ECON 2007: Quant Econ and Econometrics
Probit, Logit and Tobit

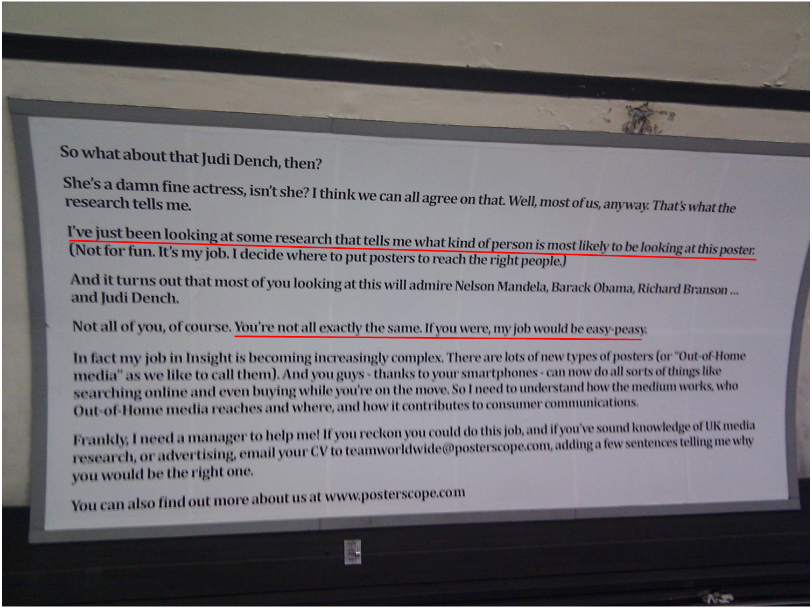
Dr. Áureo de Paula

Department of Economics
University College London

Introduction

LDV models refer to those where the dependent variable's range is restricted:

- Binary Response (e.g., Probit, Logit)
- Tobit (i.e., mixed continuous-discrete outcomes)
- Censored (e.g., top coding, duration models)
- Truncated
- Sample Selection
- Count data


So what about that Judi Dench, then?
She's a damn fine actress, isn't she? I think we can all agree on that. Well, most of us, anyway. That's what the research tells me.
I've just been looking at some research that tells me what kind of person is most likely to be looking at this poster. (Not for fun. It's my job. I decide where to put posters to reach the right people.)
And it turns out that most of you looking at this will admire Nelson Mandela, Barack Obama, Richard Branson ... and Judi Dench.
Not all of you, of course. You're not all exactly the same. If you were, my job would be easy-peasy.
In fact my job in Insight is becoming increasingly complex. There are lots of new types of posters (or "Out-of-Home media" as we like to call them). And you guys - thanks to your smartphones - can now do all sorts of things like searching online and even buying while you're on the move. So I need to understand how the medium works, who Out-of-Home media reaches and where, and how it contributes to consumer communications.
Frankly, I need a manager to help me! If you reckon you could do this job, and if you've sound knowledge of UK media research, or advertising, email your CV to teamworldwide@posterscope.com, adding a few sentences telling me why you would be the right one.
You can also find out more about us at www.posterscope.com

Food for Thought

How could you help this person out (and get hired)?

- What kind of econometric model would you use?
- What kind of data would you collect? What is the dependent variable? What would you use as explanatory variables?
- How would you estimate your model?
- How could you test whether your predictions are accurate?

Binary Response: Logit and Probit

Consider a person's decision to work or not and record this choice as $y = 1$ if the person works and $y = 0$ otherwise. Let the utility of working or not working be given by:

$$u(y; \mathbf{x}, e_y) = \beta_y^\top \mathbf{x} + e_y$$

$\mathbf{x}$: a list of personal characteristics that might affect one's preference for work (e.g., number of children, age, non-work income) (*absorb the constant term into* $\mathbf{x}$);

β_y are parameters that quantify how those personal characteristics affect the taste for work (β_1) or leisure (β_0); and

e_y represents unobserved individual taste shifters (independent of $\mathbf{x}$) that depend on whether someone works (e_1) or not (e_0) (*because* $\mathbf{x}$ *contains a constant, assume that* $E(e_y) = 0$).

