I. **Labour Supply**

1. *Neo-classical Labour Supply*

   1. Basic Trends and Stylized Facts

   2. Static Model
      
      * Decision of whether to work or not: Extensive Margin
      * Decision of how many hours to work: Intensive margin

   3. Comparative Statics

   4. Estimation of Labour Supply Functions and Elasticities
      
      * Structural and Reduced Form
      * Issues and Proposed Solutions
1.1 Basic Trends and Stylized Facts

Statistics Canada, June 2013

E
Employed (working, at work or not)
17.75m

U
Unemployed (Not employed, but looking for work)
1.36m

NLF
Not in Labor Force (students, retired persons, household workers, etc.)
9.56m

LF
Labor Force (working or actively seeking work)
19.1m

P
Population Aged 15 and Older
28.66m (10 provinces)
Labour Force Concepts:

- Labour Force = Employed + Unemployed
  - LF = E + U
  - Size of LF does not tell us about “intensity” of work

- Labour Force Participation Rate
  - LFPR = LF/P
  - P = civilian adult population 16 years or older not in institutions

- Employment: Population Ratio (percent of population that is employed)
  - EPR = E/P
  - Employed at work and not at work (e.g. maternity or sick leave) sometimes distinguished

- Unemployment Rate
  - UR = U/LF
Labour Force Participation Rate 25 to 44 Year Olds

Source: Statistics Canada, LFS and U.S. BLS, March CPS
Figure 6
Average Weekly Hours by Education Level in Canada

A. Men Working FTFY
B. Women Working FTFY

1.2. Static Model

- In neo-classical theory, the individuals’ decisions of whether or not participate in the labour market and of how many hours to work each week (and weeks per year) are modeled in static framework of consumption-leisure choice.

- From a policy point of view, this model has been very important to evaluate the potentially negative effects on labour supply of tax and transfer programs.

- From a labour econometrics viewpoint, the analysis will provide us with a classic example of correction for selection biases.
• The estimation of “the” **elasticity of labour supply** has long been an important quest for labour econometricians bearing in mind that differences across studies in labour supply estimates may come not only from differences in sampling or data differences but also in the underlying modeling assumptions.

• The more modern approaches have emphasized clearly sources of identification coming from natural and quasi-natural experiments, as well as field experiments.

  - An **identification strategy** describes the manner in which a researcher uses observational data to approximate a real experiment, i.e. a randomized trial.
The standard static, within-period labour supply model is an application of the consumer’s utility maximization problem over consumption and leisure.

Assume that each individual has a quasi-concave utility function:

\[ U(C, L, X) \]  

where \( C, L, \) and \( X \) are within-period consumption, leisure hours and individual attributes.

Then utility is assumed to be maximized subject to the budget constraint

\[ p \cdot C + w \cdot L = Y + w \cdot T \]  

where \( w \) is the hourly wage rate, \( Y \) is the non-labour income, and \( T = H + L \) is the total time \( T \) available, where \( H \) is the number of hours of work.

- \( M = Y + w \cdot T \) is sometimes called full-income.
- \( T, Y, w \) are exogenous in this model,
- \( H(L), C \) are endogenous, and \( X \) are preference-shifters.

The consumer may choose his/her hours of work \( H(L) \) by selecting across employers offering different packages of hours of work and wages.
FIGURE 2.6 The Consumer’s Labour Supply Decision

(a) Equilibrium of a nonparticipant

Income

Slope = $-W_0$

$Y_N$

O

T

Leisure (Nonmarket activity)

(b) Equilibrium of a participant

Income

Slope = $-W_0$

$W_0 h_0 + Y_N$

$Y_N$

O

$L_0$

T

Leisure (Nonmarket activity)
• Set up the Lagrangean expression:
\[ L(C, L, \lambda; p, w, X) = U(C, L, X) - \lambda[p \cdot C - w(T - L) - Y] \]

• Assuming an interior optimum, the first order conditions are
\[ L_c = U_c(C, L) - \lambda p = 0 \]
\[ L_L = U_L(C, L) - \lambda w = 0 \]
\[ L_\lambda = -p \cdot C + w(T - L) + Y = 0 \]

• In the case of an interior solution, the individual choose to participate in the labour market \( L^* < T \), the first-order conditions equates the marginal rate of substitution (\( MRS_{CL} \)) to the real wage rate
\[
\frac{U_L(C, L, X)}{U_C(C, L, X)} \bigg|_{L^*} = \frac{w}{p}
\]  

(3)
• It is important to distinguish the characteristics of the interior solution (intensive margin), for hours of work, $H > 0, (L < T)$ from the corner solution (extensive margin), $H = 0 (L = T)$.

