0.051	-0.051	-0.042
	(8.65)	(8.45)	(6.91)	(6.58)	(8.45)
Children age 0-5	0.171	0.067	0.060	0.058	0.067
	(5.33)	(6.26)	(5.1)	(4.64)	(6.26)
Children age 5-15	-0.088	-0.059	-0.063	-0.063	-0.059
	(8.41)	(16.54)	(12.08)	(11.6)	(16.54)
Health	0.017	0.002	0.090	0.001	0.002
	(5.07)	(0.6)	(0.34)	(0.25)	(0.6)
Racem1	-0.183	-0.370	-0.362	-0.361	-0.37
	(3.45)	(18.78)	(12.33)	(11.46)	(18.78)
λ_{1996}	8.647	8.817	8.873	8.942	8.817
	(14.8)	(19.91)	(13.56)	(12.81)	(19.91)
λ_{1997}	-12.739	-15.102	-14.919	-14.994	-15.102
	(22.32)	(34.85)	(24.13)	(23.21)	(34.85)
λ_{1998}	7.369	12.404	12.169	12.375	12.404
	(16.76)	(38.82)	(28.55)	(27.8)	(38.82)
λ_{1999}	-2.031	-5.292	-5.153	-5.455	-5.292
	(4.69)	(20.61)	(13.85)	(13.69)	(20.61)
λ_{2000}	-6.128	-1.381	-1.609	-1.492	-1.381
	(4.71)	(6.32)	(5.32)	(4.82)	(6.32)
λ_{2001}	-0.509	-0.028	-0.062	-0.060	-0.028
	(3.19)	(1.52)	(2.62)	-2.53)	(1.52)

Data Source: SLID (2001). Absolute value of t-statistics in parentheses

Note: 1= health variable is a binary health dummy; 2= health variable is a full scale health variable; 3= binary health variable is endogenous health; 4= full scale health variable is endogenous health. There are other variables which were included but not reported.

Table 7.Instrumental estimation results

	Random Effects		Fixed Effects	
	1[a]	2[b]	3[c]	4[d]
Good health				
Very good health				
Health	0.326	0.031	0.362	0.035
	(8.15)	(2.50)	(5.92)	(1.56)
Education	0.043	0.043	0.008	0.009
	(29.64)	(29.13)	(1.06)	(1.16)
Experience	0.030	0.028	-0.017	-0.016
	(6.74)	(6.35)	(1.52)	(1.52)
Experience squared	-0.0001	-0.0001	-0.00003	0.0002
	(1.38	(0.96)	(0.13)	(0.67)
Tenure	0.028	0.031	0.117	0.117
	(6.32)	(6.82)	(10.79)	(11.24)
Tenure squared	-0.081	-0.087	-0.087	-0.110
	(8.46	(9.12)	(3.77)	(5.20)
Visible Minority	-0.042	-0.045		
	(1.61	(1.69)		
Regional minimum wage	0.105	0.101	-0.056	-0.054
	(7.71)	(7.96)	(3.37)	(3.38)
Sigma_u	0.357	0.374	0.722	0.702
Sigma_e	0.179	0.167	0.179	0.167
Rho	0.798	0.834	0.942	0.947

Data Source: SLID (2001). Absolute value of t-statistics in parentheses
Note: Regional dummies were included in all of the models but not reported.

[a,c] health variable is binary variable. [b,d] health variable is full health scale.

For the RE and FE models the reference health category is poor health which assumes a value of zero is self reported health is poor or fair otherwise 1.

Table 8. Correlated random Effects Health Model

	Coefficient	Standard Errors	z-statistic
Father's education	0.073	0.038	1.95
Mother's education	0.143	0.043	3.29
Kids	0.108	0.151	0.72
Age	-0.066	0.023	-2.84
Age square	0.013	0.028	0.48
Logarithm of wage	0.589	0.085	6.93
Other Income	0.000	0.000	0.41
Dummy 1996	0.833	0.330	2.52
Dummy 1997	1.086	0.340	3.20
Dummy 1998	0.783	0.332	2.36
Dummy 1999	0.804	0.330	2.43
Dummy 2000	0.089	0.314	0.28
Constant	4.178	0.481	8.69
/Lnsig2u	1.575	0.062	25.27
Sigma_u	2.198	0.068	32.09
Rho	0.595	0.015	39.61
N	38689		
Groups	9208		
Wald chi2(27)	449.72		

Data Source: SLID (2001). Absolute value of t-statistics in parentheses.

Chapter 3

3. The Impact of Childcare Cost on Maternal Labor Supply: Individual Heterogeneity and Sample Selection Corrected Panel Data Evidence from Canada

Abstract

This paper estimates the impact of childcare cost on maternal labor supply decisions using panel data from Canadian Survey of Labor and Income Dynamics (SLID) spanning from 1999 to 2004. The estimated elasticity of labor supply with respect to childcare price is compared across the different panel data estimators. This paper confirms that failure to control for the individual heterogeneity and sample selection bias produces upward biased estimates of the elasticity of labour supply with respect to childcare costs. Moreover, as expected, our study confirms that labor supply decisions of single mothers are more sensitive to childcare price changes than married mothers. Panel data models that do not control for sample selection bias give a childcare price elasticity of labor supply (annual number of hours worked) that range from -0.012 to -0.113 and -0.08 to -0.166 for married and single mothers, respectively. Using a model which controls for sample selection and individual heterogeneity, elasticity of annual number of hours worked with respect to childcare cost is found to be -0.015 and -0.068 for married and single mothers respectively. Moreover using binary logit panel data models which control for individual heterogeneity and sample selection bias the elasticity of employment with respect to childcare price is estimated to be -0.01 and -0.48, for married and single mothers respectively.

3.1. Introduction

According to the Survey of Labor and Income Dynamics (SLID) the percentage of employed mothers in Canada with at least one preschool child has more than doubled from 31 percent in 1995 to 67 percent in 2004.[67] At the same time, the percentage of mothers utilizing childcare centres has increased from 42 percent in 1994 to 66 percent in 2004. This increasing labor market participation of mothers in general and that of married mothers in particular has been one of the stylized facts in labor economics in the 20th century across many developed nations (Cahuc and Zylberberg, 2004). However, paid childcare utilization did not increase for all mothers across the income distribution. While mothers in the highest income quintile have witnessed an increased usage of paid

[67] Using 2005 Canadian census data Roy (2006) reports the percentage of employed mothers with preschool children to be 67 percent.

childcare centers, mothers from lower income quintile have registered a decline in the utilization of childcare centers. For instance from 1999 to 2004 the percentage of mothers from the highest income quintile who utilized paid childcare was about double that of the mothers from the lowest income quintile. Part of the explanation for this inequality in childcare utilization could be affordability of childcare because married mothers enjoy higher average family income than single mothers. Childcare cost could be one of the key factors responsible for the dependency of single mothers on social assistance, making their fight for walking off the welfare road more difficult. Childcare costs are expected to affect labor supply decision of mothers negatively. On average from 1999 to 2004, 35 percent of mothers in SLID declared that childcare is the main reason for taking a part time job.[68]

The phenomena of increased labor force participation and increased utilization of childcare by mothers from higher income quintile on the one hand and decreased usage of paid childcare services by lower income quintile on the other hand have attracted considerable attention of researchers, policy makers and politicians. This inequality has very important implications for labor market participation of mothers, child development and public financing of childcare. Government intervention mechanisms which aim at increasing the labor force participation of mothers are usually conditional on employment and the mode of childcare choice is an irrelevant factor. Employment-based childcare support policies are expected to reduce the fiscal burden of the government by helping the transition of mothers from social assistance to the work force. However, childcare

[68]All statistics reported in this paragraph are author's calculation from SLID and are very similar to official reports from STATISTICS Canada.

polices which aim at increasing maternal employment need a clear link between childcare expenditure and labor supply of mothers[69]. Policy makers need to know not only the direction but also the magnitude of the sensitivity of maternal labor supply decisions to changes in childcare costs in designing any intervention mechanism which aims at increasing labor force participation of mothers. This forms the prime motive of this study. The main objective of this paper is to examine the relationship between childcare cost and labor supply of mothers using appropriate theory and econometric techniques to shed light on the link between childcare cost and maternal labor supply in Canada.[70]

A number of researchers from the U.S., Canada and Western Europe have empirically found that childcare costs reduce participation of mothers in the labor market and number of hours worked, but there is less agreement on the magnitude. Estimates of labor supply elasticity with respect to childcare cost range from -0.2 to -0.92 and the literature suggests that the estimates of labor supply response to changes in childcare cost are sensitive to the choice of modeling approach. It follows that a direct comparison of the results from these studies may be illegitimate because they differ in the underlying behavioural assumptions and model specification. Using nonlinear panel data models which control for individual heterogeneity and sample selection bias we find a negative childcare price elasticity of labor supply using a data set from SLID. Moreover, we also document that childcare price elasticity of labor supply of single mothers is relatively

[69] Of course we need also to know the fiscal efficacy of public financing of childcare.

[70] When we look at the child development objective of childcare polices there is no consensus. However, it is conceivable that maternal childcare could be superior from the point view of the family. It can be the case that some parents may provide quality childcare to their kids and it is hard to argue otherwise decisively. Some parents can provide a conducive and productive home care to their children. But some families might lack the financial and human capital skills to provide good home environment to boost child development.

more elastic than that of married mothers. Our results suggest that the elasticity of annual number of hours worked with respect to childcare price of married mothers to be in the range of -0.01 to -0.12 and that of single mothers to be in the range of -0.10 to -0.16.We also find the elasticity of employment with respect to childcare price for married and single mothers to be -0.01 and -0.48, respectively.