Binary Response: Logit and Probit

This person will choose to work ($y = 1$) if, and only if,

$$\begin{aligned} u(1; \mathbf{x}, e_1) > u(0; \mathbf{x}, e_0) \quad &\Leftrightarrow \quad \beta_1^\top \mathbf{x} + e_1 > \beta_0^\top \mathbf{x} + e_0 \\ &\Leftrightarrow \quad \underbrace{(\beta_1 - \beta_0)}_{\equiv \beta}{}^\top \mathbf{x} + \underbrace{e_1 - e_0}_{\equiv e} > 0 \end{aligned}$$

Consequently,

$$y = \begin{cases} 1, & \text{if } \beta^\top \mathbf{x} + e > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Or, more compactly:

$$y = 1[\underbrace{\beta^\top \mathbf{x} + e}_{\equiv y^*} > 0].$$

Binary Response: Logit and Probit

To estimate the probability that a person works ($y = 1$) we need to specify a cumulative distribution function $G(\cdot)$ for e. Typically, economists focus on

- Normal: $G(\cdot) = \Phi(\cdot)$ where

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-u^2/2)du, \text{ or}$$

- Logistic: $G(\cdot) = \Lambda(\cdot)$ where

$$\Lambda(z) = \frac{\exp(z)}{1 + \exp(z)}$$

Binary Response: Logit and Probit

These are both bell-shaped, symmetric about zero: $G(z) = 1 - G(-z)$. Consequently,

$$\begin{aligned} Pr(y = 1|\mathbf{x}) &= Pr(y^* > 0|\mathbf{x}) = Pr(e > -\beta^\top \mathbf{x}) \\ &= 1 - G(-\beta^\top \mathbf{x}) = G(\beta^\top \mathbf{x}), \end{aligned}$$

which we can use to estimate the parameter vector β.

Because $G(\cdot)$ is a nonlinear function, we cannot use OLS or GLS. Although we can use modifications of those methods, it is nevertheless preferrable to use the method of Maximum Likelihood.

Binary Response: Logit and Probit

Maximum Likelihood Principle

Out of all the possible values of the parameter, the value that makes the likelihood of the observed data largest should be chosen.

For a single observation i, to calculate the value of the likelihood at a given parameter value β we need the probability mass function of y given $\mathbf{x}_i$:

$$f(y|\mathbf{x}_i) = [G(\beta^\top \mathbf{x}_i)]^y [1 - G(\beta^\top \mathbf{x}_i)]^{(1-y)}, y = 0, 1$$

with corresponding log-likelihood

$$l_i(\beta) = y_i \log[G(\beta^\top \mathbf{x}_i)] + (1 - y_i) \log[1 - G(\beta^\top \mathbf{x}_i)].$$

Binary Response: Logit and Probit

For a random sample with n observations, the log-likelihood function to be maximized is then given by

$$\mathcal{L}(\beta) = \sum_{i=1}^{n} l_i(\beta).$$

We cannot write down a closed form solution for the estimator, which maximizes $\mathcal{L}(\beta)$. Nevertheless, a computer package will (most of the time) easily compute it and it can be shown to be consistent, asymptotically normal and efficient.

Binary Response: Logit and Probit

Probit: Female Labor Force Participation (MROZ.dta).

```
Probit regression                               Number of obs   =        753
                                                LR chi2(7)      =     227.14
                                                Prob > chi2     =     0.0000
Log likelihood = -401.30219                     Pseudo R2       =     0.2206

------------------------------------------------------------------------------
        inlf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0120237   .0048398    -2.48   0.013    -.0215096   -.0025378
        educ |   .1309047   .0252542     5.18   0.000     .0814074     .180402
       exper |   .1233476   .0187164     6.59   0.000     .0866641    .1600311
             |
     c.exper#|
     c.exper |  -.0018871      .0006    -3.15   0.002     -.003063   -.0007111
             |
         age |  -.0528527   .0084772    -6.23   0.000    -.0694678   -.0362376
     kidslt6 |  -.8683285   .1185223    -7.33   0.000    -1.100628    -.636029
     kidsge6 |    .036005   .0434768     0.83   0.408     -.049208    .1212179
       _cons |   .2700768    .508593     0.53   0.595    -.7267473    1.266901
------------------------------------------------------------------------------
```

