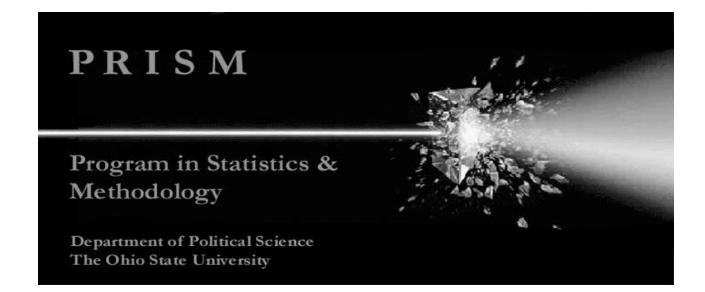
#### **Advanced Programming in Stata**

Kevin Sweeney PRISM Senior Methods Fellow

#### Brandon Bartels PRISM Junior Methods Fellow



#### Advanced Programming in Stata

- Programming your own maximum likelihood estimator.
  - Basic syntax
  - Likelihood functions
- Examples:
  - Normal regression (easy)
  - Logit and probit (easy)
  - Heteroskedastic regression (harder)
  - Split population duration model (harder)

#### Programming Likelihood Functions: The Basics

- As you will see, programming your own ML estimator is incredibly easy to do in Stata.
- From last session, we learned how to write a program in Stata using .do files, macros, looping, etc.
- In writing our own likelihood function, we need the following information:
  - An understanding of some of Stata's "ml" family of commands.
    - *Note*: The help menus provide very useful information on MLE programming; *help ml* and/or *help mlmethod*
  - Log-likelihood function
  - Syntax for how to maximize the function
  - THAT'S IT! It's so easy, it's hard to believe!

#### Programming Likelihood Functions: Brief MLE Review

- In ML, we first need to specify the data generating process for the dependent variable under examination.
- In other words, we need to specify the probability distribution that generated the dependent variable; e.g., the normal for continuous variable, logit or probit for dichotomous, poisson for count data, etc.

#### Programming Likelihood Functions: Brief MLE Review

• Then, we specify the likelihood for case *i*:

 $L_i = L(\boldsymbol{q} \mid y_i)$  $L(\boldsymbol{q} \mid y_i) \propto p(y_i \mid \boldsymbol{q})$ 

• The likelihood for the entire sample is simply the product of individual likelihoods:  $L = \prod_{i=1}^{N} L_i$ 

• MLEs are the values of the parameters for which the likelihood of observing the sample is maximized.

•  $Y \sim N(\mu, s^2)$ 

• pdf: 
$$f(y_i | \mathbf{m}_i, \mathbf{s}^2) = \frac{1}{\sqrt{2\mathbf{p}\mathbf{s}^2}} e^{-\frac{1}{2}(y_i - \mathbf{m}_i)^2}$$

- Reparameterize  $\mu_i = x_i \beta$
- Likelihood for case *i* :

$$L_{i} = \frac{1}{\sqrt{2ps^{2}}} e^{-\frac{1}{2} (y_{i} - x_{i}b)^{2}} s^{2}}$$

• Log-likelihood for case *i* (\*\*this is what Stata wants\*\*):

$$\ln L_{i} = -\frac{1}{2}\ln(2\mathbf{p}) - \frac{1}{2}\ln(\mathbf{s}^{2}) - \frac{1}{2}\left[\frac{(y_{i} - x_{i}\mathbf{b})^{2}}{\mathbf{s}^{2}}\right]$$

• The likelihood for the entire sample is simply the product of the individual likelihoods:

$$L = \prod_{i=1}^{N} \frac{1}{\sqrt{2ps^{2}}} e^{-\frac{1}{2} \frac{(y_{i} - x_{i}b)^{2}}{s^{2}}}$$

• And the log-likelihood for the entire sample is simply:

$$\ln L = -\frac{N}{2}\ln(2\mathbf{p}) - \frac{N}{2}\ln(\mathbf{s}^{2}) - \frac{1}{2}\sum_{i=1}^{N} \left[\frac{(y_{i} - x_{i}\mathbf{b})^{2}}{\mathbf{s}^{2}}\right]$$

• Again, however, Stata only needs the log-likelihood for case *i*.

# Programming Likelihood Functions: Syntax

- Goal: Write a program that Stata can use to maximize a log-likelihood function.
- First, Stata has 4 ML "evaluators": lf, d0, d1, d2.
- "lf" is the most basic evaluator; the "d" evaluators are for more advanced programs. We're only going to use "lf" in this session.

# Programming Likelihood Functions: Syntax

```
program define progname
args lnf theta1 theta2 ...
tempvar tmp1 tmp2 ...
quietly gen double `tmp1' = ...
quietly replace `lnf' = ...
end
```

- `lnf' is a variable to be filled in with values of the log-likelihood for case *i* (i.e., lnL<sub>i</sub>).
- `theta1' is associated with the first parameter, containing evaluation of the 1st equation: theta $1_i = x_{1i}b$
- `theta2' is associated with the second parameter, containing evaluation of the 2nd equation: theta $2_i = x_{2i}b$

# Programming Likelihood Functions: Syntax

• Global macros:

\$ML\_y1 is a global macro for the name of the first dependent variable.

\$ML\_y2 is a global macro for the name of the second dependent variable.

#### Onto the Machines: Start a .log File



#### Onto the Machines: Start a .log File

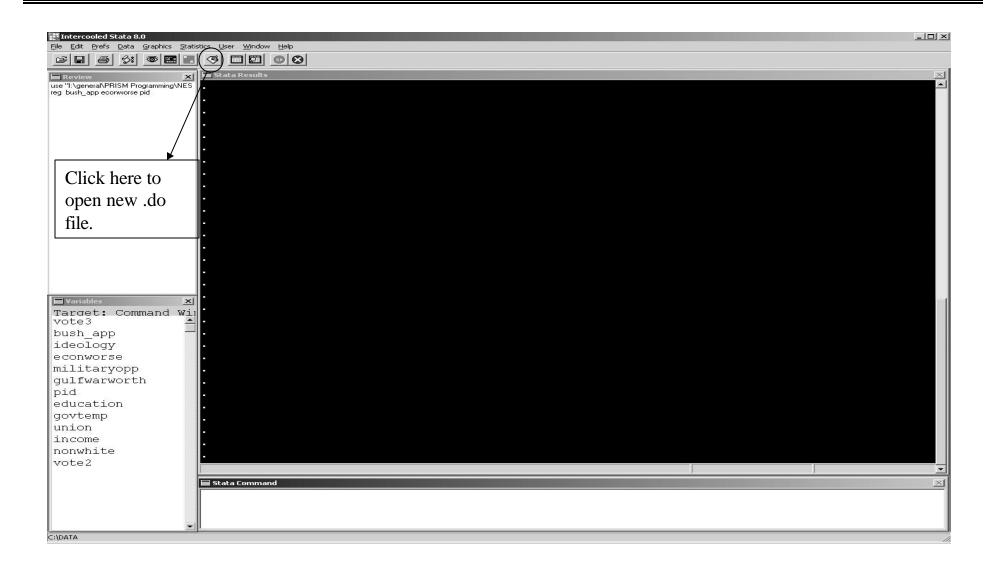
Begin Logging St	ata Output			<u>? ×</u>
Savejn:	🖵 Bartels on 'Pol1\Vol3\Gra	ads' (K:) 💽 🗢	t 🖄 🎫	
History Desktop My Documents Derby 086 My Network P	<ul> <li>506</li> <li>686</li> <li>703</li> <li>707</li> <li>709</li> <li>Change to .log</li> <li>715</li> <li>716</li> <li>717</li> <li>717 - Wright</li> <li>717 - Wright</li> <li>775</li> <li>778</li> <li>File name:</li> <li>Save as type:</li> </ul>	<ul> <li>788</li> <li>793 McGraw</li> <li>Baum Audiences book</li> <li>Baum Project</li> <li>Bayes - Franklin</li> <li>Beck et al stuff</li> <li>Chevron Data</li> <li>DeBoef Panel Data Notes</li> <li>default</li> <li>Dissertation stuff - Beck</li> <li>Econ 742</li> </ul>		àave

- Let's open some data: 1992 NES
- File, Open
  - Go to the I: drive

Double-click on "general"