The paper differs from the existing literature in a number of ways. First, we examine the impact of childcare costs on labor supply of mothers in a life cycle setting. Second, we employ nonlinear panel data models which are appropriate econometric models to account for the unobservable individual fixed effects implied by a life cycle labor supply model. We derive our sample from the SLID panel running from 1999 to 2004. This very rich panel data allows us to examine the robustness of the existing findings in the literature using structural econometric models which control for individual heterogeneity and sample selection problem. Third, unlike the widely cited previous Canadian studies on childcare and labor supply of mothers our data includes both single and married mothers and the imputed child price equations control for individual heterogeneity and sample selection bias in unison. [71]

The paper is organized as follows. Section two provides a brief literature review. Section three presents the empirical labor supply model. Section four discusses the data and

[71] One may argue that marital status is an endogenous outcome. In doing so we need to estimate a marital status equation where one of the regressors is child benefit and welfare payments and this is expected to shed light on whether the generosity of the welfare system (child benefit scheme) is associated with marital decisions.

provides some preliminary analysis. Section five presents the main empirical results from the econometric estimation. Section six concludes.

3.2. Review of Related Studies

In his seminal work Heckman (1974) shows that childcare costs reduce participation of mothers in the labor market and number of hours worked. Following Heckman (1974) there have been a number of studies, with slight alterations of the model specification and behavioral assumptions, on the sensitivity of labor supply of married mothers to childcare prices. This section presents very brief review of the main evidence from USA and Western Europe followed by Canadian literature.

As far as the evidence from the USA is concerned the literature indicates that the impact of childcare cost on female labor supply is generally weak (Blau 2003). Ribar (1995) presents a seminal paper on the impact of childcare cost on labor supply in a static framework. Ribar (1995) examines family demand for paid and unpaid childcare services and the effect of these demands on work efforts of married women. He finds that labor supply is relatively more sensitive to changes in wages but less sensitive to changes in costs.[72] Kimmel (1998) reports that childcare price elasticity of employment for single mothers and married mothers to be -0.22 and -0.92 respectively. In contrast to what we have documented in this paper her results imply that childcare price significantly impedes married mothers' labor force participation in the USA relative to single mothers. However, as she explicitly admits in her paper this is in contradiction to expectation. Blau and Robins (1998), Connely (1992), and Blau and Hagy (1998) employ similar modeling

[72] Ribar (1995) does not report elasticity estimates.

framework and report an elasticity of labor supply with respect to childcare cost to be in the range of -0.39 and -0.2. Maume (1991) employs a panel data of working mothers who paid for childcare and he reports that childcare expenditure is a key predictor of employment turnover. Besides he finds that the effect of childcare expenditure is highest for mothers with preschool children and childcare expenditure is not associated with a transition from part time to full time employment.[73] Kimmel and Powell (2006) emphasize that the nature of the jobs that mothers hold is crucial for childcare decisions and working irregular hours exacerbates the hurdle to access good quality childcare.

When we look at the European evidence we find similar results with the effect of childcare costs on labor supply being relatively weaker than those reported for North America. Wronlich (2004) and Chone *et al.* (2003) report weak relationship between childcare costs and labor supply decisions of mothers in Germany and France, respectively. Wronlich (2004) and Chone *et al.* (2003) present a very good review of studies in Germany and France before documenting their own evidence. The difference between the elasticity of labor supply with respect to childcare costs between the European and North American studies could be partially explained by the presence of highly subsidized childcare services in Europe. This reminds us that any comparison between country specific childcare and maternal labor supply studies should take into account the institutional differences among countries.

[73] Very little attention has been given to the study of the determinants of childcare expenditure in the literature and the usual practice is to estimate a childcare expenditure equation as an auxiliary regression to obtain imputed childcare prices for the employment and childcare model choice equation. Statistics Canada identifies the study of the determinants of childcare expenditure to be scanty and poorly researched area in the childcare research in Canada.

When we look at the Canadian evidence, Powell (1997) is the first attempt to employ both childcare and labor supply decisions jointly to analyze the impact of childcare cost on labor supply in Canada. Powell (1997) employs a static labor supply model to analyze the impact of childcare costs on labor supply and she documents the elasticity of employment and childcare costs with respect to number of hours to be -0.38 and -0.32, respectively. Cleveland *et al.* (1996) employ a bivariate probit equation to model the probability of employment and probability of purchasing childcare from the market. Cleveland *et al.* (1996) report the childcare elasticity of employment and paid childcare utilization to be -0.39 and -1.1 respectively. Lefebvre and Merrigan (2005) study the impact of the Québec's low fee (5/day) regulated childcare policy on labor supply of mothers with young children using difference in difference (DID) method and they find that the policy change had a very strong effect on labor supply decisions. Lefebvre and Merrigan employed the Survey of Labor and Income Dynamics (SLID). Baker et al (2005) examine the impact of the Quebec's universal childcare policy on maternal employment using DID and report the labor supply elasticity to be -0.236. Baker et al (2005, 2008) derive their sample from the National Longitudinal Survey of Children and Youth (NLSCY). However, as Lefebvre and Merrigan (2005) indicate in their paper, the results from the natural experiment need to be scrutinized using structural econometric models to ascertain their robustness.[74] White (2001) presents a lucid summary of the evolution of Canadian childcare policy in the 20th century. She argues that less attention was given to the childcare development goal of subsidizing childcare expenditures and Canadian childcare policy has mainly been motivated by increasing the labor force participation of mothers.

[74] Moreover, we need to examine if their findings could be extended to the rest of Canada.

3.3. Basic Life-Cycle Labor Supply Model

The theoretical setting of our labor supply model is standard theory of consumer behaviour where a consumer is faced with a problem of maximizing her lifetime utility subject to a wealth constraint. A basic static labor supply model of married mothers with a childcare decision was suggested by Connelly (1992) and applied to Canadian data by Powell (1992). In this paper we follow Connelly (1992) but we extend the model to a life cycle setting as in MaCurdy (1981).

Assume that a typical mother who is supposed to be a primary caregiver for her children faces a choice between leisure, childcare or work over her lifetime.[75] Assume also that the utility function is quasiconcave at age t which is given by:

$$U[c_t, l_t, q_t; x_t] \qquad [3.1]$$

Where c_t, l_t, q_t and x_t are within-period consumption, leisure time, childcare quality and a vector of individual characteristics[76] of the agent respectively. To proceed we make a number of standard assumptions. We assume that the utility function is increasing in c_t, l_t and q_t. Also assume that the utility function is additively temporally separable and within period childcare quality q_t, depends positively on the amount of time children spend under their mothers' care, s_{mt}, and the amount of time they spend in non-maternal childcare, s_{nt}. Allowing for the difference in the productivity of maternal care and non-

[75] To be more specific, this particular choice set is relevant during the time period where she has at least one child who requires childcare, otherwise her choices will only be leisure or work and no childcare.

[76] x_t contains observable and non observable individual characteristics.

maternal care let q_m and q_n denote the productivity of maternal care and non-maternal childcare respectively. Hence, average within period childcare quality over all children[77] in the family is:

$$q = q_m s_{mt} + q_n s_{nt} \qquad [3.2]$$

The mother is limited by the time path of her wealth constraint[78]:

$$a_t = (1 + r_t) a_{t-1} + d_t + w_t h_t - c_t - p_t s_{nt} \qquad [3.3]$$

where $h_t = (1 - l_t - s_{mt})$, the total number of hours available for work is normalized to 1; a_t is real assets at time t; a_{t-1} is the real asset endowment of the consumer at time $t-1$; w_t is the exogenously given within period wage level; d_t is within period non-wage income[79]; and p_t is the price of non-maternal childcare at time t. In this dynamic setting we also assume that a mother can freely borrow and save at each point in time at an interest rate r_t and that the time discount factor is ρ. Another time constraint is the amount of time spent on the child after normalizing to 1, given by $1 = s_{mt} + s_{nt}$. In effect we have two time constraints: The first one assumes that the mother can spend her time either working, at leisure or in childcare and the second assumes that the total time

[77] Total number of children is assumed to be an exogenous variable.
[78] This formulation ignores the role of taxes (progressive) which might impact the separability of choices. See Blomquist (1985) for the effect of progressive taxes on the separability of choices and Kumar (2005) for the effect of relaxing the linearity in the budget constraint on the labor supply response parameters.
[79] Strictly speaking non-wage income is $d_t + r_t a_t$.

available for the children is spent either under maternal care or non-maternal care. Hence the utility maximization problem of our consumer is one of :

$$\underset{\{c_t, l_t, a_t, s_{nt}, s_{mt}\}}{Max} \; U = \sum_{t=0}^{T} U[c_t, l_t, q_t; x_t](1+\rho)^{-t} \qquad [3.4]$$

Subject to $\sum_{t=0}^{T} \kappa_t [a_t - (1+r_t)a_{t-1} - d_t - w_t h_t + c_t + p_t s_{nt}](1+r_t)^{-t}$

This problem can be solved using the Lagrange technique. Assuming interior solutions, the first order conditions for maximization are

$$\frac{\partial \Omega}{\partial c_t} = U_{c_t}[c_t, l_t, q_t; x_t] - \frac{1+\rho}{1+r}\kappa_t = 0 \qquad [3.5]$$

$$\frac{\partial \Omega}{\partial l_t} = U_{l_t}[c_t, l_t, q_t; x_t] + \kappa_t \frac{1+\rho}{1+r}(-w_t) = 0 \qquad [3.6]$$

$$\frac{\partial \Omega}{\partial s_{mt}} = U_{q_t}[c_t, l_t, q_t; x_t](q_m - q_n) + \kappa_t \frac{1+\rho}{1+r}(p_t - w_t) = 0 \qquad [3.7]$$

$$\frac{\partial \Omega}{\partial s_{nt}} = U_{q_t}[c_t, l_t, q_t; x_t](q_n - q_m) - \kappa_t \frac{1+\rho}{1+r}(p_t - w_t) = 0 \qquad [3.8]$$

$$\kappa_t = (1+r_{t+1})\kappa_{t+1} \qquad [3.9]$$

where $U_{c_t}[.]$, $U_{l_t}[.]$, $U_{q_t}[.]$ are the marginal utilities with respect to within period consumption, leisure and childcare quality. On the optimal path, conditions [3.5] to [3.9] unambiguously suggest that the marginal rate of substitution between consumption and leisure remains the same in each time period and is equal to the net benefit of maternal childcare. The participation decision of our typical mother is given by the following condition:

$$\frac{U_{l_t}}{U_{c_t}} = w_t = \frac{U_{q_t}}{U_{c_t}}(q_m - q_n) + p_t \qquad [3.10]$$

