A paper prepared for the

Un document préparé pour le

Economic Council of Canada

Conseil économique du Canada

P.O. Box 527
Ottawa, Ontario
K1P 5V6

C.P. 527
Ottawa (Ontario)
K1P 5V6

[redacted]

The **Economic Council of Canada** was established in 1963 by Act of Parliament. The Council is a crown corporation consisting of a Chairman, two Directors and not more than twenty-five Members appointed by the Governor in Council.

The Council is an independent advisory body with broad terms of reference to study, advise and report on a very wide range of matters relating to Canada's economic development. The Council is empowered to conduct studies and inquiries on its own initiative, or if directed to do so by the Minister, and to report on these activities. The Council is required to publish annually a review of medium- and long-term economic prospects and problems. In addition it may publish such other studies and reports as it sees fit.

The Chairman is the Chief Executive Officer of the Council and has supervision over and direction of the work and staff of the Council. The expenses of the Council are paid out of money appropriated by Parliament for the purpose.

The Council as a corporate body bears final responsibility for the *Annual Review*, and for certain other reports which are clearly designated as *Council Reports*. The Council also publishes *Research Studies, Discussion Papers* and *Conference Proceedings* which are clearly attributed to individual authors rather than the Council as a whole. While the Council establishes general policy regarding such studies, it is the Chairman of the Council who bears final responsibility for the decision to publish authored research studies, discussion papers and conference proceedings under the imprint of the Council. The Chairman, in reaching a judgment on the competence and relevance of each author-attributed study or paper, is advised by the two Directors. In addition, for *authored Research Studies* the Chairman and the two Directors weigh the views of expert outside readers who report in confidence on the quality of the work. Publication of an author-attributed study or paper signifies that it is deemed a competent treatment worthy of public consideration, but does not imply endorsement of conclusions or recommendations by either the Chairman or Council members.

Établi en 1963 par une Loi du Parlement, le **Conseil économique du Canada** est une corporation de la Couronne composée d'un président, de deux directeurs et d'au plus vingt-cinq autres membres, qui sont nommés par le gouverneur en conseil.

Le Conseil est un organisme consultatif indépendant dont le mandat lui enjoint de faire des études, donner des avis et dresser des rapports concernant une grande variété de questions rattachées au développement économique du Canada. Le Conseil est autorisé à entreprendre des études et des enquêtes, de sa propre initiative ou à la demande du Ministre, et à faire rapport de ses activités. Chaque année, il doit préparer et faire publier un exposé sur les perspectives et les problèmes économiques à long et à moyen termes. Il peut aussi faire publier les études et les rapports dont la publication lui semble opportune.

Le président est le directeur général du Conseil; il en surveille les travaux et en dirige le personnel. Les montants requis pour acquitter les dépenses du Conseil sont prélevés sur les crédits que le Parlement vote à cette fin.

En tant que personne morale, le Conseil assume l'entière responsabilité des *Exposés annuels*, ainsi que de certains autres rapports qui sont clairement désignés comme étant des *Rapports du Conseil*. Figurent également au nombre des publications du Conseil, les *Études*, *Documents* et *Comptes rendus de colloques*, qui sont explicitement *attribués à des auteurs particuliers* plutôt qu'au Conseil lui-même. Celui-ci établit une politique générale touchant ces textes, mais c'est au président qu'il incombe de prendre la décision finale de faire publier, sous les auspices du Conseil économique du Canada, les ouvrages à nom d'auteur tels que les études, documents et rapports de colloques. Pour se prononcer sur la qualité, l'exactitude et l'objectivité d'une étude ou d'un document attribué à son auteur, le président est conseillé par les deux directeurs. De plus, dans le cas des *études à nom d'auteur*, le président et les deux directeurs sollicitent l'avis de lecteurs extérieurs spécialisés, qui font un rapport confidentiel sur la qualité de ces ouvrages. Le fait de publier une étude ou un document à nom d'auteur ne signifie pas que le président ou les membres du Conseil souscrivent aux conclusions ou recommandations contenues dans l'ouvrage, mais plutôt que l'analyse est jugée d'une qualité suffisante pour être portée à l'attention du public.

DISCUSSION PAPER NO. 305

Taxes and the Labour Supply of Married Women in Canada

by Eden Cloutier

Discussion Papers are working documents made available by the Council, in limited number and in the language of preparation, to interested individuals for the benefit of their professional comments.

La série "Documents" contient des documents de travail dont le Conseil fait une diffusion restreinte, dans leur version d'origine, en vue de susciter des commentaires de la part de spécialistes.

Council Secretary
Economic Council of Canada
Post Office Box 527
Ottawa, Ontario
K1P 5V6

ISSN-0225-8013

May 1986

RÉSUMÉ

Ce document présente une estimation de l'offre de travail de la part des femmes mariées au Canada, à l'aide d'un modèle statique incluant les impôts. L'offre de travail est mesurée selon les semaines travaillées au cours de l'année, estimées d'après les données des micro-fiches individuelles de l'Enquête de 1982 sur les finances des consommateurs (revenus de 1981).

Le modèle adopte d'emblée le concept du revenu virtuel. L'illusion fiscale en rapport avec le taux d'imposition marginal est établie de la manière courante, mais nous supposons néanmoins que le taux d'imposition moyen ne comporte pas d'illusion. L'équation des salaires est estimée sous forme log-linéaire et comporte les renseignements sur les heures travaillées par semaine. L'erreur quant au salaire moyen estimé a été corrigée avant le calcul des revenus annuels et de l'impôt sur le revenu. L'équation des salaires aussi bien que l'équation de l'offre de travail ont été corrigées de leurs erreurs dues à la sélection par le procédé Heckman, et les deux équations ont été estimées à l'aide de données d'échantillons non pondérées. Des coefficients de pondération des échantillons ont été utilisés pour le calcul des élasticités de l'offre de travail.

Les résultats obtenus sur les élasticités de l'offre de travail se situent dans la moyenne des résultats des études canadiennes récentes, mais ils sont considérablement inférieurs aux estimations récentes concernant les États-Unis. Des élasticités compensées relativement faibles, combinées à l'illusion sur le taux d'imposition marginal, donnent lieu à des estimations de faibles pertes de bien-être associées à l'impôt sur le revenu.

ABSTRACT

This paper presents the estimation of the labour supply of married women in Canada, using a one-period model which includes taxes. Labour supply is measured by weeks worked during the year, with the data for the estimation taken from the 1982 Survey of Consumer Finances (1981 incomes) individual micro data file.

The model fully incorporates the virtual income concept. Although tax illusion as to the marginal rate is incorporated in the standard way, it is further assumed that there is no illusion as to the average tax rate. The wage equation is estimated in log-linear form and incorporates information on hours worked per week. The bias in the estimated mean wage is corrected before the calculation of annual earnings and income taxes. Both the wage equation and the labour supply equation are corrected for selection bias using the Heckman procedure, and both are estimated using unweighted sample data. The sample weights are used to calculate labour supply elasticities.

The resulting labour supply elasticities fall about mid-range in the results of recent Canadian studies, and considerably below recent estimates for the United States. Relatively small compensated elasticities in combination with illusion as to the marginal tax rate lead to small deadweight loss estimates for income taxes.

CONTENTS

1 INTRODUCTION

In recent years, there have been substantial developments in the economic literature concerning the effect of taxes on labour force behaviour. Much of this literature, developed in the United States, has relied on data from the extensive income maintenance experiments conducted in the 1970s. The availability of such data has allowed the use of more sophisticated models and has encouraged the development of econometric methods to handle the estimation of enriched models. The result has been a general increase in the quality and reliability of the estimates produced over those available at the beginning of the 1970s, most particularly for prime aged males. Yet there remains significant variability in the estimates of labour supply response, most notably for married women.

Recent Canadian studies reveal this variability in the labour supply elasticities of married, working women. The conventional consensus was that for men, who typically work more hours and receive higher wages than women, the income effect was expected to dominate the substitution effect leading to a backward-bending labour supply schedule. Conversely, for women the dominant substitution effect would lead to a positively sloped labour supply schedule. This consensus has recently been broken by a series of papers which indicate that the labour supply elasticities of working women are negative, implying a backward-bending labour supply curve. Nakamura and Nakamura (1983) report uncompensated wage elasticities ranging from -0.087 to -0.036 for women working more than 1400 hours and from -0.083 to -0.197 for women working less than 1400 hours during the year. Similar results are reported by them for working women in the United States. Using a different sample, Robinson and Tomes (1985) also find negative uncompensated wage elasticities which offer strong support for the Nakamuras' results. Contrasted to these are the findings of Stelcner and Breslaw (1985) who report uncompensated wage elasticities for married women in Quebec ranging from 0.40 to 1.28 depending upon the estimation procedure used.

Despite the variation in the estimated elasticities, the above studies are similar in many respects including the treatment of sample selection bias by the Heckman (1976) procedure, and the treatment of spouse's income and family composition as exogenous in the determination of the wife's labour supply decision. Nakamura and Nakamura use the 1971 Census and a model with annual hours worked as the dependent variable. Taxes are included in their model as is the provision for tax illusion. Because of the simultaneity of the marginal tax rate and hours worked an iterative procedure was developed for the estimation. Their study is most often criticized for the method of obtaining hourly wages which introduces the problems of division bias and interval data.

Robinson and Tomes avoid the problem of division bias using data from the 1979 Quality of Life survey which contains a direct measure of the hourly wage rate for a subset of women. Their model, a one-period lifetime model, uses the logarithm of hours worked per week as the dependent variable, but ignores taxes. Although their estimated wage elasticities are significantly negative, the sample size is quite small and not necessarily representative of a larger population.