Binary Response: Logit and Probit

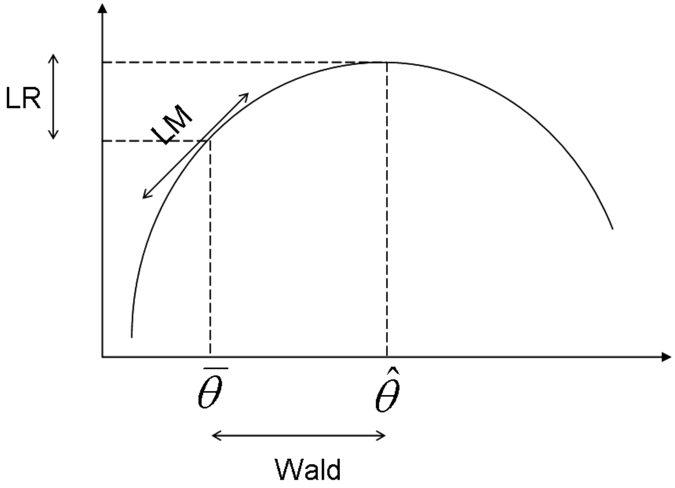
There are three possible routes to test multiple exclusion restrictions here:

1. **Lagrange multiplier or score statistic.** Only estimates restricted model as in linear case.
2. **Wald statistic.** Only estimates unrestricted model and, under H_0, is χ^2_q where q is the number of restrictions.
3. **Likelihood statistic.** Estimates both restricted and unrestricted models and relies on

$$LR = 2(\mathcal{L}_{ur} - \mathcal{L}_r)$$

which, under H_0 is χ^2_q where q is the number of restrictions.

Binary Response: Logit and Probit



Binary Response: Logit and Probit

In a LPM, since $Pr(y = 1|\mathbf{x}) = \beta^\top \mathbf{x}$, the estimated coefficient would give us the direct effect of a covariate x_j on $Pr(y = 1|\mathbf{x}) = p(\mathbf{x})$.

In a Logit or Probit, $\beta^\top \mathbf{x} = E(y^*|\mathbf{x})$ and β does not have an immediate interpretation as the **latent variable** y^* does not carry well defined measurement units.

We need to work a little further to obtain the effect of x_j on $p(\mathbf{x}) = G(\beta^\top \mathbf{x})$.

Binary Response: Logit and Probit

- If x_j is discrete, we have to evaluate $p(\mathbf{x})$ at the relevant values for x_j. For example, if x_j is a dummy (e.g. representing gender), the effect of this variable is

$$\begin{aligned} &G(\beta_0 + \beta_1 x_1 + \cdots + \beta_j + \cdots + \beta_k x_k) - \\ &G(\beta_0 + \beta_1 x_1 + \cdots + 0 + \cdots + \beta_k x_k) \end{aligned} \tag{1}$$

- If x_j is continuous, we can use calculus to obtain

$$\frac{\partial p(\mathbf{x})}{\partial x_j} = g(\beta^\top \mathbf{x})\beta_j \text{ where } g(z) \equiv G'(z). \tag{2}$$

Binary Response: Logit and Probit

A few remarks:

- (1) can be generalized for other kinds of discrete variables (see book);
- (2) can be generalized for general functional forms among regressors. For instance,

$$Pr(y = 1|\mathbf{z}) = G(\beta_0 + \beta_1 z_1 + \beta_2 z_1^2 + \beta_3 \log z_2)$$

  where the partial effect of z_1 is $\partial Pr(y = 1|\mathbf{z})/\partial z_1 = g(\beta^\top \mathbf{z})(\beta_1 + 2\beta_2 z_1)$ and of z_2 is $\partial Pr(y = 1|\mathbf{z})/\partial z_2 = g(\beta^\top \mathbf{z})\beta_3/z_2$.
- Elasticities are given by $x_j \partial Pr(y = 1|\mathbf{x})/\partial x_j / Pr(y = 1|\mathbf{x})$.
- Be careful with interactions of variables!
- Watch out for the output of computer packages (which tend to follow (2)).

Binary Response: Logit and Probit

Notice that the partial effects now depend on the value of the regressors. To evaluate those in practice there are two common alternatives:

- **PEA.** Estimate $\partial p(E[\mathbf{x}])/\partial x_j$ by $g(\hat{\beta}^\top \overline{\mathbf{x}})\hat{\beta}_j$. Two issues:
  1. What does it mean to be "47.5% female"?
  2. Averages of functions or functions of averages: $\overline{age}$ and $\overline{age^2}$ or $\overline{age}$ and $\overline{age}^2$?
- **APE.** Estimate $E[\partial p(\mathbf{x})/\partial x_j]$ by

$$\hat{\beta}_j \frac{1}{n} \sum_{i=1}^{n} g(\hat{\beta}^\top \mathbf{x}_i).$$

(Similarly for discrete regressors.)

