Lecture Notes 2: Taxation and Labor Supply

(references: Moffitt (1990), Hausman (1981). Additionally, these notes draw heavily on graduate lecture notes from Wilbert Van der Klaauw.)

Theoretical Labor Supply Model with Taxes

Suppose that the government taxes earnings at proportional rate *t*.

The worker's problem is

$$\max_H \; U(C,L)$$
$$s.t. \;\; T = L + H \;\; \text{and} \;\; C = W(1-t)H + N$$

Utility maximization yields the expression

$$W(1\text{-}t)U_C(W(1-t)H+N, T-H) = U_L(W(1-t)H+N, T-H)$$

which we solve for $H(W(1-t), N)$.

Net after-tax wage is *W(1-t)*, and we note that $\frac{\partial H}{\partial t}$ is of ambiguous sign. The tax contains positive work incentives due to the income effect of lowered net wages, and negative work incentives due to their substitution effect.

Deadweight Loss
(see graphs)

A represents the pre-tax optimum, D the post-tax optimum.
B is the point at which optimizing individuals would be as well off after the tax as they were before, and is calculated at the new (flatter) post-tax wage rate. B therefore represents the compensated labor supply, and the change from A to B is the compensated (substitution) tax effect. The distance D to R represents government revenue of *Wt*, and the **deadweight loss** is the distance above R to the dashed wage line.

As in the incidence model, we evaluate deadweight loss after compensating the decision-maker for the price change.

Convex Budget Set and Progressive Income Taxation

Suppose earnings are taxed as follows:

$$\begin{aligned} R &= t_1 WH && \text{if WH} < \text{E}_1 \\ &= t_2(WH - E_1) + t_1 E_1 && \text{if } \text{E}_1 \leq WH \end{aligned}$$

Graphing tax revenues *R* against pre-tax income *WH*, revenues increase first at rate t_1 with *WH* and then at rate t_2. The shape of the budget set depends on $t_1 > or < t_2$.

Suppose $t_1 < t_2$. Graphing consumption against leisure for the worker, the cost of the marginal hour of leisure increases as leisure consumption increases, and the tax structure creates a convex budget set for the worker with a discontinuity at *E1*.

See Moffitt (1990) figures 3(a) – 3(c).

In each of these figures the individual is facing a convex budget set with a kink point arising from a government income tax/expenditure program. Consider the manners, in figures 3(a) &(b), in which a decrease in the tax rate on earnings up to *E1* and a change in the value of *E1* can have an ambiguous effect on labor supply.

Additionally, note in 3c that in a government expenditure program a shift from price discounts on food to gifts of food valued similarly to price discounts can actually decrease food consumption.

What might such a policy change do to consumer welfare? Do consumers prefer cash or in-kind transfers? How does the answer relate to the effect of the policy change on welfare?

Econometrics of Non-linear Budget Constraints: The Convex Budget Set Case

3 possible locations of optimum:

(1) Segment 1, with tax rate t_1

(2) Segment 2, with tax rate t_2

(3) Kink

Define *V(W,N)* as the maximum indirect utility when facing a linear BC with slope W and intercept at *H = 0* of *N*.

When the individual locates on (1), the value is $V(W(1-t_1),N)$.

On (2), it's $V(W(1-t_2),\widetilde{N})$, where $\widetilde{N}=N+W(t_2-t_1)H^*$. (We find this value by noting that the N and N tilde lines intersect at the kink point H.)

On (3), it's $U(C^*,T-H^*)$, where $C^*=N+W(1-t_1)H^*$.

(see graphs)

The optimal labor supply function with this budget constraint is

$$
\begin{aligned}
H &= H[W(1-t_1),N] \text{ if } H(W(1-t_1),N)<H^* \\
&= H[W(1-t_2),\widetilde{N}] \text{ if } H(W(1-t_2),\widetilde{N})>H^* \\
&= H^* \text{ if } H(W(1-t_1),N)\geq H^* \text{ and } H(W(1-t_2),\widetilde{N})\leq H^*
\end{aligned}
$$

This is called the **conditional labor supply function**. The total labor supply function is written

$$H=D_1H[W(1-t_1),N]+D_2H[W(1-t_2),\widetilde{N}]+(1-D_1-D_2)H^*$$

where

$$D_1 = 1 \text{ if } \widehat{D_1} > 0, \; D_1 = 0 \text{ otherwise}$$

$$D_2 = 1 \text{ if } \widehat{D_2} > 0, \; D_2 = 0 \text{ otherwise}$$

$$\widehat{D_1} = H^* - H(W(1-t_1), N)$$

$$\widehat{D_2} = H(W(1-t_2), \widetilde{N}) - H^*$$

Comparative Statics

Setting aside the kink for now, & fixing D1 and D2,

$$\frac{\partial H}{\partial t_1} = D_1 H_1(W(1-t_1), N)(-W) + D_2 H_2(W(1-t_2), \widetilde{N})(-WH^*)$$

$$\frac{\partial H}{\partial t_2} = D_2 H_1(W(1-t_2), \widetilde{N})(-W) + D_2 H_2(W(1-t_2), \widetilde{N})(WH^*)$$

Note that if D2 = 0 then $\frac{\partial H}{\partial t_2} = 0$, & if D1 = 0 then $\frac{\partial H}{\partial t_1} = H_2(W(1-t_2), \widetilde{N})(-WH^*)$. Labor supply only responds through the after-tax marginal wage to the tax for the segment the worker is on (if either), and only responds through effective non-labor income to taxes on lower earnings levels.

Returning to the kink, note that if $D_1 = D_2 = 0$ then the effects of both taxes may be zero.

On the other hand, small changes in the taxes may change D1 or D2 by a full unit, causing large and unpredictable labor supply responses to the tax change.