Stelcner and Breslaw in examining the labour supply of married women in Quebec use a one-period model but with annual weeks worked as the dependent variable. The data used are from the 1979 Survey of Consumer Finances Census Family micro data file. Taxes are calculated and the estimation procedure utilizes the Nakamuras' iterative method to handle the joint determination of marginal tax rates and labour supply. Although their use of weeks worked as the dependent variable allows them to avoid the division bias in calculating an hourly wage, the lack of information on hours worked per week weakens their estimation of a weekly wage.

Common with the two studies which include taxes is an improper treatment of non-earned income. The virtual income concept introduced by Hausman (1981) is ignored. Also, the method of introducing tax illusion in the marginal rate necessarily forces individuals off their budget constraint. That is, not only is there illusion as to the slope of the budget constraint at a particular point, there is also illusion as to the location of the point as well, or equivalently, illusion in the average tax rate. By using virtual income the direct link between illusion in the marginal tax rate and the average tax rate can be broken. Furthermore, by adjusting virtual income some non-convexities in the budget curve may be removed, allowing the concept of a reservation wage to be more fully supported.

The purpose of this paper is to present yet another estimation of labour supply response of married working women in Canada. In doing so, a number of differences from the preceeding studies will be evident. First, the concept of virtual income is fully integrated into the estimation. Second, while tax illusion as to the marginal tax rate is incorporated in the traditional way, it is assumed that workers know their annual after-tax income, that is, there is no illusion as to the average tax rate. Third, the non-convexity in the budget set caused by the married exemption is removed by recalculating the husband's taxes for any point at which the exemption changes, and thus treating the effect as a change in the wife's virtual income rather than affecting her marginal tax rate. Non-convexities caused by the small number in the sample who are in receipt of social assistance payments are ignored, as is the case with the other studies. The non-convexity at zero hours of work caused by the existence of fixed costs of working is also not considered. Fourth, it is a known fact that, since the wage equation is estimated in log-linear form, the predicted wage will be a biased estimator of the mean wage. This causes no problem when entering the wage into the labour supply equation as an instrumental variable since it is generally entered in log form. However, to calculate the marginal tax rate the wage is used in non-log form and the bias should be corrected. A method proposed by Goldberger (1968) was used for this purpose and produced estimates of the mean wage 7 to 13 per cent higher than the uncorrected wage estimates. This in turn had an impact on the marginal tax rate calculations. Fifth, hours worked per week information was used in estimating the weekly market wage by entering it as an explanatory variable in the equation. Finally, all the estimations in this paper are done with unweighted sample data. The conditions required for using weighted data (see DuMouchel and Duncan, 1983) would seem very unlikely to be satisfied. Population estimates of elasticities and deadweight loss were calculated by weighting the results for individuals.

The model used, which is described in the following section, was a simple one-period model including taxes. The data source used for the study was the 1982 Survey of Consumer Finances (1981 income) individual micro data file from which 7889 observations were used. As with any such study, it is necessary to make choices and submit to compromises. In this case it was decided to use the individual file to obtain the information on normal hours worked per week at the cost of losing information on the number of children in the family if that number exceeded two. The exact number of children was calculated using Family Allowance data in the tax calculations. This calculated number was then used in the labour supply equation. In addition, special tabulations on deductions by age and income were obtained from Taxation Statistics and used in the tax calculations.

2 LABOUR SUPPLY MODEL

The model is based on the neo-classical theory of consumer choice. The labour supply function is derived from a model which assumes the maximization of a well-behaved utility function in goods and leisure, subject to constraints on time and a budget. This model is subject to the normal simplifying assumptions except that taxes are taken into account. It is further assumed that the labour supply decision of the husband precedes that of the wife and is treated as exogenous in the wife's labour supply decision. This model may be summarized as follows:

Maximize the one-period utility function

$$U = U(c,l,Z)$$

subject to a time constraint

$$T = h+l$$

and a budget constraint

$$Y = pc = wh(1-t) + E$$

where:

c = consumption of goods

p = price of goods

Y = total income (earned and non-earned) after-tax

l = leisure (non-market) time

h = market time

T = total time available

E = virtual non-earned income

w = gross market wage

t = marginal tax rate

Z = a vector of socio-economic characteristics

The solution of the maximization problem gives the labour supply function:

$$h = h(w_n, Z, E)$$

where w_n is the net wage given by

$$w_n = w(1-t) = wr$$

and r is the marginal retention rate.

The virtual income, E, is calculated in the following way. First, all non-taxable income accruing to the wife is included in its entirety. Second, all taxable non-earnings income is reduced by the average tax rate and included in virtual income. Finally, the wife's employment income at market time h is multiplied by the difference between her marginal and average tax rates and included in virtual income. This procedure results in a worker's total net income after-tax being given by the sum of employment income after-tax (at the marginal rate) plus virtual income as will be demonstrated.

Let the total gross income, both earned and non-earned, Y_g, be given by

$$Y_g = wh + I$$

where I is non-earned income and, for simplicity, is assumed to be taxable. Then,

$$Y = Y_g(1-t_a)$$

where t_a is the average tax rate at h units of work.

Substituting for Y_g yields

$$Y = w(1-t_a)h + I(1-t_a)$$

Now, introduce the marginal tax rate, t, on the earnings income and adjust the non-earnings income correspondingly,

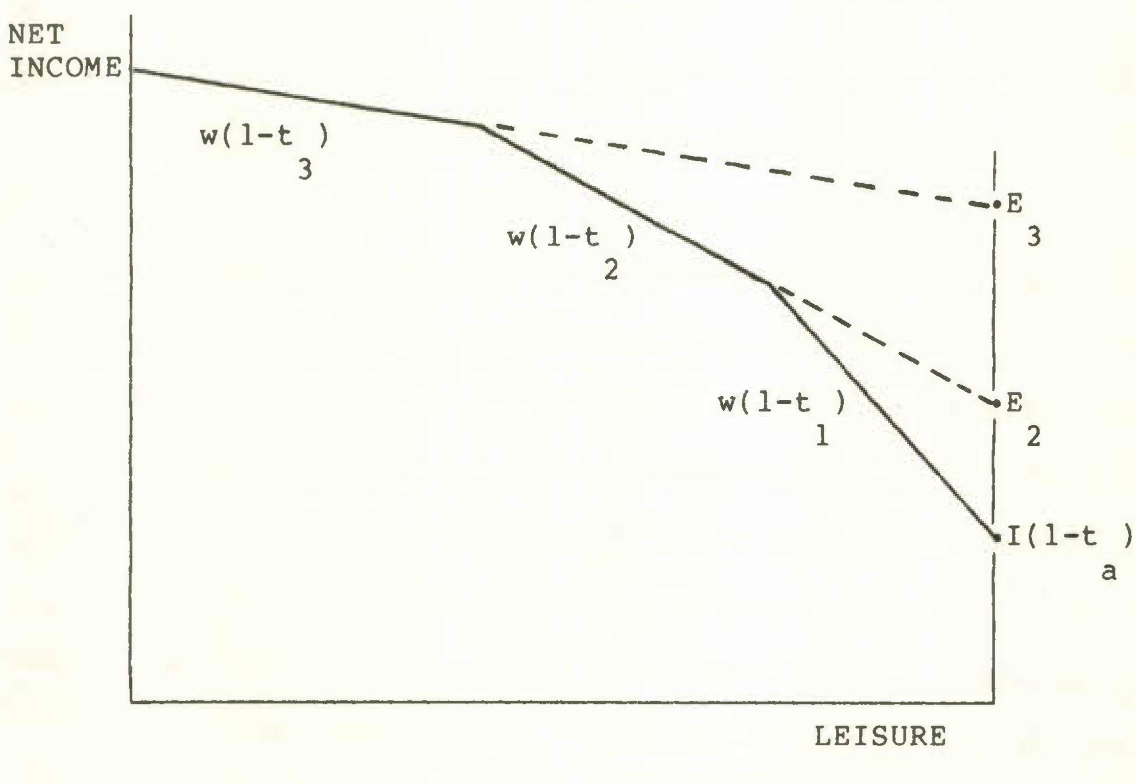
$$Y = w(1-t)h + wh(t-t_a) + I(1-t_a).$$

The effective after-tax wage is given by w(1-t), and the virtual income is given by

$$E = wh(t-t_a) + I(1-t_a)$$

When non-taxable income is included in I, it is also included directly in E with no tax adjustment, and in the calculation of t_a the non-taxable income does not enter the base.

Graphically, this procedure corresponds to linearizing the budget segments associated with different marginal tax rates as illustrated in the following diagram:



In effect, the virtual non-wage income is calculated by adjusting the after-tax non-wage income for having applied the current marginal tax rate to all wage income rather than just the wages along that segment.

Integrated with this calculation is the provision for tax illusion. In most studies a tax perception parameter is included on the marginal tax rate on the wage term, as suggested by Rosen (1976), in the following fashion,

$$w_n = w(1-t)^{\rho}$$

where ρ is an indicator of tax perception. When $\rho=0$ there is complete illusion while when $\rho=1$ there is no tax illusion. A feature of this procedure, however, is that it forces the individual off the budget constraint, which is equivalent to illusion in the average tax rate. Given the information an individual receives through such procedures as deduction at source, an equally tenable hypothesis would be that there was no tax illusion as to the average tax rate, or equivalently that a person would know at what point on the budget constraint they were located. Such an assumption is made in this study. There is illusion as to the marginal tax rate, but there is no illusion as to the average rate. To incorporate this assumption define a perceived marginal tax rate, t_{ρ}, as follows:

$$(1-t_\rho) = (1-t)^\rho$$

or
$$t_\rho = 1 - (1-t)^\rho$$

Then,
$$Y = w(1-t)^\rho h + wh(t_\rho - t_a) + I(1-t_a),$$

and
$$E = wh(t_\rho - t_a) + I(1-t_a)$$

The slope of the budget constraint is thus allowed to vary with ρ, the perception parameter, while using t_ρ in the virtual non-earned income calculation linearizes the constraint at that slope through the actual budget point. This was the form used for calculating virtual non-earned income in this study.