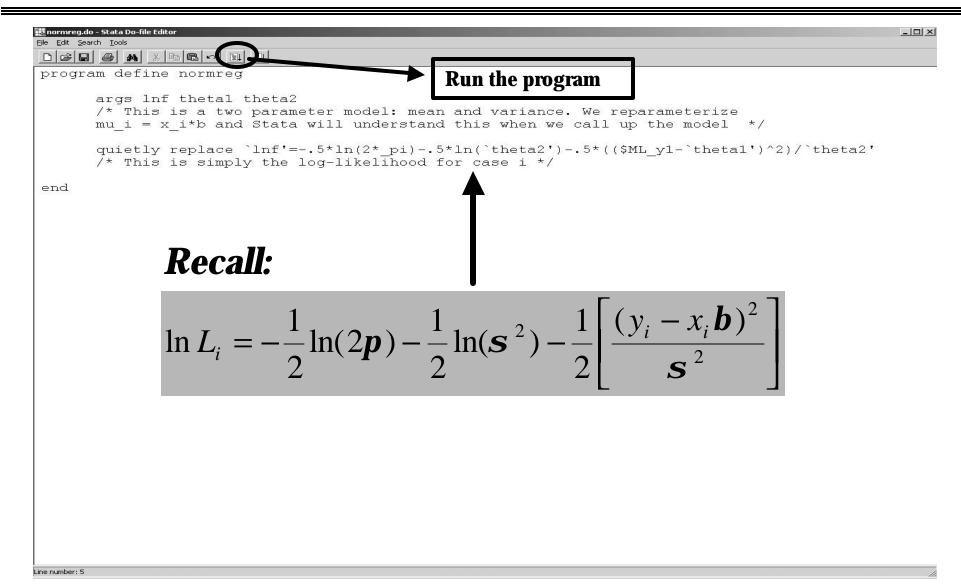
Double-click on "PRISM Programming"

Double-click on "NES 1992.dta"



 Open "normreg" program from the .do file editor: Go to the I: drive Double-click on "general" Double-click on "PRISM Programming"

Double-click on "normreg.do"



Programming Likelihood Functions: Maximizing the Likelihood Function

- Once we've written the program, we need to tell Stata to estimate it. This takes two steps:
- (1) ml model lf progname (eq1: y=x1 x2 x3)

- *or* –

ml model lf progname (eq1: y=x1 x2 x3) (eq2: y=x1 x2 x3)

#### - *or* –

*ml model lf progname (eq1: y=x1 x2 x3) / parameter* 

[If the second parameter is not reparameterized as a function of covariates, e.g., s<sup>2</sup> in ML normal regression.]

(2) *ml max* 

# Programming Likelihood Functions: Maximizing the Likelihood Function

 Other useful commands to run after *ml model*: *ml check* verifies that the program you wrote works *ml search* searches for better starting values
 *lf0(#k LL0)* reports a likelihood ratio test (included after the "ml model" command), comparing fully specified model to an intercept only (i.e., null) model. The Wald test is produced by default. For the LR test, you need to specify the LL and the number of parameters for the intercept only model.

• Let's estimate a simple model; we'll regress George H.W. Bush's approval on PID and economic perceptions.

Intercooled Stata 8.0 File Edit Prefs Data Graphics Statist	ics User Window Help
Review X use "I:\general\PRISM Programming\NES reg bush_app econworse pid do "C:\DOCUME~1\ADMINI~1\LOCALS~	Stata Results
■ Variables × Target: Command Win vote3	<pre>. do "C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\STD040000.tmp" . program define normreg 1 args lnf thetal theta2 2. /* This is a two parameter model: mean and variance. We reparamete &gt; ize &gt; mu_i = x_i*b and Stata will understand this when we call up the model &gt; */</pre>
income nonwhite	<pre>. quietly replace `lnf'=5*ln(2*_pi)5*ln(`theta2')5*((\$ML_y1-`thet &gt; 1')^2)/`theta2' 3.</pre>
	<pre>&gt;&gt; &gt;&gt; &gt;&gt; &gt;&gt; &gt;&gt; &gt;&gt; &gt;&gt; &gt;&gt; &gt;&gt; &gt;&gt;</pre>

🔣 Intercooled Stata 8.0		. D ×
Eile Edit Prefs Data Graphics Statistics User		
E B S 🕸 🖲 📰 🦉 🗖		
Review Stata	Results	$\times$
reg bush app econworse pid	ogram define normreg	
do "C:\DOCUME~1\ADMINI~1\LOCALS~		
mi model i nomreg (reg. bush_app = pid ec mi check clear		
use "I:\general\PRISM Programming\NES		
program drop normreg	of do-file	
do "C:\DOCUME~1\ADMINI~1\LOCALS~ ml model lf normreg (reg: bush_app = pid ec		
· pro	ogram drop normreg	
do	"C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\STD040000.tmp"	0.00
- 40	C. (DOCOMENT (ADMININI (DOCALDAT (TEMP (STD040000. Cmp	
, pri	ogram define normreg	
1.		
	args lnf thetal theta2	
2.	/* This is a two parameter model: mean and variance. We reparamete	er
> iz:		
>	mu_i = x_i*b and Stata will understand this when we call up the mode:	La
Variables ×		
Target: Command Will		
vote3	<pre>quietly replace `lnf'=5*ln(2*_pi)5*ln(`theta2')5*((\$ML_y1-`theta')</pre>	ca
	)^2)/`theta2'	
ideology 3.	/* This is simply the log-likelihood for case i */	
econworse .		
militaryopp . end gulfwarworth		
pid		
-	of do-file	
qovtemp		
union . ml	model lf normreg (reg: bush app = pid econworse) /sigmasq	
income		
nonwhite .		
vote2		-
_MLtu1	-omma 4	×
	Maximizes the function and produces model estimates.	
C:\DATA		1

<pre>prove will reaction to be a service of the ser</pre>	"I:\general\PRISM Programming\NES bush_app_econworse_pid	🗖 Stata Results						
Wardeles       Xumber of obs = 382 Wald chi2(2) = 257.99 Drob > chi2 = 0.0000         iVerdeles       Xumber of obs = 382 Wald chi2(2) = 257.99 Drob > chi2 = 0.0000         isignasq       bush_app         conworse       i.2781664         ild       econworse         id       .2781664         ush_app       .2781664         ots3       .2781664         .2781664       .0220503       12.62         .0000       .2349487       .3213841         econworse      245331       .05101         _cons       3.22732       .2080806       15.51         0.0000       2.81949       3.635151         id       .cons       .7088946       .0512938       13.82         onon       .cons       .7088946       .0512938       13.82	nodel if normreg (reg. bush_app = pid ec theck ar "I:sgeneral\PRISM Programming\NES "C:\D0CUME™\4DMINI™\L0CALS~ gram drop normreg "C:\D0CUME™\4DMINI™\L0CALS~ dodel if normreg (reg: bush_app = pid ec	. ml max initial: feasible: rescale: rescale eq: Iteration 0: Iteration 1: Iteration 2: Iteration 3: Iteration 4:	log likeliho log likeliho log likeliho log likeliho log likeliho log likeliho log likeliho	pod = -1872. $pod = -603.99$ $pod = -592.09$ $pod = -592.09$ $pod = -503.09$ $pod = -477.8$ $pod = -476.39$ $pod = -476.39$	1434 2563 3452 3452 3944 9919 3546 2127	ould not :	pe evaluated)	
ote3				7		Wald	chi2(2) =	257.99
conworse       pid       .2781664       .0220503       12.62       0.000       .2349487       .3213841         ilitaryopp       econworse      245331       .05101       -4.81       0.000      3453087      1453533         ulfwarworth       _cons       3.22732       .2080806       15.51       0.000       2.81949       3.635151         id	ote3 ush_app	bush_app	Coef.	Std. Err.		P>   z	[95% Conf.	Interval]
ducation ovtemp nion ncome onwhite	conworse ilitaryopp ulfwarworth	pid econworse	245331	.05101	-4.81	0.000	3453087	1453533
	ducation ovtemp nion ncome onwhite		.7088946	.0512938	13.82	0.000	.6083606	.8094286

Programming Likelihood Functions: Comparing ML Reg to OLS

• OLS and ML Normal Regression produce identical parameter estimates. It can be shown that the analytical solution for ML Normal Regression is:

 $\beta = (X'X)^{-1}X'Y$ 

which is identical to the well-known formula for the OLS estimator.

