I. **Labour Supply**

1. Problems with the OLS Estimation of Labour Supply Functions
   1. Econometric Issues
   2. Extensive vs. intensive margin responses
   3. Non-hours responses

2. Using Tax and Transfer Programs to Estimate Labour Supply
   1. Field Experiments and Randomized Trials
   2. Tax (and Transfer) Reforms
Elasticity of labour supply: Responsiveness of hours of work to changes in the wage rate

\[ \varepsilon_{HW} = \frac{\partial H}{\partial w} \cdot \frac{w}{H} \approx \frac{\Delta H}{H} \cdot \frac{w}{\Delta w} = \frac{\Delta H}{\Delta w} \cdot \frac{w}{H} \]

That is, percent change in hours worked divided by the percent change in wage rate.

- If a worker’s wage is initially 20$ per hour and she works 1,900 hours per year. After she gets a raise to 24$ per hour, she decides to work 1,938 hours per year → 2%/20%=0.1

- When we run regressions where the outcome (e.g. hours) and the treatment (e.g. wages) are in logarithms, can the resulting estimates be interpreted as elasticities?

- Or what other conditions/assumptions are needed?
2.1 Problems 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 or non-wage responses
   - Productivity responses beyond the simple labour supply model
   - Workers may work longer hours to get promotion
   - Focus on subgroups of workers for whom hours are better measured or more flexible, e.g. taxi drivers, couriers
• 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 [replication exercise] 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
     \[ H_i = \beta_0 + \beta_1 w_i + \beta_2 Y_i + z_i \delta + \epsilon_i \]  
     (1)

   • Workers with higher wages may have higher labour force attachment; put more effort, more hours,

   • More educated and more able workers get higher wages
     \( \text{Corr}(w_i, \epsilon_i) \neq 0 \) and \( \beta_1 \) will be biased

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

   • 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

\[ \ln H = \alpha + \beta \ln W + \gamma Z + \mu \]  \hspace{1cm} (2)

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

\[ \ln H^* = \alpha + \beta \ln(W^*) + \gamma Z + \mu \]  \hspace{1cm} (3)

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

- Borjas showed that the spurious negative correlation between \( \ln(H^*) \) and \( \ln(W^*) \) biased the elasticity estimate downward.
• 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)

• More recently, Barrett and Hammermesh (2016) compare CPS reported usual hours of work with those recalled through diaries in the American Time Use Survey (ATUS) found that greater salience of recalled work time among higher-earning workers instead evidence of classical measurement error (mean zero).

• They argue that usual work hours create measurement errors that are positively correlated with the respondent’s wage rate which generates upward biases in the estimated elasticities of labor supply.

• Recent surveys ask respondent more directly about their mode of remuneration which makes these issues less important.
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
    [see diagram]

- 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
• Overview: Killingsworth and Heckman (1986) identify at least eight alternative procedures to estimate static labour supply models: these are classified as first or second-generation studies

• *First-generation studies* did not make this distinction between participants and non-participants.
  o One first approach, called Procedure I, was to estimate (1) over the sample of all individuals setting the hours of non-participants to zero
  o A second approach, called Procedure II, was to estimate models using only for those observations where desired hours were actually observed, i.e. the workers only.