Returning to the main problem at hand, the traditional approach was followed by positing a reservation wage function. For this approach to be valid, the budget set must be convex at least in a region near the actual units of work supplied. There are a number of reasons why this may not be so. First, there is the married exemption to be considered. If it were assumed that the reduction in the married exemption claimed by the husband due to an increase in the wife's earnings had the effect of increasing the wife's effective marginal tax rate to that of her husband, then a non-convex budget set would occur at the origin. However, given the presence of tax illusion, an equally plausible assumption, and the one made here, was that the reduction in after-tax income caused by a decrease in the married exemption was viewed as an income effect. The wife's marginal tax rate thus remained her own while her virtual non-earned income was reduced by the amount that her husband's tax increased.

Second, non-convexities could be caused by transfer tax-backs with increasing earned income. In this study the number of cases for which this is a major consideration is relatively small and the problem was ignored. One non-convexity, that of the reduction of the child tax credit, was allowed in the budget set. Since the tax-back rate, at 5 per cent of increasing earnings, was shallow, it was felt that this would do no serious damage to the convexity requirement particularly since the non-convexity need not occur near the origin.

Finally, and potentially the most serious departure from non-convexity could be caused by the existence of fixed costs of work. This study, in common with the other Canadian studies referred to above, does not deal with the possibility of the existence of fixed costs.

Adopting the standard approach, we posit a market wage function, w, and a reservation wage function, w_r, which are expressed in log-linear form. The decision to participate in the labour market is taken if the net wage, w_n, exceeds the reservation wage with the amount of labour supplied adjusting until $w_n = w_r$.

The market wage equation is given by

$$\ln w = \alpha_0 + X\alpha_1 + u_1 \tag{1}$$

where X is a vector of market wage determining variables, and u_1 is a disturbance term.

The reservation wage equation is given by

$$\ln w_r = \beta_0 + Z\beta_1 + \beta_2 E + \beta_3 \ln w_n + \beta_4 h + u_3 \text{ if } h>0$$

and $$\ln w_r = \beta_0 + Z\beta_1 + \beta_2 E + u_3 \text{ if } h=0$$

where Z is a vector of socio-economic characteristics of the individual, and u_3 is a disturbance term. The net wage, w_n, and the virtual non-earned income, E, have been previously defined by:

$$w_n = w(1-t)^{\rho}$$

or $$\ln w_n = \ln w + \rho \ln(1-t)$$

and $$E = wh\ (t_{\rho} - t_a) + I(1-t_a)$$

Now, since an individual is a labour force participant if $w_n > w_r$, we may obtain the equilibrium labour supplied by setting $w_r = w_n$ in the reservation wage equation for h>0, and solving for h.

$$h = \gamma_0 + Z\gamma_1 + \gamma_2 E + \gamma_3 \ln w_n + u_2$$

where;

$$\gamma_0 = \frac{\beta_0}{\beta_4},\ \gamma_1 = -\frac{\beta_1}{\beta_4},\ \gamma_2 = -\frac{\beta_2}{\beta_4},\ \gamma_3 = \frac{1-\beta_3}{\beta_4},\ u_2 = \frac{1}{\beta_4} u_3$$

Substituting for the net wage w_n yields

$$h = \gamma_0 + Z\gamma_1 + \gamma_2 E + \gamma_3 \ln w + \gamma_3 \rho \ln(1-t) + u_2 \qquad (2)$$

The two equations of interest are the market wage equation (1) and the labour supply equation (2). As it stands, the labour supply equation cannot be directly estimated because of the simultaneity of h, t, and E, and both equations suffer from sample selection bias.

3 SPECIFICATION AND ESTIMATION

In this section the estimation procedure and empirical specification of the models will be discussed. First, we will consider the problem of sample selection bias as developed by Heckman (1976). Quite simply, the problem arises because the sample used in estimating the equations contains no observations on the market wage and labour supplied by individuals who did not work in the sample period. Following Heckman, the sample selection bias may be corrected by including a term in each equation resulting in

$$\ln w = \alpha_0 + \alpha_1 X + \frac{\sigma_{12}}{(\sigma_{22})^{\frac{1}{2}}} \lambda + \mu_1 \qquad (3)$$

and

$$h = \gamma_0 + Z\gamma_1 + \gamma_2 E + \gamma_3 \ln w + \gamma_3 \rho \ln(1-t) + \frac{\sigma_{22}}{(\sigma_{22})^{\frac{1}{2}}} \lambda + \mu_2 \qquad (4)$$

where

$$E \{u_1^2\} = \sigma_{11},\ E \{u_2^2\} = \sigma_{22},\ E \{u_1 u_2\} = \sigma_{12},\ \lambda = \frac{f(\phi)}{1-F(\phi)}$$

$$\phi = - \frac{\gamma_0 + Z\gamma_1 + \gamma_2 E + \gamma_3 \ln w + \gamma_3 \rho \ln(1-t)}{(\sigma_{22})^{\frac{1}{2}}}$$

$f(\phi)$ and $F(\phi)$ are, respectively, the density and distribution function of the standard normal distribution. λ is the inverse of the Mill's ratio having as its denominator the probability that an observation has data for h.

Thus, equations (3) and (4) are the equations to be estimated to take account of selectivity bias. Before estimation may proceed, however, we first need to consider the error structure, and second need to obtain λ to enter into the equation for each observation.

The error structure of equations (3) and (4) are given by

$$E\,\{\mu_1^2\} = \sigma_{11}\,([1-r^2] + r^2\,[1+\phi\lambda-\lambda^2])$$

and
$$E\,\{\mu_2^2\} = \sigma_{22}\,(1+\phi\lambda-\lambda^2)$$

where
$$r = \frac{\sigma_{12}}{(\sigma_{11}\,\sigma_{22})^{\frac{1}{2}}}$$

and
$$0 \leqslant 1 + \phi\lambda - \lambda^2 \leqslant 1$$

Clearly the error structure is heteroskedastic, and hence a weighted estimation is indicated. Also, λ is required in calculating the appropriate weight as well as for including in the equations.

The normal method of obtaining λ is to estimate a probit function for the entire sample where the dependent variable is set to one if the individual was in the labour force, and zero otherwise, and the explanatory variables are found in X and Z. For each individual one then retrieves λ and includes it in the labour supply and market wage equations.

This general procedure was modified slightly in the present study. First, since it was also desired to use data on the hours worked per week this became a consideration in the estimation of λ. In the Survey of Consumer Finances, data on the normal hours worked per week are available only for those individuals who are actually employed or awaiting recall during the reference week. The reference week is in the spring of the year following the year to which the data relate. Our sample selection rule was thus to choose individuals who were employed at some time during 1981 and in addition were employed during the sample reference week. For these wives the dependent variable was set to unity, while for those wives who either did not work during 1981 or who were not employed during the reference week the dependent variable was set equal to zero.

The second modification to the procedure was to use a logit regression to approximate a probit procedure as described by Amemiya (1981). This was done strictly for computational convenience. Thus, the predicted value from the logit regression was multiplied by 1.6, as suggested by Amemiya, and from then on normal density and distribution functions were used to calculate $\hat{\lambda}$ and the error structures of equations (3) and (4). Although the tails of the approximating distribution are slightly larger than a proper normal, provided normal density and distribution functions are used in calculating $\hat{\lambda}$, the properties of λ and $1+\phi\lambda-\lambda^2$ given previously will hold.

The market wage equation (3) was then corrected for selection bias using $\hat{\lambda}$, yielding the following equation to be estimated

$$\ln w = \alpha_0 + \alpha_1 X + \frac{\sigma_{12}}{(\sigma_{22})^{\frac{1}{2}}}\hat{\lambda} + \mu_1$$

where $$E\{\mu_1^2\} = \sigma_{11}\left([1-r^2] + r^2\,[1+\phi\hat{\lambda}-\hat{\lambda}^2]\right)$$

and $$r = \frac{\sigma_{12}}{(\sigma_{11}\sigma_{22})^{\frac{1}{2}}}$$

Because the error structure contained parameters to be estimated in the equation, this equation was estimated iteratively. First, r^2 was set to zero and the equation was estimated by OLS. From the coefficient on $\hat{\lambda}$ and the residuals a new value of r was calculated and used in the calculation of the equation error structure. The equation was then re-estimated by GLS and the iterative procedure repeated until convergence.

The results of the market wage equation were used in two ways. First, $\widehat{\ln w}$, the prediction from the market wage equation, was entered as an explanatory variable in the labour supply equation. Second, the wage was used in conjunction with the estimated labour supplied to calculate annual earnings to be used in the income tax calculation. The method used to calculate the estimated wage, $\hat{w}$, for use in the tax calculation in virtually all studies to date produces a biased estimate of the individual's mean wage. The problem of bias in log-linear regressions is well known and in this study we used a procedure developed by Goldberger (1968) to remove this bias. The effect of this procedure was to increase the wages used in the tax calculation by 7 to 13 per cent and in the process move many individuals to higher tax brackets. The procedure is outlined in an appendix.