Note that this optimal path equilibrium arises from the additive separability assumption applied to the utility function and may not be a general result.[80] The optimal path equation [3.10] predicts that the marginal rate of substitution between leisure and consumption will equal the net benefit of maternal childcare and this in turn depends on the contemporaneous wage rate. This is similar to the standard result from static labor supply models where the marginal rate of substitution between leisure and work is equal to the shadow price of leisure. The reservation wage at time t depends on the quality of maternal care relative to non-maternal care and the price per hour of non-maternal care. A rational mother will opt for a paid childcare centre as long as the net benefit of maternal childcare is not as large as the market wage (Connelly 1992, Ribar 1995). A higher expected wage in the market place is expected to increase the propensity of labor force participation. An increase in childcare cost reduces the net wage and is expected to have a negative effect on propensity of participation in a labor market. It follows that factors which increase childcare costs will have a tendency to deter maternal employment and policy initiatives which reduce childcare costs –such as childcare subsidies and childcare tax credits, will have the effect of increasing maternal employment.[81] Leisure demand (labor supply) function derived from the above first order conditions can be written in generic form as follows:

[80] The additive separability assumption is convenient and is very common in optimal control models.
[81] However there can also be disincentive effect of childcare benefits if benefits are a function of income instead of employment. In this paper we do not take into account this possibility.

$$h_t = h_t(w_t, (1+\rho)^t (1+r)^{-t} \kappa_t, (q_n - q_m), p_s, x_t) \qquad [3.11][82]$$

The labor supply expression [3.11] shows that regardless of the labor force status of the mother, labor supply depends on the functional form of the utility functions, κ_0, taste factors, the interest rate, the discount rate ρ and the contemporaneous wage rate. Most importantly the labor supply function given by [3.11] is defined for a given marginal utility of wealth (κ_t). In the literature this demand (for leisure) function is known as a κ constant demand function (Browning and Meghir 1991, Heckman and MaCurdy 1980). This has an important implication for the selection of an appropriate econometric estimation to be adopted. The presence of this unobserved individual specific effect unambiguously suggests that the econometric methodology should control for individual heterogeneity. This is the rationale behind why we prefer sample selection panel data models which control for individual heterogeneity to the models which ignore these unobserved effects.

3.4. Econometric Model Specification

Let us rewrite the generic structural labor supply function [3.11] as follows:

$$h_{it} = h_t(w_t, p_t, x_{it}, \rho, r, c_i, u_{it}) \qquad [3.12]$$

[82] The consumption demand and quality of childcare demand expressions can also be derived likewise.

Where x_{it} contains observable individual, household and regional characteristics which affect within time labor supply. Note that it is useful to mention also that x_{it} contains factors which can affect the productivity of maternal and non-maternal childcare. The discounting factor and interest rate are assumed to be constant throughout the life cycle. c_i contains individual unobserved but time constant (fixed) effects and u_{it} contains unobservable time variant (random) determinants of labor supply. We estimate a standard continuous labor supply decision where the dependent variable is annual number of hours worked and a dichotomous labor supply decision where the dependent variable is a binary labor force participation variable. We specify the structural equation as follows:[83]

$$h_{it} = \begin{cases} \beta x_{it} + \eta p_t + c_i + u_{it} & \text{if working} \\ 0 & \text{if not working} \end{cases} \qquad [3.13]$$

Also note that in expression [3.13] age directly enters as an argument in the labor supply model as part of x_{it}. We include wage and price of childcare separately because we believe that the response of labor to a dollar spent on childcare is different from a dollar change in wages.[84] Equation [3.13] assumes exogenously given wages which may not be always the case. Wage is determined by observable and unobservable individual characteristics such as education, experience, ability and the like. Moreover, wages are not observed for individuals who do not work or who report zero annual hours worked. It

[83] Note that in our first essay we derived a similar labor supply model by employing a specific form of utility function. This formulation is very common in the literature because of its tractability.
[84] Using some additively separable utility functions we can show that age will enter directly as an argument if and only if $\rho \neq r$. See Jackubson (1988) MaCurdy (1981) or Hum, Simpson and Fissuh (2007) for details.

is not appropriate to assume wages to be zero if annual number of hours worked is zero. To deal with this data problem and endogeneity of wages, we propose a wage offer equation with the following structure:

$$\ln w_{it}^{*} = \lambda_1 y_{it} + \varpi_i + \varepsilon_{it} \qquad i = 1,...,n \qquad t = 1,...,T \qquad [3.14]$$

$$s_{it}^{*} = \lambda_2 z_{it} + \varsigma_i + v_{it} \qquad v_{it} / z_i \sim N(0,1) \qquad [3.15]$$

$$s_{it} = 1 \quad \text{if} \quad s^{*}_{it} > 0 \qquad [3.16]$$

$$\ln w_{it} = (\ln w_{it}^{*})(s_{it}) \qquad [3.17]$$

where $\ln w^{*}_{it}$ is a latent wage with an observable counterpart $\ln w_{it}$ and s^{*}_{it} is the latent participation decision with an observable counterpart s_{it}. Equation [3.14] is the wage offer equation which is the equation of interest and equation [3.15] is a reduced form model for the propensity to participate in the labor market. y_{it} and z_{it} contain vectors of exogenous individual characteristics, such as years of experience, years of schooling, marital status, number of kids in a family, other family income, and immigration status. It is conceivable that most of the variables that enter the wage equation will also determine participation in the labor market. In our empirical models we will impose some standard exclusion restriction[85]. λ_1 and λ_2 are vectors of unknown parameters and ε_{it} and v_{it} are random error terms with $E(\varepsilon_{it} / v_{it}) \neq 0$. We assume that (ε, v) is independent of z_i

[85] It is conceivable that most of the variables which influence participation in the labor market will also affect wages hence we can expect that z contains most of elements of y. Ideally, we would like to have some exclusion restriction for efficiency reason. We do not require exclusion restriction for identification purpose. The equation will be identified at least by the inverse Mills ratios.

(where z_i might contain elements of y_i), and that ϖ_i and ς_i are individual fixed effects which are time invariant. We use this sample selection model to impute wage offers for all mothers and hope that this circumvents the problem of sample selection and endogeneity of wage in our labor supply equation.

The second auxiliary regression to be estimated is the price of childcare equation. Akin to the missing data problem for wages we only observe the childcare price if a mother decides to work and send her children to paid childcare centers. This is a double selection problem and to account for this selection problem, as in Kimmel (1998) and Powell (1997), we include two selection terms in our childcare price model. We compute an inverse Mill's ratio from a probit model for participation in the labor market to use as a correction term in the price of childcare equation. We also impute a probit model of paid childcare and impute an inverse Mill's ratio. However, unlike Kimmel (1998), Powell (1997) and most other researchers, our econometric model controls for individual heterogeneity by estimating a sample selection model of childcare price.[86] Moreover, we propose that the parameters of the price equation for married and single mothers differ and hence estimate two separate childcare price models.[87] In our empirical estimation we test for this presupposition using a likelihood ratio test. The estimated childcare price equations are then used to impute childcare cost per hour for all mothers. The childcare price predictor variables include the number of parents (grandparents of the children)

[86] We experimented if the estimation of a simple linear model or a fixed effects model matters to the imputation of childcare price and thereby the implied elasticity of labour supply with respect to childcare cost. Our results reveal that the results of the price elasticities are very similar and do not seem to make a material difference. Because of its theoretical appeal we use the fixed effects model to compute predicted childcare price.

[87] We also test this presupposition using simple interaction variable addition test akin to Chow test. See footnote 94 for an explanation of the test.

living in the family, the number of adults in the family and the number of unemployed people in the family. These variables are expected to capture the availability of alternative caregivers or the access to low cost childcare. We also include family income other than the wage of the mother and the amount of childcare benefit to capture variations in quality of childcare that children receive.[88] Number of children in a family: including preschool children, children ages 6 to 16 and 6 dummy variables for the age of the youngest preschool child in a family are included to allow for scale effects in childcare costs of school children. We expect a negative relationship between the age of the youngest child in a family and total hourly childcare cost. Also included are immigration status of the mother and provincial childcare regulation variables such as teacher children ratio in childcare centers and average wage for childcare workers which will further help to identify the childcare equation.

By derivation, equation [3.10] represents the life cycle labor supply decision of every mother of working age whether or not she was working at the time of the survey. However, the annual number of hours worked is bottom coded (censored) at zero (because of participation decision). This selection problem in labor supply models has been long identified in the literature (Wales and Woodland 1980, Gronau 1974, Lewis 1974, Heckman 1978, Vella 1998). To deal with this famous problem of nonrandom sampling we propose that the labor supply equation [3.13] follows the following sample selection panel data structure [89]:

---

[88] Child benefit and family income may be correlated as child benefit is a function of family income. For this reason we experimented with one variable at a time but it did not make any material difference.
[89] See Baltagi (2005), Hsiao (2003), Greene (2003), or Wooldridge (2002) for a textbook discussion on panel data modeling. See Vella (1998) for a readable survey of sample selection models.

$$h_{it}^* = \beta x_{it} + c_i + u_{it} \qquad i = 1,...,n \qquad t = 1,...,T^{90} \qquad [3.18]$$

$$s^*_{it} = \lambda z_{it} + \varsigma_i + v_{it} \qquad v_{it} \mid z_i \sim N(0,1) \qquad [3.19]$$

$$s_{it} = 1 \quad \text{if} \quad s^*_{it} > 0 \qquad [3.20]$$

$$\ln h_{it} = (\ln h_{it}^*)(s_{it}) \qquad [3.21]$$

where h^*_{it} is a latent endogenous number of hours with an observable counterpart h_{it} and s^*_{it} is a latent participation decision with an observable counterpart s_{it}. Note that, unlike expression [3.13] β now includes π — x_{it} now includes p_t. Equation [3.18] is a general form—type II Tobit or sample selection model—the labor supply function from [3.13] which is the equation of interest and equation [3.19] is a reduced form model of propensity to participate in the labor market—more correctly the probability of employment. x_i and z_i contain vectors of exogenous individual characteristics such as, age, years of schooling, health, imputed wages, the price of non-maternal childcare and others. It is conceivable that most of the variables that enter the hours worked equation will also determine participation in the labor market. For efficiency reasons in our empirical models we impose some fairly standard exclusion restrictions. β and λ are vectors of unknown parameters and u_{it} and v_{it} are random error terms with $E[u_{it} \mid v_{it}] \neq 0$. We assume that (u,v) is independent of z_i (where z_i might contain elements of x_i), and c_i and ς_i are individual fixed effects which are time invariant.