• *Second generation studies* explicitly recognized that $H_i$ is a censored variable (negative values not observed), and that the decision to participate may be different from the decision involving the choice of hours of work. They explicitly modeled the self-selection decision.
  o Killingsworth and Heckman (1986) distinguish the procedure as using either the standard Tobit model for censored variable or the Heckit model for selection correction.
<table>
<thead>
<tr>
<th>Issues:</th>
<th>Desired hours of non-participant not observed</th>
<th>Wage of non-participant not observed</th>
<th>Endogeneity of wage rate/Selection bias</th>
<th>Estimation Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First-generation studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure I</td>
<td>Set to zero</td>
<td>Use imputed wages</td>
<td>Use predicted wage</td>
<td>OLS</td>
</tr>
<tr>
<td>Procedure II</td>
<td>Use only workers</td>
<td>Use only workers</td>
<td>Use predicted wage</td>
<td>OLS</td>
</tr>
<tr>
<td><strong>Second-generation studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure III</td>
<td>Use Tobit</td>
<td>Use imputed wages</td>
<td>Use predicted wage</td>
<td>Tobit on hours</td>
</tr>
<tr>
<td>Procedure IV</td>
<td>Set to zero</td>
<td>Use only workers</td>
<td>Use reduced form</td>
<td>OLS</td>
</tr>
<tr>
<td>Procedure V</td>
<td>Use Tobit</td>
<td>Use only workers</td>
<td>Use predicted wage with selection correction from Tobit</td>
<td>Tobit on hours and OLS on wages</td>
</tr>
<tr>
<td>Procedure VI</td>
<td>Joint ML of hours and wages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure VII</td>
<td>Estimate LFP with Probit</td>
<td>Use only workers with selection correction</td>
<td>Use reduced form for wages with selection correction</td>
<td>Heckit</td>
</tr>
<tr>
<td>Procedure VIII</td>
<td>Estimate LFP with Probit</td>
<td>Use only workers with selection correction</td>
<td>Use predicted wage from wage equation with selection correction</td>
<td>Heckit</td>
</tr>
</tbody>
</table>
• 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:

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

(2')

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

with reference to the theoretical model, the latter individuals are the ones for which the offered wage is below their reservation wage ($w_i < w^*_i$).
• 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 $x_i = [w_i, p_i, X_i, Y_i, \beta, z_i]$ be a vector of covariates with an associate vector of coefficients $\beta$.

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

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

$$
= \frac{1}{\sigma} \phi \left( \frac{H_i - x_i \beta}{\sigma} \right), \text{ where } \phi(\cdot) \text{ is the PDF of a N}(0,1).
$$
• 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
\[ L = \prod_{\{i|H_i>0\}} \frac{1}{\sigma} \phi \left( \frac{H_i - x_i\beta}{\sigma} \right) \prod_{\{i|H_i=0\}} \Phi \left( \frac{-x_i\beta}{\sigma} \right) \] (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 (who he got the Nobel Prize in 2002). This procedure is known as the Heckit: it focused only participants removing all non-participants for the hours estimation regression.

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

\[ E(H_i \mid H_i > 0) = x_i \beta + E(\varepsilon_i \mid \varepsilon_i > x_i \beta) = x_i \beta + \sigma \cdot \frac{\phi(x_i \beta / \sigma)}{\Phi(x_i \beta / \sigma)} = x_i \beta + \sigma \cdot \lambda(x_i \beta / \sigma), \]  

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 \textbf{inverse Mills ratio} or the sample selection correction term. (see Appendix B in Supporting TF)
• 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, \(\hat{\lambda}(x_i \hat{\beta} / \hat{\sigma})\).

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

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

• When using the Heckit, it is better to use pre-programmed procedure, as in STATA, that computes the standard errors correctly (allowing for heteroscedasticity) and gives us the correct t-statistics on \(\hat{\lambda}\), that serves as a test for selection bias.
• 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 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:
  \[
  \max U(w^pH - T(w^pH), H) \implies \text{FOC: } U_c(w^p(1 - T')) - U_H = 0
  \]
  where \(w^p\) is the pre-tax wage, \(T(w^pH)\) 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.
\[ W_3 = (1-t_3) W \]
\[ W_2 = (1-t_2) W \]
\[ W_1 = (1-t_1) W \]

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 [see Hausman exhibit]

- 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
2. Using Tax and Transfer Programs to Estimate Labour Supply

1. Field Experiments and Randomized Trials

   a. Negative Income Tax (NIT) experiments

   - NIT were designed to test replacement of existing welfare and income support policies with a single income support program operating through the tax and transfer system

   - Conducted in late 1960s, early 1970s in New Jersey, Seattle/Denver (SIME/DIME), and Gary, Indiana and other rural sites (RIME) in the US and in Manitoba (Mincome), Canada

   - The experiments were designed to improve work incentives for those on welfare, but could potentially have adverse work incentives for this group

   - NIT experiments designed to provide evidence on the magnitudes of any adverse work incentives for working poor
• Basic design of the program was a lump sum transfer \( (G) \) with a 50-70% phase-out rate \( (\tau) \).

