

Omitted Variable Test

A regression of the log of hourly pay on years of work experience (exper) and dummy variable for gender (female:1=female, 0 otherwise) gives the following

```
. u psldat
```

```
. reg lhw exper female grad union
```

Source	SS	df	MS			
Model	1338.97647	4	334.744119	Number of obs = 17321		
Residual	4827.18625	17316	.278770285	F(4, 17316) = 1200.79		
Total	6166.16272	17320	.356014014	Prob > F = 0.0000		
				R-squared = 0.2171		
				Adj R-squared = 0.2170		
				Root MSE = .52799		

lhw	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
exper	.0063046	.0003415	18.46	0.000	.0056351	.006974
female	-.2710643	.0080369	-33.73	0.000	-.2868174	-.2553112
grad	.5703866	.011521	49.51	0.000	.5478044	.5929688
union	.2079973	.0087041	23.90	0.000	.1909364	.2250582
_cons	1.803265	.0099212	181.76	0.000	1.783819	1.822712

```
. ovtest
```

Ramsey RESET test using powers of the fitted values of lhw

```
Ho: model has no omitted variables
      F(3, 17313) = 16.04
      Prob > F = 0.0000
```

The Ramsey reset test rejects null that functional form is correctly specified

Similarly, saving the residuals from this (restricted) specification

```
. predict urest, resid
```

and regressing them on the original X vector and an additional variable (the square of experience) gives

```
. reg urest exper female grad union exper2
```

Source	SS	df	MS			
Model	292.546211	5	58.5092422	Number of obs = 17321		
Residual	4534.64003	17315	.261890848	F(5, 17315) = 223.41		
Total	4827.18624	17320	.278705903	Prob > F = 0.0000		
				R-squared = 0.0606		
				Adj R-squared = 0.0603		
				Root MSE = .51175		

urest	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
exper	.0386276	.0012022	32.13	0.000	.0362712	.0409841
female	-.0174888	.0078073	-2.24	0.025	-.0327919	-.0021856
grad	.007176	.0111688	0.64	0.521	-.0147159	.0290679
union	-.031009	.0084873	-3.65	0.000	-.0476451	-.014373
exper2	-.0008684	.000026	-33.42	0.000	-.0009194	-.0008175
_cons	-.283922	.012831	-22.13	0.000	-.3090721	-.258772

and the LM test is $NR^2_{auxillary} \sim \chi^2_{(no. restrictions)}$ (asymptotically)

$$LM = 17321 * 0.0606 = 1049.6$$

No. restrictions = 1 and $\chi^2_{(1)}$ critical value = 3.83 (at 5% level)

So estimated $\chi^2 > \chi^2_{(1)}$ critical value

So reject null that restricted equation is acceptable. Omitted variable is relevant

However

. reg lhw exper exper2 female grad union

Source	SS	df	MS			
Model	1631.52268	5	326.304537	Number of obs =	17321	
Residual	4534.64004	17315	.261890848	F(5, 17315) =	1245.96	
				Prob > F	= 0.0000	
				R-squared	= 0.2646	
				Adj R-squared	= 0.2644	
Total	6166.16272	17320	.356014014	Root MSE	= .51175	

lhw	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
exper	.0449322	.0012022	37.37	0.000	.0425757	.0472887
exper2	-.0008684	.000026	-33.42	0.000	-.0009194	-.0008175
female	-.2885531	.0078073	-36.96	0.000	-.3038562	-.27325
grad	.5775626	.0111688	51.71	0.000	.5556707	.5994546
union	.1769883	.0084873	20.85	0.000	.1603523	.1936243
_cons	1.519343	.012831	118.41	0.000	1.494193	1.544493

. ovtest

Ramsey RESET test using powers of the fitted values of lhw

Ho: model has no omitted variables

$$F(3, 17312) = 20.04$$

$$Prob > F = 0.0000$$

A RESET test on this new regression model still suggests functional form needs adjusting

Functional form still rejected. (Note standard errors on original variables are little changed by addition of extra **relevant** variables and/or uncorrelated with original right hand side variables)