

## 8) Generalized Least Squares (GLS)

Vitor Kamada

January 2018

## Known and Nonsingular $V(X)$

$$E(\epsilon\epsilon'|X) = \sigma^2 V(X)$$

$(n \times n)$

$$V^{-1} = C' C$$

$$\tilde{y} = Cy, \quad \tilde{X} = CX, \quad \tilde{\epsilon} = C\epsilon$$

$$\tilde{y} = \tilde{X}\beta + \tilde{\epsilon}$$

$$E(\tilde{\epsilon}|\tilde{X})$$

$$E(\tilde{\epsilon}|X)$$

$$E(C\epsilon|X)$$

$$CE(\epsilon|X)$$

$$0$$

# Variance of the Transformed Error $\tilde{\epsilon}$ is Spherical

$$E(\tilde{\epsilon}\tilde{\epsilon}'|\tilde{X})$$

$$E(\tilde{\epsilon}\tilde{\epsilon}'|X)$$

$$C\sigma^2VC'$$

$$\sigma^2CVC'$$

$$\sigma^2I_n$$

# Generalized Least Squares (GLS) Estimator

$$\begin{aligned}\hat{\beta}_{GLS} &= (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y} \\ &= [(CX)'(CX)]^{-1}(CX)'Cy \\ &= (X'C'CX)^{-1}(X'C'Cy) \\ &= (X'V^{-1}X)^{-1}X'V^{-1}y\end{aligned}$$

$$\text{Var}(\hat{\beta}_{GLS}|X)$$

$$(X'V^{-1}X)^{-1}X'V^{-1}\text{Var}(y|X)V^{-1}X(X'V^{-1}X)^{-1}$$

$$(X'V^{-1}X)^{-1}X'V^{-1}(\sigma^2V)V^{-1}X(X'V^{-1}X)^{-1}$$

$$\sigma^2(X'V^{-1}X)^{-1}$$

# Weighted Least Squares (WLS)

$$\text{Var}(\epsilon_i|X) = E(\epsilon_i^2|X) = \sigma^2 v_i(X)$$

$$\tilde{y}_i = \frac{y_i}{\sqrt{v_i(X)}}, \quad \tilde{X} = \frac{X_i}{\sqrt{v_i(X)}},$$
$$\tilde{\epsilon} = \frac{\epsilon_i}{\sqrt{v_i(X)}}$$

$$\tilde{y} = \tilde{X}\beta + \tilde{\epsilon}$$

# Papke (1995)

```
library(foreign); library(lmtest); library(car); library(stargazer)
d401k<-read.dta("C:/RData/401ksubs.dta")
summary(d401k); stargazer(d401k)
```

Statistic	N	Mean	St. Dev.	Min	Max
e401k	9,275	0.392	0.488	0	1
inc	9,275	39.255	24.090	10.008	199.041
marr	9,275	0.629	0.483	0	1
male	9,275	0.204	0.403	0	1
age	9,275	41.080	10.300	25	64
fsize	9,275	2.885	1.526	1	13
nettfa	9,275	19.072	63.964	-502.302	1,536.798
p401k	9,275	0.276	0.447	0	1
pira	9,275	0.254	0.436	0	1
incsq	9,275	2,121.192	3,001.469	100.160	39,617.320
agesq	9,275	1,793.653	895.649	625	4,096



# R Code for WLS and Heteroskedasticity

```
OLS <- lm(nettfa ~ inc + I((age-25)^2) + male + e401k,  
  data=d401k, subset=(fsize==1))  
OLSRobRef <- coeftest(OLS,hccm)  
OLSRob <- coeftest(OLS, vcov=hccm(OLS, type="hc0"))  
  
wlsreg <- lm(nettfa ~ inc + I((age-25)^2) + male + e401k,  
  weight=1/inc, data=d401k, subset=(fsize==1))  
WLS <- coeftest(wlsreg)  
WLSRob <- coeftest(wlsreg,hccm)
```

## R Code for Table of Results

```
library(stargazer)

stargazer(OLSRob,OLSRobRef,WLS,WLSRob,
  title="Regression Results",
  dep.var.labels="Net Financial Wealth",
  column.labels=c("OLSRob","OLSRobRef",
    "WLS","WLSRob"),
  no.space=TRUE,
  column.sep.width = "1pt",
  omit.stat=c("ser","f"))
```

# Regression Results

	<i>Dependent variable:</i>			
	Net Financial Wealth			
	OLSRob	OLSRobRef	WLS	WLSRob
	(1)	(2)	(3)	(4)
inc	0.771*** (0.099)	0.771*** (0.100)	0.740*** (0.064)	0.740*** (0.075)
l((age - 25)^2)	0.025*** (0.004)	0.025*** (0.004)	0.018*** (0.002)	0.018*** (0.003)
male	2.478 (2.056)	2.478 (2.065)	1.841 (1.564)	1.841 (1.313)
e401k	6.886*** (2.284)	6.886*** (2.292)	5.188*** (1.703)	5.188*** (1.574)
Constant	-20.985*** (3.491)	-20.985*** (3.520)	-16.703*** (1.958)	-16.703*** (2.248)

**Note:**

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01