[90] Note also that $t = t_i,...,T_i$ implies that the panel structure could be unbalanced.

Testing and correcting for sample selection

There are a number of suggestions in the literature concerning the detection and correction of sample selection bias in panel data models (Wooldridge 1995, Verbeek and Nijiman 1996, Vella 1998, Vella and Verbeek 1994). In this paper, we follow Wooldridge (1995). The basic premise of this approach is that it parameterises the conditional expectation required for the consistency of the pooled estimator. Under the null hypothesis of $E[u_{it} \mid x_{it}, s_{it}, c] = 0,\ \ t = 0,1,...,T$ the inverse Mill's ratios from equation [3.19] (for each cross-section) should not be significant in an equation estimated by fixed effects (least squares dummy) method. In effect, this test involves two steps. In the first step we estimate the inverse Mill's ratio (λ_t)[91] from [3.19] for each cross-section. The next step is to estimate equation [3.18] using the fixed effects model on the selected sample by including the Inverse Mill's ratios as additional regressors and then testing for sample selection bias using t tests for the Inverse Mill's ratios in this fixed effects model. Wooldridge (1995) shows that the limiting distribution of t under the assumption $E[u_{it} \mid x_i, s_{it}, c_i] = 0$ is not affected whether we estimate a pooled model, random effects model or a fixed effects model of participation equation [3.19]. As long as the standard errors are robust and adjusted for hetroskedasticity, we can rely on the student t test. If λ_t are found to be statistically significant we should make the necessary correction as follows:

$$E[y_{it} \mid x_i, v_{it}] = \beta x_{it} + \pi x_i + \xi_t \lambda(z_i \psi) \qquad [3.22]$$

91 $\lambda \equiv \frac{\phi(.)}{\Phi(.)}$ where is $\phi(.)$ is the density function and $\Phi(.)$ is the cumulative density function.

It is possible to obtain a consistent estimate of μ by first estimating a labor force participation equation using probit regression of s_{it} on x_i for each panel j and saving the Inverse Mill's ratios ($\hat{\lambda}_{it}$). The next crucial step is to run pooled OLS regression using the selected sample: y_{it}, on $x_{it}, x_i, \hat{\lambda}_{it}, d_t\hat{\lambda}_{it}, ..., d_T\hat{\lambda}_{it}$ for all $s_{it} = 1$ where $d_t - d_T$ are the time dummies. We can get consistent estimates of equation [3.22] using Ordinary Least Squares method (Wooldridge 1995).[92] Note this approach allows for the correlation between the unobservable components in the selection equation, v_{it} and the unobservable components in the wage offer equation (c_i, u_{it}) since the selection process might operate via both the error term from the main equation u_{it} and the unobservable individual effect c_i. However, we need to adjust the standard errors for general hetroskedasticity and autocorrelation and for the first stage estimation.[93] Note that we also estimate a dichotomous labor supply decision with a binary dependent variable with the inclusion of inverse Mill's ratios from the participation and childcare utilization equations. We estimate this model using three specifications: pooled, random effects and fixed effects models.

3.5. Data

The data employed in this study is drawn from SLID. Our sample contains mothers between the ages of 15 and 50 with at least one preschool child. We focus on 6 years: 1999, 2000, 2001, 2002, 2003 and 2004. Our sample consists of an unbalanced panel of

[92] See Wooldridge (1995, 2002) for the detailed derivation of the model.
[93] We estimate our selection using Baltagi and Cheng's (1996) approach.

2396 individuals and 7819 total observations of both those who reported positive and zero hours. The minimum number of times each individual is observed is two time periods. About 83% of the sample are married and the rest are single mothers: never married, widowed, separated and others. The SLID is a continuing panel of Canadian households which began in 1993. It combines the former Labor Force Activity Survey, an intermittent series of panel surveys conducted during the 1980s, with the Survey of Consumer Finance, a regular cross-sectional survey conducted annually. The SLID design is a series of overlapping 6-year panels, with a new panel enrolled every three years.

The dependent variable in our continuous labor supply decision model is the annual number of hours worked. In SLID the annual number of hours and the composite hourly wage are calculated from an extensive annual interview with detailed questions on each job and payment that individuals get in the survey period. Questions are posed on the number of jobs held and the hours worked pay by pay, the number of weeks worked the number of weeks absent from work and others. Respondents are taken through a detailed questionnaire to retrieve the relevant information or provide access to appropriate income files to produce reliable information on the number of hours worked and the hourly wage. The dependent variable in our dichotomous labor supply model is a dummy variable for labor market participation which assumes a value of one if a mother works positive hours and zero otherwise. The other variables included in our labor supply model include: age, number of children age 0 to 5, number of children age 6 to 16, number of parents living in a family immigration status, marital status, total child benefit received per family,

family income other than the total wages received by the mother, imputed childcare price and imputed hourly wages.

----------------------Figure 1 here--------------------------

To examine the relationship between annual hours worked and childcare expenditure we first produce a scatter plot of these variables. Figure 1 shows a strong positive relationship between hours worked and childcare expenditure per child. While on average a mother from the lower hours quintile spends 57 dollars per child per year, a mother from the upper hours quintile spends more than 2000 dollars per child per year. This may be explained by a number of socio economic factors and the task of the multivariate analysis is to examine the impact of childcare costs on annual number of hours worked after controlling for these socio economic factors. Table 1 reports the descriptive statistics of the variables employed in our study by marital status.

----------------------------Table 1 here-------------------

3.6. Estimation Results

This section presents the estimation results by marital status.[94] Before presenting the sample selection model results, we report the results of the pooled OLS and the

[94] As it was mentioned above the estimation is conducted by marital status. One crude way of testing the independence of the models for single and married mothers is to conduct a Chow test using dummy variables technique. We conducted a Cow test using an F test and the test decisively rejected the null hypothesis of no interdependence. To highlight the difference in the sensitivity of labor supply decision for childcare changes we report a model which includes an interaction variable of marital status and childcare price as a regressor in the sample which includes all mothers. In the models estimated the interaction variable was significant which implies that the effect of childcare price is not the same across single and married mothers. Table 8 reports the regression output which contains the interaction variable between

traditional panel data estimators: FE and RE models. Table 2 reports the results from a pooled OLS both for single and married mothers in the first three columns. The pooled OLS model corresponds to standard cross-sectional results in the literature that assume that the regressors in the model are not correlated with individual heterogeneity. The estimates seem to produce the common results reported in the literature. We discuss only the relevant variables for our study. The coefficients of imputed childcare price and wage conform to expectations. In all models the coefficient of childcare price is negative, as expected, and statistically significant at the 5 percent level of significance. The coefficient of the childcare price variable for single mothers is in absolute terms larger than that of married mothers suggesting that single mothers are relatively more sensitive to childcare price than married mothers. The coefficient of the imputed wage variable is positive, as expected, and statistically significant at 5 percent. The implied wage elasticity of married mothers is larger than that of single mothers which suggests that married mothers are more responsive to changes in hourly wage than single mothers.[95] The implied childcare price elasticity of labor supply (annual number of hours worked) for married mothers and single mothers range from -0.012 to -0.113 and -0.08 to -0.166. The implied intertemporal labor elasticity with respect to wage is calculated to be in the range of 0.7 to 0.9 and from 0.7 to 0.8 for married and single mothers . [96] The coefficient of the variable for number of preschool children is negative and significant at 5 percent both in the model for married mothers and single mothers.

childcare price and marital status. The size and sign of the interaction variable is in line with the estimates from the other model.

[95] We test for the possible difference by estimating pooled model for married and single mothers using an interaction term of marital status and childcare price, and the interaction term was found to be statistically significant. Table 8 reports the result of this test.

[96] Note that this elasticity measure keeps the marginal utility of wealth constant.

-------Table 2 here-----------------

The general result of a negative impact of childcare price on labor supply is in line with our expectation. However, since in pooled OLS model it is assumed that the individual effects are not correlated with the explanatory variables we may suspect that OLS estimator could be biased and inconsistent. In fact, an LM test has decisively rejected the null hypothesis of non correlation of the error terms with the explanatory variables at less than 1 percent level of significance for all the models reported. Hausman tests were also conducted to compare the fixed effects and random effects models and in all cases the test decisively rejected the null hypotheses of non correlation of the error terms with the explanatory variables (consistency of RE model) at less than 1 percent level of significance. Due to the fact that in all cases the results rejected the RE specification in favour of Fixed Effects, we discuss the FE estimates only.[97]

The results from the fixed effects model are reported in the final three columns of Table 2. As was mentioned above, regardless of the correlation between the error terms and regressors in our model the fixed effects estimator produces unbiased and consistent estimates provided that there is no time varying variable which is driving both the participation equation and the reduced number of hours equation. Looking at Table 2 we find that the coefficient for the childcare price for married mothers' changes marginally from -7.25 to -6.75 and remains statistically significant at 1 percent. However, the coefficient of the childcare price for single mothers' model changes dramatically from -

[97] The RE results are provided in Table 2 for the interested reader but are not discussed in the paper.