• With linear tax system with tax rate \( \tau \) and grant \( G \), if the pre-tax wage rate \( w^p \), the budget constraint is:
  \[
  C = w^p (1 - \tau) + G
  \]
  that is a reduction in wages and an increase in unearned income.

• Main finding was that NIT would have small but \textit{non-trivial negative effects} on labour supply (hours worked) of working poor.

• This early attempt at experimentation in the US (and Canada) was not ultimately successful

• Experiments were poorly designed and made the estimation of ATE difficult (Ashenfelter and Plant, 1990).
  - Non-random selection into experiment (selected on income)
  - Non-random assignment to Treatment (T) and Control (C) groups
  - Self-reported earnings with incentives for T to underreport so that they got NIT payment (Lesson: need to match to administrative records: UI, SS, firm tax records.)
  - Sample attrition
• Non-randomness undoes the simple T/C comparison that is so powerful in randomized studies. So much statistical modeling was used here.

• Basic results were actually estimated with structural models (Burtless, 1986).
  o Husbands \( (\varepsilon_{Hw} \approx 0.08, \eta_{HY} \approx -0.10) \)
  o Wives \( (\varepsilon_{Hw} \approx 0.13, \eta_{HY} \approx -0.16) \)
  o Single female heads \( (\varepsilon_{Hw} \approx 0.17, \eta_{HY} \approx -0.06) \)

• Other issues that arose in interpreting the NIT data included:
  o external validity, extent to which one can generalize from the sites chosen to country as a whole
  o whether short duration experiments provide evidence on long term responses
b. Self-Sufficiency Project (SSP)

- SSP demonstration carried out in provinces of BC and NB in 1990s
  - Experimental sample randomly assigned between Nov 1992 and Feb 1995
  - Sample size: program group 2859, control group 2827
- SSP was designed to test a financial incentive to leave welfare – a policy to “make work pay”
  - Targeted single parents on welfare for at least 12 months
  - Intended effect on increasing long-term employment was through accumulated work experience

- SSP Earnings Supplement (ES) (see complete description Card and Hyslop (2005))
  - Supplement payment is ½ the difference between actual earnings and an earnings benchmark, equal to $2,500 per month in NB and $3,083 per month in BC in 1993, and adjusted for inflation in later years
  - Supplement payments not affected by unearned income, or income of spouse/partner, ES treated as regular income for income tax purposes
  - Eligibility: 12 months following random assignment to obtain FT employment (FT = at least 30 hours per week on average over the month)
  - Receipt: up to 3 years of ES providing FT employment maintained
    - Supplement quite generous relative to U.S. programs tested in early 1990s
• SSP effects
  ▪ Because of more careful randomization, the comparison between T and C provides more useful results in this case, but the program effects are still estimated within a D-D framework with controls because of differences between T and C in work experience and education upgrading.

  ▪ The effects on full-time employment and reduced welfare participation, although quite spectacular during the program, faded away after the end of the program.
    ○ At peak, SSP generated a 14 percentage point reduction in welfare participation.

  ▪ Given the requirements of the program of maintaining FT employment, there was a clearer positive impact on accumulated labor market experience
In terms of wage growth, it turns out that the program requirement of maintaining full-time employment had some unintended consequences:

- In the short run, it forced some program participants to accept lower wages to get full-time employment (rather than part-time) (Zabel et al, 2004)
- There is evidence of a relative wage progression of approximately 10 percentage points for the program participants who would not have found full-time employment in the first year without such an incentive.
- In the longer run, it prevented some program participants to upgrade their education yielding less wage growth than among the control group (Riddell and Riddell, 2006)

In terms of the cost of the program, the tax receipts from the income tax on earnings mitigated the program’s costs (Michalpoulos et al, 2005)

Overall, some features of the program in particular the 12-month window to obtain full-time employment dominated the analysis, which limited the external validity of the experiments.
Card and Hyslop (2005) built a simple optimizing model that incorporates the eligibility rules which created an "establishment" incentive to find a job and leave welfare within a year of random assignment, and an "entitlement" incentive to choose work over welfare once eligibility was established.