• Standard errors will be different, though, because:

- In ML: 
$$\mathbf{s}^{2} = \frac{\sum_{i=1}^{N} e_{i}^{2}}{N}$$
  
- In OLS: 
$$\mathbf{s}^{2} = \frac{\sum_{i=1}^{N} e_{i}^{2}}{\mathbf{s}^{2}}$$

# Programming Likelihood Functions: Comparing ML Reg to OLS

Ele Edit Prefs Data 8.0 Ele Edit Prefs Data Graphics Statistics User Window Help							_[]×
	8						
■ Review       ★         Use "I:\general\PRISM Programming\NES reg bush_app convose pid do "C\DDCUME~I\ADMINI~I\LDCALS~ mI model in nomreg (reg: bush_app = pid ec mi check.       ml max         .       ml max         .	log likeliho log likeliho log likeliho log likeliho	$\begin{array}{rcl} \text{bod} &=& -1872.\\ \text{bod} &=& -603.9\\ \text{bod} &=& -592.0\\ \text{bod} &=& -592.0\\ \text{bod} &=& -503.0\\ \text{bod} &=& -477.8\\ \text{bod} &=& -476.3\\ \text{bod} &=& -476.3\\ \text{bod} &=& -476.3 \end{array}$	1434 2563 3452 3452 8944 9919 3546 2127	ould not	be evaluated)		×
	od = -476.3212			Wald Prob	r of obs = chi2(2) = > chi2 =	382 257.99 0.0000	
□ Variables 🗵 bush_app	Coef.	Std. Err.	Z	P> z	[95% Conf.	Interval]	
Target: Command Will vote3 bush_app ideology econworse militaryopp	245331	.0220503 .05101 .2080806	12.62 -4.81 15.51	0.000 0.000 0.000	.2349487 3453087 2.81949	.3213841 1453533 3.635151	
gulfwarworth sigmasq	.7088946	.0512938	13.82	0.000	.6083606	.8094286	
education govtemp union · income nonwhite vote2							
Cibata	app pid ecor	nworse		Est	imate this C	DLS.	×

# Programming Likelihood Functions: Comparing ML Reg to OLS

e <b>rcooled Stata 8.0</b> dit <u>P</u> refs <u>D</u> ata <u>G</u> raphics <u>S</u> tat	istics <u>U</u> ser <u>W</u> indow <u>H</u> elp							
<b>-</b> 4 9: • • • •								
view 🔀 general\PRISM Programming\NES		STREET, ST						
sh_app econworse pid DOCUME~1\ADMINI~1\LOCALS~								
el lf normreg (reg: bush_app = pid ec sk	bush_app	Coef.	Std. Err.	Z	₽> z	[95% Conf.	Interval]	
general\PRISM Programming\NES DOCUME~1\ADMINI~1\LOCALS~	reg							
n drop normreg .DOCUME~1\ADMINI~1\LOCALS~ el lf normreg (reg: bush_app = pid ec		.2781664 245331	.0220503 .05101	12.62 -4.81	0.000 0.000	.2349487 3453087	.3213841 1453533	
h_app pid econworse	econworse _cons	3.22732	.2080806	-4.81 15.51	0.000	2.81949	3.635151	
	sigmasq							
	_cons	.7088946	.0512938	13.82	0.000	.6083606	.8094286	
	. reg bush_app	pid econwors	e					
	Source	SS	df	MS		Number of obs		
riables 🔀	Carles Contractor	182.888133	2 91 4	440663		F( 2, 379) Prob > F	= 127.98 = 0.0000	
cet: Command Wi	Residual	270.797731		505887		R-squared	= 0.0000	
h_app —						Adj R-squared		
ology nworse	Total	453.685864	381 1.19	077655		Root MSE	84528	
taryopp								
Twarworth	bush_app	Coef.	std. Err.	t	P> t	[95% Conf.	Interval]	
cation	pid	.2781664	.0221374	12.57	0.000	.234639	.3216938	
temp	econworse	245331	.0512114	-4.79	0.000	3460251	1446368	
on		3.22732	.2089025	15.45	0.000	2.816567	3.638073	
ome nite								
nite 2								
_								]
	Stata Command							
		<b>Coefficie</b>	ıts are ide	ntical.	SEs an	e different.		
	1							

- In binary response models, we want to model the probability of "success" for case *i*, i.e.,  $Pr(y_i=1) = ?_i$
- We parameterize ?  $_i$  as a cumulative distribution function (cdf) of a particular distribution, i.e.,  $F(x_i\beta)$ 
  - For logit, we use the logistic cdf:

$$\Pr(y_i = 1) = F(x_i \boldsymbol{b}) = \frac{\exp(x_i \boldsymbol{b})}{1 + \exp(x_i \boldsymbol{b})}$$

- For probit, we use the normal cdf:  $Pr(y_i = 1) = F(x_i \boldsymbol{b}) = \Phi(x_i \boldsymbol{b})$ 

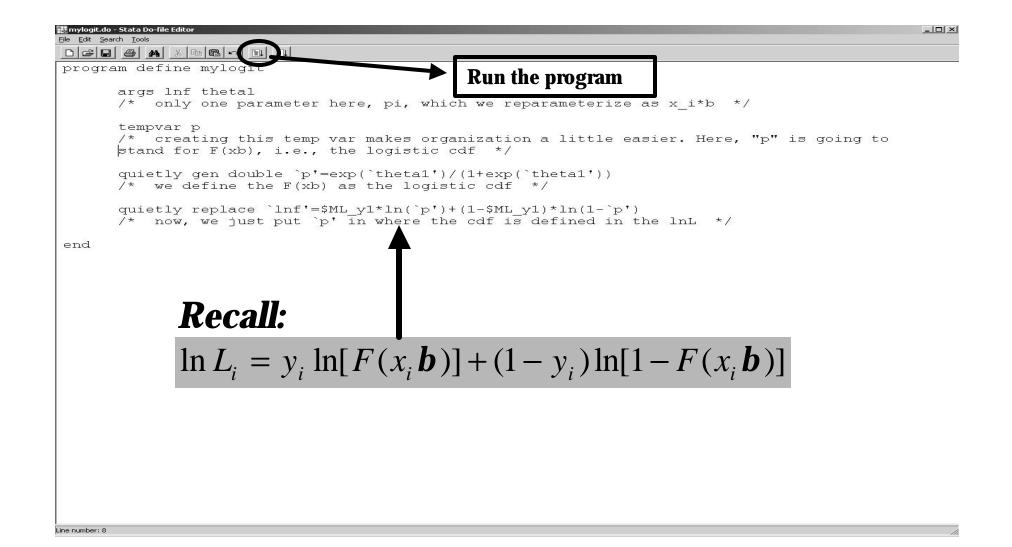
- The likelihood for case *i* is:  $L_i = [F(x_i \boldsymbol{b})]^{y_i} [1 - F(x_i \boldsymbol{b})]^{1-y_i}$
- The log-likelihood for case *i* is:

 $\ln L_i = y_i \ln[F(x_i \boldsymbol{b})] + (1 - y_i) \ln[1 - F(x_i \boldsymbol{b})]$ 

\*\*Again, this is what we're going to give Stata\*\*

• For logit, we'll replace  $F(x_i \beta)$  with the logistic cdf, and for probit, the normal cdf.

 Open "mylogit.do" from the .do file editor. Go to the I: drive Double-click on "general" Double-click on "PRISM Programming" Double-click on "mylogit.do"



والتاليج المحاد	ata Results				_			
iew X Loss app econworse pid DCUME~1\ADMINI~1\DCALS~ If normeg freg: bush app = pid ec	Total	453.685864	381 1.19	077655		Root MSE	84528	
eneral\PRISM Programming\NES	bush_app	Coef.	Std. Err.	t	₽> t	[95% Conf.	Interval]	
OCUME~1\ADMINI~1\LOCALS~ drop normreg OCUME~1\ADMINI~1\LOCALS~ If normreg (reg: bush_app = pid ec	pid econworse	.2781664 245331	.0221374 .0512114	12.57 -4.79	0.000	.234639 3460251	.3216938 1446368	
_app pid econworse OCUME~1\ADMINI~1\LOCALS~	_cons	3.22732	.2089025	15.45	0.000	2.816567	3.638073	
		~1\ADMINI~1\	LOCALS~1\To	mp\ 970//	0000 ±mp"			
			LOCALS~I(Ie	mp\51D04(	p			
	rogram defir •	e mylogit						
•		lnf thetal	arameter he	re ni 1	which we	reparameteriz	⊳ as v i*h	*/
-				re, pr, «	ont cn. we	reparameterrz		
•	tempy		his temp va	r makes d	organizat	ion a little	easier. Here	. "p" is ao
iables 🔀 >		F(xb), i.e.						
et: Command Wil	guiet	ly gen doubl	.e `p'=exp(`	theta1')/	/(1+exp(`	theta1'))		
		we define						
ology . nworse .	quiet	ly replace `	lnf'=\$ML v1	*ln(`p')-	⊦(1-\$ML y	1)*ln(1-`p')		
itaryopp						is defined in	the lnL */	
fwarworth	nd							
cation								
temp •								
on enc	of do-file							
white								
e2 -								
	La command					,		
				L 0		nworse)		