Binary Response: Logit and Probit

Probit: Partial Effect at the Average (PEA)

```
Conditional marginal effects                      Number of obs   =        753
Model VCE    : OIM

Expression   : Pr(inlf), predict()
dy/dx w.r.t. : nwifeinc educ exper age kidslt6 kidsge6
at           : nwifeinc        =    20.12896 (mean)
               educ            =    12.28685 (mean)
               exper           =    10.63081 (mean)
               age             =    42.53785 (mean)
               kidslt6         =    .2377158 (mean)
               kidsge6         =    1.353254 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0045448   .0018286    -2.49   0.013    -.0081288   -.0009607
        educ |   .0494796   .0095876     5.16   0.000     .0306883    .0682709
       exper |   .0314576   .0031229    10.07   0.000     .0253368    .0375784
         age |  -.0199773   .0032404    -6.17   0.000    -.0263284   -.0136263
     kidslt6 |  -.3282122   .0452473    -7.25   0.000    -.4168953    -.239529
     kidsge6 |   .0136092    .016439     0.83   0.408    -.0186107    .0458291
------------------------------------------------------------------------------
```

Remark: $PEA_{\texttt{exper}} = \phi(\overline{\mathbf{x}}^{\top}\hat{\beta}) \times (\hat{\beta}_{\texttt{exper}} + 2\hat{\beta}_{\texttt{exper}^2}\overline{\texttt{exper}})$.

Binary Response: Logit and Probit

Probit: Average Partial Effect (APE)

```
Average marginal effects                          Number of obs   =        753
Model VCE    : OIM

Expression   : Pr(inlf), predict()
dy/dx w.r.t. : nwifeinc educ exper age kidslt6 kidsge6

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0036162   .0014414    -2.51   0.012    -.0064413   -.0007911
        educ |   .0393703   .0072216     5.45   0.000     .0252161    .0535244
       exper |   .0255825   .0022272    11.49   0.000     .0212172    .0299478
         age |  -.0158957   .0023587    -6.74   0.000    -.0205186   -.0112728
     kidslt6 |  -.2611542   .0318597    -8.20   0.000    -.3235982   -.1987103
     kidsge6 |   .0108287   .0130584     0.83   0.407    -.0147654    .0364227
------------------------------------------------------------------------------
```

Remark: $APE_{\texttt{exper}} = \sum_{i=1}^{N} \phi(\mathbf{x}_i^\top \hat{\beta}) \times (\hat{\beta}_{\texttt{exper}} + 2\hat{\beta}_{\texttt{exper}^2}\texttt{exper}_i)/N.$

Binary Response: Logit and Probit

Probit: PEA. Be careful with the derivatives!

```
Conditional marginal effects                    Number of obs    =        753
Model VCE    : OIM

Expression   : Pr(inlf), predict()
dy/dx w.r.t. : nwifeinc educ exper age kidslt6 kidsge6
at           : nwifeinc        =    20.12896 (mean)
               educ            =    12.28685 (mean)
               exper           =    10.63081 (mean)
               expersq         =    178.0385 (mean)
               age             =    42.53785 (mean)
               kidslt6         =    .2377158 (mean)
               kidsge6         =    1.353254 (mean)

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0046962   .0018903    -2.48   0.013    -.0084012   -.0009913
        educ |   .0511287   .0098592     5.19   0.000     .0318051    .0704523
       exper |   .0481771   .0073278     6.57   0.000     .0338149    .0625392
         age |  -.0206432   .0033079    -6.24   0.000    -.0271265   -.0141598
     kidslt6 |  -.3391514   .0463581    -7.32   0.000    -.4300117   -.2482911
     kidsge6 |   .0140628   .0169852     0.83   0.408    -.0192275    .0473531
------------------------------------------------------------------------------
```

These are different from the previous one because (1) they use $\overline{\text{exper}^2}$ instead of $\overline{\text{exper}}^2$ and (2) the derivative for exper does not account for exper^2 (i.e., $PEA_{\texttt{exper}} = \phi(\bar{\mathbf{x}}^\top \hat{\beta}) \times \hat{\beta}_{\texttt{exper}}$)

Binary Response: Logit and Probit

Probit: APE. Be careful with the derivatives!