• In the case of the corner solution, $L^* = T$,

$$\frac{w}{p} \leq \frac{w_R}{p} = \frac{U_L(C, L, X)}{U_C(C, L, X)} \bigg|_{L=0}$$ (4)

where the reservation wage, $w_R$, is equal to the negative of $MRS_{CL}$ of working hours for commodities at $h = 0 (L = T)$.

• Solving the FOC (3) or (4) yield the Marshallian demand functions

$$C = C(w, Y, X) \text{ and } L = L(w, Y, X)$$

or equivalently $H = H(w, Y, X)$ (5)
arises only among persons who are already allocating some time to the labor market. As we have seen, a higher wage increases the probability that the market wage exceeds the reservation wage, and hence makes a person more likely to enter the labor force. Increases in the wage, therefore, unambiguously increase a group's labor force participation rate, but hours of work for a particular worker may increase or decrease depending on whether substitution or income effects dominate.

Figure 2-13 uses data drawn from the 1991 Current Population Survey (CPS) to illustrate the labor supply curve actually observed among working men and women in
FIGURE 2.7 The Effect of an Increase in Nonlabour Income on Labour Supply

(a) Leisure a normal good

Income

(b) Leisure an inferior good

Income

Leisure
(Nonmarket activity)

Source: Benjamin, Gunderson and Riddell (1998)

FIGURE 2.8 Income and Substitution Effects of a Wage Increase

Income

Slope = \(-W_1\)

Income

Slope = \(-W_1\)

Income

Slope = \(W_0\)

Substitution effect
Income effect
Net effect
Leisure
(Nonmarket activity)
1.3 Comparative Statics

- The **comparative statics** of the impact of changes in income and wage rate of the labour supply functions $H = H(w,Y,X)$ are best illustrated in a diagram of consumption-leisure choice.

- The general effects are the following.

- **An increase in non-labour income**: An increase in non-labour income will shift the budget line outwards without changing the slope of the line: this is a pure income effect. The effect on the optimal amount of leisure consumed or hours worked can then be summarized as:
  - L will rise and H will fall if leisure is a normal good
  - L will fall and H will rise if leisure is an inferior good.

  - There are very strong reasons to believe that leisure is a normal good, e.g. those who win the lottery (a large increase in non-labour income) are more likely to work less afterwards than before and certainly not the other way round. Hence, it is likely that an increase in non-labour income will reduce hours of work.
• **An increase in the real hourly wage:** An increase in the real hourly wage will pivot the budget line about the point where L=T making the line steeper: here there are two effects:
  - An *income effect.* Individuals are better-off than before so there is a positive income effect that, because leisure is a normal good, makes individuals work fewer hours than before.
  - A *substitution effect.* An hour of work now buys more consumption than previously so that there is an incentive to increase consumption and reduce leisure. Hours of work will rise as a result.

  ▪ Hence, the impact of a change in the wage on hours of work is theoretically ambiguous. They may rise or fall.

  ▪ There is one exception to this: for non-participants there is no income effect as they have no labour income so nobody can be induced to reduce hours of work to zero as a result of an increase in the wage.
How can we quantify these effects?

- Recall that the Hicksian labour supply function is the solution to the expenditure minimization problem

\[ E(w, p, \bar{U}) = \min(p \cdot C - wH) \text{ subject to } U(C, H) \geq \bar{U} \]

and correspond to the following uncompensated labour supply function

\[ H(w, p, \bar{Y}) = H^c(w, p, \bar{U}) \text{ where } \bar{Y} = E(w, p, \bar{U}) \]

- Differentiating with respect to \( w \)

\[ \frac{\partial H}{\partial w} + \frac{\partial H}{\partial Y} \frac{\partial E}{\partial w} = \frac{\partial H^c}{\partial w} \bigg|_{\bar{u}} \]
• With the application of Sheppard Lemma and because $H$ is a factor (reverses the sign), we get the Slutsky equation

$$\frac{\partial H}{\partial w} = \frac{\partial H^C}{\partial w} \left( \frac{w}{H} \right) + H \cdot \frac{\partial H}{\partial Y}$$

(6)

substitution effect

income effect

where the overall effect of a wage change is decomposed into a substitution effect plus an income effect.