The labour supply equation (4) adjusted for sample selection bias was estimated iteratively by GLS using a procedure similar to that proposed by Nakamura and Nakamura. First, the marginal tax rate was set equal to zero resulting in both the perceived marginal tax, t_ρ, and the average tax rate, t_a, also being zero. The tax perception parameter, ρ, was arbitrarily set equal to one. Virtual income and the error structure were calculated and the equation estimated by GLS. From the resulting parameter estimates, by dividing the coefficient of $\ln(1-t)$ by the coefficient of $\ln w$ a value for ρ was obtained. The predicted labour supply was multiplied by the estimated wage, $\hat{w}$, and income taxes were calculated on this earnings base yielding values for t and t_a. It was then possible to re-calculate t_ρ and virtual income. All these new values were used in a re-estimation of the labour supply equation by GLS. The procedure was continued until covergence. Unlike the market wage equation the error structure of the labour supply equation, once calculated, remained constant. The iterative procedure was required to treat the endogeneity of labour supply with taxes, tax perception, and virtual income.

The empirical specification of the estimated equations generally followed those found in the literature. The logistics regression, used to correct for sample selection bias, included as regressors age, education, children, location, homeownership, and husband's income variables in the following specification.

$$\begin{aligned} PART_i = {} & \beta_0 + \beta_1 \cdot AGE_i + \beta_2 \cdot EDUCA_i + \beta_3 \cdot EDUCB_i + \beta_4 \cdot EDUCC_i + \\ & \beta_5 \cdot EDUCD_i + \beta_6 \cdot EDUCE_i + \beta_7 \cdot C\_U\_6_i + \beta_8 \cdot C\_6\_17_i + \\ & \beta_9 \cdot AREA_i + \beta_{10} \cdot HOME_i + \beta_{11} \cdot SINC_i \end{aligned}$$

The variables were defined in the following way.

$PART_i$ = was set equal to one if the i^{th} individual worked during 1981 and during the reference week, and was set equal to zero otherwise.

AGE_i = was the age of the i^{th} individual divided by 10.

The next set of variables were dummy variables indicating the level of education attained. The reference group consisted of those individuals with up to eight years of elementary schooling.

$EDUCA_i$ = 9, 10 or 11 years of elementary or secondary schooling

$EDUCB_i$ = 12 or 13 years of elementary or secondary schooling

$EDUCC_i$ = some post-secondary education

$EDUCD_i$ = post-secondary certificate or diploma

$EDUCE_i$ = university degree

$C\_U\_6_i$ = number of children under six years of age. This variable has permitted values of 0, 1, 2, or 3, where 3 is used for 3 or more children.

$C\_6\_17_i$ = number of children 6 to 17 years of age. This variable has the same characteristics as the previous variable.

$AREA_i$ = this is a dummy variable for the size of the place of residence, with a value of one for those individuals living in a large urban centre with a population of 100,000 or more, and zero otherwise.

$HOME_i$ = this is a dummy variable for home tenure, with value of zero if the home was owned without a mortgage, and one otherwise.

$SINC_i$ = spouse's total taxable income in thousand of dollars.

A number of different specifications were tried for the age variable, however, in the final regressions the linear form was used since it performed as well as any other. Although there are a number of off-setting influences affecting the coefficient of the age variable, it is generally expected that participation decreases with age.

To the extent that education is correlated with earnings, it is expected that participation is positively related to education, and that the relationship strengthens with increasing levels of education.

Dependent children reflect family responsibilities of the wife and are thus expected to have a negative coefficient. In addition, separating dependent children into two age groups allows pre-school and school aged children to have a differential impact on participation.

With respect to the remaining variables, participation is expected to be positively related to both the size of place of residence and the home tenure variables and negatively related to the spouse's taxable income.

The market wage equation was specified in log-linear form in the following fashion:

$$\ln w_i = \alpha_0 + \alpha_1 AGE_i + \alpha_2 (AGE_i)^2 + \alpha_3 EDUCA_i + \alpha_4 EDUCB_i +$$

$$\alpha_5 EDUCC_i + \alpha_6 EDUCD_i + \alpha_7 EDUCE_i + \alpha_8 CC\_U\_6_i +$$

$$\alpha_9 CC\_6\_17_i + \alpha_{10} AREA_i + \alpha_{11} HRS_i + \alpha_{12} \hat{\lambda}_i + \mu_{1i}$$

In this equation most of the variables retain the definitions previously given for the logistics regression. The variables for dependent children, however, have been replaced by values calculated using Family Allowance data to obtain the exact number of children rather than using the value 3 to represent 3 or more children. In this particular equation the number of dependent children is used as a proxy for labour force interruptions and hence it is desired to use the exact number of children. The age variable is entered in quadratic form to accommodate a humped age-earnings profile.

The dependent variable, $\ln w_i$, is the natural logarithm of the weekly wage calculated by dividing annual earnings by the actual number of weeks worked during 1981. Clearly, the weekly wage will depend upon the hours worked during the week and it would be preferable to estimate an hourly wage if the data were available. However, as has been pointed out before, such a procedure with the available data could lead to division bias. Suppose the proper measure of hours worked per week were available. Then, the hourly wage could be calculated as

$$\text{hourly wage} = \frac{\text{annual earnings}}{\text{weeks worked}} \cdot \frac{1}{\text{hours per week}} = \frac{\text{weekly wage}}{\text{hours per week}}$$

When the logarithm is taken before estimating the equation the ratio on the right hand side becomes a difference which may then be split resulting in a weekly wage equation of the form

$$\ln w_i = X_i\alpha + \ln(\text{hours per week})_i + \mu_i$$

The procedure used in the estimation in this paper was to use the logarithm of the hours worked per week during the reference week as an instrument with a coefficient to be estimated. This is the HRS_i variable in the equation.

Finally, $\hat{\lambda}_i$ is the inverse of the Mill's ratio calculated from the logistics regression and included to correct for sample selection bias.

The labour supply equation specification allowed for differences between full-time and part-time workers in certain variables where significant differences in response were expected. The split between full- and part-time was done using a separate variable in the survey which indicated the main full- or part-time status during 1981, and did not use the hours worked during the reference week. The wage, tax and virtual income variables were included separately for each group, and in addition pre-school aged children were also included separately.

The labour supply equation was specified in linear form in the following fashion:

$$h_i = \gamma_{F0} + \gamma_{P0} + \gamma_{F1} \ln w_{Fi} + \gamma_{F2} \ln(1-t)_{Fi} + \gamma_{F3}E_{Fi} +$$

$$\gamma_{P1} \ln w_{Pi} + \gamma_{P3}E_{Pi} + \gamma_4 AGE_i + \gamma_{F5}CC\_U\_6_{Fi} + \gamma_{P5}CC\_U\_6_{Pi} +$$

$$\gamma_7 CC\_6\_17_i + \gamma_8 AREA_i + \gamma_9 HOME_i + \gamma_{10}\hat{\lambda}_i + \mu_{2i}$$

where the subscript F and P on variables and coefficients refer to full-time and part-time respectively.

As described in the previous section, ln w is the logarithm of the estimated weekly market wage, ln(1-t) is the logarithm of one minus the estimated marginal tax rate, E is the virtual income, and h is the weeks worked. The dependent children variables reflect the household responsibilities of the wife, while the HOME variable is included to reflect financial responsibility. The AREA variable is included to allow for higher levels of labour supply of those living in large urban areas. Finally, $\hat{\lambda}$ is the inverse of the Mill's ratio included to correct for sample selection bias. In this case it may be seen from equation (4) that this coefficient should be positive.

This concludes the estimation procedure and the empirical specification of the equations to be estimated. Some detail of the tax calculations and the calculation of the number of dependent children from Family Allowances data is found in an appendix at the end of this paper. We next turn to the results of the estimations and calculated labour supply elasticities.

4 RESULTS

The results of the logistics regression are given in Table 1. The sample consisted of 15,914 observations, from which had been removed those observations where: 1) the wife was not aged 20 to 54 years inclusive; 2) the wife was a member of a farm family; 3) the wife was an immigrant in 1981-82; 4) the wife was a full-time student; 5) the wife was an unpaid family worker or self-employed; and 6) data were missing on earnings, labour supply, or hours worked per week. Of the 15,914 observations 7911 were not labour force participants during 1981 or were not working during the reference week, while 8003 did participate in the labour force at some time during 1981 and were employed during the reference week.

All variables entered the regression significantly and all had the expected signs. The results show that the probability of participation decreases with age. This is the general result found by most labour market researchers.

The education variables as a group are highly significant. The results show that an increasing level of schooling leads to an increasing probability of participation. This, no doubt, is a reflection of the correlation between schooling and earnings potential.

Table 1

Logistics Estimates of the Participation Equation Married Women, Aged 20 to 54 years, Canada, 1981

Variable	Coefficient	Standard Error	Chi-Square
Intercept	0.4630	0.1132	16.72
AGE	-0.2187	0.0225	94.38
EDUCA	0.4636	0.0547	71.88
EDUCB	0.9865	0.0587	282.26
EDUCC	1.0765	0.0798	182.18
EDUCD	1.5261	0.0670	518.10
EDUCE	1.6663	0.0841	392.93
C_U_6	-0.9504	0.0294	1047.04
C_6_17	-0.1617	0.0175	85.57
AREA	0.3925	0.0374	110.38
HOME	0.2848	0.0412	47.81
SINC	-0.0097	0.0013	59.72

Unweighted Data

Dependent Variable = $\begin{cases} 0 \text{ for } 7911 \text{ observations} \\ 1 \text{ for } 8003 \text{ observations} \end{cases}$

Model Chi-Square = 2434.75

Fraction of concordant pairs of predicted probabilities and responses = 0.711.

