

GVPT 722

Examples of Detecting and Dealing with Heteroskedasticity in Stata

(1) OLS regression of percentage of state legislative and congressional seats held by African-Americans on the percentage of the state population that is African-American:

```
. reg blackleg black, beta
```

Source	SS	df	MS		Number of obs =	50
Model	2393.38215	1	2393.38215		F(1, 48) =	366.12
Residual	313.781991	48	6.53712482		Prob > F =	0.0000
					R-squared =	0.8841
					Adj R-squared =	0.8817
					Root MSE =	2.5568
Total	2707.16414	49	55.2482478			

blackleg	Coef.	Std. Err.	t	P> t	Beta
black	.7419118	.0387739	19.13	0.000	.9402617
_cons	.465191	.5139555	0.91	0.370	.

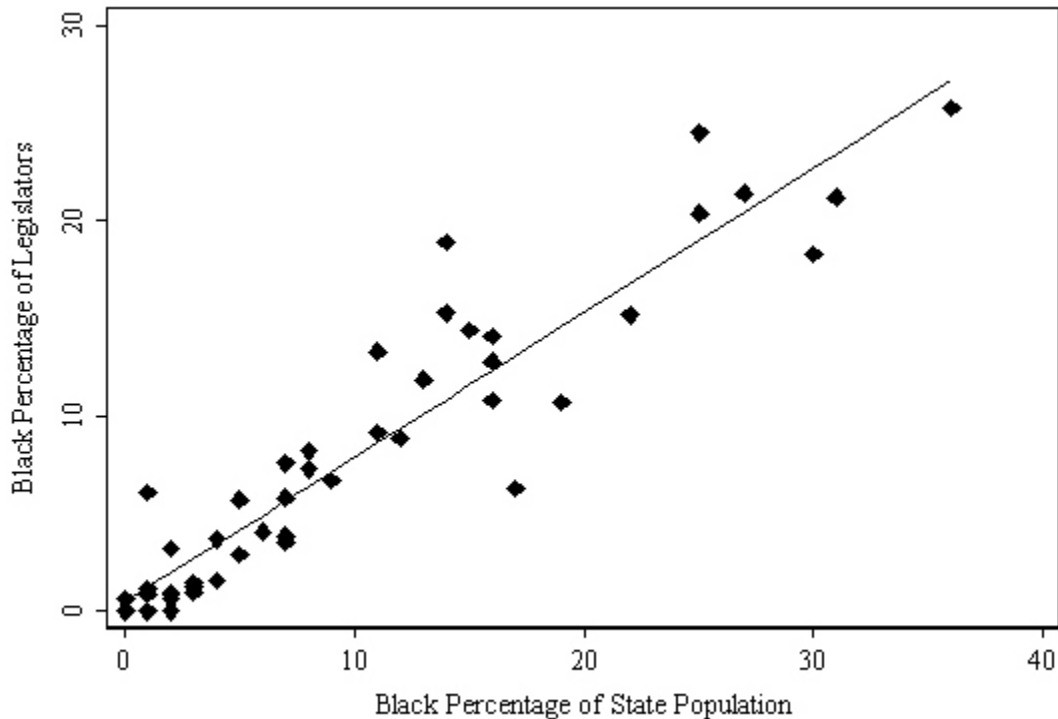
(2) Predicting predicted values of Y and residuals

```
. predict blackleg_hat  
(option xb assumed; fitted values)
```

```
. predict errorhat, residuals
```