• Multiplying the entire equation (6) by $\frac{w}{H}$ and the last term (income effect) by $\frac{Y}{Y}$

$$\frac{\partial H}{\partial w} \cdot \frac{w}{H} = \frac{\partial H^C}{\partial w} \left( \frac{w}{H} \right) + \frac{w}{H} \cdot \frac{w}{Y} \cdot \frac{\partial H}{\partial Y} \cdot \frac{Y}{H}$$

or in terms of elasticities

$$\varepsilon_{Hw} = \varepsilon_{Hw}^C + s_L \cdot \eta_{HY}$$

(7)
• Thus, there are three “sufficient statistics” of labour supply
  o the uncompensated wage elasticity: the % change in labour supply resulting from 1% change in the wage rate; sign is theoretically ambiguous as the positive substitution effect can sometimes be dominated by the negative income effect
    \[ \varepsilon_{Hw} = \frac{\partial H}{\partial w} \cdot \frac{w}{H} > 0(\approx 0)? \]
  o the compensated wage elasticity: the % change in labour supply resulting from 1% change in the wage rate, after compensation for the wage change; sign is positive as it reflects a pure substitution effect
    \[ \varepsilon_{Hw}^C = \frac{\partial H^C}{\partial w} \cdot \frac{w}{H} > 0 \]
  o the income elasticity: the % change in labour supply resulting from 1% change in non-labor income; sign is expected to be negative
    \[ \eta_{HY} = \frac{\partial H}{\partial Y} \cdot \frac{Y}{H} < 0 \]
• The simple consumption-leisure model has been extended (altered) to analyze labour supply under various conditions:

  o introducing the fixed (money) cost of working or time cost (commuting) of working
  o moonlighting (2nd job) and overtime pay
  o should a firm offer flexible hours (part-time) or hire only full-time workers
  o family labour supply (actually more than a simple extension)
Effect of a Cash Grant on Work Incentives

A take-it-or-leave-it cash grant of $500 per week moves the worker from point P to point G, and encourages the worker to leave the labor force.

Figure 2-15

Effect of a Welfare Program on Hours of Work

A welfare program that gives the worker a cash grant of $500 and imposes a 50 percent tax on labor earnings reduces work incentives. In the absence of welfare, the worker is at point P. The income effect resulting from the program moves the worker to point Q; the substitution effect moves the worker to point R. Both income and substitution effects reduce hours of work.
Figure 1.7. Labor supply with money costs of working

Figure 1.8. Labor supply with time costs of working
1.4 Estimating Labour Supply Functions and Elasticities

a. Structural and Reduced-Form Estimation

- We can proceed by assuming that the individuals have a direct utility function of the form:

\[ U(C, L) = C^\alpha L^\beta, \]

- The FOC will become \( \frac{\beta / L}{\alpha / C} = \frac{w}{p} \). Combining that equation with the budget constraint and using the fact that \( H^* = 1 - L^* \), we obtain

\[ H^* = 1 - \gamma - \gamma (Y / w) \]

\[ C^* = (1 - \gamma)[(w + Y) / p], \text{ where } \gamma \equiv \beta / (\alpha + \beta). \]

- For example, see Abbott and Ashenfelter (1976) for the results of the estimation of a Stone-Geary utility function.
A. A Direct Utility Function

We have chosen an augmented Stone-Geary [37] utility function,

\[ u = B_0 \ln (1 - \gamma_l) + \Sigma B_i \ln (x_i - \gamma_i), \quad (i = 1, \ldots, n), \] ... (12)

with \( \Sigma B_i = 1 \quad (i = 0, \ldots, n) \) and \( (x_i - \gamma_i) > 0 \quad (i = 1, \ldots, n) \), \( (l - \gamma_l) > 0 \), as one method for deducing a complete set of commodity demand and labour supply functions. This choice is not entirely arbitrary. It is based on the practical fact that (12) is the only utility function consistent with linear expenditure and labour earnings functions, and that the resulting commodity demand system has seen very extensive practical use. For this case the first-order conditions are

\[ B_i/(x_i - \gamma_i) = \lambda p_i, \quad (i = 1, \ldots, n) \] ... (13)

\[ B_l/(1 - \gamma_l) = \lambda w, \]

and the budget constraint (2). Summing the \((n + 1)\) equations in (13) gives the solution for \( \lambda \) as

\[ \lambda = (wT + y - \Sigma \gamma_i p_i - \gamma_l w)^{-1} \quad (i = 1, \ldots, n). \] ... (14)

Substitution for \( \lambda \) from (14) into the first-order conditions (13) then gives the expenditure functions

\[ p_i x_i = \gamma_i p_i + B_i (wT + y - \Sigma \gamma_i p_i - \gamma_l w), \quad (i = 1, \ldots, n) \] ... (15)

\[ w l = \gamma_l w + B_0 (wT + y - \Sigma \gamma_i p_i - \gamma_l w). \]

\[ \ldots \]

\[ T \text{ need never enter the picture if we first recognize that } T - \gamma_l = \gamma_h \text{ is maximum feasible working hours. Substituting } l = T - h \text{ and } \gamma_l = T - \gamma_h \text{ into (15) then gives} \]

\[ p_i x_i = \gamma_i p_i + B_i (y + \gamma_h w - \Sigma \gamma_i p_i), \quad (i = 1, \ldots, n) \] ... (16a)

\[ -wh = -\gamma_h w + B_0 (y + \gamma_h w - \Sigma \gamma_i p_i), \] ... (16b)

which, after the addition of error terms, are susceptible of direct estimation with the data
### TABLE IV