Rank correlation between predicted probability and response = 0.434.

Data Source Statistics Canada, 1982 Survey of Consumer Finances, Micro Data File, Individuals Age 15 and Over, With or Without Income, 1981.

The family responsibilities of the wife, as reflected by the number of dependent children, are a major deterrent to labour force participation. This is especially true with pre-school aged children where the demand on the mother's time is not broken by school attendance. The financial responsibilities as reflected by the HOME variable increases the probability of participation, while the probability is reduced with an increase in the spouse's income.

As previously mentioned, since a logistics regression was used the output was modified, following Amemiya (1981), by multiplying the predicted value from the equation by 1.6 and then using this as the standardized value of a normal distribution. The inverse of the Mill's ratio was then calculated for inclusion as a variable in the wage and labour supply equations.

The market wage equation was estimated iteratively by GLS where the iteration was required because parameters in the error term were estimated in the equation itself. The results of the estimation are given in Table 2. These results were obtained after five iterations when the procedure had converged.

All variables entered the regression significantly and all had the expected signs. The AGE variables indicate a peak at 33 years of age, a very low estimate given what is generally expected from life-cycle earnings. However, it must be emphasized that the data

Table 2

GLS Estimates of the Market Wage Equation
Married Women, Aged 20 to 54 Years, Canada, 1981

Variable	Coefficient	Standard Error	t-Value
Intercept	2.8785	0.1065	27.03
AGE	0.2827	0.0519	5.44
$(AGE)^2$	-0.0430	0.0070	-6.18
EDUCA	0.1829	0.0221	8.29
EDUCB	0.4155	0.0319	13.02
EDUCC	0.4801	0.0365	13.14
EDUCD	0.6055	0.0431	14.06
EDUCE	0.9680	0.0457	21.17
CC_U_6	-0.1757	0.0263	-6.68
CC_6_17	-0.0668	0.0069	-9.68
AREA	0.1502	0.0142	10.57
HRS	0.4738	0.0094	50.14
$\hat{\lambda}$	0.4116	0.0680	6.05

Unweighted Data

MSE = 0.2900 $\qquad r = \dfrac{\sigma_{12}}{(\sigma_{11}\sigma_{22})^{\frac{1}{2}}} = 0.7643$

F Ratio = 391.76

R^2 = 0.3738 $\qquad$ Number of observations = 7889

Data Source Statistics Canada, 1982 Survey of Consumer Finances, Micro Data File, Individuals Age 15 and Over, With and Without Income, 1981.

are cross-sectional and not life-cycle. Available evidence indicates an individual life-cycle profile peaking somewhere over age 50, but each successive age group starting at higher and higher levels and remaining higher throughout the cycle. What we have estimated above is thus not a life-cycle profile, but rather the locus of points on different life-cycle profiles at a point in time with relatively more weight given to younger age cohorts.

The education variables taken as a group show a strong positive relationship between education and the market wage, with each successive coefficient entering the equation with a larger value and more strongly.

The dependent children variables, entered as a proxy for labour market interruptions, depress the market wage. The most recent interruptions, measured by younger children, have a noticeable impact which appears to decay over time as indicated by the coefficient on older children.

The normal hours worked as measured by hours during the reference week is, as one would expect, strongly and positively related to the weekly market wage. The coefficient is significantly different from unity, which is what the coefficient should be if a proper measure of hours were used.

Finally, there is a finding of significant sample selection bias as indicated by the coefficient of $\hat{\lambda}$. By contrast, the majority of Canadian studies find no significant sample selection bias in the market wage equation.

The predicted market wage in logarithmic form was then entered as an instrumental variable in the labour supply equation. In addition, an unbiased estimate of the weekly market wage was calculated as outlined in Appendix A for use in the tax calculations.

The final equation to be estimated was the labour supply equation. This equation was estimated by GLS and the results are presented in Table 3.

In this equation, while all the coefficients had the expected sign, not all entered the regression significantly. This was particularly true for the part-time sample. The sample division between full-time and part-time was accomplished using a variable in the survey which indicated whether an individual worked mostly full-time or part-time during 1981. This variable was used to avoid misclassifying full-time part-year workers as part-time as would be done if an annual labour supply variable had been used. Even then, from the point of view of labour supply, a misclassification could occur if an individual had been demand constrained. In fact, in the sample of 2159 part-time workers, 405 indicated

Table 3

GLS Estimates of the Labour Supply Equation
Married Women, Aged 20 to 54 Years, Canada, 1981

Variable	Coefficient	Standard Error	t-Value
Intercept_F	-6.9722	5.9696	-1.17
Intercept_P	19.1933	5.9238	3.24
$\ln w_F$	9.0384	1.0636	8.50
$\ln(1-t)_F$	6.0367	3.0691	1.97
E_F	-0.0671	0.0188	-3.57
$\ln w_P$	4.6441	0.7508	6.19
E_P	-0.0063	0.0225	-0.28
AGE	1.2345	0.1936	6.38
CC_U_6_F	-1.9172	0.4185	-4.58
CC_U_6_P	-2.4395	0.4577	-5.33
CC_6_17	-0.6516	0.1423	-4.58
AREA	0.6580	0.2977	2.21
HOME	0.9328	0.3509	2.66
$\hat{\lambda}$	2.8834	0.9398	3.07

Unweighted Data

MSE = 349.6742
F Ratio = 46.88
R^2 = 0.0718

Tax Illusion Parameter

$\rho = 0.6679$

Part time observations = 2159
Full-time observations = 5730

Data Source Statistics Canada, 1982 Survey of Consumer Finances, Micro Data File, Individuals Age 15 and Over, With and Without Income, 1981.

that they could only find part-time work. When these workers were moved to the full-time sample, there was virtually no change in the full-time results and a slight deterioration in the part-time results. From this, and a few other variations, it would appear that the classification used to separate full-time and part-time had no major impact on the results presented for the two groups.

In examining the coefficients, the gross wage, for both groups, is a strong determinant of labour supplied. An increase in the gross wage increases labour supplied with the effect being roughly twice as large for full-time workers as for part-time. The marginal retention rate (one minus the marginal tax rate) for full-time workers is just significant at the 5 per cent level with a coefficient which, when combined with the coefficient on the gross wage, indicates that existence of tax illusion with the tax illusion parameter, ρ, being estimated as 0.67. This level of tax illusion means, for example, that a marginal tax rate of 40 per cent would be perceived as about 29 per cent, which is above the average tax rate but well short of the actual marginal rate.

It will be observed that no tax variable has been included for part-time workers. The initial specification for part-time workers was identical to that for full-time workers. The tax variable, however, did not enter the equation significantly in any of the regressions conducted. Subsequently, two alternatives were tried. The first assumed no tax illusion and used the net wage as the wage variable in the regression, while the second alternative assumed complete tax illusion and used the gross wage with no tax variable. The second alternative was preferred on statistical grounds. This alternative leads, of course, to there being no dead weight loss of taxation with respect to the part-time workers since it is the perceived tax rate that matters for this purpose.

The virtual income variable had the proper sign for both groups of workers, however, it was significant only for full-time workers whose coefficient was roughly ten times the size of that for part-time workers. In summary, full-time workers appeared to behave as postulated by the theoretical model whose estimated coefficients for the economic variables were all of the proper sign and significant. The results for part-time workers with respect to the economic variables were much more mixed, with only the wage variable being significant and the tax and income variables seemingly having little effect.

The age variable shows an increasing labour supply with age, while all the variables for dependent children have significantly negative coefficents. Here, the variable for pre-school aged children was split for full-time and part-time workers to allow for a differential impact due to differences in being able to arrange day-care on a full-time as against a part-time basis. As expected, young children have a greater impact on part-time workers.

Residence in a large metropolitan area and increased financial responsibility as measured by the HOME variable both increase labour supplied. Finally, there is significant sample selection bias. The coefficient of $\hat{\lambda}$ is positive, as required by theory, and significant.

Because the equation was estimated using unweighted data, the computation of population elasticities cannot be done directly from the equation. Instead, individual labour supply responses given by the equation to wage, tax, and income changes were calculated and weighted by sample weights to calculate population elasticities. In the case of a wage change, the change enters directly into the wage variable and indirectly through the calculation of virtual income. A change in non-earned income enters only into virtual income. From these two calculations it is then possible to calculate a compensated wage elasticity. For these calculations it was assumed that marginal and average taxes remained constant. Tax elasticities were calculated by assuming a 1 per cent change in both marginal and average taxes, and furthermore that such tax changes would also affect the spouse's income tax. Such tax changes thus affected the calculation of virtual income, both directly, and indirectly through the spouse's after-tax income. The elasticities are presented in Table 4.

Table 4

Calculated Labour Supply Elasticities
Married Women, Aged 20 to 54 Years, Canada, 1981

Elasticity	Full-Time	Part-Time	Total
Uncompensated Wage	0.1895	0.1079	0.1684
Income	-0.0260	-0.0031	-0.0201
Compensated Wage	0.2054	0.1093	0.1806
Income Tax	-0.0557	-0.0288	-0.0487

Source Statistics Canada, 1982 Survey of Consumer Finances, and calculations from Table 3 equation results.

The elasticities calculated from this study fall in about the mid-range of elasticities calculated in recent Canadian studies, and considerably below comparable elasticities for the United States. The uncompensated wage elasticity of about 0.2, and lower for part-time than for full-time, shows a mild degree of labour supply responsiveness to wage changes. The small income elasticities are again in line with recent Canadian estimates, and combined with the low wage elasticities yield small substitution elasticities. The tax elasticities are also small. However, it must be recalled that for part-time workers the specification excludes any effect of the marginal tax rate and includes only the average tax rate operating on non-earned income.