Estimation

If the optimal hours function $H(W,N)$ is such that the hours equation is linear of the form

$$H = \beta_0 + \beta_1 W + \beta_2 N + \varepsilon$$

then for segments 1 & 2, respectively,

$$H = \beta_0 + \beta_1 W(1-t_1) + \beta_2 N + \varepsilon$$
$$H = \beta_0 + \beta_1 W(1-t_2) + \beta_2 \widetilde{N} + \varepsilon$$

If we knew each person *i*'s segment then we could estimate

$$H_i = \beta_0 + \beta_1 W(1-t_i) + \beta_2 N_i + \varepsilon$$

where

$t_i = t_1$ and $N_i = N$ if on segment 1

$= t_2$ and $N_i = \widetilde{N}$ if on segment 2

What problems are we likely to run into in estimating this OLS model?

Could they be corrected using an IV estimator? Why or why not?

Maximum Likelihood Estimation

(Note that we are taking a particularly **structural** approach to estimating the response of labor hours to taxes. Later in Lecture Notes 2 we will look at a paper by Eissa (1995) which uses a **nonparametric** technique to estimate this relationship.)

Our empirical specification is

$$H = D_1(H[W(1-t_1), N] + \alpha) + D_2(H[W(1-t_2), \widetilde{N}] + \alpha) + (1 - D_1 - D_2)H^* + \varepsilon$$

where

$$D_1 = 1 \text{ if } H(W(1-t_1), N) + \alpha < H^*, \text{ 0 otherwise}$$

$$D_2 = 1 \text{ if } H(W(1-t_2), \widetilde{N}) + \alpha > H^*, \text{ 0 otherwise}$$

There are two error terms.

<u>Measurement error ε</u>
The assumption of measurement error in data collection implies that we don't observe the segment the individual is located on with certainty. Additionally, ε spreads people away from the kink point, since those choosing the kink point might be observed some distance above or below it due to ε.

<u>Individual Heterogeneity α</u>
Individuals with the same *W* and *N* locate at different places on the constraint because of heterogeneity in preferences. α will also lead to more clustering at the kink, as variations in indifference curves make the kink the preferred point for many (*W*,*N*) pairs.

Thus we get identification of the distribution parameters of the error terms from the degree of clustering at the kink; the greater the clustering, the greater the relative variance of the heterogeneity error.

To estimate this model we must assume parametric distributions for our error terms. For example, $\varepsilon \sim N(0,\sigma_\varepsilon{}^2)$.

If we assume for the time being that the ε distribution is degenerate at 0, then the likelihood function is

$$
\begin{aligned}
M = & \prod_{seg1} \Pr[\alpha = H - H(W(1-t_1), N)] \\
& \times \prod_{seg2} \Pr[\alpha = H - H(W(1-t_2), \widetilde{N})] \\
& \times \prod_{kink} \Pr[H^* - H(W(1-t_1), N) < \alpha < H^* - H(W(1-t_2), \widetilde{N})]
\end{aligned}
$$

We can write it this way because with only the individual heterogeneity term α we are certain that the segment we observe is the true segment.

[To board for likelihood assuming $\alpha \sim N(0,\sigma_\alpha{}^2)$.]

Next allow ε to have a non-degenerate distribution. Then we don't know that the observed segment is the true one, and the likelihood is

$$M = \prod \Pr(H)$$

where

$$
\begin{aligned}
\Pr(H) = & \Pr[\alpha + \varepsilon = H - H(W(1-t_1), N), \alpha < H^* - H(W(1-t_1), N)] \\
& + \Pr[\alpha + \varepsilon = H - H(W(1-t_2), \widetilde{N}), \alpha > H^* - H(W(1-t_2), \widetilde{N})] \\
& + \Pr[\varepsilon = H - H^*, H^* - H(W(1-t_1), N) < \alpha < H^* - H(W(1-t_2), \widetilde{N})]
\end{aligned}
$$

Since we don’t know the true segment in the likelihood, we have to account for the possibility of each being the true case when evaluating the probability of any *H*.

Application to the US Tax System by Hausman (1981)

Hausman estimates

$$H = \alpha W + \beta N + \varepsilon$$

with $\varepsilon \sim N(0, \sigma^2)$ and $\beta \sim N(0, \sigma_\beta^2)$.

However, since we can’t have positive effects of nonlabor income on employment, Hausman assumes the beta distribution is a truncated normal, with positive density only on $\mathbb{R}^-$.

Hausman does not reject H_0 that $\alpha = 0$, indicating a lack of substitution of leisure for consumption in response to tax increases. He does find a significant nonlabor income influence on hours.

Overall, he finds that the US tax system induces an 8% decrease in labor supply, that this effect is greater for women and that the ratio of the deadweight loss resulting from the US tax system to GNP is .2.

The observation that women’s labor supply is more responsive to taxes than men’s would provide a nice transition to the Eissa (1995) paper, but first we must cover the case of non-convex budget sets with discontinuities created by government benefits that are taxed away with labor income.

Non-Convex Budget Sets and 'Negative Income Taxes'

See Moffitt Figure 3(d)

Consider a government transfer program to support the poor in which a transfer of *G* dollars is given, and taxed away dollar-for-dollar as the recipient earns income in the labor market. Individuals I, II and III locate on the budget constraint as depicted in the figure.

Now suppose that the government decides to decrease the rate at which it taxes away the benefit to *t* for every dollar earned. Does this increase the labor supply of all recipients depicted in Figure 3(d)? Why or why not?

What US government programs create these sorts of non-convex budget sets for recipients?

If we have time remaining, we will construct the graph, conditional labor supply function, comparative statics and econometric approach for the case of the non-convex budget constraint. If not, look for some or all of these to show up in Problem Set 1.