ntercooled Stata 8.0 Edit Prefs Data Graphics Statis	tics <u>U</u> ser <u>W</u> indow <u>H</u> elp	_	_	-	-			_
; 🖬 🎒 🛠 🚳 🔜 📰								
AS/150	Stata Results							
"I:\general\PRISM Programming\NES bush_app_econworse_pid C:\DDCUME~1\ADMINI~1\LDCALS~ odel If normreg (reg: bush_app = pid_ec	bush_app	Coef.	std. Err.	t	₽> t	[95% Conf	. Interval]	
eck	pid	.2781664	.0221374	12.57	0.000	.234639	.3216938	
1:\general\PRISM Programming\NES 2:\DOCUME~1\ADMINI~1\LOCALS~	econworse	245331	.0512114	-4.79	0.000	3460251	1446368	
am drop normreg ::\DOCUME~1\ADMINI~1\LOCALS~ odel lf normreg (reg: bush_app = pid ec		3.22732	.2089025	15.45	0.000	2.816567	3.638073	
ax ush_app pid econworse :\DDCUME~1\ADMINI~1\LDCALS~ :del lf mylogit (logit: vote2 = pid econwi	. do "C:\DocuM	IE~1\ADMINI~1	LOCALS~1\Te	mp\STD04	0000.tmp"			
	. program defi	ne mylogit.						
	1.							
		lnf thetal * onlv one i	parameter be	re ni i	which we	renarameteri	ze as x i*b	*/
		onry one j	Sarameter ne	, pr, ,	mitten we	reparameterr		· · · · · · · · · · · · · · · · · · ·
	. temp	ovar p						
						ion a little	easier. Here	e, "p" is goin
	> to stand for	or F(xb), i.e	., the logis	tic cdf	*/			
Variables X	. auie	tly gen doub	le `n'=eyn(`	thetal')	/(1+eyn(`	thetal'))		
arget: Command Wi		* we define						
ote3 🔺	•		ne veze r <u>e</u> ntra a		100 - 1000 C			
ish_app —		tly replace * now, we از					n + b = 1 n T + t	,
leology :onworse	5. /	^ now, we ji	ust put pr	ru wuere	the car	ra derrued r	n une ini ^/	
litaryopp	. end							
lfwarworth								
.d								
lucation	•							
vtemp	end of do-file							
ion	ml model lf	mylogit (log:	it. voto? -	nid ocen	Jonse)			
come	. Mit moder II	myrogre (10g.	LL. VOLEZ -		MOLSE,			
nwhite								
ote2								
ILtul	🖻 əcata Lomma 🖪						,	
4LW1								
	ml max	- May	imizos tho	functi	on and	araducas r	nodel estin	natas
			mizes die		un and j	produces I	uvuei estili	iales.
(mark)	1							

Eile Edit Prefs Data Graphics Statis								
use "It/general/PRISM Programming/NES treg bush_app econvorse pid do "C:\DOCUME "I\ADMINI" I\LOCALS" mi model if nomreg (reg: bush_app = pid eci mi check clear use "I:/general/PRISM Programming/NES do "C:\DOCUME "I\ADMINI" I\LOCALS" mi model if nomreg do "C:\DOCUME "I\ADMINI" I\LOCALS" mi model if nomreg (reg: bush_app = pid eci mi max reg bush_app pid econvorse mi model if mylogit (logit: vote2 = pid econwing max	end of do-file . ml model lf r . ml max initial: alternative: rescale: Iteration 0:	log likeliho log likeliho log likeliho log likeliho	od = -264.7 od = -254.0 od = -254.0 od = -254.0 od = -254.0	8222 9741 9741 9741 9741	vorse)			
Target: Command Wij vote3 bush_app	Iteration 1: Iteration 2: Iteration 3: Iteration 4: Iteration 5: Log likelihood	<pre>log likeliho log likeliho log likeliho log likeliho log likeliho </pre>	d = -162.4 d = -162.4 d = -162.4 d = -162.4 d = -162.4	9286 5214 5212	Wald	er of obs = chi2(2) = > chi2 =	382 105.87 0.0000	
gulfwarworth pid education govtemp union	vote2 pid econworse _cons	Coef. .8437178 3076029 .2923819	std. Err. .0878516 .1606638 .6554855	2 9.60 -1.91 0.45	P> z  0.000 0.056 0.656	[95% Conf. .6715318 6224982 992346	Interval] 1.015904 .0072923 1.57711	
income nonwhite vote2	• Stata Command							

Review X								
e "I:'sgeneral\PRISM Programming\NES g bush app econverse pid ."C:\DDCUME~1\ADMINI~1\LDCALS^ model If nomreg (reg: bush_app = pid ec check	Log likelihood	1 = -162.4521	2			chi2(2) = > chi2 =	105.87 0.0000	
kar e ''I:\general\PRISM Programming\NES ''C:\DOCUME~1\ADMINI~1\LOCALS^ igram drop normreg	vote2	Coef.	Std. Err.	Z	P> z	[95% Conf.	Interval]	
"C:\DOCUME~1\ADMINI~1\LOCALS~ model If normreg (reg: bush_app = pid ed		.8437178	.0878516	9.60	0.000	.6715318	1.015904	
max j bush_app pid econworse ''C:\DDCUME~1\ADMINI~1\LDCALS^ model If mylogit (logit: vote2 = pid econw max	econworse 	3076029 .2923819	.1606638 .6554855	-1.91 0.45	0.056 0.656	6224982 992346	.0072923 1.57711	
vote2 pid econworse	. logit vote2	pid econwors	9					
	Iteration 0:	log likelih	-	0798				
	Iteration 1:		-169.3					
	Iteration 2:	log likelih	-162.8	4667				
	Iteration 3:	log likelih	-162.4	5449				
	Iteration 4:	log likelih	-162.4	5212				
	Iteration 5:	log likelih	pod = -162.4	5212				
	Will Logit estimates					r of obs = i2(2) =	382 183.26	
oush_app .deology conworse	Log likelihood	1 = -162.4521	2		Prob Pseud			
nilitaryopp gulfwarworth	vote2	Coef.	std. Err.		P> z	[95% Conf.	Interval]	
oid education	pid	.8437178	.0878516	9.60	0.000	.6715318	1.015904	
jovtemp	econworse	3076029	.1606638	-1.91	0.056	6224982	.0072923	
inion	_ <sup>cons</sup>	.2923819	.6554855	0.45	0.656	992346	1.57711	
ncome onwhite							18	
onwhile vote2	-							
otez	J							
	🖬 Stata Command							
	1							

### Programming Likelihood Functions: The Likelihood Ratio Test

- By default, "ml model" produces a Wald test for overall goodness of fit test (which tests that the coefficients are jointly equal to zero).
- To get an LR test, we need to:
  - Estimate an intercept only model to get LL0, the initial LL.
  - We need to specify *k* for the intercept-only model, which in this case is 1.
  - After the "ml model" command, we enter *lf0(k LL0)*.

#### Programming Likelihood Functions: The Likelihood Ratio Test

Intercooled Stata 8.0 File Edit Prefs Data Graphics Stati	tictics Ticer Window Help	_02
ml model If normreg (reg: bush_app = pid ec ml check	Wald chi2(2)=105.87Log likelihood = -162.45212Prob > chi2=0.0000	×
clear use "I\general\PRISM Programming\NES do "C:\DDCUME~1\ADMINI~1\LDCALS~ program drop nomreg do "C:\DDCUME~1\ADMINI~1\LDCALS~	Vote2 Coel. Std. Brr. Z P> Z  [95% Conl. Interval]	
mi model if normeg (reg: bush_app = pid ec mi max reg bush_app pid econworse do "C\DDCUME"\ADMINI"\LDCALS" <u>mi model if mylogit (logit vote2 = pid econwi</u> max logit vote2 pid econworse	econworse3076029 .1606638 -1.91 0.0566224982 .0072923	
n na 🖉 dae ze da ko	. logit vote2 pid econworse	
	Iteration 0: log likelihood = -254.0798 Iteration 1: log likelihood = -169.39683 Iteration 2: log likelihood = -162.84667 Iteration 3: log likelihood = -162.45449 Iteration 4: log likelihood = -162.45212 Iteration 5: log likelihood = -162.45212	
Target: Command Wi vote3 bush_app ideology		
econworse militaryopp gulfwarworth	vote2 Coef. Std. Err. z P> z  [95% Conf. Interval]	
pid education govtemp union income	pid .8437178 .0878516 9.60 0.000 .6715318 1.015904 econworse3076029 .1606638 -1.91 0.0566224982 .0072923 _cons .2923819 .6554855 0.45 0.656992346 1.57711	
nonwhite vote2	•	
	ml model lf mylogit (logit: vote2 =) Estimates intercept only model	
C:\DATA		