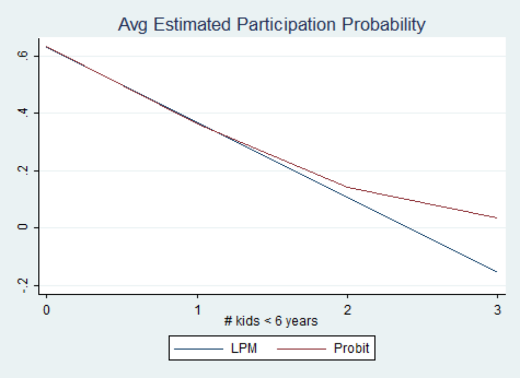
```
Average marginal effects                        Number of obs     =        753
Model VCE      : OIM

Expression     : Pr(inlf), predict()
dy/dx w.r.t.   : nwifeinc educ exper age kidslt6 kidsge6

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0036162   .0014414    -2.51   0.012    -.0064413   -.0007911
        educ |   .0393703   .0072216     5.45   0.000     .0252161    .0535244
       exper |   .0370974   .0051522     7.20   0.000     .0269993    .0471956
         age |  -.0158957   .0023587    -6.74   0.000    -.0205186   -.0112728
     kidslt6 |  -.2611542   .0318597    -8.20   0.000    -.3235982   -.1987103
     kidsge6 |   .0108287   .0130584     0.83   0.407    -.0147654    .0364227
------------------------------------------------------------------------------
```

Remark: $APE_{\texttt{exper}} = \sum_{i=1}^{N} \phi(\mathbf{x}_i^\top \hat{\beta}) \times \hat{\beta}_{\texttt{exper}}/N$.

Notice that only the marginal effect for exper is different than the previous APE. Why?

Binary Response: Logit and Probit

We can also take into account the fact that kidslt6 is a discrete variable.



Binary Response: Logit and Probit

How well do we fit the data? In binary response models there are few alternatives:

- Percent Correctly Predicted: predict 1 if $G(\hat{\beta}^\top \mathbf{x}_i) \geq 0.5$ and 0 otherwise. Check proportion of correct predictions (overall and for each outcome).
- Use fraction of successes in-sample as threshold (especially if number of successes in data is small).
- Pseudo R-squared: $1 - \mathcal{L}_{ur}/\mathcal{L}_0$ or compare $\hat{P}_i$ to y_i as in usual R^2.

Binary Response: Logit and Probit

Pseudo-R^2 is 22.06% in our example.

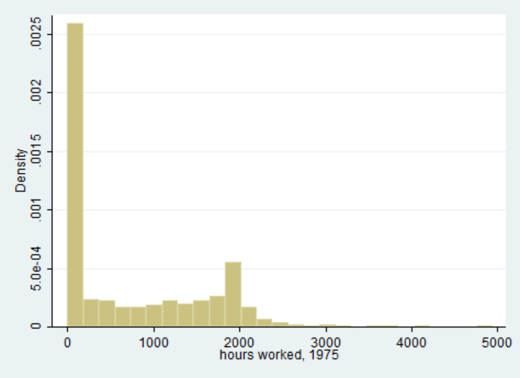
For prediction percentages:

=1 if in lab frce, 1975	1(p_probit>=0.5) 0	1	Total
0	205	120	325
1	80	348	428
Total	285	468	753

Tobit

There are many cases in Economics where corner solutions arise.

For example, consider the case of a person deciding how many hours of work to supply.



Tobit

Let y represent the number of hours worked. It is (a) continuous over positive values and (b) is zero with positive probability.

To model $E(y|\mathbf{x})$ we rely (again) on a latent variable y^*:

$$y^* = \beta^\top \mathbf{x} + u \quad u \sim \mathcal{N}(0, \sigma^2)$$
$$y = \max(c, y^*)$$

where c is a real number (and again we absorb the constant into $\mathbf{x}$. Notice that, as $c \to -\infty$, $y^* \to y$.

Here, we will assume that $c = 0$.

Tobit

Then, we can deduct that

$$\begin{aligned} Pr(y = 0|\mathbf{x}) &= Pr(y^* < 0|\mathbf{x}) = Pr(u < -\beta^\top \mathbf{x}|\mathbf{x}) \\ &= Pr(u/\sigma < -\beta^\top \mathbf{x}/\sigma|\mathbf{x}) \\ &= \Phi(-\beta^\top \mathbf{x}/\sigma) = 1 - \Phi(\beta^\top \mathbf{x}/\sigma) \end{aligned}$$

Consequently, the density of y conditional on $\mathbf{x} = \mathbf{x}_i$ is given by

$$(1/\sigma)\phi[(y - \beta^\top \mathbf{x}_i)/\sigma], y > 0 \qquad 1 - \Phi(\beta^\top \mathbf{x}_i/\sigma), y = 0$$

which we can use to construct a MLE for β!

