



User: Stephen Sheppard
Project: IPEC

```
name: <unnamed>
log: /home/steve/Documents/IndianPoint/Contention 17/TolleyOriginalLinearSquareRoot.smcl
log type: smcl
```

```
1 . use "/home/steve/Documents/IndianPoint/Contention 17/MLS Data STATA Format.dta"
2 . * Tolley MLS model from report
3 . regress askprice ipdist ipdistsq med_income house_age attached rail_dist pilotpay_2011
```

Source	SS	df	MS	Number of obs =	296
Model	7028465.17	7	1004066.45	F(7, 288) =	26.44
Residual	10937065.5	288	37975.922	Prob > F =	0.0000
				R-squared =	0.3912
				Adj R-squared =	0.3764
Total	17965530.7	295	60900.1041	Root MSE =	194.87

askprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ipdist	-79.32316	69.82126	-1.14	0.257	-216.7478 58.1015
ipdistsq	19.90054	10.78365	1.85	0.066	-1.324214 41.12529
med_income	2.382077	.68255	3.49	0.001	1.038658 3.725496
house_age	-5.319204	.7153479	-7.44	0.000	-6.727177 -3.911231
attached	-263.2139	31.24132	-8.43	0.000	-324.7042 -201.7236
rail_dist	-38.50069	11.35376	-3.39	0.001	-60.84756 -16.15383
pilotpa~2011	10.38756	15.72596	0.66	0.509	-20.56482 41.33994
_cons	577.5368	121.6523	4.75	0.000	338.0965 816.9771

Dr. Tolley's original model

```
4 . * Tolley model with IPEC impact modeled as linear in distance
5 . regress askprice ipdist med_income house_age attached rail_dist pilotpay_2011
```

Source	SS	df	MS	Number of obs =	296
Model	6899132.97	6	1149855.5	F(6, 289) =	30.03
Residual	11066397.7	289	38292.0337	Prob > F =	0.0000
				R-squared =	0.3840
				Adj R-squared =	0.3712
Total	17965530.7	295	60900.1041	Root MSE =	195.68

askprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ipdist	46.8946	14.10106	3.33	0.001	19.14079 74.6484
med_income	2.169724	.6755748	3.21	0.001	.8400531 3.499394
house_age	-5.422697	.7161083	-7.57	0.000	-6.832146 -4.013248
attached	-276.5584	30.51919	-9.06	0.000	-336.6265 -216.4904
rail_dist	-34.94915	11.23594	-3.11	0.002	-57.0638 -12.8345
pilotpa~2011	19.19545	15.04643	1.28	0.203	-10.41904 48.80993
_cons	409.3269	80.90215	5.06	0.000	250.0948 568.559

Model with IPEC impact proportional to linear distance.

```
6 . * Tolley model with IPEC impact modeled as proportional to square root of distance
7 . generate distsquareroot=ipdist^0.5
8 . regress askprice distsquareroot med_income house_age attached rail_dist pilotpay_2011
```

Source	SS	df	MS	Number of obs =	296
Model	6843474.45	6	1140579.08	F(6, 289) =	29.64
Residual	11122056.3	289	38484.6238	Prob > F =	0.0000
				R-squared =	0.3809
				Adj R-squared =	0.3681
Total	17965530.7	295	60900.1041	Root MSE =	196.17

askprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
distsquare~t	149.1657	48.24852	3.09	0.002	54.20262	244.1287
med_income	2.131092	.6773751	3.15	0.002	.7978782	3.464306
house_age	-5.439066	.7180836	-7.57	0.000	-6.852403	-4.025729
attached	-277.9586	30.69289	-9.06	0.000	-338.3686	-217.5487
rail_dist	-32.94105	11.14021	-2.96	0.003	-54.86727	-11.01482
pilotpa~2011	20.22589	15.26558	1.32	0.186	-9.819931	50.27171
_cons	296.0152	105.2117	2.81	0.005	88.9368	503.0936

Model with
IPEC impact
proportional
to square root
of distance.