#### Programming Likelihood Functions: The Likelihood Ratio Test

view 🔀	Stata Results									
sgeneral/PRISM Programming/NES sh_app econworse pid DDCUME~1/ADMINI~1/LDCALS~ el lf normreg [reg: bush_app = pid ec sk sgeneral/PRISM Programming/NES	Log likelihood	= -162.4521	2		LR ch Prob Pseud	> chi2 =	183.26 0.0000 0.3606			
DOCUME~1\ADMINI~1\LOCALS~ n drop normreg .DOCUME~1\ADMINI~1\LOCALS~ el lf normreg (reg: bush_app = pid ec		Coef.	Std. Err.		₽> z	[95% Conf.	Interval]			
h_app pid econworse DOCUME*1\ADMINI*1\LOCALS* al li mylogit (logit: vote2 = pid econw te2 pid econworse el li mylogit (logit: vote2 =)	pid econworse <sup>cons</sup>	.8437178 3076029 .2923819	.0878516 .1606638 .6554855	9.60 -1.91 0.45	0.000 0.056 0.656	.6715318 6224982 992346	1.015904 .0072923 1.57711			
	. ml model lf mylogit (logit: vote2 =)									
riables Xi rget: Command Win e3 App	and the second comparison and the second	log likelin log likelin log likelin log likelin log likelin log likelin	pod = -254.0 pod = -254.0 pod = -254.0 pod = -254.0 pod = -254.0	9741 9741 9741						
ology nworse itaryopp fwarworth	Log likelihood	= -254.079	3		Wald	r of obs = chi2(0) = > chi2 =	382 •			
l Ication	vote2	Coef.	Std. Err.	Z	P> z	[95% Conf.	Interval]			
rtemp .on :ome white	cons	4802252	.1052929	-4.56	0.000	6865956	2738548			
e2	• Stata Command		ci		0 <b>0</b> 0 000 00					

#### Programming Likelihood Functions: The Likelihood Ratio Test

<u>e</u> % • • • • •								
	itata Results				LR ch:	i2(2) =	183.26	
eral/PRISM Programming/NES UME~1\4DMINI~1\LOCALS~ hormreg (reg: bush_app = pid ec eral/PRISM Programming/NES	g likelihood	= -162.4521	2			> chi2 =	0.0000	
CUME ~1\ADMINI~1\LOCALS~ p normreg CUME~1\ADMINI~1\LOCALS~ normreg (reg: bush_app = pid ec	vote2	Coef.	Std. Err.	Z	₽> z	[95% Conf.	Interval]	
pp pid econworse DME~1\4DMINI~1\LDCALS~ mylogit (logit: vote2 = pid econwor pid econworse mylogit (logit: vote2 =)	pid econworse _ <sup>cons</sup>	.8437178 3076029 .2923819	.0878516 .1606638 .6554855	9.60 -1.91 0.45	0.000 0.056 0.656	.6715318 6224982 992346	1.015904 .0072923 1.57711	
	ml model lf	mylogit (log:	it: vote2 =)					
les ∑I t: Command Wij	itial: ternative: scale: eration 0: eration 1: eration 2:	log likelih log likelih log likelih log likelih	pod = -264.7 $pod = -254.0$	9741 9741 9741 0798				
app – .ogy worse aryopp Lo warworth	g likelihood	= -254.079	8		Wald	r of obs = chi2(0) = > chi2 =	382 •	
ation	vote2	Coef.	Std. Err.	Z	P> z	[95% Conf.	Interval]	
emp n ne	_cons	4802252	.1052929	-4.56	0.000	6865956	2738548	
lite		• • • • • • • •			<u></u>			
	tata Command . model lf	mylogit	(logit: vo	te2 = 1	pid eco	nworse), l:	EO(1 -254.	0798)

#### Programming Likelihood Functions: The Likelihood Ratio Test

2 B 6 9: 0 E F	<u>s dd os</u>							_
	Stata Results							
e "I:\general\PRISM Programming\NES g_bush_app_econworse_pid "C:\DDCUME~1\ADMINI~1\LOCALS~							29	
cholcome Tradmini Trebcals model If normreg (reg: bush_app = pid ec check ar	vote2	Coef.	Std. Err.		₽> z	[95% Conf.	Interval]	
9"I:\general\PRISM Programming\NES "C:\DOCUME~1\ADMINI~1\LOCALS~ gram drop normreg	_cons	4802252	.1052929	-4.56	0.000	6865956	2738548	
"C:\DOCUME~1\ADMINI~1\LOCALS~ nodel If normreg (reg: bush_app = pid ec								
nax bush_app pid econworse "C:\DDCUME~1\ADMINI~1\LOCALS~	. ml model lf	mylogit (log:	it: vote2 =	pid econ	worse), l	f0(1 -254.079	3)	
nodel If mylogit (logit: vote2 = pid econw nax t vote2 pid econworse nodel If mylogit (logit: vote2 =)	. ml max							
	initial:	log likelihe	ood = -264.7	8222				
	alternative:	log likelih						
	rescale:	log likelih						
	Iteration 0:	log likelih						
	Iteration 1:	log likelih						
	Iteration 2:	log likelih						
	Iteration 3: Iteration 4:	log likelih log likelih						
	Iteration 5:	log likelih						
	iteration 5:	IOG IIKEIIN	JUU162.4	5212				
arget: Command Wi					Numbe	er of obs =	382	
ush app						$i_{12}(2) =$	183.26	
	Log likelihood	= -162 4521	2			> chi2 =	0.0000	
conworse	log iikeiinood	- 102.4021	2		FLOD	> CHIZ -	0.0000	
ilitaryopp								
ulfwarworth	vote2	Coef.	std. Err.	Z	P> z	[95% Conf.	Intervall	
id						1970 TATA - TATA MANA		
ducation	pid	.8437178	.0878516	9.60	0.000	.6715318	1.015904	
ovtemp	econworse	3076029	.1606638	-1.91	0.056	6224982	.0072923	
nion	_cons	.2923819	.6554855	0.45	0.656	992346	1.57711	
ncome								
onwhite								
onwhile ote2								
olez								
i.	Stata Command							

# Programming Likelihood Functions: Probit Program

🔣 myprobit.do - Stata Do-file Editor	×
program define myprobit	
args lnf thetal	
tempvar p	
quietly gen double `p'=norm(`thetal')	
quietly replace `lnf'=\$ML_y1*ln(`p')+(1-\$ML_y1)*ln(1-`p')	
end	
Line number: 5	
and homeon to	lle lle

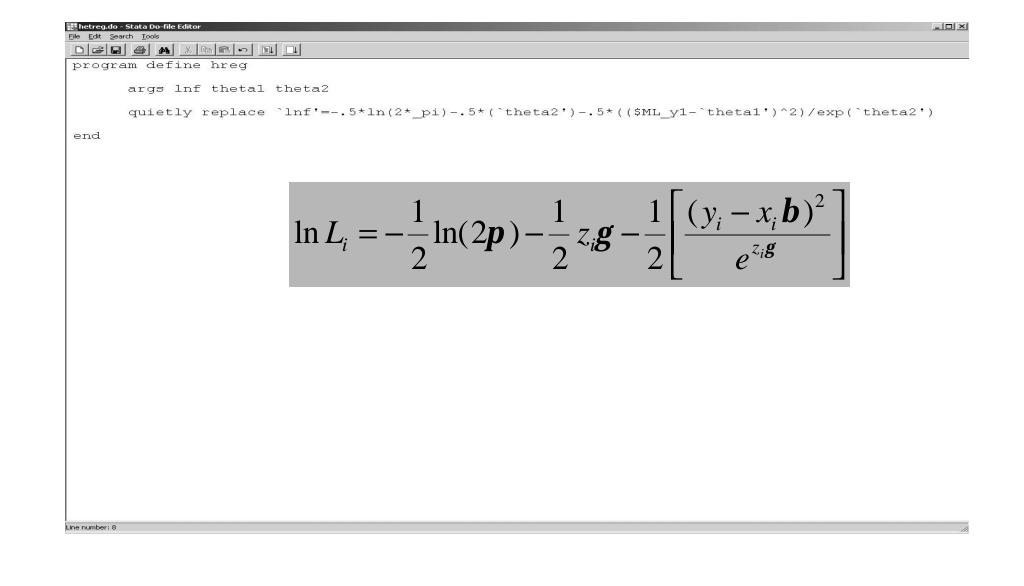
- Heteroskedastic regression allows us to model the factors that influence both the expected value of Y *and* the factors that affect the *variability* around that expected value (see Franklin 1991; Alvarez and Brehm 1995, 1997, 1998).
- In regression, we always assume homoskedasticity: s<sup>2</sup>
- With het. reg., we're explicitly interested in modeling the factors that influence s<sup>2</sup><sub>i</sub>.
- Good pedagogical example: it's more complicated, it generates two sets of simultaneously generated coefficients. *But*, bottom line: all you have to do is know the likelihood function, and you can program it in Stata.

