

Pooled Cross Sections – Looking at the evolution of gender wage gap over time.

Source: Compiled from the 1996 and 2006 Current Population Surveys

union	float	%8.0g	=1 if respondent is union member
wage	float	%9.0g	average hourly earnings (in \$)
female	float	%9.0g	1=female, 0=male
educ	float	%9.0g	years of education
exper	float	%9.0g	years potential experience
year	float	%9.0g	1996 or 2006

Variable	Obs	Mean	Std. Dev.	Min	Max
union	4000	.98975	.9251264	0	2
wage	4000	15.59488	10.43896	.05	144.25
female	4000	.50125	.5000609	0	1
educ	4000	13.5125	2.113637	9	18
exper	4000	19.7475	12.54303	0	64
year	4000	2001	5.000625	1996	2006

1. To compare wages across different years, need to use real wages

Use the Consumer Price Index (CPI) 1996: 156.9, 2006: 201.6, i.e., inflation has been 28.5% over 10 years, which is in average 2.53% per year

```
. g rwage=wage
. replace rwage=wage*201.6/156.9) if year==1996
. label variable rwage "average real hourly earnings (in 2006$)"
. bysort year: sum wage rwage;
```

-> year = 1996

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	2000	12.69577	7.907586	.125	62.5
rwage	2000	16.31273	10.16042	.1606119	80.30593

-> year = 2006

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	2000	18.49399	11.77498	.05	144.25
rwage	2000	18.49399	11.77498	.05	144.25

```
. gen yr2006=(year==2006);
```

```
. reg rwage educ exper union female yr2006;
```

Source	SS	df	MS	Number of obs =	4000
Model	131177.324	5	26235.4647	F(5, 3994) =	293.43
Residual	357107.093	3994	89.4108896	Prob > F =	0.0000
Total	488284.417	3999	122.10163	R-squared =	0.2686
				Adj R-squared =	0.2677
				Root MSE =	9.4557

rwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ	2.327535	.071202	32.69	0.000	2.187939 2.467131
exper	.1504879	.0119518	12.59	0.000	.1270557 .1739202
union	-.7340047	.4264605	-1.72	0.085	-1.570105 .102096
female	-4.748954	.299762	-15.84	0.000	-5.336655 -4.161253
yr2006	-.2317524	.7904708	-0.29	0.769	-1.781516 1.318012
_cons	-13.79644	1.249616	-11.04	0.000	-16.24639 -11.3465

2. What happened to the gender gap?

```
. gen yr2006female=yr2006*female;

. reg rwage educ exper union female yr2006 yr2006female;
```

Source	SS	df	MS	Number of obs =	4000
Model	131623.93	6	21937.3217	F(6, 3993) =	245.60
Residual	356660.487	3993	89.3214342	Prob > F =	0.0000
Total	488284.417	3999	122.10163	R-squared =	0.2696
				Adj R-squared =	0.2685
				Root MSE =	9.451

rwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ	2.332599	.0712024	32.76	0.000	2.193003 2.472196
exper	.1507326	.0119463	12.62	0.000	.1273111 .1741541
union	-.7856355	.4268721	-1.84	0.066	-1.622543 .0512721
female	-4.078936	.4237362	-9.63	0.000	-4.909695 -3.248176
yr2006	.3487309	.8316317	0.42	0.675	-1.281731 1.979193
yr2006female	-1.339191	.5989048	-2.24	0.025	-2.513378 -.165003
_cons	-14.10789	1.256733	-11.23	0.000	-16.57178 -11.64399

Pooled Cross Sections – Measuring the impact of a policy with the difference in difference estimator.

Source: K.A. Kiel and K.T. McClain (1995), “House Prices During Sitting Decision Stages: The Case of an Incinerator from Rumor Through Operation,” *Journal of Environmental Economics and Management* 28, 241-255.