10.77 to -136.2. This may suggest that the assumption of no correlation between the individual effect and the explanatory variables could be potentially damaging for the sub sample of single mothers.[98] Note, however, that just as for the pooled OLS results the selected sample includes only those who reported positive hours and our results in Table 2 do not control for the sample selection problem.

We next present the sample selection corrected models. However, before making any sample selection correction it is useful and appropriate to conduct a diagnostic test for sample selection. We conduct two types of test. First, we estimate cross sectional wage equations with standard Heckit type models and assess the statistical significance of the correction term (λ_t). In the wage offer equations λ_t is imputed after imposing some fairly standard exclusion restrictions. Other family income, child benefits, and the presence of young adults in a family were included in the selection equation but not the wage equation. Education, experience, experience squared, marital status and immigration status were included in both equations. These restrictions should improve the reliability of our test. The last column of Table 3 reports the λ_t from cross sectional wage offer equations by year. All the inverse Mill's ratios (λ_t), as expected, are negative and statistically significant at 1 percent indicating that the null hypothesis of no selection for each year can be decisively rejected. The second test involves estimating a fixed effects model, as suggested by Wooldridge (2002), using the selected sample only but including the inverse Mill's ratios as additional regressors. According to Table 3, most of the λ_t turn out to be statistically significant at 5 percent in all models. Moreover, in all

[98] This may also suggest that the sub sample of single mothers is highly heterogeneous.

cases we decisively reject the null hypotheses that all λ_t are jointly zero ($H_0 : \lambda_t = 0$, $\forall t$). These tests confirm that sample selection bias can be potentially damaging and should not be ignored.[99]

--------------------------------Table 3 here-------------

In view of a sample selection problem the fixed effects model which includes λ_t seems a natural candidate. However, Wooldridge (2002) shows that this procedure can render OLS to be an inconsistent estimator. The bias from the presence of a common time varying variable which affects the structural equation and the selection equation could be exacerbated with the employment of FE model (Dustman and Rochinna-Barrachina 2007). Table 4 presents the results from the fixed effects model estimated with the inclusion of sample selection terms. Table 4 shows that the coefficient of the imputed childcare price is negative and significant at 5 percent in all models estimated. The size of the childcare price from the models in Table 4 and that of the linear fixed effects model in Table 2 are very similar. However the coefficient of childcare price for single mothers has increased from -136 to -105 from the linear fixed effects model in Table 2. If we look at the coefficient of the imputed wage it does not seem to change a lot and the implied elastcicities are very similar. This may indicate that the fixed effect model with the inverse Mill's ratios included is picking up some selection effects which are driving the differences. However, as is discussed in the model development section, this model is not without its problems. For this reason we estimate the Wooldridge (1995) model because

[99] Given our sample is unbalanced we tested for attrition. Following Vella and Verbeek (1198) we conduct a variable addition test for attrition by including s_{t-1} as a regressor in equation [3.22] with and without πx_i and the null hypothesis of attrition bias was decisively rejected after controlling for sample selection bias.

it is expected to control for individual heterogeneity and the selection problem in a unified framework.

-------------------------------Table 4 here-----------------------------------

Table 5 reports results from estimating [3.22], which is expected to control for sample selection and individual heterogeneity in unison. According to the results in Table 5 the coefficient of childcare price has slightly increased in absolute value for married mothers. On the other hand the size of the price coefficient in the model for single mothers is only about one third of the estimate from fixed effects model with selection terms included and about one fourth of the fixed effects model with no selection terms included. This result may imply that accounting for the individual heterogeneity using the FE technique is not enough to control for sample selection bias.[100] This result may suggest that the selection process may be operating via time variant variables and hence the FE model may not suffice. The results in this paper reinforce the findings we reported on the impact of health on wages in a previous paper. This is potentially a very important result as it may imply proper accounting for sample selection bias is essential and studies which do not take into account individual heterogeneity and sample selection could be misleading.

---

[100] We also tried to estimate the fixed effects sample selection model. However, there were computational problems with this specification. The sample selection model is not globally concave and hence the iteration procedure in the second stage of the estimation broke down in all the specifications that were attempted as the Hessian matrix was singular and indefinite. This is not uncommon in panel data sample selection estimations. Also note that the fixed effects model is subject to the infamous incidental parameter problem. Given that the estimation procedure involves two stages, the estimates from this model are also inefficient. For this reason, we employed 30 bootstrap replications to improve on the results. However, the estimation was terminated during the initial stage of the bootstrapping hence the models could not be estimated. For technical details of estimating a sample selection model with fixed effects see LIMDEP (2003) manual.

---------------------------Table 5 here----------------------

To further examine the impact of childcare cost on labor supply we have also estimated binary choice variable models where the dependent variable is dichotomous labor force participation. The dependent variable assumes a value of 1 if a mother reports positive hours and 0 otherwise.[101] In other words we estimate equation [3.18] with a binary dependent variable. However, because we are interested in computing elasticity of employment with respect to childcare price, we need to calculate the marginal effects of the variables which are different from the estimated coefficients. The marginal effects are given by the product of coefficient estimates and value of a density function. We calculate the elasticity of participation with respect to the child care price at sample mean values. If a typical mother gives more weight to the quality component of childcare, then we would expect this mother to be relatively less responsive to changes in the price of child care (Kimmel 1998). Thus, we expect the labor supply elasticity of childcare services to be relatively smaller for married mothers than single mothers.

-------------------------------Table 6 here----------------------

Table 6 reports the results from three traditional variants of panel logit models by marital status. The pooled logit model treats the panel data as an extended cross-section and our results confirm previous cross-sectional results in the literature (Powell, 1997, Connelly 1992). The elasticity of employment with respect to child care in the random effects and fixed effects model are all negative and significant at 5 percent. The employment

[101] It is well established in the econometrics literature that this class of models are consistent with random utility maximization.

elasticities range from -0.016 to -0.48 for single mothers but range only from -0.012 to -0.10 for married mothers. On average, if the childcare price increases by 1 percent, the probability of participation in the labor market decreases by 0.5 percent for single mothers but only by 0.01 percent for married mothers.[102] Note that the fixed effects model gives a relatively higher, in absolute terms, coefficient estimate than the random effects model. However the fixed effects model should be viewed with caution because it may suffer from the incidental parameter problem associated with limited panel length. With fixed T the estimates for the constants in the system cannot be consistent and this problem carries over to other coefficients in the model. But with T above 5 Greene (2003) shows using a Monte Carlo experiment that the bias is not as bad as what Heckman (1981) reports. Comparison of our estimates with the literature indicates that our estimates for married mothers are very conservative in comparison with previously reported in the literature but for single mothers the estimates of this study are within the range that has been reported for married mothers.

----------------------Table 7----------------------------

[102] Akin to the test described in foot note 94 for the reduced hours equation we conducted a test using a pooled regression model with an interaction term and the interaction term was statistically significant.

3.7. Conclusion

This study uses panel data from the SLID second panel from 1999 to 2004 to examine the impact of childcare cost on the maternal labor supply decision. The study aimed to demonstrate, among other things, that the combined effects of unmeasured attributes implied by a life cycle labor supply model and sample selection bias in labor supply estimation can be accounted for using appropriate nonlinear panel data models. This was achieved through the use of the Wooldridge's (1995) suggested estimation method. The general finding is that the elasticity of annual number of worked with respect to childcare cost is found to be very weak. The implied elasticity of childcare price with respect of annual number of hours worked is in the range of -0.01 and -0.012 for married mothers and in the range of –0.016 to -0.16. As far as the employment elasticity with respect to childcare price is concerned our estimates range from -0.012 to -0.10 for married mothers and from -0.016 to -0.48 for single mothers. Table 7 reports that the range of cross sectional estimates in the literature for using Canadian data is between -0.2 to -0.38. We estimated a cross sectional labor supply model using the data for the year 1999 and employment elasticity for married and single mothers was found to be -0.18 and -1.84 respectively. The implied smaller reaction of labour supply decision of mothers to childcare costs suggest that the cost of creating an incentive for mothers to revise their labour supply decision via childcare policies which lower childcare prices does not seem to be promising.

Many of the usual caveats still apply and our results should be interpreted with caution. The most obvious limitation of this research is that it does not distinguish among the different modes of paid childcare and this may confound the differences in consumer preferences towards these different modes of paid childcare and quality. It is also evident that such a crude grouping of the childcare modes might not uncover the differences in marginal and fixed costs for childcare service and future research should account for the choices towards the different modes of childcare. This study implicitly assumes that mothers make their labor supply decisions freely and independently of the choice of childcare mode. This may be a problem given the fact that individual labor supply decisions are intertwined with the choice of childcare mode. Another avenue that could be explored with future research is explicit modelling of family labor supply decision of married mothers in examining the impact of childcare cost on the joint labor supply decisions of mothers and fathers. Such studies would also benefit from the inclusion of the quality aspect of childcare at household level as opposed to provincial level. We also assume in this study that mothers face no liquidity constraint and there is no human capital accumulation which could affect the amount of hours worked and earned wages. It would be useful to experiment with different econometric estimators, such as the ones suggested by Kyrizidou(1997) and Honore and Kyrizidou(2000), that do not require any distributional assumptions but attempt to control for sample selection and individual heterogeneity. Future research might also take into account the possibility of a nonlinear budget constraint and its implications for econometric estimation. Lastly our study does not take into account the errors in variables such as childcare price and wages. It is left for further research to develop explicit models with errors in variables and see if our results are robust to this line of inquiry.

References

Baker, M., *et al.* (2005) "Universal Childcare, Maternal Labor Supply and Family Well-Being," Working Paper No.11832 (December), NBER, Cambridge, MA.