The calculation of deadweight loss from the income tax for full-time workers led to an estimate of 3.5 per cent of tax revenues raised, a dramatically different result from comparable United States results. For part-time workers no calculation of deadweight loss was done since complete tax illusion as to the marginal rate was incorporated into the estimation. Since it was assumed at the same time that there existed no tax illusion as to the average rate, it would have been preferable in the estimation to have set the marginal rate equal to the average rate in the case of complete illusion rather than zero. Nonetheless, the resulting deadweight loss would have remained small being the result of small substitution elasticities in the presence of tax illusion.

Overall, the results for married women working full-time indicate strong support for the theoretical model. The results for part-time workers, however, are much weaker indicating both a lack of the model in capturing the underlying structure and in a greater degree of heterogeneity among part-time workers.

5 SUMMARY AND CONCLUSIONS

This paper has examined the labour supply response of married women in Canada using a one-period model in which the tax system was considered. The data used were cross-sectional from the Survey of Consumer Finances for the 1981 income year. The estimation procedure used the Heckman technique to correct for selectivity bias. This paper differs from previous Canadian studies in a number of ways. First, virtual income is fully and correctly calculated in the estimated equations. Second, tax illusion was allowed to exist only for the marginal tax rate and not the average tax rate. Third, some non-convexities in the budget set are removed and treated as income effects. Fourth, information on hours worked per week is included in the wage equation. Fifth, the wage predictions are purged of their bias. Finally, all of the equations are estimated using unweighted data with the sample weights being used subsequently in the elasticity calculations. The model, in common with most other Canadian studies, assumes away non-convexities in the budget set caused by fixed costs and transfer reductions.

The results of the estimations indicate that the model fits the full-time sample much more strongly than it does part-time workers. The aggregate elasticity estimates are about in the middle range of those found in other Canadian studies. For the full-time sample there is a finding of tax illusion with respect to the marginal tax rate. For part-time workers no estimate of tax illusion was obtained and the final form of the labour supply equation assumed complete tax illusion. The existence of tax illusion, which lowered the perceived marginal tax rate, in combination with a small substitution substitution elasticity produced estimates of deadweight loss of three to three and one-half per cent of revenues raised.

In a recent survey article, Rea (1983) gave best guess estimates for labour supply elasticities of married women of 0.50 for the uncompensated wage, -0.25 for income and 0.75 for the substitution elasticity. These elasticities are quite heavily weighted by results from the United States although they do contain some Canadian results. From this it would appear that the labour supply of married women in Canada is much less sensitive to wages, income and taxes than for their American counterparts. This being said, care must be exercised in attempting to draw conclusions at this stage.

Many of the American results are based on work using data from income maintenance experiments and as such deal with lower income groups. It is generally thought that elasticities for low income groups facing relatively high effective tax rates are often substantially higher than for the population as a whole. Preliminary findings by other researchers using the Mincome data from the Manitoba guaranteed annual income experiment indicate that this is also true for Canada. Thus, one must be sure when looking at best guesses from groups of studies that the underlying populations are relatively similar between comparisons. Yet, set against this reservation are the results of individual American studies of the elasticities for the entire population of married women in excess of most Canadian results. Hausman (1981), for example, has estimated the uncompensated wage elasticity as 0.91, and the substitution elasticity as 1.42.

There are substantial differences Canada and the United States in both transfer and tax systems. Canadian families, on average, receive a higher level of transfers than do similar American families, but face generally higher tax schedules. The requirement of joint filing in the United States, however, could easily mean that the tax burden faced by a married woman in the United States was much higher than a similarly paid married woman in Canada. These and other structural differences could go a long way towards explaining the differences in labour supply behaviour.

To the extent that there are such differences in labour supply between the two countries, why would such findings be important? Aside from being interesting to research in their own right, such results could also lead to different conclusions about the efficiency costs of taxes, transfers and redistribution. Results such as those of Browning and Johnson (1984) which suggest a high marginal cost of redistribution based upon labour supply effects, primarily of individuals in the upper end of the income distribution, need not be directly relevant for Canada, or at least could be substantially modified.

APPENDIX A

OBTAINING A MINIMUM VARIANCE UNBIASED ESTIMATE OF THE CONDITIONAL MEAN WAGE FROM THE MARKET WAGE EQUATION

This appendix closely follows the development given in Goldberger (1968) adapted to the specific problem at hand.

1. Preliminaries

Let $$u \sim N(\mu, \sigma^2)$$

and let $$U = e^u$$

Then, U is lognormally distributed with mean

$$E\{U\} = e^{E\{u\}+\frac{1}{2}V(u)} = e^{\mu+\frac{1}{2}\sigma^2}$$

That is, $E\{U\}$ is the moment generating function of u.

Lemma: Let vw/σ^2 be distributed as χ^2_v, where w is a random variable, v a positive integer, and σ^2 a positive parameter. For a given constant c, an unbiased estimator of $e^{c\sigma^2}$ is given by the function

$$F(w; v, c) = \sum_{j=0}^{\infty} \frac{f_j (cw)^j}{j!}$$

where $$f_j = \frac{(\frac{1}{2}v)^j \Gamma(\frac{1}{2}v)}{\Gamma(\frac{1}{2}v+j)}$$

and $\Gamma(\alpha)$ is the gamma function. If α is a positive integer then $\Gamma(\alpha) = (\alpha-1)!$

2. Market Wage Equation

The market wage equation may be written as

$$W = e^{X\beta}U$$

To estimate the equation, the logarithm of the equation is taken yielding

$$\ln W = X\beta + u \qquad \text{where } U = e^u$$

Since in the case of the market wage equation u is non-spherical, let the error structure be described by

$$u \sim N(o, \sigma^2 \Sigma)$$

where $\Sigma = dg(\omega_1, \omega_2, \ldots, \omega_T)$ is a diagonal matrix.

The mean and variance of ln W are given by

$$E\,\{\ln W\} = X\beta$$

$$\operatorname{var}(\ln W) = \sigma^2\,\Sigma$$

Thus $$\ln W \sim N(X\beta,\ \sigma^2\,\Sigma)$$

However, if we are interested in W rather than ln W, then

$$W = e^{\ln W} = \exp\,\{X\beta + u\}$$

$$E\,\{W\} = E\,\{e^{\ln W}\} = \exp\,\{X\beta + \tfrac{1}{2}\,\sigma^2\,dg(\Sigma)\} \qquad (1)$$

Since the error structure is heteroskedastic a weighting procedure is used before estimating the equation.

Define a matrix $\Phi = dg(\omega_1^{-\frac{1}{2}},\ \omega_2^{-\frac{1}{2}},\ldots,\ \omega_T^{-\frac{1}{2}})$

such that $$\Sigma = \Phi^{-1}\,\Phi'^{-1}$$

Weight the market wage equation by Φ

$$\Phi \ln W = \Phi X\beta + \Phi u$$

Then, $$\Phi u \sim N(o,\ \sigma^2\,I_T)$$

Estimate the weighted market wage equation by OLS.

$$\hat{\beta} = (X'\Sigma^{-1}X)^{-1} X'\Sigma^{-1} \ln W$$

$$\hat{\sigma}^2 = \frac{1}{T-k} \sum_{t=1}^{T} (\ln W_t - X_t\hat{\beta})^2$$

Then $$\widehat{\ln W} = X\hat{\beta} = X\beta + X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}u$$

The mean and variance of $\widehat{\ln W}$ are given by

$$E\{\widehat{\ln W}\} = X\beta$$

$$\mathrm{var}\,(\widehat{\ln W}) = \sigma^2 X(X'\Sigma^{-1}X)^{-1}X'$$

Thus $$\widehat{\ln W} \sim N(X\beta,\ \sigma^2 X(X'\Sigma^{-1}X)^{-1}X')$$

Now, if we are interested in $\hat{W}$ rather than $\widehat{\ln W}$ we must transform back to the original form

$$\hat{W} = e^{\widehat{\ln W}} = \exp\{X\beta + X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}u\}$$

and $$E\{\hat{W}\} = E\{e^{\widehat{\ln W}}\} = \exp\{X\beta + \tfrac{1}{2}\sigma^2 dg(X[X'\Sigma^{-1}X]^{-1}X')\} \qquad (2)$$

Comparing (2) with (1), the bias of $\hat{W}$ as an estimator of the mean of W is evident.

To obtain a minimum variance, unbiased estimator of the mean wage the lemma given previously is invoked. Set $w=\hat{\sigma}^2$, $v=T-k$, and $c=\frac{1}{2}dg\ (\Sigma-X[X'\Sigma^{-1}X]^{-1}X')$.

Then $$F(w;\ v,c) = F(\hat{\sigma}^2;\ T-k,\ \tfrac{1}{2}\ dg(\Sigma-X[X'\Sigma^{-1}X]^{-1}X'))$$

$$= \sum_{j=0}^{\infty} \frac{f_j}{j!} \left(\tfrac{1}{2}\ \hat{\sigma}^2\ dg\ (\Sigma-X[X'\Sigma^{-1}X]^{-1}X')\right)^j$$

$$= F_E$$

and $$E\ \{F_E\} = \exp\ \{\tfrac{1}{2}\ \sigma^2\ dg(\Sigma-X[X'\Sigma^{-1}X]^{-1}X')\}$$

Now, set $\hat{W} = e^{\widehat{\ln W}}\ F_E$ and take the expected value

$$E\ \{\hat{W}\} = E\ \{e^{\widehat{\ln W}}\}\ E\ \{F_E\}$$

The expectation on the right hand side was decomposed into the expression shown because of the independence of $\widehat{\ln W}$ and F_E. Substituting for the terms on the right hand side results in

$$E\ \{\hat{W}\} = e^{X\beta+\frac{1}{2}\ \sigma^2\ dg\ \Sigma} = E\ \{W\}$$

This then is the desired minimum variance unbiased estimator of the mean of W.