• We parameterize  $s_i^2$  as:

 $\boldsymbol{s}_{i}^{2}=e^{z_{i}\boldsymbol{g}}$ 

- Where the z<sub>i</sub>'s exogenous variables that influence the variability around the expected value, and gamma is a vector of parameters.
- Log-likelihood:

$$\ln L_{i} = -\frac{1}{2}\ln(2\mathbf{p}) - \frac{1}{2}z_{i}\mathbf{g} - \frac{1}{2}\left[\frac{(y_{i} - x_{i}\mathbf{b})^{2}}{e^{z_{i}\mathbf{g}}}\right]$$



38 51 0 🖼	
Review se "K:VMLE Franklin\Exercise 2b\vo "C:\DOCUME=1\QBMIN=1\LOC model if hreg (slopes: reppct00= rep max	
	48-student Stata for Windows (network) perpetual license: Serial number: 1980523842 Licensed to: Ohio State University Ohio State University
	Notes: 1. (/m# option or -set memory-) 1.00 MB allocated to data . use "K:\MLE Franklin\Exercise 2b\vote2000.dta", clear (Written by R. )
Variables Parget: Command Peppct00 Peppct96	<pre>will do "C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\STD040000.tmp" . program define hreg</pre>
olatile dsgore dsbush .ogvotes .e	<pre>1.     args lnf theta1 theta2     guietly replace `lnf'=5*ln(2*_pi)5*(`theta2')5*((\$ML_y1-`theta1')^2)/exp(`thet &gt; ')     3.</pre>
	. end end of do-file
	<pre>Instate Command Instate C</pre>

Review X												
"K:\MLE Franklin\Exercise 2b\vote20 "C:\D0CUME~1\ADMINI~1\L0CALS^		log likelih			ot concav							
model If hreg (slopes: reppct00= reppct9	iteration 3:	log likelih			ot concav							
nax	Iteration 4:	log likelih			ot concav							
	Iteration 5:	log likelih			ot concav							
	Iteration 6: Iteration 7:	log likelih log likelih			ot concav ot concav							
	Iteration 8:	log likelih			ot concav							
	Iteration 9:				ot concav	e,						
	Iteration 9: log likelihood = -3590.76 Iteration 10: log likelihood = -3439.2212 (backed up)											
	Iteration 10: log likelihood = $-3439.2212$ (backed up) Iteration 11: log likelihood = $-3188.4392$											
	Iteration 12:	log likelih										
	Iteration 13: log likelihood = $-3184.6537$											
	Iteration 14:	log likelih	-3184	.6537								
					Numbe	r of obs =	1184					
						chi2(4) =	10562.78					
	Log likelihood	l = -3184.653	7		Prob	> chi2 =	0.0000					
Variables X	reppct00	Coef.	std. Err.		₽> z	[95% Conf.	Interval]					
arget: Command Wi												
eppct00	a Tobos											
eppct96	reppct96	.9604168	.0103549	92.75	0.000	.9401216	.9807121					
olatile	logvotes	-4.795379	.1706207	-28.11	0.000	-5.129789	-4.460968					
dsgore	adsgore	6220649	.1298877	-4.79	0.000	8766402	3674896					
lsbush	adsbush	1.057464 29.13514	.165257 .9552502	6.40 30.50	0.000 0.000	.7335665 27.26288	1.381362 31.00739					
ogvotes		29.13514	. 9552502	30.30	0.000	21.20200	51.00759					
2	var											
	volatile	.0818051	.0174115	4.70	0.000	.0476792	.1159309					
	cons	2.278007	.0695479	32.75	0.000	2.141696	2.414319					
	🔚 Stata Command											

- Standard duration models, which model the hazard of an event occurring, assume that all cases will eventually experience the event of interest.
- This assumption may not hold for the process under examination; if not, will produce incorrect inferences.
  - The timing of congressional overrides of Supreme Court decisions (Hettinger and Zorn 2004).
  - Corporate and labor PAC contributions to congressional candidates (Box-Steffensmeier et al. 2004).

- The split population duration (SPD) model relaxes the assumption that all cases will eventually experience the event of interest (Schmidt and Witte 1988, 1989; Forster and A. Jones 2001; Box-Steffensmeier and B. Jones 2004).
- Simultaneously estimates two sets of coefficients:
  - 1. Explaining the likelihood of the event occurring (i.e., the censoring indicator is the DV).
  - 2. Explaining the timing of the event occurring, conditional on the event having occurred in the first place.

- LIMDEP is the only package that has a canned routine for the SPD. Great example of an advanced model that hasn't made its way into a lot of stat packages. But you can program it yourself!
- Acknowledgements to Forster and Jones...

File Edit S	do - Stata Do-file Editor	_ 8 >
progr	ram define splitpop	
	args lnf thetal theta2 theta3 /* thetal: XB for duration equation; theta2: XB for logit equation; theta3: shape parameter (delta) */	
	tempvar p s d l	
	quietly gen double `l'=exp(-`thetal') /* lambda of log-logistic distribution */	
	quietly gen double `d'=exp(`theta2')/(1+exp(`theta2')) /* logistic cdf for probability of failure */	
	quietly gen double `s'=1-`d'+`d'*(1/(1+(`l'*\$ML_y1)^(1/`theta3'))) /* survival function */	
	quietly gen double `p'=ln(`d')-ln(`theta3')+((1/`theta3')-1)*ln(\$ML_y1)+(1/`theta3')* */ ln(`l')-2*ln(1+(`l'*\$ML_y1)^(1/`theta3')) /* pdf */	/*
	quietly replace `lnf'=\$ML_y2*(`p')+(1-\$ML_y2)*ln(`s')	
end		
	$\ln L_{i} = R_{i} [\ln d_{i} + \ln g(t_{i}, q)] + (1 - R_{i}) \ln[1 - d_{i} + d_{i}G(t_{i}, q)]$	
cen	soring indicator pdf survivor function	

- Example: Explaining the incidence and timing of labor PAC contributions to incumbent House members, 1993-1994 (Box-Steffensmeier et al. 2004).
- We're interested in the timing of contributions in an election cycle. Early money is "seed money" for a campaign effort, and it helps candidates raise more down the line (Jacobson 1992).
- We don't expect labor PACs to contribute to every House incumbent, though. E.g., people trying to reform OSHA, or investigating the Teamsters.

Intercooled Stata 8.0		
<u>Eile Edit Prefs Data Graphics Statis</u>	tics User Window Help	
G G G G G G G G G G G G G G G G G G G		
use "C:\Documents and Settings\Administr- set mem 100m use "C:\Documents and Settings\Administr- do "C:\DOCUME~1\ADMINI~1\LOCALS~	> ", clear	Ē
ml model if splitpop (duration: T = educatn p	<pre>. do errorent frimmer friedrice friedrice from p . program define splitpop 1.</pre>	
■Variables 🛛 💌 Target: Command Wij	<pre>. quietly gen double `s'=1-`d'+`d'*(1/(1+(`l'*\$ML_y1)^(1/`theta3'))) 6. /* survival function */ quietly gen double `p'=ln(`d')-ln(`theta3')+((1/`theta3')-1)*ln(\$ML_y1)+(1/`theta3')* /* &gt; */ ln(`l')-2*ln(1+(`l'*\$ML_y1)^(1/`theta3')) 7. /* pdf */</pre>	
rawdays pactrecp totcnt pair pairtot	<pre>// put / / put / / quietly replace `lnf'=\$ML_y2*(`p')+(1-\$ML_y2)*ln(`s') 8 end</pre>	
trecp bcash paccnt large lagprecp pacshare stshare	end of do-file . ml model lf splitpop (duration: T = educatn prestige seniorty dleader rleader republic cope votepc > ality pquality bcash lagprecp recpsqrd totpct stshare homecand dcpac) (gave = educatn prestige sen > y dleader rleader republic cope votepct quality pquality bcash lagprecp recpsqrd totpct stshare hom > nd dcpac) /shape if large==1	iort
nfrompac candnum subrank subchair comrank dleader	State Command ml model lf splitpop (duration: T = educatn prestige seniorty dleader rleader republic cope votepct quality pquality bcash lagprecp recpsqrd totpct stshare homecand dcpac) (gave = edu prestige seniorty dleader rleader republic cope votepct quality pquality bcash lagprecp recpsqrd totpct stshare homecand dcpac) /shape if large==1	