Baker, M., *et al.* (2008) "Universal Childcare, Maternal Labor Supply and Family Well-Being," *Journal of Political Economy,* 116(4), pp. 709-745.

Baltagi, B.H. and Chang, Y.J. (2000) "Simultaneous Equations with Incomplete Panels," *Econometric Theory,* **16**, pp 269-279.

Blau, D. M., and Hagy, P.A. (1998) "The Demand for Quality Childcare," *Journal of Political Economy*, 106(1), pp. 104-146.

Cleveland, G., Gundrson, M. and Hyatt, D. (1996) "Childcare Costs and the Employment Decisions of Women: Canadian Evidence," *Canadian Journal of Economics*, 29(1), pp. 132-148.

Connelly, R. and Kimmel, J. (2003) "The Effect of Childcare Costs on the Employment and Welfare Recipience of Single Mothers," *Southern Economic Journal*, 69(3), pp. 498-519.

Heckman, J. J. (1974) "Effects of Childcare Programmes on Women's Work Effort," *Journal of Political Economy*, 82(2), pp. 136-163.

Heckman, J.J. (1976) "A Life Cycle Model of Earnings, Learning and Consumption," *Journal of Political Economy*, 84(4), pp. S9 -S44.

Honore, B. and Kyrizidou, E. (2000) "Panel Data Discrete Choice Models with Lagged Dependent Variables," *Econometrica*, 68, pp. 839-874.

Hum, D. Simpson, W. and Fissuh, E. (2006) "The Impact of Health on Labor Supply in Panel Data," *Global Business and Economics Anthology*, selected papers from the 2006 Business and Economics Society International Conference, Florence Italy, July 15-19, 2006.

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(3), pp. 302-329.

Kimmel, J. and Powell, M.L. (2001) " Nonstandand Work and Childcare Choices of Married mothers," *Upjohn Institute Staff Working Paper N.01-74*, Upjohn Institute, USA.

Kimmel, J. (1998) "Childcare as a Barrier to Employment for Single and Married Mothers," Review *of Economics and Statistics*, 80(2), pp. 287-299.

Kumar, A. (2005) "Life cycle Consistent Estimation of Effect of Taxes on Female Labor Supply in the U.S.: Evidence from Panel Data," Research Department Working Paper, Unpublished Manuscript, Federal Reserve Bank of Dallas, USA.

Kyriazidou , E. (1997) "Estimation of a Panel Data Sample Selection Model," *Econometrica*, 65(6), pp.1335-1364.

Lefebvre, P. and Merrigan, P. (2005) " Low –fee ($5/dayt/child) Regulated Childcare Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada," Working Paper 05-08, Centre Interuniversitaire Sur Le risqué, les politiques economiques e l'emploi, Montreal, Canada.

MaCurdy, T. E. (1981) "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy*, 89, pp. 1059-85.

MaCurdy, T., Green, D. and Paarsch, H. (1990) "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *The Journal of Human Resources*, 25(3), pp. 415-490.

Maddala, G.S. (1983) Limited Dependent and Qualitative variables in Econometrics, Econometric Society Monographs No.3, Cambridge: Cambridge University Press, UK.

Meyer, B. D. and Dan T. Rosenbaum, T.D. (2001) "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers," *Quarterly Journal of Economics*, pp. 1063-1114.

Michalopoulos, C. and Philip K. Robins, P.K. (2002) "Employment and Childcare Choices of Single Parent families in Canada and United States," *Journal of Population Economics*, 15, pp. 465-493.

Nijman, T. and Verbeek, M. (1992) "Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function," *Journal of Applied Econometrics*, 7, pp. 243-57.

Powell, L. M. (1995) "A Structural Model of Childcare and Labor Supply of Married Women," *Journal of Labor Economics*, 13(3), pp. 558-597.

Powell, L.M. (1997) "The Impact of Childcare Costs on the Labor Supply of Married Mothers: Evidence from Canada," *Canadian Journal of Economics*, 30(3), pp. 577-594.

Powell, L. M. (2002) "Joint Labor Supply and Childcare Choice Decisions of Married Mothers," *The Journal of Human Resources*, 37(1), pp. 106-128.

Ribar, C. D. (1995) "A Structural Model of Childcare and Labor Supply of Married Women," *Journal of Labor Economics*, 13(3), pp.558-597.

Roy, F. (2006) "From She to She: Changing Patterns of Women in the Canadian Labor Force," *Canadian Economic Observer*, Statistics Canada, 3, pp.1-3.10.

Vella, F. and Verbeek , M. (1999)"Two-step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias ," *Journal of Econometrics,* 90, pp. 239-263.

Vella, F. (1998) "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, XXXIII (1), pp. 127-169.

Verbeek, M. (1990) "On the Estimation of a Fixed Effects Model with Selectivity Bias," *Economics Letters*, 34, pp. 267-70.

Wales, T.J., Woodland, A.D. (1980) "Sample Selectivity and the Estimation of Labor Supply Functions," *International Economic Review*, 1980, 21(2), pp. 437-468.

White, L. A. (2001) "Childcare, Women's Labor Market Participation and Labor Market Policy Effectiveness in Canada," *Canadian Public Policy*, 27(4), pp. 385-405.

Wooldridge, J.M. (2002) Econometric Analysis of Cross-Section and Panel Data, 2nd ed., London, England: MIT Press.

Wooldridge, J. M. (1995) "Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions," *Journal of Econometrics*, 68, pp. 115-32.

Wronlich, K. (2004) "Childcare Costs and Mothers' Labor Supply: An Empirical Analysis from Germany," *Working Paper*, unpublished manuscript, German Institute for Economic Research (DIW Berlin).

Table 1. Descriptive Statistics of the key variables in this study

Variable		Single N	Single Mean	Single S. D	Married N	Married Mean	Married S. D
Hours	Annual Number of hours worked	1454	942	930	7722	1079	872
LFP	Lfp= 1 if a mother works positive number of hours, 0 otherwise.	1632	0.7	0.5	8254	0.79	0.41
Price	Imputed childcare cost per hour	1302	1.4	1.4	7283	1.91	5.13
Kids	Number of children ages 0 to 5 years	1632	1.2	0.5	8254	1.38	0.58
Bkids	Number of children ages 6 to 16	1632	0.6	0.8	8254	0.67	0.89
Parent2	Dummy variable if one parent or grandparent lives in the family.	1632	0.1	0.2	8254	0.02	0.12
Parent3	Dummy variable if two grand parents live in the family	1632	0.1	0.3	8254	0.01	0.11
Imigrant	Dummy variable for an immigrant	1469	0.1	0.3	7636	0.17	0.37
Other income	Other family non wage income	1632	12141	25717	8241	46751	40785
Married	Dummy variable for marital status	1632	0	0	8254	1	0
Childage1	Dummy variable for the youngest child in a family age 1 year	1623	0.1	0.3	8222	0.18	0.39
Childage2	Dummy variable for the youngest child in a family age 2 year	1623	0.1	0.4	8222	0.2	0.4
Childage3	Dummy variable for the youngest child in a family age 3 year	1623	0.2	0.4	8222	0.18	0.38
Childage4	Dummy variable for the youngest child in a family age 4 year	1623	0.2	0.4	8222	0.16	0.36
Childage5	Dummy variable for the youngest child in a family age 5 year	1623	0.2	0.4	8222	0.14	0.35
Impwage	Imputed hourly wage	1409	2.7	0.2	7480	2.9	0.21
Childben	Child benefit	1632	4064	2744	8252	1949	2443
Lambda2	Imputed inverse Mill's ratio from paid childcare utilization	1032	1.2	0.3	5593	0.89	0.23
Lambda1	Imputed inverse Mill's ratio from participation in the labor market	1442	0.6	0.4	7550	0.43	0.3

Note: The minimum and maximum of the variables could not be reported because of the privacy issues at statistics Canada.

Table 2. Women Labor supply model: dependent variable number of hours worked

	Pooled			RE			FE		
	All	Married	Single	All	Married	Single	All	Married	Single
Imputed Childcare Price	-7.578***	-7.256***	-10.77	-6.672***	-6.805***	-65.18**	-6.432***	-6.752***	-136.2***
	(1.886)	(1.884)	(32.66)	(1.487)	(1.507)	(32.4)	(1.558)	(1.614)	(47.72)
Age	139.2***	109.9***	148.8***	158.7***	125.7***	184.2***	214.1***	165.3***	264.0***
	(14.64)	(17.84)	(30.7)	(17.8)	(21.48)	(36.58)	(29.79)	(35.19)	(76.79)
Age Squared	(2.128***	-1.752***	-1.996***	-2.267***	-1.824***	-2.567***	-2.478***	-1.822***	-2.903**
	(0.223)	(0.268)	(0.487)	(0.271)	(0.322)	(0.584)	(0.444)	(0.516)	(1.275)
Number of Kids age 0-5	-154.0***	-161.1***	-76.71	-189.0***	-187.0***	-117.2*	-231.1***	-244.7***	-2.975
	(20.35)	(21.58)	(61.24)	(19.73)	(21.11)	(60.56)	(30.09)	(33.42)	(103.3)
Number of kids age 6-13	28.05*	33.38**	46.88	-45.57***	-33.73*	-30.51	-173.4***	-177.6***	-32.26
	(14.94)	(15.82)	(46.68)	(17.13)	(18.7)	(49.96)	(29.06)	(33.74)	(97.71)
Parent2	166.4**	260.6***	-52.64	74.93	167.6*	-23.71	-1.621	56.26	2.666
	(64.85)	(85.11)	(108.4)	(71.89)	(96.28)	(113.7)	(96.74)	(131.8)	(165)
Parent3	142.8**	335.2***	-126.2	-52.58	45.83	-98.02	-258.1***	-254.8*	-30.86
	(61.71)	(111.6)	(105)	(65.88)	(113.3)	(107.1)	(91.27)	(149.3)	(160.5)
Immigrant	-190.0***	-182.9***	-289.5***	-196.3***	-181.7***	-313.0**	0	0	0
	(30.99)	(32.45)	(102.9)	(46.91)	(49.67)	(136.3)	0	0	0
Other Income	-0.004***	-0.004***	0.002	-0.001***	-0.001***	-0.00002	0.0004	0.001*	-0.004*
	(0.0003)	(0.0004)	(0.001)	(0.0004)	(0.0004)	(0.001)	(0.0005)	(0.0005)	(0.002)
Child Benefit	-0.117***	-0.120***	-0.119***	-0.081***	-0.082***	-0.091***	-0.069***	-0.064***	-0.086***
	(0.00622)	(0.00682)	(0.0158)	(0.006)	(0.007)	(0.015)	(0.008)	(0.009)	(0.022)
Imputed Wage	988.3***	1025***	808.9***	915.6***	892.4***	1039***	749.7***	802.5***	707.4
	(55.78)	(59.51)	(166.7)	(76.42)	(82.29)	(198.4)	(191.8)	(210.3)	(527.8)
Married	-112.5***			-180.0***			-219.3***		
	(29.94)			(32.69)			(46.56)		
Constant	-3195***	-2850***	-3210***	-3476***	-3015***	-4283***	-4484***	-4016***	-5444***
	(243.50)	(301.50)	(540.40)	(303.30)	(371.90)	(637.10)	(606.90)	(725.10)	(1502.00)
Child care price elasticity	-0.012*	-0.012*	-0.013	-0.011*	-0.011*	-0.08**	-0.104*	-0.113*	-0.166*