The MLE maximizes $\mathcal{L}(\beta) = \sum_{i=1}^{n} l_i(\beta)$ where

$$\begin{aligned} l_i(\beta) &= 1[y_i = 0] \log[1 - \Phi(\beta^\top \mathbf{x}_i/\sigma)] + \\ & \quad 1[y_i > 0] \log\{(1/\sigma)\phi[(y - \beta^\top \mathbf{x}_i)/\sigma]\} \end{aligned}$$

Tobit

The coefficient β_j gives the partial effect of x_j on $E(y^*|\mathbf{x})$, but the observed outcome is instead y.

The conditional expectation of y is given by

$$\begin{aligned} E[y|\mathbf{x}] &= Pr(y > 0|\mathbf{x}).E(y|\mathbf{x}, y > 0) + 0.Pr(y = 0|\mathbf{x}) \\ &= \Phi(\beta^\top \mathbf{x}/\sigma).[\beta^\top \mathbf{x} + \sigma\lambda(\beta^\top \mathbf{x}/\sigma)] \\ &= \Phi(\beta^\top \mathbf{x}/\sigma)\beta^\top \mathbf{x} + \sigma\phi(\beta^\top \mathbf{x}/\sigma) \end{aligned}$$

where $\lambda(z) = \phi(z)/\Phi(z)$ is called the **inverse Mills' ratio.**

Tobit

There are two partial effects of potential interest:

- The 'conditonal' partial effect:

$$\begin{aligned}\partial E(y|\mathbf{x}, y > 0)/\partial x_j &= \beta_j \left\{1 + \lambda'(\beta^\top \mathbf{x}/\sigma)\right\} \\ &= \beta_j \left\{1 - \lambda(\beta^\top \mathbf{x}/\sigma)[\beta^\top \mathbf{x}/\sigma + \lambda(\beta^\top \mathbf{x}/\sigma)]\right\}\end{aligned}$$

- The 'unconditional' partial effect:

$$\begin{aligned}\partial E(y|\mathbf{x})/\partial x_j &= \partial Pr(y > 0|\mathbf{x})/\partial x_j . E(y|\mathbf{x}, y > 0) \\ &\quad + Pr(y > 0|\mathbf{x}).\partial E(y|\mathbf{x}, y > 0)/\partial x_j \\ &= \beta_j \Phi(\beta^\top \mathbf{x}/\sigma)\end{aligned}$$

These can be estimated as **PEA** or **APE** in the Probit or Logit cases.

Tobit

```
Tobit regression                                  Number of obs   =        753
                                                  LR chi2(7)      =     271.59
                                                  Prob > chi2     =     0.0000
Log likelihood = -3819.0946                       Pseudo R2       =     0.0343

------------------------------------------------------------------------------
       hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -8.814243   4.459096    -1.98   0.048    -17.56811   -.0603724
        educ |   80.64561   21.58322     3.74   0.000     38.27453    123.0167
       exper |   131.5643   17.27938     7.61   0.000     97.64231    165.4863
             |
     c.exper#|
     c.exper |  -1.864158   .5376615    -3.47   0.001    -2.919667   -.8086479
             |
         age |  -54.40501   7.418496    -7.33   0.000    -68.96862    -39.8414
     kidslt6 |  -894.0217   111.8779    -7.99   0.000    -1113.655   -674.3887
     kidsge6 |    -16.218   38.64136    -0.42   0.675    -92.07675    59.64075
       _cons |   965.3053   446.4358     2.16   0.031     88.88528    1841.725
-------------+----------------------------------------------------------------
      /sigma |   1122.022   41.57903                      1040.396    1203.647
------------------------------------------------------------------------------
  Obs. summary:        325  left-censored observations at hours<=0
                       428     uncensored observations
                         0 right-censored observations
```

Tobit

Tobit: Average Partial Effect (APE) on hours conditional on hours>0

```
Average marginal effects                          Number of obs   =        753
Model VCE    : OIM

Expression   : E(hours|hours>0), predict(e(0,.))
dy/dx w.r.t. : nwifeinc educ exper age kidslt6 kidsge6

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -3.968784   2.007582    -1.98   0.048    -7.903573   -.0339953
        educ |   36.31225   9.703038     3.74   0.000     17.29465    55.32986
       exper |    37.5935   2.965955    12.68   0.000     31.78034    43.40667
         age |  -24.49691   3.362492    -7.29   0.000    -31.08728   -17.90655
     kidslt6 |  -402.5507   50.74877    -7.93   0.000    -502.0164   -303.0849
     kidsge6 |  -7.302468   17.40427    -0.42   0.675     -41.4142    26.80927
------------------------------------------------------------------------------
```

Tobit

Tobit: Average Partial Effect (APE) on uncensored probability

```
Average marginal effects                          Number of obs   =        753
Model VCE    : OIM

Expression   : Pr(hours>0), predict(pr(0,.))
dy/dx w.r.t. : nwifeinc educ exper age kidslt6 kidsge6

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -.0024212   .0012202    -1.98   0.047    -.0048128   -.0000297
        educ |    .022153   .0058285     3.80   0.000     .0107294    .0335766
       exper |   .0261601   .0019735    13.26   0.000      .022292    .0300281
         age |  -.0149448   .0019298    -7.74   0.000    -.0187271   -.0111625
     kidslt6 |  -.2455841   .0282462    -8.69   0.000    -.3009457   -.1902225
     kidsge6 |   -.004455   .0106216    -0.42   0.675     -.025273     .016363
------------------------------------------------------------------------------
```

Tobit

Tobit: Average Partial Effect (APE) on observed hours

```
Average marginal effects                          Number of obs   =        753
Model VCE    : OIM

Expression   : E(hours*|hours>0), predict(ystar(0,.))
dy/dx w.r.t. : nwifeinc educ exper age kidslt6 kidsge6

------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -5.188622    2.62141    -1.98   0.048    -10.32649   -.0507525
        educ |   47.47311    12.6214     3.76   0.000     22.73562    72.21061
       exper |   48.79312   3.587271    13.60   0.000      41.7622    55.82404
         age |  -32.02624   4.292112    -7.46   0.000    -40.43862   -23.61385
     kidslt6 |  -526.2779   64.70622    -8.13   0.000    -653.0997    -399.456
     kidsge6 |   -9.54694   22.75225    -0.42   0.675    -54.14054    35.04665
------------------------------------------------------------------------------
```

Tobit

Compare with OLS results. . .

```
      Source |       SS       df       MS              Number of obs =     753
-------------+------------------------------           F(  7,   745) =   38.50
       Model |  151647606     7  21663943.7           Prob > F      =  0.0000
    Residual |  419262118   745  562767.944           R-squared     =  0.2656
-------------+------------------------------           Adj R-squared =  0.2587
       Total |  570909724   752  759188.463           Root MSE      =  750.18

------------------------------------------------------------------------------
       hours |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    nwifeinc |  -3.446636      2.544    -1.35   0.176    -8.440898    1.547626
        educ |   28.76112   12.95459     2.22   0.027     3.329283    54.19297
       exper |   65.67251   9.962983     6.59   0.000     46.11365    85.23138
     expersq |  -.7004939   .3245501    -2.16   0.031    -1.337635   -.0633524
         age |  -30.51163   4.363868    -6.99   0.000    -39.07858   -21.94469
     kidslt6 |  -442.0899    58.8466    -7.51   0.000    -557.6148    -326.565
     kidsge6 |  -32.77923   23.17622    -1.41   0.158     -78.2777    12.71924
       _cons |   1330.482   270.7846     4.91   0.000     798.8906    1862.074
------------------------------------------------------------------------------
```

Tobit

The Tobit model has a few important limitations.

- The effect of x_j on $E(y|\mathbf{x}, y > 0)$ and $Pr(y > 0|\mathbf{x})$ are both proportional to β_j.

  This rules out situation in which x_j affects both objects in different directions. For instance, the effect of a person's age on the amount of life insurance. Younger people may be less likely to buy life insurance, so $Pr(y > 0|\mathbf{x})$ increases with age. Conditional on having a policy, its value may nonetheless decrease with age, as people get near the end of their lives. In this case, $E(y|\mathbf{x}, y > 0)$ would decrease with age. (For a informal specification test using probits, see book.)

These slides covered:

Wooldridge Chs. 17.1 and 17.2, Stock and Watson Ch.11.