🔣 Intercooled Stata 8.0		
Eile Edit Prefs Data Graphics Statis	tics User Window Help	
Review use "C:\Documents and Settings\Administr. set men 100m use "C:\Documents and Settings\Administr. do "C:\DocuME"1\ADMINI"1\L0CALS" mI modell i splitpop (duration: T = educatn p	■ Stata Results . use "C:\Documents and Settings\Administrator\My Documents\PACs Data\labor - with dc pacs.dta > ", clear . do "C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\STD040000.tmp"	
	<pre>. program define splitpop 1 args lnf thetal theta2 theta3 2. /* theta1: XB for duration equation; theta2: XB for logit equation; &gt; theta3: shape parameter (delta) */ . tempvar p s d l</pre>	
	<pre>3. quietly gen double `l'=exp(-`theta1') 4.     /* lambda of log-logistic distribution */ .        quietly gen double `d'=exp(`theta2')/(1+exp(`theta2')) 5.</pre>	
	. quietly gen double `s'=1-`d'+`d'*(1/(1+(`l'*\$ML_y1)^(1/`theta3'))) 6. /* survival function */	
Target: Command Wij rawdays	<pre>. quietly gen double `p'=ln(`d')-ln(`theta3')+((1/`theta3')-1)*ln(\$ML_y1)+(1/`theta3')* /* &gt; */ ln(`l')-2*ln(1+(`1'*\$ML_y1)^(1/`theta3')) 7. /* pdf */</pre>	
pactrecp totcnt pair pairtot trecp	. [quietly replace `lnf'=\$ML_y2*(`p')+(1-\$ML_y2)*ln(`s') [8. . end	
bcash paccnt large lagprecp pacshare stshare	end of do-file . ml model lf splitpop (duration: T = educatn prestige seniorty dleader rleader republic cope votepct > ality pquality bcash lagprecp recpsqrd totpct stshare homecand dcpac) (gave = educatn prestige sen: > y dleader rleader republic cope votepct quality pquality bcash lagprecp recpsqrd totpct stshare hom > nd dcpac) /shape if large==1	iort
nfrompac candnum subrank subchair comrank dleader	Inl search Searches for better starting values	

🔣 Intercooled Stata 8.0		_ U ×
<u>File Edit Prefs Data Graphics Stati</u>	tics User Window Help	
Review	🗖 Stata Results	×
use "C:\Documents and Settings\Administr- set mem 100m use "C:\Documents and Settings\Administr- do "C:\DOCUME~1\ADMINI~1\LDCALS~ ml model if splitpop (duration: T = educath p	. args lnf thetal theta2 theta3 2. /* theta1: XB for duration equation; theta2: XB for logit equation; > theta3: shape parameter (delta) */	
ml search	. tempvar p s d l 3.	
	quietly gen double `l'=exp(-`thetal') 4. /* lambda of log-logistic distribution */	
	. quietly gen double `d'=exp(`theta2')/(1+exp(`theta2')) 5. /* logistic cdf for probability of failure */	
	. quietly gen double `s'=1-`d'+`d'*(1/(1+(`l'*\$ML_y1)^(1/`theta3'))) 6. /* survival function */	
	quietly gen double `p'=ln(`d')-ln(`theta3')+((1/`theta3')-1)*ln(\$ML_y1)+(1/`theta3')* /* > */ ln(`l')-2*ln(1+(`l'*\$ML_y1)^(1/`theta3')) 7. /* pdf */	
Target: Command Wil rawdays	end of do-file	
pactrecp totcnt pair	. ml model lf splitpop (duration: T = educatn prestige seniorty dleader rleader republic cope votepct > ality pquality bcash lagprecp recpsqrd totpct stshare homecand dcpac) (gave = educatn prestige seni > y dleader rleader republic cope votepct quality pquality bcash lagprecp recpsqrd totpct stshare hom	iort
pairtot	> nd dcpac) /shape if large==1	
trecp	. ml search	
bcash	. mi searcn initial:         log likelihood =       - <inf>   (could not be evaluated)</inf>	
paccnt —	feasible: log likelihood = -73241.643	
large	improve: log likelihood = -41736.789	
lagprecp	rescale: log likelihood = -40900.402 rescale eq: log likelihood = -40014.636	
pacshare	rescare eq: log likelinood40014.636	
stshare		
nfrompac		
candnum	Stata Command	
subrank		
subchair		
comrank		
dleader 🗸		
C:\DATA		
		1

😂 S: 👁 🔜 🔳	3 🗆 🖸 🤇						
	Stata Results						
euments and Settings\Administr،					Numbe	r of obs =	10168
m					Wald	chi2(17) =	534.57
uments and Settings\Administr. UME~1\ADMINI~1\LOCALS~	og likelihood	= -33857.72	6		Prob	> chi2 =	0.0000
olitpop (duration: T = educatn p							
		Coef.	Std. Err.	Z	P >  z	[95% Conf.	Intervall
olitpop (duration: T = educatn p		COEL.	SCG. ELL.	2	F / 2	[50% CONT.	incervalj
d	luration						
	educatn	.0657934	.0465943	1.41	0.158	0255297	.1571166
	prestige	.0375843	.0296507	1.27	0.205	0205299	.0956986
	seniorty	.2655538	.1936916	1.37	0.170	1140749	.6451824
	dleader	2008474	.0426538	-4.71	0.000	2844473	1172475
	rleader republic	2369234 0596892	.1366631 .060448	-1.73 -0.99	0.083 0.323	5047783 178165	.0309314 .0587866
	cope	2930568	.108651	-2.70	0.323	5060088	0801048
	votepct	.3198829	.1227685	2.61	0.009	.0792612	.5605047
	quality	.0164541	.043222	0.38	0.703	0682594	.1011677
	pquality	.0376763	.0765442	0.49	0.623	1123476	.1877002
	bcash	.351434	.0716984	4.90	0.000	.2109078	.4919603
	lagprecp	2789671	.0200107	-13.94	0.000	3181872	2397469
	recpsqrd	.0259331	.0023862	10.87	0.000	.0212563	.03061
	totpet	5041463	.2263586	-2.23	0.026	9478011	0604915
	stshare homecand	.4819196 2420607	.1913304 .1037068	2.52 -2.33	0.012	.106919 4453224	.8569203 0387991
	dcpac	.4026476	.0298946	13.47	0.020	.3440552	.4612399
ies X	_cons	5.717728	.1143268	50.01	0.000	5.493652	5.941804
t: Command Wi							
vs	ogit						
-	educatn	.1092988	.1184172	0.92	0.356	1227946	.3413922
ecp	prestige	.2806098	.0635753	4.41	0.000	.1560045	.4052151
t	seniorty	7225828	.3885568	-1.86	0.063	-1.48414	.0389745
	dleader	.4483005 .2813031	.117503 .1753403	3.82 1.60	0.000	.2179989 0623576	.6786021 .6249638
ot	rleader republic	-1.798423	.1000995	-17.97	0.000	-1.994614	-1.602231
	cope	3.250226	.2013426	16.14	0.000	2.855602	3.64485
	votepct	-2.374621	.2654199	-8.95	0.000	-2.894835	-1.854408
	quality	.2355944	.0915296	2.57	0.010	.0561997	.414989
t —	pquality	4067091	.149639	-2.72	0.007	6999962	113422
	bcash	4111929	.1515213	-2.71	0.007	7081691	1142167
ecp	lagprecp	.4141835	.0447867	9.25	0.000	.3264032	.5019639
are	recpsqrd	0333042	.0054234	-6.14	0.000	0439338	0226746
22.22.224	totpet	3.878042	.4972922	7.80	0.000	2.903368	4.852717
re	stshare homecand	.8182422 .298608	.4648726 .2603486	1.76 1.15	0.078 0.251	0928913 2116658	1.729376 .8088819
pac	dcpac	.0935252	.0631602	1.15	0.139	0302664	.2173169
um	cons	-1.331899	.2248491	-5.92	0.000	-1.772595	8912029
nk					0.000		
	hape						
air	_cons	.480325	.0074244	64.70	0.000	.4657735	.4948765
nk							
er 🔤							