*Maximum likelihood estimates of (16), the augmented linear expenditure system*

<table>
<thead>
<tr>
<th>Commodity group</th>
<th>Estimates (and estimated asymptotic standard errors) of $\gamma_1 \text{ and } \gamma_k$</th>
<th>$B_i$</th>
<th>$R^2 \ast$</th>
<th>$DW \ast$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durables</td>
<td>$-0.922 \quad (0.481)$</td>
<td>0.238 (0.0325)</td>
<td>0.579</td>
<td>2.34</td>
</tr>
<tr>
<td>Food</td>
<td>0.699 \quad (0.128)</td>
<td>0.163 (0.0193)</td>
<td>0.866</td>
<td>1.37</td>
</tr>
<tr>
<td>Clothing</td>
<td>0.466 \quad (0.068)</td>
<td>0.134 (0.0151)</td>
<td>0.870</td>
<td>1.11</td>
</tr>
<tr>
<td>Other non-durables</td>
<td>0.651 \quad (0.091)</td>
<td>0.0254 (0.00207)</td>
<td>0.913</td>
<td>1.99</td>
</tr>
<tr>
<td>Housing services</td>
<td>0.763 \quad (0.057)</td>
<td>0.0755 (0.00512)</td>
<td>0.946</td>
<td>2.45</td>
</tr>
<tr>
<td>Transportation services</td>
<td>0.510 \quad (0.084)</td>
<td>0.0997 (0.00704)</td>
<td>0.937</td>
<td>2.22</td>
</tr>
<tr>
<td>Other services</td>
<td>0.424 \quad (0.167)</td>
<td>0.142 (0.0190)</td>
<td>0.847</td>
<td>1.12</td>
</tr>
<tr>
<td>Labour supply</td>
<td>2357 \quad (128.7)</td>
<td>0.121 (0.0368)</td>
<td>0.970</td>
<td>1.64</td>
</tr>
</tbody>
</table>

$\ast$ The $R^2$ and Durbin-Watson (DW) statistics are reported for the data in first difference form, and are computed as if each equation were separate.

To get more insight into the implications of the estimates of (16) it is useful to compute the price and wage-rate elasticities of commodity demand and labour supply. From equations (16) it is easy to see that we can identify $p_i(x_1/\partial p_j) = B_i$ and $-w(\partial h/\partial y) = B_0$ as parameters in this model. The uncompensated and compensated price elasticities of commodity demand are, respectively,

\[
(p_j/x_i)(\partial x_i/\partial p_j) = -\delta_{ij} + (\delta_{ij} - B_i)[(\gamma_j p_j)/(x_i p_j)]
\]

and

\[
(p_j/x_i)S_{ij} = (B_i - \delta_{ij})\frac{[(x_j p_j - \gamma_j p_j)/(x_i p_i)]}{1}
\]

where the Kronecker $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The effects of a wage change on commodity demands are only slightly different:

\[
(w/x_i)(\partial x_i/\partial w) = B_i[(\gamma_h w)/(x_i p_i)]
\]

and

\[
(w/x_i)S_{io} = B_i[(\gamma_h w - hw)/(x_i p_i)]
\]

Finally, the uncompensated and compensated labour supply elasticities are:

\[
(w/h)(\partial h/\partial w) = -1 + (1 - B_0)[(\gamma_h w)/(hw)]
\]

and

\[
-(w/h)(S_{oo}) = (1 - B_0)[(\gamma_h w - hw)/(hw)]
\]