9 . * Tolley model with IPEC impact modeled as proportional to square of distance

10. regress askprice ipdistsq med_income house_age attached rail_dist pilotpay_2011

Source	SS	df	MS	Number of obs =	296
Model	6979449.73	6	1163241.62	F(6, 289) =	30.60
Residual	10986081	289	38014.1211	Prob > F =	0.0000
Total	17965530.7	295	60900.1041	R-squared =	0.3885
				Adj R-squared =	0.3758
				Root MSE =	194.97

askprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ipdistsq	7.8997	2.169942	3.64	0.000	3.628806	12.17059
med_income	2.254154	.6735366	3.35	0.001	.9284944	3.579813
house_age	-5.385319	.7133353	-7.55	0.000	-6.78931	-3.981328
attached	-272.1092	30.25947	-8.99	0.000	-331.6661	-212.5523
rail_dist	-37.46601	11.32286	-3.31	0.001	-59.75174	-15.18029
pilotpa~2011	16.62003	14.74548	1.13	0.261	-12.40211	45.64217
_cons	467.6234	73.78841	6.34	0.000	322.3925	612.8542

Model with
IPEC impact
proportional
to the square
of distance.

```

name: <unnamed>
log: /home/steve/Documents/IndianPoint/Contention 17/RepeatSalesAnalysisOfTolleyQuestions.sm
log type: smcl

```

```

1 . * Here is the model that is the basis of Dr. Sheppard's analysis:
2 . regress nomreturn salepre74post76 distkm if nomreturn>-1 & nomreturn<1, vce(cluster id)

```

```

Linear regression
Number of obs = 1511
F( 2, 506) = 9.07
Prob > F = 0.0001
R-squared = 0.0076
Root MSE = .19957

```

(Std. Err. adjusted for 507 clusters in id)

nomreturn	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
salepre74~76	-.0292563	.0084169	-3.48	0.001	-.0457926	-.01272
distkm	-.0180762	.0055988	-3.23	0.001	-.029076	-.0070764
_cons	.1585513	.0196089	8.09	0.000	.1200264	.1970762

Dr. Sheppard's model.

```

3 . * Dr. Tolley raises a question about inclusion of data where one or more sale
4 . * involved a vacant lot. Consider the impact of excluding these observations
5 . * from the data used for the estimates:

```

```
6 . generate salewithlot=0
```

```
7 . replace salewithlot=1 if lot==1
(325 real changes made)
```

```
8 . regress nomreturn salepre74post76 distkm if (salewithlot==0 & nomreturn>-1 & nomreturn<1), vce(cluster id)

```

```

Linear regression
Number of obs = 1222
F( 2, 414) = 11.33
Prob > F = 0.0000
R-squared = 0.0085
Root MSE = .18151

```

(Std. Err. adjusted for 415 clusters in id)

nomreturn	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
salepre74~76	-.0423157	.0090465	-4.68	0.000	-.0600985	-.0245328
distkm	-.0138921	.0058859	-2.36	0.019	-.025462	-.0023222
_cons	.1519886	.0211621	7.18	0.000	.1103901	.1935871

Model with any vacant lot data excluded.

```

9 . * This shows that excluding these sales strengthens the results used in Dr.
10 . * Sheppard's analysis. The negative impact of the treatment group relative to the
11 . * control is significantly larger. The results are estimated with greater
12 . * precision.

```

```

13 . * Dr. Tolley objects to including sales that occurred during one of the times of rapid
14 . * house price increase. Dropping these observations altogether is unwarranted.
15 . * If these time periods are different the preferred approach is to include an indicator

```

- 16. * or dummy variable in the model to account for any excess returns.
- 17. * We use indicator variables for the 1984Q2 to 1988Q1 time period and a separate
- 18. * indicator for the time from 1999 through 2009.

19. regress nomreturn salepre74post76 distkm dummy_80sbubble after98 if (salewithlot==0 & nomreturn>-1
> , vce(cluster id)

Linear regression

Number of obs = 1222
 F(4, 414) = 11.19
 Prob > F = 0.0000
 R-squared = 0.0251
 Root MSE = .18013

Model with
 vacant lot
 data excluded
 and indicator
 variables for
 1984-88 and
 1999-2009.

(Std. Err. adjusted for 415 clusters in id)

nomreturn	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
salepre74~76	-.0300483	.0117221	-2.56	0.011	-.0530906	-.007006
distkm	-.0150295	.0057691	-2.61	0.010	-.0263699	-.003689
dummy_80sb~e	.0653683	.023289	2.81	0.005	.0195888	.1111477
after98	.0538824	.0133926	4.02	0.000	.0275563	.0802085
_cons	.1123363	.0226813	4.95	0.000	.0677514	.1569212

- 20. *This shows that the result of Dr. Sheppard's analysis remains essentially
- 21. * unaffected by accounting for the two time periods with rapid house price appreciation.
- 22. * The time periods DO show unusually high returns to owning housing, but the impact of
- 23. * the IPEC treatment is not statistically different from Dr. Sheppard's original result.
- 24. *