They are then able to simulate the impact of alternative eligibility on welfare entry and exit rates among those who achieved eligibility and the simulation suggests that allowing more time for people to establish eligibility would have led to at most a 20% larger program impact in months 18-45 and also would have shifted the peak program impact to the right somewhat.
2. Tax (and Transfer) Reforms

- Using tax reforms as a “natural experiment” to evaluate the effect of taxes on labor supply (and other outcomes) has been quite popular in the UK (Blundell, Duncan and Meghir, 1998) and the US (especially the tax reform (TRA86) of 1986).

- Using a unique quasi-natural experiment, Tax Holiday in Iceland, Bianchi et al. (2001) have shown large labor supply responses (extensive margins)
  - In 1987, Iceland transitioned from paying taxes on previous year's income to current income
  - To avoid double taxation during transition, no tax charged over 1987 incomes
  - Average tax rate of 14.5% in 1986, 0% in 1987, 8% in 1988
  - Reform was announced in late 1986, so it was an unanticipated temporary tax change

- Bianchi et al. (2001) look at labour supply and employment effects and found substantially large effects
  - Despite the fact that more people are employed in 1987 than in 1986, their supply of weeks falls from the benchmark case due to entry and exit
• The average elasticity for both sexes combined is 0.42 (0.07) when weeks worked are used and 0.67 (0.25) when earnings are used.
• Differences come a composition effect: the ones who quit employment worked more on average than those who joined.
• More generally, there was substantial heterogeneity across groups, e.g. self-employed

| Table 3—Changes in Labor Supply (Percent) Relative to the Average of 1986 and 1988 |
|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | Weeks           | Earnings        |                 |
|                 | Male | Female | Male | Female |
| Entry and exit in 1987 | -1.4 | -0.6 | 0.0 | 0.2 |
| Δweeks (Δearnings)  | 6.6  | 2.0  | 8.9  | 2.9  |
| Entry in 1988      | -0.2 | -0.8 | 0.0 | 0.0 |
| Sum                | 5.0  | 0.6  | 8.9  | 3.1  |
| Both sexes         | 5.6  | 12.0 | 

Source: Bianchi, Gundmundsson, and Zoega (2001)
• In Canada, recent work (Baker, Gruber, and Milligan, 2008) has exploited differences across provinces in maternity leave benefits, child care subsidies, and child tax credits.

• **Advantage** of tax reform: policies can affect some groups and not others, creating **natural treatment** and **control** groups.

• **Disadvantages**: changes in tax laws need to be sufficiently large to generate some behavioural responses.
  
  o Perhaps for that reasons, reforms at the lower end affecting **single mothers** (EITC in the US and WTFC in the UK) have received lots of attention (Blundell and Hoynes, 2004).
  
  o A difficulty here is that single mothers may benefits from a combination of transfers and in-kind programs. [see Blundell (2015)]
  
  o More generally, it is unclear that participants are always informed about the changes in the features of the benefit schedule (Chetty and Saez, 2009)

1) A first approach is a **difference-in-difference strategy** that compares the average behaviour **before** and **after** the reform for the eligible group with the before and after contrast for the comparison group.
• This approach removes unobservable individual effects and common macro effects to recover the average effect of the reform, under the assumptions of common time effects across groups and no composition changes with the group. But this is not without problems.

• For example, Eissa (1995) uses the Tax Reform Act (TRA) of 1986 to identify the effect of MTRs on labor force participation and hours of married women.

• This tax reform TRA 1986 cut top income MTR from 50% to 28% from 1986 to 1988 but did not significantly change tax rates for the middle class.
  ➢ Substantially increased incentives to work of wives of high income husbands relative wives of middle income husbands.
  ➢ Results are not as robust as desired

• These effects might be similar to those of “income splitting” which were in effect for the 2014-2015(?) tax year in Canada.
Difference-in-Difference Methodology in a Nutshell

Step 1: Simple Difference
- Outcome: LFP (labor force participation)
- Two groups: Treatment group (T) which faces a change [women married to high income husbands] and control group (C) which does not [women married to middle income husband]
- Simple Difference estimate: $D = \text{LFP}_T - \text{LFP}_C$ captures treatment effect, if in the absence of treatment, LFP equal across 2 groups
- Note: this assumption always holds when T and C status is randomly assigned. To test for this assumption, we can compare LFP before the treatment:
  $$D_B = \text{LFP}_B^T - \text{LFP}_B^C$$