Table 5 contains values of some of the elasticities given by (17) evaluated at the sample means, and we comment on them below.
<table>
<thead>
<tr>
<th>Commodity group</th>
<th>Marginal propensity to consume unearned income</th>
<th>Uncompensated own price elasticity</th>
<th>Compensated own price elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rotterdam Separable Rotterdam Linear Addilog</td>
<td>Rotterdam Separate Rotterdam Linear Addilog</td>
<td>Rotterdam Separate Rotterdam Linear Addilog</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durables</td>
<td>0.278  0.292  0.238  0.342</td>
<td>-1.067  -1.255  -0.955</td>
<td>-0.789  -1.290  -0.612</td>
</tr>
<tr>
<td>Food</td>
<td>0.206  0.208  0.163  0.427</td>
<td>-0.378  -0.468  -0.605</td>
<td>-0.172  -0.261  -0.442</td>
</tr>
<tr>
<td>Clothing</td>
<td>0.053  0.060  0.134  0.156</td>
<td>-0.529  -0.270  -0.581</td>
<td>-0.476  -0.211  -0.447</td>
</tr>
<tr>
<td>Other non-durables</td>
<td>0.097  0.096  0.025  0.210</td>
<td>-0.399  -0.496  -0.581</td>
<td>-0.302  -0.400  -0.556</td>
</tr>
<tr>
<td>Housing services</td>
<td>0.082  0.078  0.076  0.275</td>
<td>-0.047  -0.074  -0.518</td>
<td>0.035   0.003   0.442</td>
</tr>
<tr>
<td>Transportation services</td>
<td>0.024  0.025  0.098  0.051</td>
<td>-1.216  -1.439  -0.637</td>
<td>-1.192  -1.414  -0.537</td>
</tr>
<tr>
<td>Other services</td>
<td>0.088  0.091  0.142  0.241</td>
<td>-0.156  -0.414  -0.653</td>
<td>-0.068  -0.323  -0.511</td>
</tr>
<tr>
<td>Labour supply</td>
<td>-0.173  -0.152  -0.121  0.703</td>
<td>-0.143  -0.070  -0.084</td>
<td>0.030   0.081   0.037</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Commodity group</th>
<th>Uncompensated (cross) wage elasticity</th>
<th>Compensated (cross) wage elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rotterdam Separable Rotterdam Linear Addilog</td>
<td>Rotterdam Separate Rotterdam Linear Addilog</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durables</td>
<td>2.593  1.829  1.501  1.405</td>
<td>1.004  0.160  0.060  0.055</td>
</tr>
<tr>
<td>Food</td>
<td>0.374  0.558  0.479  1.405</td>
<td>-0.131  0.049  0.019  0.357</td>
</tr>
<tr>
<td>Clothing</td>
<td>0.685  0.431  0.540  1.405</td>
<td>0.335  0.038  0.022  0.378</td>
</tr>
<tr>
<td>Other nondurables</td>
<td>0.571  0.555  0.637  1.405</td>
<td>0.059  0.048  0.026  0.296</td>
</tr>
<tr>
<td>Housing services</td>
<td>0.077  0.305  0.590  1.405</td>
<td>-0.217  0.027  0.024  0.420</td>
</tr>
<tr>
<td>Transportation services</td>
<td>0.296  0.597  0.608  1.405</td>
<td>-0.218  0.052  0.025  0.304</td>
</tr>
<tr>
<td>Other services</td>
<td>0.0192  0.446  0.733  1.405</td>
<td>-0.376  0.039  0.030  0.323</td>
</tr>
<tr>
<td>Labour supply</td>
<td>-0.143  -0.070  -0.084  1.069</td>
<td>0.030  0.081  0.037  0.366</td>
</tr>
</tbody>
</table>
• Because of the identification problems above, many studies focus more directly on the wage elasticity of the Marshallian supply function and on the associated utility-constant Hicksian wage elasticity.

• Suppose that we have individual data on hours of work $H$, the wage rate $w$, and on non-labor income $Y$, we could estimate a simple OLS regression

$$H_i = \beta_0 + \beta_1 \frac{w_i}{p_i} + \beta_2 \frac{Y_i}{p_i} + \epsilon_i$$  \hspace{1cm} (8)

• Then the estimated effects,
  
  - $\hat{\beta}_1$ will be the overall (uncompensated) effect $\frac{\partial H}{\partial w}$,
  - $\hat{\beta}_2 \bar{H}$ will be the income effect (evaluated at mean hours),
  - $\hat{\beta}_1 - \hat{\beta}_2 \bar{H}$ will be the substitution effect (evaluated at mean hours)
  - $\hat{\beta}_2 \frac{\bar{Y}}{\bar{H}}$ will be the income elasticity of labour supply.
• We may normalize $p_i = 1$ and control for individual attributes
\[ H_i = \beta_0 + \beta_1 w_i + \beta_2 Y_i + \beta_3 X_i + \varepsilon_i \] (9)
and we usually assume that the distribution of the $\varepsilon_i$ would be a normal distribution.

• There have many studies estimating labour supply and income elasticities of labour supply, and there have been many meta-analysis of these studies (e.g. Hansson and Stuart (1985), Killingsworth (1983), Killingsworth and Heckman (1986), Pencavel (1986), and Evers, de Mooij and van Vuuren (2008).