3. Estimation Procedure

In this section the steps followed to obtain the unbiased estimator of the mean wage following the estimation of the market wage equation, as presented in the main text, are outlined.

I. From the last iteration in the estimation of the market wage equation obtain $(X'\Sigma^{-1}X)^{-1}$.

II. In addition save the calculated weights, ω_i, used to weight the equation on the last iteration, where

$$\omega_i = [1 + \rho^2(\phi_i\hat{\lambda}_i - \hat{\lambda}_i^2)]^{-1}$$

III. Finally, save the calculated M.S.E., $\sigma_{11} = \hat{\sigma}^2$

IV. Define the following matrices:

$$(X'\Sigma^{-1}X)^{-1} = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1k} \\ M_{21} & M_{22} & \cdots & M_{2k} \\ \vdots & \vdots & & \vdots \\ M_{k1} & M_{k2} & \cdots & M_{kk} \end{bmatrix} \; ; \; X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & & \vdots \\ X_{T1} & X_{T2} & \cdots & X_{Tk} \end{bmatrix}$$

For the t^{th} observation calculate

$$\tfrac{1}{2}\hat{\sigma}^2 \left[\omega_t^{-1} - \sum_{i=1}^{k}\sum_{j=1}^{k} x_{ti}\, x_{tj}\, m_{ij}\right],\ t = 1, 2, \ldots, T$$

These are the T diagonal elements of $\frac{1}{2}\hat{\sigma}^2\ (\Sigma - X[X'\Sigma^{-1}X]^{-1}X')$ required in the calculation of F_E.

V. Next, calculate $f_j/j!$. Note that these are independent of the observation and thus need be calculated only once.

In the example in the text $v = T - k = 7876$ and thus $\frac{1}{2}v = 3938$ is a positive integer. This allows us to write the gamma function in terms of a factorial.

$$\Gamma(\tfrac{1}{2}v + j) = (\tfrac{1}{2}v + j - 1)!$$

Then,

$$\frac{f_j}{j!} = \frac{(\frac{1}{2}v)^j\ (\frac{1}{2}v-1)!}{(\frac{1}{2}v+j-1)!\ j!}$$

Further $f_j/j!$ may be calculated recursively since

$$\frac{f_j}{j!} = \frac{(\frac{1}{2}v)}{(\frac{1}{2}v+j-1)} \cdot \frac{f_{j-1}}{(j-1)!},\ j = 1, 2, \ldots$$

Now, $\frac{f_0}{0!} = 1$ and calculate $\frac{f_j}{j!}$, $j = 1, 2, \ldots$

The question arises as to how fast the series converges and thus how many terms will be needed to closely approximate the infinite sum. The answer will depend upon the properties of both terms in F_E, the one calculated in step IV as well as in step V.

For the term $f_j/j!$, with the parameters given above, the speed of convergence may be seen in the following table:

j	$f_j/j!$	j	$f_j/j!$	j	$f_j/j!$
0	1.0	7	1.9736×10^{-4}	14	1.1209×10^{-11}
1	1.0	8	2.4626×10^{-5}	15	7.4462×10^{-13}
2	4.9987×10^{-1}	9	2.7307×10^{-6}	16	4.6362×10^{-14}
3	1.6654×10^{-1}	10	2.7244×10^{-7}	17	2.7162×10^{-15}
4	4.1603×10^{-2}	11	2.4705×10^{-8}	18	1.5025×10^{-16}
5	8.3122×10^{-3}	12	2.0530×10^{-9}	19	7.8718×10^{-18}
6	1.3836×10^{-3}	13	1.5744×10^{-10}	20	3.9170×10^{-19}

VI. For the t^{th} observation calculate

$$(F_E)_t = \sum_{j=0}^{20} \frac{f_j}{j!} \left[\tfrac{1}{2} \hat{\sigma}^2 \left(\omega_t^{-1} - \sum_{i=1}^{k} \sum_{l=1}^{k} x_{ti} x_{tl} m_{il}\right)\right]^j$$

and $$\hat{W}_t = e^{\widehat{\ln W_t}} (F_E)_t \qquad t = 1, 2, \ldots, T$$

In the case of the market wage, since the term calculated in step IV was also less than unity, convergence occurred in about seven or eight summation terms.

This, then, was the unbiased estimate of the mean wage used in conjunction with the number of weeks to generate annual earnings.

APPENDIX B

CALCULATION OF INCOME TAXES AND THE NUMBER OF DEPENDENT CHILDREN

The tax calculations followed the detailed 1981 income tax return as closely as possible. Income items and the division of income into that subject to tax and non-taxable components was done using SCF data. Reported earnings were replaced by earnings calculated using the estimated labour supply and estimated wage. Family Allowances were allocated to the parent with the highest net income (excluding Family Allowances) based on calculations for both husband and wife. The net income of the Family Allowance recipient was then recalculated.

The following table, from a special tabulation of taxation data, was used to obtain deductions from total income to arrive at net income. In the case of the husband, since only broad age groups were used, simple averages of the appropriate columns were used.

Personal exemptions were calculated using SCF data and the tax calculations to arrive at net income. Exemptions included the basic personal exemption, the married exemption for the husband or wife where applicable, the age exemption in the case of the husband only since our sample excluded older married women, and the dependent child exemption.

Average Deductions by Income Subject to Tax, Age, and Sex, Canada, 1980

Income Subject to Tax	Under 20	20-24	25-34	35-44	45-54	55-64	65-69	70 and over
					Males			
Under $4,000	121	259	141	100	59	38	17	113
$ 4,000 - $ 7,999	250	350	276	237	211	151	8	8
8,000 - 10,999	380	467	459	451	436	357	75	19
11,000 - 13,999	529	713	747	678	724	663	208	55
14,000 - 16,999	683	929	975	962	1,009	1,065	393	77
17,000 - 19,999	777	1,142	1,243	1,233	1,295	1,420	726	148
20,000 - 22,999	797	1,274	1,501	1,446	1,526	1,865	1,027	237
23,000 - 25,999	943	1,398	1,704	1,725	1,928	2,288	1,245	406
26,000 - 29,999	1,644	1,619	1,996	2,114	2,361	2,650	1,903	422
30,000 - 34,999	1,343	1,976	2,377	2,560	2,885	3,338	2,344	958
35,000 - 39,999	1,197	2,331	3,057	3,143	3,389	4,008	3,394	935
40,000 - 99,999	7,149	5,432	6,024	6,016	5,923	6,293	6,097	1,906
100,000 - 199,999	4,789	16,910	13,713	12,794	12,267	12,705	11,917	6,075
200,000 and over	14,862	77,523	44,859	62,029	55,485	46,592	45,904	30,566
					Females			
Under $4,000	131	155	69	42	26	22	18	9
$ 4,000 - $ 7,999	274	335	343	225	185	126	16	9
8,000 - 10,999	433	553	656	484	496	417	82	12
11,000 - 13,999	693	862	1,042	845	891	869	181	35
14,000 - 16,999	938	1,214	1,394	1,173	1,354	1,251	292	50
17,000 - 19,999	757	1,540	1,786	1,551	1,701	1,676	469	87
20,000 - 22,999	792	2,001	2,229	2,104	2,219	2,026	895	67
23,000 - 25,999	598	2,005	2,616	2,537	2,710	2,578	780	71
26,000 - 29,999	1,277	3,186	3,115	3,099	3,183	2,855	1,170	264
30,000 - 34,999	1,240	4,289	3,498	3,652	3,694	3,162	1,509	183
35,000 - 39,999	451	2,213	3,780	3,462	4,201	4,039	2,111	329
40,000 - 49,999	420	2,613	4,556	4,079	3,943	3,821	1,947	632
50,000 - 99,999	1,978	3,553	5,510	5,814	5,335	4,660	3,066	1,341
100,000 - 199,999	3,942	9,197	10,637	12,167	9,235	7,593	6,127	2,760
200,000 and over	5,701	42,834	38,169	45,298	32,890	33,646	29,492	14,019

Since the number of children in each age group in the family beyond two were reported in one size category, Family Allowance data were used to calculate the exact number, both for use in calculating the exemption and for direct use in the estimated equations. The formulae used for this purpose were as follows, depending upon province and whether the children were under 6 or 6 to 17 years of age.

1 Quebec

a) Children under 6 > 3 and Children 6 to 17 = 0

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 1303.32}{832.32}\right) + 3$$

b) Children under 6 > 3 and Children 6 to 17 = 1

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 1596.78}{832.32}\right) + 3$$

c) Children under 6 > 3 and Children 6 to 17 = 2

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 2005.20}{832.32}\right) + 4$$

d) Children under 6 = 0 and Children 6 to 17 > 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 1411.14}{868.26}\right) + 3$$

e) Children under 6 = 1 and Children 6 to 17 > 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 1668.66}{868.26}\right) + 3$$

f) Children under 6 = 2 and Children 6 to 17 > 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 2041.14}{868.26}\right) + 4$$

g) Children under 6 > 3 and Children 6 to 17 > 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 4015.92}{868.26}\right) + 6$$

h) In all other cases for Quebec the actual number of children is given directly by the SCF data.