### Programming Likelihood Functions: The Predict Command and Est. Split

	istics <u>U</u> ser <u>W</u> indow	Help								
3888	a second s	0 8								
Review	🗖 Stata Results									
et mem 100m	seniorty	.2655538	.1936917	1.37	0.170	114075	.6451826			
se ''K:\Jan - PACs\Data\large labor PACs o ''C:\DOCUME~1\ADMINI~1\LOCALS~	dleader rleader	2008474 2369235	.0426538 .1366632	-4.71 -1.73	0.000 0.083	2844473 5047784	1172475 .0309314			
o "C:\DOCUME~1\ADMINI~1\LOCALS~		0596892	.060448	-0.99	0.323	1781651	.0587866			
il model If splitpop (duration: T = educatn p en T=rawdays	cope	2930568	.108651	-2.70	0.007	5060089	0801047			
I model If splitpop (duration: T = educatn p	votepct	.3198831	.1227685	2.61	0.009	.0792613	.560505			
il search Il max	quality	.0164541	.043222	0.38	0.703	0682594	.1011677			
i nan	pquality	.0376763	.0765443	0.49	0.623	1123477	.1877002			
	bcash	.3514341	.0716984	4.90	0.000	.2109078	.4919604			
	lagprecp	2789671	.0200107	-13.94	0.000	3181873	2397469			
	recpsqrd	.0259331 5041463	.0023862 .2263587	10.87 -2.23	0.000 0.026	.0212563 9478013	.03061 0604914			
	totpct stshare	.4819198	.1913305	2.52	0.026	9478013	.8569207			
	homecand	2420607	.1037069	-2.33	0.012	4453224	038799			
	dcpac	.4026476	.0298946	13.47	0.000	.3440552	.4612399			
	_cons	5.717728	.1143268	50.01	0.000	5.493652	5.941804			
	-									
	logit educatn	.1092988	.1184172	0.92	0.356	1227947	.3413923			
	prestige	.2806099	.0635753	4.41	0.000	.1560045	.4052152			
	seniorty	7225828	.3885569	-1.86	0.063	-1.48414	.0389747			
	dleader	.4483005	.117503	3.82	0.000	.2179988	.6786022			
	rleader	.2813031	.1753403	1.60	0.109	0623576	.6249638			
	republic	-1.798423	.1000995	-17.97	0.000	-1.994614	-1.602231			
Variables X	cope	3.250226 -2.374621	.2013426	16.14 - 8.95	0.000 0.000	2.855602 -2.894835	3.64485 -1.854408			
		.2355944	.0915296	2.57	0.010	.0561997	.414989			
Target: Command Wi	pquality	4067092	.1496391	-2.72	0.007	6999964	113422			
ignored 🔶	bcash	4111928	.1515213	-2.71	0.007	7081691	1142164			
signal	lagprecp	.4141835	.0447867	9.25	0.000	.3264031	.5019638			
first	recpsqrd	0333042	.0054234	-6.14	0.000	0439338	0226746			
copecase	totpet	3.878043	.4972923	7.80	0.000	2.903368	4.852718			
rrawdays	stshare	.818243	.4648729	1.76	0.078	0928911	1.729377			
orevious	homecand dcpac	.298608	.2603486 .0631602	1.15 1.48	0.251 0.139	211666 0302663	.8088819 .217317			
	cons	-1.331899	.2248491	-5.92	0.000	-1.772595	8912029			
nomecand										
lndays	shape									
recpsgrd	_cons	.4803252	.0074244	64.70	0.000	.4657737	.4948767			
republic										
senior2										
lndays2										
sample								1	)	
dcpac	🖬 Stata Command									
lagrecp2	predict xh	), eq(logi	t)							
		co <del>7</del> .810 <del></del> 1								
recsqrd2										
recsqrd2 ocash2 F										

## Programming Likelihood Functions: The Predict Command and Est. Split

intercooled Stata 8.0									
<u>E</u> dit <u>P</u> refs <u>D</u> ata <u>G</u> raphics <u>S</u> tatis	itics <u>U</u> ser <u>W</u> indow	Help							
F 🖬 🎒 🛠 📾 📰 📰		I 🔁							
Review X	🗖 Stata Results								
mem 100m	rleader	2369235	.1366632	-1.73	0.083	5047784	.0309314		
"K:\Jan - PACs\Data\large labor PACs. "C:\DOCUME~1\ADMINI~1\LOCALS~	republic	0596892	.060448	-0.99	0.323	1781651	.0587866		
C:\DOCUME ~1\ADMINI ~1\LOCALS~	cope	2930568	.108651	-2.70	0.007	5060089	0801047		
nodel If splitpop (duration: T = educatn p	votepct quality	.3198831 .0164541	.1227685	2.61 0.38	0.009	.0792613 0682594	.560505		
nodel If splitpop (duration: T = educatn p		.0376763	.0765443	0.38	0.623	1123477	.1877002		
earch	bcash	.3514341	.0716984	4.90	0.000	.2109078	.4919604		
nax dict xb, eq(logit)	lagprecp	2789671	.0200107	-13.94	0.000	3181873	2397469		
	recpsqrd	.0259331	.0023862	10.87	0.000	.0212563	.03061		
	totpet	5041463	.2263587	-2.23	0.026	9478013	0604914		
	stshare	.4819198	.1913305	2.52	0.012	.106919	.8569207		
	homecand	2420607	.1037069	-2.33	0.020	4453224	038799		
	dcpac	.4026476	.0298946	13.47	0.000	.3440552	.4612399		
		5.717728	.1143268	50.01	0.000	5.493652	5.941804		
	logit								
	educatn	.1092988	.1184172	0.92	0.356	1227947	.3413923		
	prestige	.2806099	.0635753	4.41	0.000	.1560045	.4052152		
	seniorty	7225828	.3885569	-1.86	0.063	-1.48414	.0389747		
	dleader	.4483005	.117503	3.82	0.000	.2179988	.6786022		
	rleader	.2813031	.1753403	1.60	0.109	0623576	.6249638		
	republic cope	-1.798423 3.250226	.1000995 .2013426	-17.97 16.14	0.000	-1.994614 2.855602	-1.602231 3.64485		
	votepct	-2.374621	.26542	-8.95	0.000	-2.894835	-1.854408		
	quality	.2355944	.0915296	2.57	0.010	.0561997	.414989		
¥ariables X	pquality	4067092	.1496391	-2.72	0.007	6999964	113422		
arget: Command Wi	bcash	4111928	.1515213	-2.71	0.007	7081691	1142164		
opecase	lagprecp	.4141835	.0447867	9.25	0.000	.3264031	.5019638		
rawdays	recpsqrd	0333042	.0054234	-6.14	0.000	0439338	0226746		
	totpet	3.878043	.4972923	7.80	0.000	2.903368	4.852718		
revious	stshare homecand	.818243	.4648729 .2603486	1.76 1.15	0.078 0.251	0928911 211666	1.729377 .8088819		
omecand	dcpac	.0935253	.0631602	1.48	0.139	0302663	.217317		
ndays	cons	-1.331899	.2248491	-5.92	0.000	-1.772595	8912029		
ecpsgrd									
epublic	shape								
-	_cons	.4803252	.0074244	64.70	0.000	.4657737	.4948767		
enior2									
ndays2									
ample	. predict xb,	ed(logic)							
cpac									
agrecp2									
ecsqrd2	📰 Stata Command				_				
cash2	-								
	gen prob=e	exp(xb)/(1	+exp(xb)	)					
b									1
						1 1 • 1 • .	6.1		1
				aicula	tes ni	opapiiit	v of the	exchange of a contribution.	

## Programming Likelihood Functions: The Predict Command and Est. Split

Edit Prefs Data Graphics Stati:							
teview X	Stata Results						
nem 100m	bcash	4111928	.1515213	-2.71	0.007	7081691	1142164
"K:\Jan - PACs\Data\large labor PACs. C:\DOCUME~1\ADMINI~1\LOCALS~	laqprecp	.4141835	.0447867	9.25	0.000	.3264031	.5019638
C:\DOCUME~1\ADMINI~1\LOCALS~ odel If splitpop (duration: T = educatn p		0333042	.0054234	-6.14	0.000	0439338	0226746
present splitpop (duration: T = educatin p odel If splitpop (duration: T = educatin p	totpat	3.878043	.4972923	7.80	0.000	2.903368	4.852718
arch	stshare	.818243	.4648729	1.76	0.078	0928911	1.729377
ax ct xb, eq(logit)	homecand	.298608	.2603486	1.15	0.251	211666	.8088819
prob=exp(xb)/(1+exp(xb)) prob	dcpac	.0935253	.0631602	1.48	0.139	0302663	.217317
gave	_cons	-1.331899	.2248491	-5.92	0.000	-1.772595	8912029
	shape						
	_cons	.4803252	.0074244	64.70	0.000	.4657737	.4948767
	. predict xb, . gen prob=exp	eq(logit) >(xb)/(1+exp(:	xb))				
tariables 🔀	. sum prob						
arget: Command Wij rawdays revious	Variable	Obs	Mean	Std. Dev.		Min	Max
evious mecand days	prob	10168	.4719391	.317901 .00		84166 .972	5916
cpsqrd public	. sum gave						
nior2 days2	Variable	Obs	Mean	Std. De	ev.	Min	Max
mple	gave	10168	.4353855	.49583	18	0	1
pac							
grecp2	-						
csqrd2					-		
ash2	🔙 Stata Command						38. 
OD							
dor		(	<b>Observed sp</b>	lif			

## Programming Likelihood Functions: Conclusion

#### • Bottom line:

- If you need to estimate a model that is not canned in a popular software package, you can probably program it in Stata.
- All you need to know is the likelihood function!