• Evers, de Mooij and van Vuuren (2008) conclude that an uncompensated elasticity of 0.5 for women and 0.1 for men is a good reflection of what the literature reveals, although for the US it may be negative for men, due to the income effect.
  - For male workers, relatively small wage effects
  - For female workers, much larger elasticities with larger variations across studies and declining over time as women have become more attached to the labour market.
where \( e \) is the uncompensated wage elasticity, \( c \) is a constant and \( X \) is a matrix of moderator variables (see below). Study characteristics are assumed to affect the elasticity in a linear way, with slope parameters in the vector \( \beta \). The error term \( \eta \) is assumed to be asymptotically normally distributed, independently so across different observations.³ An OLS estimator with White heteroskedasticity-consistent standard errors will be used, as e.g. in De Mooij

³ It should noted that some elasticities are estimated from the same data sets, and therefore, error terms may be correlated. This argument is of limited importance for our main findings as the included empirical studies are based on many different data sources. Yet, it means that the reported standard errors may be somewhat underestimated.
• Identifying the “correct” wage elasticity is important not only to satisfy our understanding of labour supply model in the static partial equilibrium model, but it is used as an important component of general equilibrium models.

• It is also important in public finance: optimal tax rate depends inversely on compensated wage elasticity of labor supply.

• It is used in the design of income support schemes as an important concern of policy makers is that some social assistance programs have adverse work incentives, that is, some programs cause people to reduce their working hours or forego working altogether.

• Indeed, negative labour supply effects found in the earlier literature (Negative Income Tax – NIT) experiments have lead to policy makers to better design policies that supplement earnings (e.g. Earned Income Tax Credit – EITC).
b. Issues with the OLS Estimation of Labour Supply Functions

1) Econometric issues [potential solutions]
   a) Unobserved heterogeneity/endogeneity of wage rate [tax instruments]
   b) Measurement error in wages and division bias [tax and other instruments]
   c) Selection into labor force/unobserved wage rate for non-labour market participants [selection models]
   d) Endogenous tax rates [non-linear budget set methods]

2) Extensive vs. intensive margin responses
   • With fixed costs of work, individuals may jump from non-participation to part time or full time work, this requires discrete choice models of participation

3) Non-hours responses
   • Productivity responses beyond the simple labour supply model
   • Focus on subgroups of workers for whom hours are better measured or more flexible, e.g. taxi drivers, courriers
• Mroz (1987) is an earlier paper attempting to address some of the econometric issues with the estimation of female labour supply functions
  
  o Uses background variables as “credibly exogenous” instruments [Parents' education, age, education polynomials] in the selection equation
  o Tests validity of labor market experience, average hourly earnings, and previous reported wages as instruments for wages
    ▪ Rejects validity of all three
  o Shows that earlier estimates are highly fragile and unreliable

• The influential Mroz (1987) study contributed to the idea that some source of exogenous variation, such as tax and transfer reforms, was necessary to identify the parameters of interest.
1) Econometric issues
   a) Unobserved heterogeneity/endogeneity of wage rate

   • Recall the labour supply function
     \[ H_i = \beta_0 + \beta_1 w_i + \beta_2 Y_i + z_i \delta + \varepsilon_i \]  
     \[ (1) \]

   • Workers with higher wages may have higher labour force attachment, put more effort, more hours, etc.
   • More educated and more able workers get higher wages
     \[ \text{Corr}(w_i, \varepsilon_i) \neq 0 \] and \( \beta_1 \) will be biased

   • If \( w_i \) was the only the covariate, the bias would be
     \[ \lim_{p \to \infty} (\hat{\beta}_1^{OLS}) = \beta_1 + \frac{Cov(w_i, \varepsilon_i)}{Var(w_i)} \]

   • Controlling for X’s help, but it may not be sufficient to remove all omitted variables bias
b) Measurement error in wages and division bias

- For many workers, the wage rate is computed as earnings divided by hours, this generates a spurious negative correlation in hours, called the division bias (Borjas, 1980).

- Borjas (1980) considers a standard labour supply equation

\[
\ln H = \alpha + \beta \ln W + \gamma Z + \mu \quad \text{(B-1)}
\]

where only earnings (annual or weekly are observed) and usual hours of work (annual or weekly), so that \( W = E / H \).

- So that the observed equation is

\[
\ln H^* = \alpha + \beta \ln(W^*) + \gamma Z + \mu \quad \text{(B-2)}
\]

where \( W^* = E / H^* \).

- Spurious negative correlation between \( \ln(H^*) \) and \( \ln(W^*) \) will bias the elasticity estimate downward.
• More precisely, if there is measurement error in the hours of work,
\[ \ln H^* = \ln H + \varepsilon^* \iff \varepsilon^* = \ln H^* - \ln H, \]

Substituting in equation (1), it becomes

\[ \ln H^* = \alpha + \beta \ln\left(E / H^*\right) + \gamma Z + \mu + \varepsilon^* - \beta \varepsilon^* \]

• It can be shown that
\[ p \lim \hat{\beta} = \beta - \frac{\sigma^2_\varepsilon (1 + \beta)}{\sigma^2_w} \]