2 <u>Alberta</u>

a) Children under 6 > 3 and Children 6 to 17 = 0

$$\text{Number of Children} = \frac{\text{Family Allowance}}{218.40}$$

b) Children under 6 > 3 and Children 6 to 17 = 1

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 328.15}{218.40}\right) + 1$$

c) Children under 6 > 3 and Children 6 to 17 = 2

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 656.29}{218.40}\right) + 2$$

d) Children under 6 = 0 and Children 6 to 17 $\geq$ 3

$$\text{Number of Children} = \frac{\text{Family Allowance}}{328.20}$$

e) Children under 6 = 1 and Children 6 to 17 $\geq$ 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 218.40}{328.20}\right) + 1$$

f) Children under 6 = 2 and Children 6 to 17 $\geq$ 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 436.80}{328.20}\right) + 2$$

g) Children under 6 $\geq$ 3 and Children 6 to 17 $\geq$ 3

$$\text{Number of Children} = \frac{\text{Family Allowance}}{287.52}$$

h) In all other cases for Alberta the actual number of children is given directly by the SCF data.

3 All other provinces

a) Children under 6 $\geq$ 3 and Children 6 to 17 $\leq$ 2

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 287.52 \times \text{Children 6 to 17}}{287.52}\right)$$

$$+ \text{ Children 6 to 17}$$

b) Children under 6 $<$ 2 and Children 6 to 17 $>$ 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance} - 287.52 \times \text{Children under 6}}{287.52}\right)$$

$$+ \text{ Children under 6}$$

c) Children under 6 $>$ 3 and Children 6 to 17 $>$ 3

$$\text{Number of Children} = \left(\frac{\text{Family Allowance}}{287.52}\right)$$

d) In all other case the actual number of children is given directly by the SCF data.

In some cases, particularly for Quebec, the final form of the formula was arrived at after a number of simulations. The exemption for dependent children was the calculated and assigned to the recipient of the Family Allowance for tax purposes.

The deduction for medical expenses and charitable donations was calculated using the following table.

Average Medical Expenses and Charitable Donations by Income Subject to Tax, Canada, 1980

Income Subject to Tax	Deduction
Under $4,000	101
$ 4,000 - $ 7,999	123
8,000 - 11,999	151
12,000 - 15,999	168
16,000 - 19,999	183
20,000 - 22,999	194
23,000 - 25,999	210
26,000 - 29,999	238
30,000 - 34,999	292
35,000 - 39,999	358
40,000 - 49,999	417
50,000 - 99,999	730
100,000 - 199,999	1,715
200,000 and over	5,872

For those whose main activity in 1981 when neither working nor looking for work was going to school, an education deduction was calculated using the following formula.

$$\text{education deduction} = \$50 \times 0.23 \times (52 - \text{weeks worked} - \text{weeks looking})$$

Finally, other deductions from net income were calculated using the following table.

Average Deductions from Net Income by Income Subject to Tax, Age, and Sex, Canada, 1980

Income Subject to Tax	Under 20	20-24	25-34	35-44	45-54	55-64	65-69	70 and over
					Males			
Under $4,000	47	77	86	143	214	308	285	462
$ 4,000 - $ 7,999	70	99	130	215	341	724	896	1,293
8,000 - 10,999	92	118	178	284	443	942	1,918	2,121
11,000 - 13,999	80	121	211	321	456	945	2,054	2,459
14,000 - 16,999	98	148	210	332	480	869	2,169	2,241
17,000 - 19,999	118	175	241	371	497	962	2,113	2,262
20,000 - 22,999	152	204	297	412	530	951	2,174	2,302
23,000 - 25,999	313	249	366	483	650	1,031	2,018	2,382
26,000 - 29,999	348	319	436	584	770	1,179	2,204	2,429
30,000 - 34,999	324	447	562	699	887	1,344	2,341	2,185
35,000 - 39,999	442	680	746	849	1,053	1,544	2,420	2,493
40,000 - 49,999	746	1,007	1,007	1,152	1,322	1,835	2,636	3,130
50,000 - 99,999	5,854	2,368	1,991	2,296	2,473	2,966	3,602	3,130
100,000 - 199,999	4,831	11,117	8,102	8,081	7,961	8,666	7,871	6,861
200,000 and over	14,501	74,810	37,508	43,227	50,444	42,161	40,050	32,988
					Females			
Under $4,000	36	49	103	151	254	398	289	314
$ 4,000 - $ 7,999	87	96	148	232	439	833	965	1,079
8,000 - 10,999	114	125	193	273	470	819	1,460	1,404
11,000 - 13,999	147	173	216	320	523	862	1,580	1,587
14,000 - 16,999	132	230	266	396	590	985	1,611	1,634
17,000 - 19,999	336	281	325	492	739	1,057	1,561	1,633
20,000 - 22,999	308	342	356	531	807	1,223	1,660	1,662
23,000 - 25,999	523	419	458	666	969	1,327	1,668	1,599
26,000 - 29,999	1,089	1,212	528	806	1,163	1,511	1,750	1,727
30,000 - 34,999	1,747	895	730	869	1,363	1,681	1,830	1,627
35,000 - 39,999	506	713	937	1,007	1,637	2,264	1,786	1,848
40,000 - 49,999	766	1,495	1,096	1,185	1,561	1,748	1,854	2,049
50,000 - 99,999	1,152	2,673	2,338	2,769	2,764	3,295	2,740	2,616
100,000 - 199,999	3,455	7,776	9,125	8,365	7,206	6,708	5,388	4,029
200,000 and over	6,076	41,788	37,031	44,179	31,255	32,354	19,689	15,408

Taxable income was calculated by subtracting all personal exemptions, the medical expenses and charitable donations deduction, the education deduction and other deductions from net income.

Federal taxes payable were calculated adjusting for any provincial abatements and the federal tax reduction. Provincial taxes were calculated adjusting for provincial tax reductions and surtaxes. In the case of Quebec, detailed calculations were performed following the Quebec tax return using the above information.

The Child Tax Credit was calculated using both husband's and wife's net income, and used to reduce the wife's tax payable.

The marginal tax rate was calculated by incrementing earnings by $500 and recalculating total taxes payable following the procedure given above. The size of the increment was dictated by the occasional reversals found in the tables of deductions from total income and deductions from net income. Too small an increment near a reversal could lead to very high tax rates.

Aside from the information given in the tables above, the tax calculations relied only on data available in the SCF and information and rules contained in the federal (for each province) and Quebec tax returns.

List of Tables

REFERENCES

Amemiya, Takeshi (1981), "Qualitative Response Models: A Survey," Journal of Economic Literature, Vol. XIX, pp. 1483-1536.

Browning, Edgar K. and William R. Johnson (1984), "The Trade-Off between Equality and Efficiency," Journal of Political Economy, Vol. 92, No. 2, pp. 175-203.

DuMouchel, William H. and Greg J. Duncan (1983), "Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples," Journal of the American Statistical Association, Vol. 78, No. 383, pp. 535-543.

Goldberger, Arthur S. (1968), "The Interpretation and Estimation of Cobb-Douglas Functions," Econometrica, Vol. 35, No. 3-4, pp. 464-472.

Hausman, J.A. (1983), "Taxes and Labour Supply," Working Paper No. 1102, National Bureau of Economic Research, Cambridge.

______, (1981), "Labour Supply," in H.J. Aaron and J.A. Pechman (eds.) How Taxes Affect Economic Behaviour, The Brookings Institution, Washington, pp. 27-72.

Heckman, J.J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for such Models," Annals of Economic and Social Measurement, Vol. 5, No. 4, pp. 475-492.

______, (1979), "Sample Selection Bias as a Specification Error," Econometrica, Vol. 47, No. 1, pp. 153-61.

Killingsworth, M.R. (1983), Labour Supply, Cambridge University Press, Cambridge.

Nakamura, Alice and Masao Nakamura (1983), "Part-time and full-time work behaviour of married women: a model with a doubly truncated dependent variable," Canadian Journal of Economics, Vol. XLI, No. 2, pp. 229-57.

______, (1981), "A Comparison of the Labor Force Behavior of Married Women in the United States and Canada, with Special Attention to the Impact of Income Taxes," Econometrica, Vol. 49, No. 2, pp. 451-89.

Prescott, David, Robert Swidinsky and David Wilton (1986), "Labour supply estimates for low-income female heads of household using Mincome data," Canadian Journal of Economics, Vol. XIX, No. 1, pp. 134-41.

Rea, Samuel A. Jr. (1983), "The Impact of Taxes and Transfers on Labour Supply: A Review of the Evidence," a report prepared for the Ontario Economic Council, mimeo, 45 pp.

Robinson, Chris and Nigel Tomes (1985), "More on the labour supply of Canadian women," Canadian Journal of Economics, Vol. XVIII, No. 1, pp. 156-63.

Rosen, Harvey S. (1976), "Taxes in a Labor Supply Model with Joint Wage-Hours Determination," Econometrica, Vol. 44, No. 3, pp. 485-507.

Smith, James P. (ed.) (1980), Female Labour Supply: Theory and Estimation, Princeton University Press, Princeton.

Statistics Canada, Micro Data File 1982 Survey of Consumer Finances, Individuals Age 15 and Over, With and Without Income 1981; Household Surveys Division, Ottawa.

Stelcner, Morton and Jon Breslaw (1985), "Income Taxes and the Labour Supply of Married Women in Quebec," Southern Economic Journal, Vol. 51, No. 4, pp. 1053-72.

______ and J. Barry Smith (1985), "Labour Supply of Married Women in Canada: Non-convex Budget Constraints and the CES Utility Function," Working Paper No. 1985-9, Concordia University Working Papers, Concordia University, Montreal.