Construction of the incinerator in 1981

- | | |
|------------|---------------------------------------|
| 1. year | 1978 or 1981 |
| 2. rprice | real selling price (in 1978 dollars) |
| 3. age | age of house |
| 4. rooms | # rooms in house |
| 5. area | square footage of house |
| 6. dist | dist. from house to incinerator, feet |
| 7. nearinc | =1 if dist <= 15840 |

Does the proximity to the incinerator reduce the price of houses?

Using data in 1981, after the construction

```
. reg rprice nearinc if year==1981
```

Source	SS	df	MS	Number of obs =	142
Model	2.7059e+10	1	2.7059e+10	F(1, 140) =	27.73
Residual	1.3661e+11	140	975815069	Prob > F =	0.0000
Total	1.6367e+11	141	1.1608e+09	R-squared =	0.1653
				Adj R-squared =	0.1594
				Root MSE =	31238
rprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
nearinc	-30688.27	5827.709	-5.27	0.000	-42209.97 -19166.58
_cons	101307.5	3093.027	32.75	0.000	95192.43 107422.6

Using data from 1978, before the construction.

```
. reg rprice nearinc if year==1978
```

Source	SS	df	MS	Number of obs =	179
Model	1.3636e+10	1	1.3636e+10	F(1, 177) =	15.74
Residual	1.5332e+11	177	866239953	Prob > F =	0.0001
				R-squared =	0.0817
				Adj R-squared =	0.0765
				Root MSE =	29432
Total	1.6696e+11	178	937979126		

rprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
nearinc	-18824.37	4744.594	-3.97	0.000	-28187.62 -9461.117
_cons	82517.23	2653.79	31.09	0.000	77280.09 87754.37

Using data from 1978 (before the construction) and 1981 (after the construction)

House average real prices

	Near the incinerator (nearinc=1)		Difference	
	0	1		
1978	82517.23	63692.86	-18824.37	
1981	101307.50	70619.24	-30688.26	
Difference	18790.27	6926.38	-11863.89	

In a regression framework:

```
. gen y8lnrinc = y81*nearinc
. reg rprice y81 nearinc y8lnrinc
```

Source	SS	df	MS	Number of obs =	321
Model	6.1055e+10	3	2.0352e+10	F(3, 317) =	22.25
Residual	2.8994e+11	317	914632749	Prob > F =	0.0000
				R-squared =	0.1739
				Adj R-squared =	0.1661
				Root MSE =	30243
Total	3.5099e+11	320	1.0969e+09		

rprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y81	18790.29	4050.065	4.64	0.000	10821.88 26758.69
nearinc	-18824.37	4875.322	-3.86	0.000	-28416.45 -9232.293
y8lnrinc	-11863.9	7456.646	-1.59	0.113	-26534.67 2806.866
_cons	82517.23	2726.91	30.26	0.000	77152.1 87882.36

Then add control variables to improve the precision

```
. reg rprice y81 nearinc y8lnrinc rooms area intst age agesq
```

Source	SS	df	MS	Number of obs =	321
Model	2.2016e+11	8	2.7520e+10	F(8, 312) =	65.63
Residual	1.3084e+11	312	419346545	Prob > F =	0.0000
				R-squared =	0.6272
				Adj R-squared =	0.6177
				Root MSE =	20478
Total	3.5099e+11	320	1.0969e+09		

rprice	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
y81	12593.47	2908.527	4.33	0.000	6870.664 18316.28
nearinc	3680.788	4647.894	0.79	0.429	-5464.392 12825.97
y8lnrinc	-14540.97	5157.338	-2.82	0.005	-24688.53 -4393.413
rooms	4531.718	1664.595	2.72	0.007	1256.466 7806.97
area	22.21028	2.064422	10.76	0.000	18.14833 26.27223
intst	-.298511	.1950392	-1.53	0.127	-.6822695 .0852474
age	-895.3766	120.8499	-7.41	0.000	-1133.16 -657.5927
agesq	4.25266	.7878684	5.40	0.000	2.702453 5.802867
_cons	17384.18	11558.97	1.50	0.134	-5359.219 40127.57