• Solution: when “usual hours last year” is the dependent variable, instrument average hourly earnings (earnings last year/usual hours last year) with alternative measure of average hourly earning (earnings last year/hours last week)
c) Selection into labor force/unobserved wage rate for non-labour market participants

- When there are some fixed costs of working, some individuals choose not to work
- Wages are unobserved for non-labor force participants
- Thus, OLS regression on workers only includes observations with $H_i > 0$
  - This can bias OLS estimates if LFP is non-random: low wage earners must have very high unobserved propensity to work to find it worthwhile
- Requires a selection (on observables) correction pioneered by Heckman in the 1970s (e.g. Heckit, Tobit, or ML estimation): problem is that identification is based on strong functional form assumptions, does not resolve issue of instrument
- Alternative current approach: use panel data to distinguish entry/exit from intensive-margin changes and approximate wage using past data
Figure 5
The Effect of an Increase in the Real Wage on the Budget Constraint

Source: Manning E333e

Figure 6
A scatterplot of desired hours against x

Source: Manning E333e
• One problem with the estimation (1) with OLS is that hours of work cannot be negative.

• If they could, the budget line would extend past the kink point at \( L = T \) to the horizontal axis and individuals choosing hours to maximize utility subject to this imaginary budget line.

• People who were choosing \( H = 0 \), would prefer a point on the imaginary part of the budget line with \( H < 0 \).

• Defining desired hours, \( H_i^* \), we will have a latent relationship for desired hours of work

\[
H_i^* = \beta_0 + \beta_1 w_i + \beta_2 Y_i + z_i \delta + \varepsilon_i
\]  

(2)

• We will have the following relationship between actual hours and desired hours:

\[
H_i = H_i^* \text{ if } H_i^* \geq 0
\]

\[
H_i = 0 \text{ if } H_i^* < 0
\]  

(2′)

with reference to the theoretical model, the latter individuals are the ones for which the offered wage is below their reservation wage (\( w_i < w_i^R \)).
The Tobit model is the standard way to deal with a model with censoring such (2)-(2'). Censoring occurs here because the dependent variables is simply recorded as being below a certain level (desired hours below zero are recorded as zeros.) More generally, censoring can also happen from above (top-coding of earnings).

It starts with the principle of using all information, i.e. computing the likelihood contributions of both participants and non-participants.

Letting $H_i^*$ satisfy the classical linear model assumptions of a normal, homoskedastic distribution with a linear conditional mean. Letting $x_i = [w_i, p_i, Y_i / p_i, z_i]$ be a vector of covariates with an associate vector of coefficients $\beta$.

The contribution to the likelihood function of a worker will be,

$$\Pr(H_i > 0 \mid x_i) = \Pr(H_i^* = H_i \mid x_i) = \Pr\left(\frac{\varepsilon_i}{\sigma} = \frac{H_i - x_i \beta}{\sigma} \mid x_i\right)$$

$$= 1/\sigma \phi\left(\frac{H_i - x_i \beta}{\sigma}\right),$$

where $\phi(\cdot)$ is the PDF of a N(0,1).
• The contribution to the likelihood function of a non-worker will be,

$$\Pr(H_i = 0 \mid x_i) = \Pr(H^*_i < 0 \mid x_i) = \Pr(\varepsilon_i < -x_i \beta)$$

$$= \Pr(\varepsilon_i / \sigma < -x_i \beta / \sigma) = \Phi(-x_i \beta / \sigma),$$

where $\Phi(\cdot)$ is the CDF of a N(0,1), since $\varepsilon_i / \sigma$ is N(0,1).

• So the likelihood looks like this

$$L = \prod_{\{i \mid H_i > 0\}} \frac{1}{\sigma} \phi\left( \frac{H_i - x_i \beta}{\sigma} \right) \prod_{\{i \mid H_i = 0\}} \Phi\left( \frac{-x_i \beta}{\sigma} \right)$$ (3)

• The marginal effects can be shown to be (see, for example, Wooldridge, chap. 17)

$$\frac{\partial E(H \mid x)}{\partial x_j} = \beta_j \Phi\left( \frac{x \beta}{\sigma} \right)$$

• The Tobit model (3) will yield consistent estimates of the parameter of the labour supply model, but it crucially depends on the normality and homoskedasticity assumptions in the underlying latent variable model.
• An alternative way of getting consistent estimates of the parameters of the labour supply model is due to Heckman (and arguably is the main reason why he got the Nobel Prize in 2002). This procedure is known as the Heckit.