β	0.878*	0.887*	0.848*	0.815*	0.773*	1.10*	0.666*	0.695*	0.742
	(0.05)	(0.052)	(0.176)	(0.068)	(0.072)	(0.212)	(0.171)	(0.182)	(0.554)
Observations	7819	6665	1154	7819	6665	1154	7819	6665	1154
R-squared	0.158	0.152	0.187				0.062	0.057	0.097
Number of persons				2396	2054	516	2396	2054	516

Data Source: (SLID 2004). The price equations were estimated using pooled OLS with the selection terms included.
Note: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3. Inverse Mill's ratios from the fixed effects sample selection model

	Fixed Effects			Cross Section
	ALL	Married	Single	All sample
λ_{1999}	-2635***	-1408	-2734***	-0.381***
	(502.6)	(869.9)	(1021	(0.076)
λ_{2000}	5356***	6200***	3898	-0.254***
	(970)	(1181)	(2538)	(0.091)
λ_{2001}	-1296*	-2942***	470.6	-0.434***
	(682.3)	(811.5)	(2072)	(0.091)
λ_{2002}	-1959***	-2029***	-2852	-0.416***
	(621.8)	(724.6)	(2024)	(0.089)
λ_{2003}	64.86	-29.63	-525.1	-0.299***
	(672.8)	(737.5)	(2191)	(0.079)
λ_{2004}	2281***	2239***	3708***	-0.281***
	(387.2)	(418.0)	(1290)	(0.079)
F Test for $H_0 : \lambda_i = 0$	19.51(p=0)	13.25(p=0)	4.79(p=0.00)	

Note: Full results of the Fixed effect models are part of Table 3 in Appendix A.

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4. Labor supply models-selection corrected-fixed effects model[a]

	ALL	Married	Single
Price	-6.379***	-6.19	-105.4**
	(1.542)	(11.90)	(48.11)
Age	429.1***	313.1***	592.9***
	(41.05)	(49.92)	(133.5)
Age2	-5.925***	-5.124***	-8.407***
	(0.634)	(0.736)	(2.235)
Kids 0-5	-457.2***	-213.5***	-477.2***
	(38.66)	(56.60)	(146.6)
Kids age 6-16	-170.3***	16.69	-46.6
	(29.69)	(54.01)	(101.3)
Parent2	10.87	82.39	36.13
	(95.83)	(130.7)	(162.5)
Parent3	-209.7**	-204.0	-5.293
	(90.9)	(147.4)	(158)
Other income	-0.004***	-0.00507***	-0.008***
	(0.001)	(0.000920)	(0.003)
Child benefit	-0.227***	-0.264***	-0.262***
	(0.028)	(0.0353)	(0.068)
Imputed wage	784.8***	852.4***	977.6*
	(193.6)	(211.5)	(568.7)
Marital1	-448.4***		
	(86.76)		
	(387.2)		(1290)
Constant	-7283***	-5247***	-10079***
	(781.9	(991.5)	(2264)
Child care price elasticity	-0.10	-0.011	-0.14
β	1.04	0.79	1.310
Observations	7819	6616	1154
Groups	2396	2040	516
R-squared	0.124	0.088	0.137

Data Source(SLID 004)

*** P<0.01, ** p<0.05, * p<0.1 Standard errors in parentheses

[a] Fixed effects model which includes the inverse Mills's ratios which are reported in Table 4 below.

Table 5. WE [Wooldridge (1995)] model estimates of maternal labor supply

	All	Married	Single
Price	-9.292***	-9.250***	-35.63
	(1.759)	(1.750)	(34.13)
Age[a]	0	0	46.74
	(0)	(0)	(43.80)
Age2	-3.899***	-4.225***	-0.864
	(0.289)	(0.330)	(0.696)
Education	-69.21***	-70.17***	-67.51**
	(9.933)	(10.53)	(28.02)
kids	-253.4***	-273.1***	59.13
	(25.96)	(27.17)	(91.69)
Kids age 6-16	0.618	70.47***	84.58*
	(15.87)	(19.99)	(50.98)
2 parents	213.6***	364.6***	-65.04
	(62.63)	(81.84)	(105.0)
3 parents	171.8***	306.7***	-60.11
	(59.68)	(106.6)	(103.6)
Immigrant dummy	-58.48	-77.18*	63.19
	(37.98)	(40.72)	(116.9)
Other income	-0.007***	-0.008***	0.003**
	(0.0004)	(0.0005)	(0.0015)
Child benefit	-0.214***	-0.279***	-0.067**
	(0.013)	(0.016)	(0.033)
Imputed Wage	1861***	1881***	1282***
	(109.8)	(115.3)	(342.3)
marital1	-339.0***	0	0
	(40.03)	(0)	(0)
Age of youngest child			
1	-351.6***	-371.8***	-458.1***
	(34.08)	(36.94)	(97.08)
2	-283.0***	-266.0***	-357.0***
	(32.80)	(35.42)	(85.88)
3	-78.93**	-76.48**	-98.94
	(32.34)	(34.89)	(83.63)
4	-70.42**	-56.34	-133.3
	(31.86)	(34.38)	(81.64)
5	-21.92	-24.23	-73.48
	(31.92)	(34.73)	(78.63)
Child care price elasticity		-0.068	-0.015
β			
Observations	7775	6632	1143
R-squared	0.225	0.231	0.265

Note: The inverse Mill's ratios and other variable were included in the estimation but not reported.
[a] dropped during estimation process due to collinearity
Dependent variable: Annual number of hours worked

Table 6. Logit models-participation equations with selection corrections

VARIABLES	Pooled All	Pooled Married	Pooled Single	RE All	RE Married	RE Single	FE All	FE Married	FE Single
Price	-0.03	-0.008	-0.265	-0.102***	-0.095**	-0.655*	-0.106**	-0.195***	-1.483***
	(0.019)	(0.020)	(0.191)	(0.036)	(0.040)	(0.373)	(0.051)	(0.071)	-0.891)
Lambda1	-2.042***	-2.511***	-1.147***	-2.529***	-3.019***	-1.330*	-1.322**	-0.754	1.801
	(0.246)	(0.266)	(0.406)	(0.472)	(0.61)	(0.804)	(0.674)	(0.948	(2.05)
Lambda2	-0.166	0.021	-0.703	-1.472**	-1.866**	-1.713	-2.579**	-5.052***	-4.104
	(0.304	(0.348)	(0.57)	(0.631)	(0.832)	(1.164)	(1.015)	(1.576	(3.104)
Other income	-0.00002*	-0.00001***	-0.0001***	-0.000001	-0.00001*	-0.00002***	-0.00001	-0.00001	0.00001
	(0.00001)	(0.00001)	(0.00003)	(0.00002)	(0.00003)	(0.00006)	(0.00003)	(0.00006)	(0.00003)
Childage1	-0.223*	-0.258*	-0.419	-0.684***	-0.774***	-0.670*	-0.937***	-0.879***	-0.508
	(0.121)	(0.14)	(0.264)	(0.199)	(0.237	(0.407)	(0.238)	(0.298	(0.547)
Childage2	-0.407***	-0.425***	-0.541**	-0.971***	-1.047***	-0.911**	-1.149***	-1.177***	-0.56
	(0.118)	(0.133)	(0.257)	(0.189)	(0.224	(0.393)	(0.226)	(0.272	(0.55)
Childage3	-0.092	-0.121	-0.0477	-0.375**	-0.434**	-0.216	-0.578***	-0.604**	-0.148
	(0.122)	(0.136)	(0.264)	(0.185)	(0.217	(0.388)	(0.214)	(0.256	(0.518)
Childage4	-0.0911	-0.11	-0.0319	-0.321*	-0.344	-0.148	-0.502**	-0.503**	-0.265
	-0.122)	-0.136	-0.259)	-0.181)	-0.213	-0.373)	-0.208)	-0.246	-0.493)
Childage5	0.0303	0.0427	-0.01	-0.009	-0.0226	0.0414	-0.0984	-0.135	0.112
	(0.123)	(0.142)	(0.254)	(0.179)	(0.212	(0.357)	(0.196)	(0.234	(0.454)
Kids age 6-17	-0.0983**	-0.0683	-0.127	-0.188**	-0.14	-0.344	-0.0888	-0.239	-0.00591
	(0.0463)	(0.0477)	(0.112)	(0.087)	(0.103	(0.212)	(0.121)	(0.163	(0.389)
Imputed wage	2.403***	2.462***	2.682***	4.718***	4.644***	5.548***	5.190***	5.437***	8.370***
	(0.245)	(0.245)	(0.538)	(0.473)	(0.549)	(1.079)	(0.801)	(1.001)	(1.994)
Married	0.153			(0.212)			(0.187		
	-0.0954			-0.202			-0.28		
Sigma				2.033***	2.207***	1.550***			
				(0.096)	(0.106	(0.232			
Constant	-3.905***	-4.087***	-4.467***	-7.234***	-6.325***	-9.636***			
	(0.784)	(0.799)	(1.527)	(1.419)	(1.679	(2.915			
Observations	6503	5501	1002	6503	5501	1002	2037	1540	371