Step 2: Difference-in-Difference (DD)
- If $D_B \neq 0$, we can estimate DD:
  $$DD = D_A - D_B = [\text{LFP}_A^T - \text{LFP}_A^C] - [\text{LFP}_B^T - \text{LFP}_B^C]$$
  where A = after reform, B = before reform
- DD is unbiased if the parallel trend assumption holds: absent the treatment, the difference across T and C would have stayed the same before and after
• DD can be estimated by OLS to control for additional covariates
  \[ LFP_{it} = \beta_0 \text{After} + \beta_1 \text{Treat} + \gamma \text{After} \times \text{Treat} + \beta_x X_{it} + \epsilon_{it} \]

• It is easy to show that \( \hat{\gamma} = \bar{DD} \).
• DD most convincing when groups are very similar to start with [closer to randomized experiment]
• Should always test DD using data from more periods and plot the two time series to check parallel trend assumption [e.g. Imbens et al.]
• Use alternative control groups [not as convincing as potential control groups are many]
• In principle, can create a DDD as the difference between actual DD and DD^{Placebo} (DD between 2 control groups)

• Results in Eissa (1995):
  1) Participation elasticity around 0.4 but large standard errors
  2) Hours elasticity of 0.6
  3) Total elasticity (unconditional hours) is 0.4+0.6 = 1
• Caveats:
  1) Does the common trends assumption hold? Potential story biasing the result: Trend toward “power couples” and thus DD might not be due to taxes:
  In 1983-1985, high income husbands had non-working spouses,
  In 1989-1991, high income husbands married to professionals [and no change for middle class]

  2) LFP before the reform is very different across T and C groups: DD sensitive to functional form assumption [such as levels vs logs]

  1) Liebman and Saez (2006) plot full time-series CPS plot and show that Eissa's results are not robust using admin data (SSA matched to SIPP)
• In a follow-up study, Eissa and Liebman (1996) consider the labour force participation of *single* mothers using CPS data
  o used a difference-in-differences strategy
  o to examine the relatively small 1986 EITC expansion in the US
  o comparing single mothers (treatment) to single women without children (control)
  o in the periods 1984-1986 (before) vs. 1988-1990 (after)

\[
\Pr(lfp_{it} = 1) = \Phi(\alpha + \beta Z_{it} + \gamma_0 \text{treat}_i + \gamma_1 \text{post86}_i + \gamma_2 (\text{treat} \times \text{post86})_{it}),
\]

• They find a small but significant DD effect: 2.4% (larger DD effect 4% among women with low education)
  o Translates into substantial participation elasticities above 0.5
  o Conventional standard errors probably overstate precision
  o LFP for women with/without children are not great comparison groups (70% LFP vs. +90%): time series evidence is only moderately convincing

• Subsequent studies have used much bigger EITC expansions of the mid 1990s and also find positive effects on labour force participation of single women/single mothers
As we saw previously, earning income tax credit generates a piece-wise linear budget constraint with kinks.

2) Other more recent approaches adopt a structural model by using a flexible functional form for the utility function or for the hours function, and construct a likelihood function of the choice of hours corresponding to each segment (and kink) of the budget constraint.

For example, Blundell and Shephard (2008) develop a structural model

$$\max U(c, h_j; X, \varepsilon) \text{ subject to } c \equiv wh_j - T(wh_j, h_j; X) - C \cdot 1\{j > 0\},$$

that allows for discrete choices over non-linear budget constraints and fixed costs of work and allows them to reexamine the problem of the optimal design of the transfer system.

They also use the differences-in-differences to assess the validity of the parameters from their structural estimation which explicitly allows for different labour supply responses at the intensive and extensive margins.
**Take-away from the Static Labour Supply Literature**

1) Small elasticities for prime-age males
   - Probably institutional restrictions, need for at least one income, etc. prevent a short-run response

2) Larger responses for workers who are less attached to labor force
   - Married women, low incomes, retirees

3) Responses essentially driven by extensive margin
   - Extensive margin (participation) elasticity around 0.2-0.5
   - Intensive margin (hours) elasticity close 0

2) Self-employed are much more responsive than employees, in particular the highest earners
   - Hours constraints likely important for employees
Basic readings:

Chapter 2. pp.1-64 (see on-line papers on course web-site)


Further readings:
