

WWS 508b

Precept 10

John Palmer

April 27, 2010

LPM vs. Logit vs. Probit

Example: married women's labor force participation

The MROZ.dta data set has information on labor force participation and other characteristics of married women in 1975.

- ▶ $inlf = 1$ if respondent reported working for a wage outside home at some point during the year (1975); zero otherwise.
- ▶ $nwifeinc =$ family income excluding respondent's income (in thousands of dollars).
- ▶ $city = 1$ if respondent lived in standard metropolitan statistical area; zero otherwise.
- ▶ $educ =$ respondent's education (in years).
- ▶ $age =$ respondent's age.
- ▶ $kidslt6 =$ number of kids less than 6 years old.

LPM vs. Logit vs. Probit

One dependent variable

How to regress *inlf* on *city* in Stata?

LPM: `regress inlf city, r`

Logit model: `logit inlf city`

Probit model: `probit inlf city`

LPM vs. Logit vs. Probit

One dependent variable

Estimates:

	(1) LPM	(2) logit	(3) probit
city	-0.00638 (0.0377)	-0.0260 (0.154)	-0.0162 (0.0959)
_cons	0.572*** (0.0302)	0.292* (0.123)	0.183* (0.0769)
N	753	753	753

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

LPM vs. Logit vs. Probit

One dependent variable

What is the probability of being in the workforce predicted by each model for city-dwellers? Non-city-dwellers?

$$\text{LPM: } \Pr\{inlf = 1 | city\} = \hat{\beta}_0 + \hat{\beta}_1 city$$

$$\text{Logit model: } \Pr\{inlf = 1 | city\} = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 city}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 city}}$$

$$\text{Probit model: } \Pr\{inlf = 1 | city\} = \Phi(\hat{\beta}_0 + \hat{\beta}_1 city)$$

LPM vs. Logit vs. Probit

One dependent variable

So for city-dwellers, we have:

$$\text{LPM: } \Pr\{inlf = 1 | city=1\} = 0.572 - 0.00638(1) = 0.5661157$$

$$\text{Logit model: } \Pr\{inlf = 1 | city=1\} = \frac{e^{0.292 - 0.0260(1)}}{1 + e^{0.292 - 0.0260(1)}} = 0.5661157$$

$$\text{Probit model: } \Pr\{inlf = 1 | city=1\} = \Phi(0.183 - 0.0162(1)) = 0.5661157$$

LPM vs. Logit vs. Probit

One dependent variable

To do these calculations in Stata:

```
regress inlf city
disp "LPM: Pr{inlf=1|city=1}=" _b[_cons] + _b[city]
```

```
logit inlf city
disp "Logit: Pr{inlf=1|city=1}=" exp(_b[_cons] + _b[city])/(1+exp(_b[_cons] + _b[city]))
```

```
probit inlf city
disp "Probit: Pr{inlf=1|city=1}=" normal(_b[_cons] + _b[city])
```

(Note that the stuff I place in quotation marks in these commands is optional—it's just so that I can keep track of what is being displayed.)

LPM vs. Logit vs. Probit

One dependent variable

Why are all three results the same? Because we have only one independent variable and it's dichotomous. Note that we could get the same result simply with a two-by-two table:

```
. tab inlf city, col
```

```
+-----+
| Key          |
|-----|
|      frequency      |
| column percentage |
+-----+
```

```
=1 if in |
lab frce, | =1 if live in SMSA
1975 |      0      1 |      Total
-----+-----+-----+
0 |      115      210 |      325
  |      42.75      43.39 |      43.16
-----+-----+-----+
1 |      154      274 |      428
  |      57.25      56.61 |      56.84
-----+-----+-----+
Total |      269      484 |      753
      |      100.00      100.00 |      100.00
```

LPM vs. Logit vs. Probit

adding more variables

Now try this:

```
regress inlf city nwifeinc educ age kidslt6  
estimates store LPM
```

```
logit inlf city nwifeinc educ age kidslt6  
estimates store logit
```

```
probit inlf city nwifeinc educ age kidslt6  
estimates store probit
```

```
esttab LPM logit probit, se mtitles
```

LPM vs. Logit vs. Probit

interpreting the results

```
. esttab LPM logit probit, se mtitles
```

	(1) LPM	(2) logit	(3) probit
city	-0.00325 (0.0364)	0.0131 (0.175)	0.0101 (0.106)
nwifeinc	-0.00694*** (0.00154)	-0.0356*** (0.00804)	-0.0214*** (0.00466)
educ	0.0534*** (0.00779)	0.262*** (0.0407)	0.158*** (0.0239)
age	-0.0108*** (0.00233)	-0.0522*** (0.0114)	-0.0315*** (0.00685)
kidslt6	-0.294*** (0.0356)	-1.458*** (0.195)	-0.881*** (0.114)
_cons	0.583*** (0.143)	0.354 (0.691)	0.216 (0.417)
N	753	753	753

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?
 - ▶ Each additional year of age is associated with approximately a 1 percentage point decrease in the probability of participating in the labor force.

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?
 - ▶ Each additional year of age is associated with approximately a 1 percentage point decrease in the probability of participating in the labor force.
- ▶ What is the partial effect of age in the Logit model?

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?
 - ▶ Each additional year of age is associated with approximately a 1 percentage point decrease in the probability of participating in the labor force.
- ▶ What is the partial effect of age in the Logit model?
 - ▶ Each additional year of age decreases the *odds* of participating in the labor force by $100 \times (1 - e^{-0.0522}) \approx 5\%$.

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?
 - ▶ Each additional year of age is associated with approximately a 1 percentage point decrease in the probability of participating in the labor force.
- ▶ What is the partial effect of age in the Logit model?
 - ▶ Each additional year of age decreases the *odds* of participating in the labor force by $100 \times (1 - e^{-0.0522}) \approx 5\%$.

huh?

LPM vs. Logit vs. Probit

understanding the Logit interpretation

To understand why we can interpret the Logit estimates this way, consider the model in terms of the predicted odds of labor force participation:

$$\ln(\widehat{\text{odds}}) = \hat{\beta}_0 + \hat{\beta}_1 \text{city} + \hat{\beta}_2 \text{nwifeinc} + \hat{\beta}_3 \text{educ} + \hat{\beta}_4 \text{age} + \hat{\beta}_5 \text{kidslt6}$$

So that means:

$$\widehat{\text{odds}} = e^{\hat{\beta}_0 + \hat{\beta}_1 \text{city} + \hat{\beta}_2 \text{nwifeinc} + \hat{\beta}_3 \text{educ} + \hat{\beta}_4 \text{age} + \hat{\beta}_5 \text{kidslt6}}$$

or

$$\widehat{\text{odds}} = e^{\hat{\beta}_0} e^{\hat{\beta}_1 \text{city}} e^{\hat{\beta}_2 \text{nwifeinc}} e^{\hat{\beta}_3 \text{educ}} e^{\hat{\beta}_4 \text{age}} e^{\hat{\beta}_5 \text{kidslt6}}$$

LPM vs. Logit vs. Probit

understanding the Logit interpretation

Now compare the ratio of two predicted odds: $\widehat{\text{odds}}_0$ with all variables set to any given values, and $\widehat{\text{odds}}_1$ with *age* increased by 1:

$$\frac{\widehat{\text{odds}}_1}{\widehat{\text{odds}}_0} = \frac{e^{\hat{\beta}_0} e^{\hat{\beta}_1 \text{city}} e^{\hat{\beta}_2 \text{nwifeinc}} e^{\hat{\beta}_3 \text{educ}} e^{\hat{\beta}_4 (\text{age}+1)} e^{\hat{\beta}_5 \text{kidslt6}}}{e^{\hat{\beta}_0} e^{\hat{\beta}_1 \text{city}} e^{\hat{\beta}_2 \text{nwifeinc}} e^{\hat{\beta}_3 \text{educ}} e^{\hat{\beta}_4 \text{age}} e^{\hat{\beta}_5 \text{kidslt6}}}$$

Notice that everything cancels out so that we get:

$$\frac{\widehat{\text{odds}}_1}{\widehat{\text{odds}}_0} = \frac{e^{\hat{\beta}_4 (\text{age}+1)}}{e^{\hat{\beta}_4 \text{age}}} = e^{\hat{\beta}_4}$$

In other words, the odds ratio is equal to $e^{\hat{\beta}_4}$.

LPM vs. Logit vs. Probit

understanding the Logit interpretation

How do we get from the odds ratio to talking about a percentage decrease or increase?

The odds ratio tells us that when we increase age by 1, we get new predicted odds that are $e^{\hat{\beta}_4}$ times our initial predicted odds. If $\hat{\beta}_4$ is negative, $e^{\hat{\beta}_4}$ is less than one, so we can express the change as a *decrease* of $100 \times (1 - e^{\hat{\beta}_4})$ percent. If $\hat{\beta}_4$ is positive, then $e^{\hat{\beta}_4}$ is greater than one, so we can express the change as an *increase* of $100 \times (e^{\hat{\beta}_4} - 1)$ percent.

LPM vs. Logit vs. Probit

interpreting the results

	(1) LPM	(2) logit	(3) probit
city	-0.00325 (0.0364)	0.0131 (0.175)	0.0101 (0.106)
nwifeinc	-0.00694*** (0.00154)	-0.0356*** (0.00804)	-0.0214*** (0.00466)
educ	0.0534*** (0.00779)	0.262*** (0.0407)	0.158*** (0.0239)
age	-0.0108*** (0.00233)	-0.0522*** (0.0114)	-0.0315*** (0.00685)
kidslt6	-0.294*** (0.0356)	-1.458*** (0.195)	-0.881*** (0.114)
_cons	0.583*** (0.143)	0.354 (0.691)	0.216 (0.417)
N	753	753	753

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?
 - ▶ Each additional year of age is associated with approximately a 1 percentage point decrease in the probability of participating in the labor force.
- ▶ What is the partial effect of age in the Logit model?
 - ▶ Each additional year of age *decreases* the odds of participating in the labor force by $100 \times (1 - e^{-0.0522}) \approx 5\%$.

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?
 - ▶ Each additional year of age is associated with approximately a 1 percentage point decrease in the probability of participating in the labor force.
- ▶ What is the partial effect of age in the Logit model?
 - ▶ Each additional year of age *decreases* the odds of participating in the labor force by $100 \times (1 - e^{-0.0522}) \approx 5\%$.
- ▶ What is the partial effect of education in the Logit model?

LPM vs. Logit vs. Probit

interpreting the results

- ▶ What is the partial effect of age in the LPM?
 - ▶ Each additional year of age is associated with approximately a 1 percentage point decrease in the probability of participating in the labor force.
- ▶ What is the partial effect of age in the Logit model?
 - ▶ Each additional year of age *decreases* the odds of participating in the labor force by $100 \times (1 - e^{-0.0522}) \approx 5\%$.
- ▶ What is the partial effect of education in the Logit model?
 - ▶ Each additional year of education *increases* the odds of participating in the labor force by $100 \times (e^{0.262} - 1) \approx 30\%$.

LPM vs. Logit vs. Probit

interpreting the results

Can we interpret the Probit model in terms of odds? Not easily.

How about if we want to interpret the Logit or Probit models in terms of the partial effect on probability? Now we need to specify the values of all the variables at which we are interested in the effect.

One simple approach is to set all variables to their average values in the sample.

LPM vs. Logit vs. Probit

interpreting the results

In Stata, to calculate partial effects for each variable with all variables set to the average, use the following command after running the regression:

```
mf
```

```
Marginal effects after logit  
y = Pr(inlf) (predict)  
= .57420164
```

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
city*	.0032034	.0427	0.08	0.940	-.080481	.086888	.642762	
nwifeinc	-.0087062	.00197	-4.42	0.000	-.012564	-.004848	20.129	
educ	.0640876	.00995	6.44	0.000	.044591	.083584	12.2869	
age	-.0127634	.00279	-4.58	0.000	-.01823	-.007297	42.5378	
kidslt6	-.3565846	.04782	-7.46	0.000	-.450309	-.26286	.237716	

(*) dy/dx is for discrete change of dummy variable from 0 to 1

LPM vs. Logit vs. Probit

interpreting the results

But often we would prefer to know the partial effects for specific values of certain variables. For instance, in our example, the partial effect with *city* set to its average isn't particularly useful.

LPM vs. Logit vs. Probit

interpreting the results

We do this by adding additional the `at` option. Here will will specify that all partial effects are to be evaluated with *city* set to 1 and *age* set to 34. All other variables stay set to their average values. So we are looking at partial effects for 34-year old city-dwellers with average family income, education and number of kids under 6:

```
mfx, at(city=1 age=34)
```

```
warning: no value assigned in at() for variables nwifeinc educ kidslt6;  
means used for nwifeinc educ kidslt6
```

```
Marginal effects after logit  
y = Pr(inlf) (predict)  
= .67904757
```

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
city*	.0028614	.03815	0.08	0.940	-.071903	.077626		1
nwifeinc	-.0077607	.00177	-4.38	0.000	-.011234	-.004288		20.129
educ	.0571276	.00945	6.05	0.000	.038614	.075641		12.2869
age	-.0113773	.00209	-5.44	0.000	-.015479	-.007276		34
kidslt6	-.3178594	.04039	-7.87	0.000	-.397024	-.238694		.237716

(*) dy/dx is for discrete change of dummy variable from 0 to 1

LPM vs. Logit vs. Probit

interpreting the results

Another useful approach is to calculate the average partial effects—meaning the average of the partial effects predicted at all values within the sample. *The questions in the problem set asking you to do this are optional.* However, if you are curious, I have included in the .do file posted with these slides an example of how to do it in Stata, drawing on Wooldridge's equations 17.15 and 17.17.

Tobit basics

To fit a Tobit model in Stata use same basic syntax but add a comma and specify the lower limit (ll) or upper limit (ul) of the data—i.e., where it is censored. For example:

```
tobit hours educ age, ll(0)
```

To test joint significance or linear hypotheses:

```
test educ age  
test educ + age = 0
```

To calculate the average partial effect scale factor:

```
gen effect = normal((_b[_cons] + _b[educ]*educ + _b[age]*age)/_b[/sigma])  
mean(effect)  
  
scalar APEscalar = _b[effect]
```

Tobit basics

To obtain estimates of $E(\text{hours}|\mathbf{x})$:

```
tobit hours educ age, ll(0)
```

```
gen hourshat = normal((_b[_cons] + _b[educ]*educ + _b[age]*age)/_b[/sigma])*(_b[_cons] ///  
+ _b[educ]*educ + _b[age]*age) + _b[/sigma]*normalden((_b[_cons] + _b[educ]*educ + ///  
_b[age]*age)/_b[/sigma])
```

```
sum hourshat
```