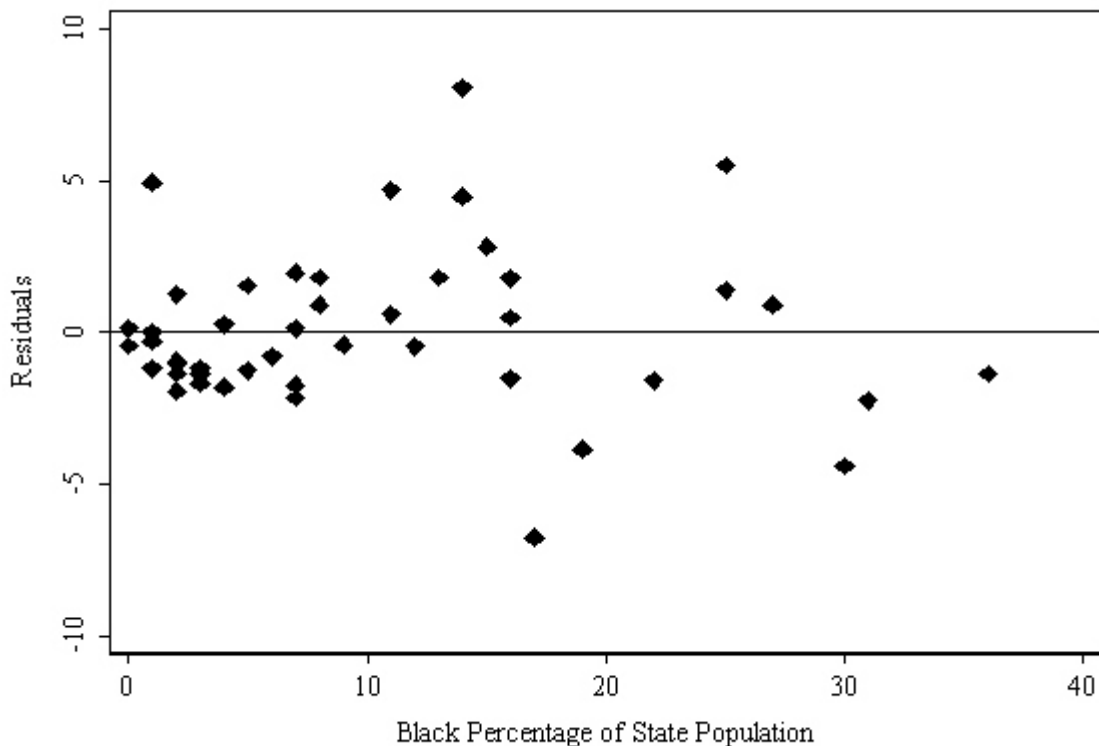
(3) Graphing predicted values and actual values of Y on X

```
. twoway (scatter blackleg black, sort msymbol(diamond) mcolor(black)) (connected  
blackleg_hat black, sort msymbol(none) clcolor(black) clpat(solid)), ytitle(Black  
Percentage of Legislators, margin(medsmall)) xtitle(Black Percentage of State  
Population, margin(medsmall)) legend(off)
```



(4) Graphing residuals on X:

```
. twoway (scatter errorhat black, sort msymbol(diamond) mcolor(black)),
yttitle(Residuals, margin(medsmall)) yline(0, lcolor(black) lpattern(solid))
xttitle(Black Percentage of State Population, margin(medsmall)) legend(off)
```



(5) OLS regression of black percentage of state legislators and members of Congress on percentage black in population and state opinion liberalism:

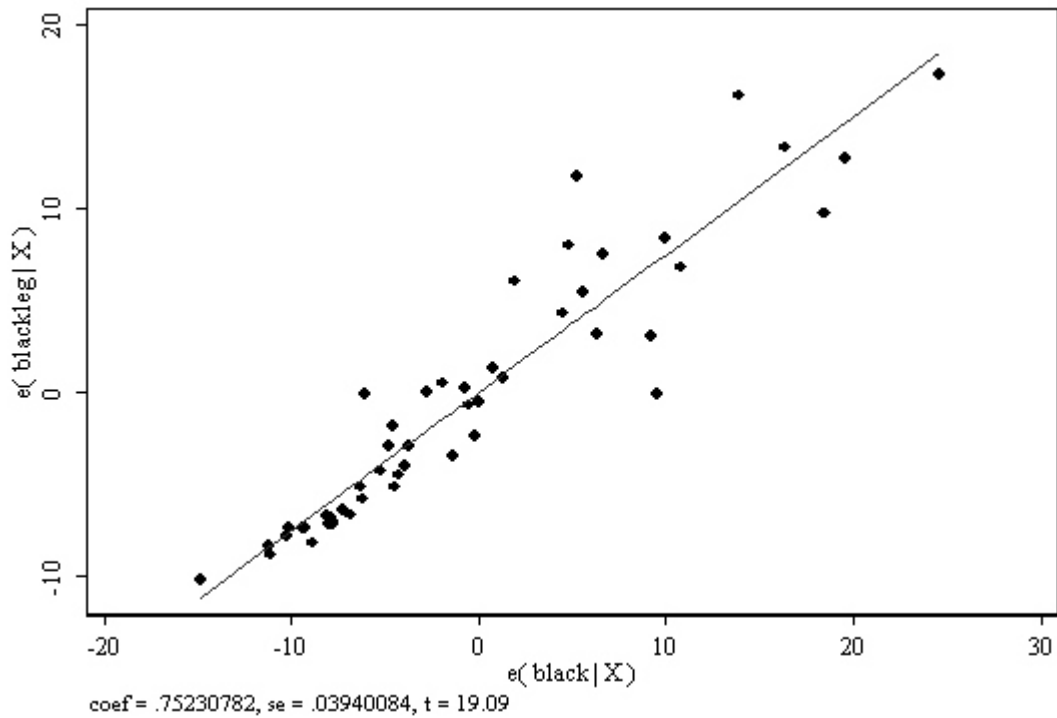
```
. reg blackleg black opinionlib, beta
```

Source	SS	df	MS		
Model	2403.7105	2	1201.85525	Number of obs =	50
Residual	303.45364	47	6.45646044	F(2, 47) =	186.15
Total	2707.16414	49	55.2482478	Prob > F =	0.0000
				R-squared =	0.8879
				Adj R-squared =	0.8831
				Root MSE =	2.541

	blackleg	Coef.	Std. Err.	t	P> t	Beta
	black	.7523078	.0394008	19.09	0.000	.9534371
	opinionlib	4.346963	3.436906	1.26	0.212	.0631568
	_cons	.9414105	.6345539	1.48	0.145	.

(6) Added-variable plot from regression in (5):

```
. avplot black, mcolor(black)
```



(7) Testing for heteroskedasticity using White's test:

```
. imtest, white
```

White's test for H_0 : homoskedasticity
against H_a : unrestricted heteroskedasticity

```
chi2(2)      =      7.64  
Prob > chi2  =      0.0219
```

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	7.64	2	0.0219
Skewness	1.42	1	0.2336
Kurtosis	3.01	1	0.0827
Total	12.07	4	0.0168

(8) Testing for heteroskedasticity using the Breusch-Pagan test:

```
. hettest
```

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
 H_0 : Constant variance
Variables: fitted values of blackleg

```
chi2(1)      =      8.42  
Prob > chi2  =      0.0037
```

(9) Weighted least squares regression with σ_i computed based on standard deviation of Y across values of X

```
. egen sigmai=sd(blackleg), by(black)
(12 missing values generated)

. egen sigmai_b=sd(blackleg), by(black6)

. reg blackleg black [aweight=1/sigmai_b]
(sum of wgt is 2.1706e+01)
```

Source	SS	df	MS			
Model	2196.80747	1	2196.80747	Number of obs	=	50
Residual	252.302188	48	5.25629557	F(1, 48)	=	417.94
				Prob > F	=	0.0000
				R-squared	=	0.8970
				Adj R-squared	=	0.8948
Total	2449.10966	49	49.9818297	Root MSE	=	2.2927

blackleg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
black	.7501341	.036693	20.44	0.000	.6763579	.8239102
_cons	.2902348	.4278891	0.68	0.501	-.5700938	1.150563

(10) Regression with heteroskedasticity-consistent standard errors (White/Huber):

```
. reg blackleg black
```

Source	SS	df	MS			
Model	2393.38215	1	2393.38215	Number of obs	=	50
Residual	313.781991	48	6.53712482	F(1, 48)	=	366.12
				Prob > F	=	0.0000
				R-squared	=	0.8841
				Adj R-squared	=	0.8817
Total	2707.16414	49	55.2482478	Root MSE	=	2.5568

blackleg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
black	.7419118	.0387739	19.13	0.000	.6639516	.819872
_cons	.465191	.5139555	0.91	0.370	-.5681858	1.498568

```
. reg blackleg black, robust
```

Regression with robust standard errors

blackleg	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
black	.7419118	.0405427	18.30	0.000	.6603953	.8234283
_cons	.465191	.3512043	1.32	0.192	-.2409526	1.171335