• Given model in (2), when we assume normality of the residuals, we have

\[
E(H_i \mid H_i > 0) = x_i \beta + E(\varepsilon_i \mid \varepsilon_i > x_i \beta) = x_i \beta + \sigma E[\varepsilon_i / \sigma > -x_i \beta / \sigma]
\]

\[
= x_i \beta + \sigma \frac{\phi(x_i \beta / \sigma)}{\Phi(x_i \beta / \sigma)} = x_i \beta + \sigma \lambda(x_i \beta / \sigma),
\]

since \( E(c \mid c > 0) = \frac{\phi(c)}{1 - \Phi(c)} \), \( \phi(c) = \phi(-c) \), and \( 1 - \Phi(c) = \Phi(c) \),

where \( \Phi(\cdot) \) is the cdf of a N(0,1), since \( \varepsilon_i / \sigma \) is N(0,1), and where \( \lambda(c) = \frac{\phi(c)}{\Phi(c)} \) is called the inverse Mills ratio or the sample selection correction term.

• One can interpret (4) as saying that one could estimate the model by OLS if one could control for the sample selection correction. The problem of course is that the selection correction term depends on the parameters.
• So the procedure known at the Heckit is actually a two-step procedure.
  1) Estimate a Probit model for whether the dependent variable is observed \((H_i > 0)\) or not \((H_i = 0)\) to obtain consistent estimates of \(\beta/\sigma\). Use these estimates to compute an estimate of the inverse Mills ratio, \(\hat{\lambda}(x, \hat{\beta}/\hat{\sigma})\).

 2) Estimate by OLS the hours regression that include the correction term

\[
H_i = \beta_0 + \beta_1 w_i + \beta_2 Y_i + z_i \delta + \beta_3 \hat{\lambda} + \epsilon_i
\]

• When using the Heckit, it is better to use pre-programmed procedure, as in STATA, that computes the standard errors correctly (allowing for heteroscedasticity).

• One advantage of the Heckit over the Tobit is robustness. If the parameters from the participation equation are different from those of the hours regression, the Heckit will still have nice properties. But a sort of test of specification is whether the two models give similar answers.
• One potential problem with this form of the Heckit is that the inverse Mills ratio is simply of function, albeit non-linear, of the same $x_i$ used in the hours regression, which can be extremely collinear with the linear function of the $x_i$ in the rest of the equation. There is thus a feeling that adding excluded instruments would work better.

d) Endogenous tax rates

• Actual tax system is not linear but piece-wise linear with varying marginal tax rate due to (a) means-tested transfer programs, (b) progressive individual income tax, (c) ceiling in payroll tax

• Utility maximization problem becomes:
  \[
  \max \ U(w^p H - T(w^p H), H) \Rightarrow \text{FOC} \quad U_c(w^p (1 - T')) - U_H = 0
  \]
  where $w^p$ is the pre-tax wage, $T(w^p H)$ are taxes, and $w = w^p (1 - T')$. 

• Main complications: (a) $w$ become endogenous to choice of $H$, (b) FOC may not hold if individual bunches at a kink, (c) FOC may not characterize the optimum choice

• Non-linear budget set creates two problems:
  1) Model misspecification: OLS regression no longer recovers structural elasticity parameter of interest
  Two reasons: (a) underestimate response because people pile up at kink [see diagram exhibit] and (b) mis-estimate income effects
Marginal Tax Rates in Denmark in 1995

\[\Delta \log(\text{NTR}) = -29\%\]
\[\Delta \log(\text{NTR}) = -6\%\]
\[\Delta \log(\text{NTR}) = -9\%\]

Note: $1 \cong 6$ DKr

Source: Chetty et al. (2009)
Income Distribution for Wage Earners Around Top Kink (1994-2001)

Source: Chetty et al. (2009)
Single Men

Excess mass = 1.83%
Standard error = 0.34%

Source: Chetty et al. (2009)
Married Women

Excess mass = 14.1%
Standard error = 0.90%

Source: Chetty et al. (2009)
$$W_3 = (1-t_3) W$$

$$W_2 = (1-t_2) w$$

$$W_1 = (1-t_1) W$$

Source: Hausman (Hbk 1985)
2) Econometric bias: \( \tau_i \) depends on income and hence on \( H_i \). Tastes for work are positively correlated with \( \tau_i \) downward bias in OLS regression of hours worked on net-of-tax rates. 
Solution to problem #2: only use reform-based variation in tax rates.

- But problem #1 requires fundamentally different estimation method

- Issue addressed by non linear budget set studies pioneered by Hausman in late 1970s (Hausman, 1985 Public Economics Handbook Chapter)

- Method uses a structural model of labour supply where the likelihood of being of each segment of the non-linear budget set is computed.

- Key point: the method uses the standard cross-sectional variation in pre-tax wages for identification. Taxes are seen as a problem to deal with rather than an opportunity for identification.

- New literature identifying labor supply elasticities using tax changes has a totally different perspective: taxes are seen as an opportunity to identify labor supply
Additional readings:

