

Selection of variables, interpreting the results of a regression

Source: D. Harrison and D.L. Rubinfeld (1978), "Hedonic Housing Prices and the Demand for Clean Air," *Journal of Environmental Economics and Management* 5, 81-102. (data Hprice2 in Wooldridge)

Unit of analysis census tract in the Boston area – Most data 1970 U.S. Census.

1. price median housing price, \$
2. crime crimes committed per capita
3. nox nitrous oxide, parts per 100 million. (EPA standard 5.3)
4. rooms avg number of rooms per house
5. dist weighted dist. to 5 employ centers
6. stratio average student-teacher ratio
7. lowstat % of people 'lower status'

Variable	Obs	Mean	Std. Dev.	Min	Max
price	506	22511.51	9208.856	5000	50001
nox	506	5.549783	1.158395	3.85	8.71
crime	506	3.611536	8.590247	.006	88.976
dist	506	3.795751	2.106137	1.13	12.13
rooms	506	6.284051	.7025938	3.56	8.78
stratio	506	18.45929	2.16582	12.6	22
lowstat	506	12.70148	7.238066	1.73	39.07

1. Comparing the effect of different variables: the standardized coefficients

. reg price nox crime dist rooms lowstat stratio, beta

Source	SS	df	MS	Number of obs =	506
Model	3.0150e+10	6	5.0250e+09	F(6, 499) =	197.82
Residual	1.2675e+10	499	25401468.7	Prob > F =	0.0000
				R-squared =	0.7040
				Adj R-squared =	0.7005
Total	4.2826e+10	505	84803032	Root MSE =	5040

price	Coef.	Std. Err.	t	P> t	Beta
nox	-1757.656	331.4642	-5.30	0.000	-.2210981
crime	-80.57672	30.47786	-2.64	0.008	-.0751639
dist	-1202.372	170.5011	-7.05	0.000	-.2749917
rooms	4412.584	415.8469	10.61	0.000	.3366601
lowstat	-519.7665	48.41627	-10.74	0.000	-.4085311
stratio	-998.834	115.819	-8.62	0.000	-.2349146
_cons	34431.7	4732.075	7.28	0.000	.

2. Functional forms for explanatory variables

Variables in logarithm?

. reg lprice lnox lcrime ldist rooms lowstat stratio

Source	SS	df	MS	Number of obs =	506
Model	61.5328377	6	10.2554729	F(6, 499) =	222.02
Residual	23.0493873	499	.046191157	Prob > F =	0.0000
				R-squared =	0.7275
				Adj R-squared =	0.7242
Total	84.582225	505	.167489554	Root MSE =	.21492

lprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnox	-.4813406	.1092641	-4.41	0.000	-.6960151 -.2666662
lcrime	-.0227803	.0082673	-2.76	0.006	-.0390232 -.0065373
ldist	-.2232397	.0355563	-6.28	0.000	-.2930982 -.1533812
rooms	.1126329	.0177255	6.35	0.000	.077807 .1474588
lowstat	-.0312046	.0020751	-15.04	0.000	-.0352816 -.0271277
stratio	-.0370554	.0050761	-7.30	0.000	-.0470286 -.0270822
_cons	11.37608	.29297	38.83	0.000	10.80047 11.95168

Using quadratic terms?

```
. g dist2=dist*dist
. reg lprice lnox lcrime dist dist2 rooms lowstat stratio
```

Source	SS	df	MS	Number of obs =	506
Model	61.5588301	7	8.79411858	F(7, 498) =	190.22
Residual	23.0233949	498	.046231717	Prob > F =	0.0000
				R-squared =	0.7278
				Adj R-squared =	0.7240
Total	84.582225	505	.167489554	Root MSE =	.21502

lprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnox	-.4828057	.1112795	-4.34	0.000	-.7014409 -.2641704
lcrime	-.0203814	.0082682	-2.47	0.014	-.0366261 -.0041366
dist	-.0886245	.0257773	-3.44	0.001	-.1392702 -.0379788
dist2	.0035394	.0021225	1.67	0.096	-.0006307 .0077096
rooms	.1092026	.0177895	6.14	0.000	.0742508 .1441543
lowstat	-.0308096	.0020745	-14.85	0.000	-.0348855 -.0267338
stratio	-.0386724	.0050914	-7.60	0.000	-.0486757 -.0286691
_cons	11.43129	.307844	37.13	0.000	10.82645 12.03612

3. Choosing among non-nested models: using ldist or dist and dist2 ?

Compare adjusted R-squared

4. Interactions

Source: Wooldridge HPRICE1: Collected from the real estate pages of the *Boston Globe* during 1990. These are homes selling in the Boston, MA area.

```
price          house price, $1000s
bdrms          number of bedrooms
sqrft         size of house in square feet
lotsize       size of lot in square feet
colonial      =1 if home is colonial style
```

```
. sum price bdrms lotsize sqrft colonial
```

Variable	Obs	Mean	Std. Dev.	Min	Max
price	88	293.546	102.7134	111	725
bdrms	88	3.568182	.8413926	2	7
lotsize	88	9019.864	10174.15	1000	92681
sqrft	88	2013.693	577.1916	1171	3880
colonial	88	.6931818	.4638161	0	1

```
. gen bdrms_sqrft=bdrms*sqrft;
. reg price bdrms sqrft bdrms_sqrft lotsize
```

Source	SS	df	MS	Number of obs =	88
Model	634546.045	4	158636.511	F(4, 83) =	46.48
Residual	283308.461	83	3413.35495	Prob > F =	0.0000
				R-squared =	0.6913
				Adj R-squared =	0.6765
Total	917854.506	87	10550.0518	Root MSE =	58.424

price	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
bdrms	-33.71534	22.82291	-1.48	0.143	-79.10919 11.67852
sqrft	.0337926	.0414616	0.82	0.417	-.0486728 .1162579
bdrms_sqrft	.0218268	.0096631	2.26	0.027	.0026074 .0410462
lotsize	.0019927	.0006279	3.17	0.002	.0007439 .0032416
_cons	165.4265	87.73015	1.89	0.063	-9.065246 339.9182