Table 7. Summary of labor supply elasticity with respect to childcare

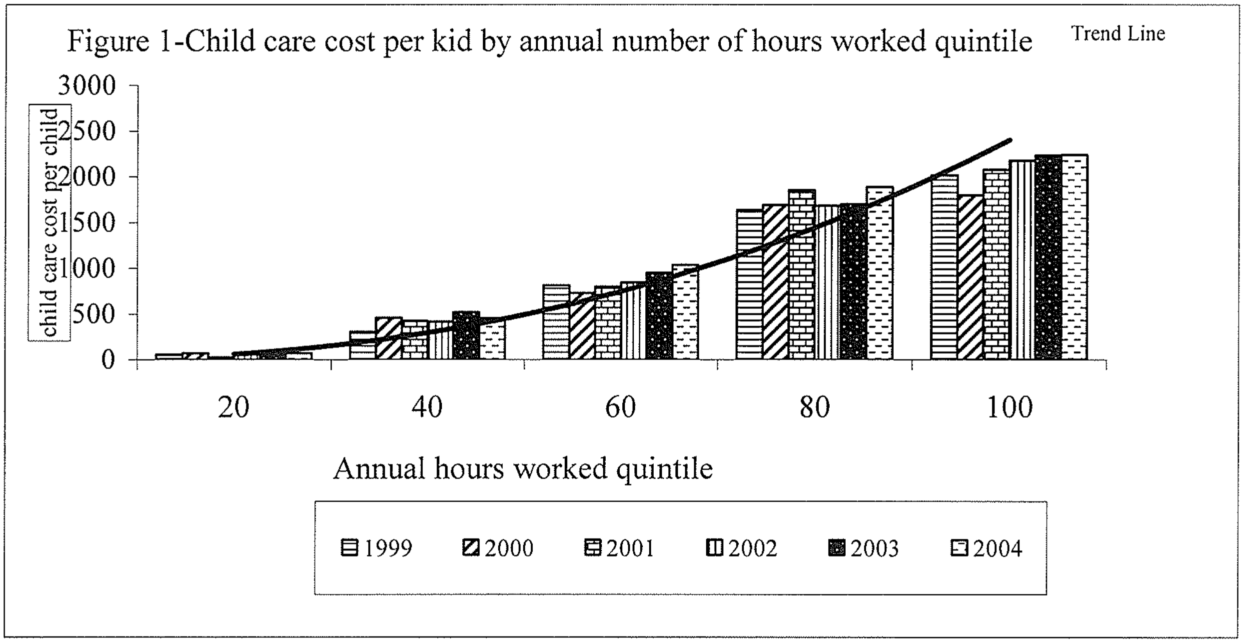
	Childcare cost elasticity with respect to				Country	Methodology	Data
	Labor force participation	Hours worked	Wage	Income		OLS and bivariate Probit	
Barrow(?)	-0.23		0.21	-0.04	USA	OLS and bivariate Probit	cross-section
Connelly(1992a)	-0.2				USA	OLS and bivariate Probit	cross-section
Blau and Robins(1998)	-0.38				USA	OLS and bivariate Probit	cross-section
Kimmel(1993)	-0.31		0.58		USA	OLS and bivariate Probit	cross-section
Kimmel(1998)	-0.22 for single mothers -0.92 for married mothers				-0.92	OLS and bivariate Probit	cross-section
Powell(1997)	-0.38	-0.32	0.85		Canada	Heckit and bivariate Probit	cross-section
Anderson and Levine(1995)	-0.46		0.58		USA	OLS and bivariate Probit	cross-section
Ribar (1992)	-0.74		0.68		USA		cross-section
Michalopoulos, Robins and Garfinkel(1992)(			0.04	-0.01	USA	OLS and bivariate Probit	cross-section
Cleveland at al (1996)	-0.39				Canada	OLS and bivariate Probit	cross-section
Lefebvre and Merrigan(2005)					Canada	Difference indifference method	Panel data
Baker et al(2005)					Canada	Difference in difference method	Panel Data
Ribar(1995)				-0.05	USA	OLS and bivariate Probit	cross-section
This study		-0.015 for married mothers -0.068 for single mothers			Canada	Sample selection panel data models	Panel data

Table 8. Testing for the independence of models using an interaction dummy

	Pooled		RE		FE	
	Coef.	P-value	Coef.	P-value	Coef.	P-Value
Child care price						
Childcare Price	-14.324	0.599	-46.589	0.058	-61.106	0.031
Childcare pricem	4.837	0.859	39.553	0.109	55.078	0.052
Mother's Age						
Age	120.298	0.000	138.059	0.000	154.088	0.000
Age Squared	-1.940	0.000	-2.156	0.000	-2.392	0.000
Education						
High school Grad	-34.368	0.329	3.008	0.949	-63.578	0.552
Non Uni Grad	67.322	0.078	83.934	0.099	-10.866	0.924
Degree	58.045	0.203	77.658	0.208	143.123	0.403
Children						
Ages 0 to 5	-64.663	0.003	-86.246	0.000	-68.315	0.069
Ages 6-16	17.082	0.260	-20.939	0.235	-15.754	0.664
Grand parents						
One	184.310	0.005	77.609	0.280	-32.928	0.736
Two	169.108	0.006	2.521	0.969	-220.307	0.015
Immigrant	-177.031	0.000	-173.572	0.000		
other income	-0.004	0.000	-0.002	0.000	0.000	0.798
Child benefit	-0.122	0.000	-0.093	0.000	-0.073	0.000
impwage	960.432	0.000	901.549	0.000	740.760	0.000
Marital	-80.876	0.089	-189.667	0.000	-288.973	0.000
Age of youngest Child						
One	-353.194	0.000	-347.110	0.000	-350.727	0.000
Two	-304.134	0.000	-300.294	0.000	-307.278	0.000
Three	-94.162	0.005	-109.053	0.000	-126.391	0.001

Four	-78.795	0.016	-105.595	0.000	-125.227	0.000
Five	-24.454	0.457	-44.477	0.062	-59.845	0.022
Constant	-2690.3	0.000	-2871.283	0.000	-2685.287	0.000
Over all R-squared	0.18		0.18		0.15	
N	7759		7759.00		7759.00	

Data Source: SLID (2004)
Note: Childcare pricem= Price*Marital status



Data Source: SLID (2004)

IV. Summary and Conclusions

The first essay attempts to estimate a life cycle labour supply model for men in Canada using sample selection corrected panel data models that also control for individual heterogeneity. The results confirm that life cycle labor supply estimates are very sensitive to the treatment of unobservable individual heterogeneity and sample selection and to changes in the econometric methods employed. Failure to control for individual heterogeneity, as in cross-sectional models, and sample selection in most cases produces upward biased intertemporal labour substitution elasticity estimates. After correcting for the sample selection problem and individual heterogeneity, we estimate the intertemporal labour substitution elasticity to be 0.16. Moreover we find the wealth effect to be statistically insignificant implying that the intertemporal elasticity is a very good approximation to both the compensated and uncompensated labour supply elasticity. The empirical results show that the effect of a temporary (one year) 100 percent increase in wages at a typical age would induce an increase in labor supply at that particular age by 16 percent but would leave all labour supply decisions in all other ages unchanged. On the other hand, a permanent 100 percent increase in wages over all ages (parallel shift of the wage profile) would induce a 16% increase in labor supply in all ages.

The second essay examines the effect of health on wages in a manner which accounts for a number of possible problems which might render estimates inconsistent: the problem of unobservable individual fixed effects such as genetic endowment; the problem of measurement error from the employment of self assessed health; and the problem of sample selection associated with endogenous choice of participation into the labour market. We utilize the estimator suggested by Wooldridge (1995) and some IV panel data estimators to take care of the aforementioned problems. The results indicate that the sample selection problem causes upward biased estimates of health effects on wages. After controlling for endogeneity of health and for sample selection bias our results reveal that the effect of health on wages is, as expected, positive but not statistically significantly different from zero. The key message of this paper is that accounting for sample selection bias and endogeneity of health is crucial in any attempt to uncover the health-wage nexus.

The third essay examines the impact of childcare cost on the maternal labour supply decision. The study aims to demonstrate, among other things, that the combined effects of unmeasured attributes implied by a life cycle labour supply model and sample selection bias in labour supply estimation can be accounted for using appropriate nonlinear panel data models. The general finding is that the elasticity of the annual number of hours worked with respect to childcare cost is very weak. The implied elasticity of the childcare price with respect to the annual number of hours worked is in the range of -0.01 to -0.012 for married mothers and in the range of –0.016 to -0.16 for single mothers. As far as the employment elasticity with respect to the childcare price is concerned, our estimates range from -0.012 to -0.10 for married mothers and from -0.016 to -0.48 for single mothers. The implied smaller labour supply reaction of mothers to childcare costs suggest that policies which lower childcare costs to enhance labour market opportunities for mothers do not seem to be promising.

In a nutshell, this dissertation emphasises the significance of controlling for individual heterogeneity and for sample selection when selection bias is not time invariant. In such cases, our research results imply that accounting for the individual heterogeneity using the traditional fixed effects technique may not be enough to control for the bias introduced by the sample selection problem.

Lastly, we need to be cautious in reading these conclusions and need to remember that they are subject to the usual caveats explicitly mentioned in each essay.