

ECON 2320 – Economics of Labor and Population

Problem Set 4

Samuel Brown

Fall 2011

Due: December 13

- (a) The estimated relationship between changes in air pollution and housing prices can be seen in Table 1. Column 1 of Table 1 does not adjust for any control variables; Column 2 adjusts for the main effects of selected control variables; and Column 3 adjusts for the main effects, polynomials, and interactions of the control variables included in Column 2. In Column 1, the unadjusted correlation between changes in housing prices and TSPs has a “perverse” positive sign and is statistically significant—this suggests that pollution is an amenity rather than a disamenity, which does not correspond with common sense. When controls are added in Columns 2 and 3, however, the partial correlation between changes in housing prices and TSPs becomes economically small and statistically insignificant. This is consistent with other observables driving both pollution and housing prices, and suggests that there may be unobservables that do likewise.

Table 1: First-Difference Estimates of the Effect of TSPs Pollution on Log Housing Values

	(1)	(2)	(3)
Mean TSPs ($\times 1/100$)	0.100*** (0.031)	0.023 (0.019)	−0.003 (0.016)
Controls	No	Yes	Yes
R^2	0.017	0.553	0.854
F	10.083	40.201	.
N	1000	995	995

Equation (3) includes quadratics, cubics, and interactions of the controls.

Huber-White standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

If there are unobservable shocks that are positively correlated with both pollution and housing prices, as one would expect, then this will introduce a positive bias in the estimates in Table 1. In this case, reductions in pollution and in housing prices due to reduced economic activity are misidentified as reductions in housing prices due to reductions in pollution. Table 2 provides some evidence that this may be the case. In Column 1 of Table 2, mean economic indicators are compared between counties with above- and below-median reduction in TSPs. Above-median reductions in TSPs are associated with larger reductions in observable measures of economic activity.

- (b) In order for 1975–1976 regulatory status to be a valid instrument for pollution changes when the outcome of interest is changes in housing prices, two conditions must be met. First, being regulated must be associated with significant reductions in TSPs between 1970 and 1980. Secondly, whether or not a county is regulated must not be correlated with other determinants of changes in housing prices (such as endogenous sorting of homeowners across counties, or changes in economic activity).

Table 2 provides some evidence that 1975–1976 regulatory status may be a valid instrument. In Column 2 of Table 2, mean economic indicators are compared between regulated and unregulated counties. Observable measures of economic activity are better balanced between these two groups than those of Column 1, suggesting that 1975–1976 regulatory status may be uncorrelated with economic shocks.

- (c) The “first-stage” relationship between regulation and air pollution changes can be seen in Table 3. The “reduced-form” relationship between regulation and changes in housing prices can be seen in Table 4. The specifications used in Columns 1–3 of these tables correspond to those of Table 1. Tables 3 and 4 show that being regulated is associated with significant reductions in TSPs and significant increases in housing prices, respectively. Moreover, these estimates are largely insensitive to model specification.

Table 2: Differences in Sample Means Between Groups of Counties, Defined by TSPs Levels or Nonattainment Status

	(1) First Difference 1980–1970	(2) TSPs Nonattainment in 1975 or 1976
Housing value	−3166.017*** (705.745)	2621.042*** (789.447)
Mean TSPs	−30.956*** (1.020)	−9.854*** (1.544)
Income per capita (1982–84 dollars)	−158.767*** (40.525)	48.012 (45.448)
Unemployment rate ($\times 100$)	0.522*** (0.129)	0.031 (0.144)
% employment in manufacturing ($\times 10$)	−0.112*** (0.026)	−0.005 (0.029)
Population density	−66.859*** (24.562)	−19.050 (27.446)
% urban ($\times 10$)	−0.007 (0.005)	−0.001 (0.006)
% houses build in last 10 years	−0.034*** (0.007)	−0.007 (0.008)
<i>N</i>	1000	1000

Mean values; standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Estimates of the Impact of Mid-Decade TSPs Nonattainment on 1970–80 Changes in TSPs Pollution

	(1)	(2)	(3)
TSPs nonattainment in 1975 or 1976	−9.854*** (1.778)	−10.240*** (1.896)	−9.433*** (2.238)
Controls	No	Yes	Yes
R^2	0.039	0.098	0.457
F	30.722	4.619	.
N	1000	995	995

Equation (3) includes quadratics, cubics, and interactions of the controls.

Huber-White standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Estimates of the Impact of Mid-Decade TSPs Nonattainment on 1970–80 Changes in Log Housing Values

	(1)	(2)	(3)
TSPs nonattainment in 1975 or 1976	0.036*** (0.012)	0.023*** (0.009)	0.019** (0.009)
Controls	No	Yes	Yes
R^2	0.009	0.556	0.856
F	8.748	40.472	.
N	1000	995	995

Equation (3) includes quadratics, cubics, and interactions of the controls.

Huber-White standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The two-stage least-squares estimate of the causal effect of reductions in air pollution on housing prices can be obtained by dividing the reduced-form estimate by the first-stage estimate. (Note that this procedure does not yield estimates with a perverse sign, as was the case with the first-difference estimates.) Two-stage least-squares proceeds by replacing the instrument in the reduced-form equation with the predicted values generated by the estimated first-stage equation, and then estimating this “second stage” equation. Two-stage least-squares estimates can be seen in Table 5.

Table 5: 2SLS Estimates of the Effect of 1970–80 Changes in TSPs Pollution on Changes in Log Housing Values, Using TSPs Nonattainment in 1975 or 1976 as an Instrument

	(1)	(2)	(3)
Mean TSPs ($\times 1/100$)	−0.369*** (0.125)	−0.224*** (0.087)	−0.203** (0.094)
Controls	No	Yes	Yes
R^2	0.009	0.556	0.856
F	8.748	40.472	.
N	1000	995	995

Equation (3) includes quadratics, cubics, and interactions of the controls.

Changes in TSPs instrumented for with 1975/76 TSPs nonattainment status.

Huber-White standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The two-stage least-squares estimates in Table 5 suggest that a $1\mu\text{g}/\text{m}^3$ reduction in TSPs results in a 0.2–0.4 % increase in housing prices, with the adjusted estimates at the lower end of this range. As can be seen in Table 6, similar results are obtained when only 1975 regulatory status is used as an instrument. This suggests that the results are robust to alternative measurements of regulation status.

- (d) Since the treatment assignment is discontinuous, we can identify the program effect off of discontinuities in outcomes (changes in TSPs pollution/changes in housing values) at the threshold. This can be done with a regression discontinuity design. The resulting estimates will be valid as long as: (i) control functions can be constructed to control for confounding effects away from the threshold (or, alternatively, only observations within a narrow band of the threshold are used), (ii) there are no other discontinuities determining outcomes at the threshold, (iii) there is perfect program compliance, and (iv) there is no endogenous sorting at the threshold.

Plotting county nonattainment status against the metrics used to determine regulation status would allow one to verify the discontinuity in treatment and to check for program compliance. Plotting the change in outcomes against the metrics used to determine regulation status would allow one to verify a discontinuity in outcomes, and to identify possible confounding effects away from the threshold (for example, mean reversion bias in TSPs pollution). Plotting other observables against the metrics used to determine regulation status would allow one to verify whether there are any other *observable* discontinuities that

Table 6: 2SLS Estimates of the Effect of 1970–80 Changes in TSPs Pollution on Changes in Log Housing Values, Using TSPs Nonattainment in 1975 Only as an Instrument

	(1)	(2)	(3)
Mean TSPs ($\times 1/100$)	−0.348*** (0.121)	−0.203** (0.088)	−0.224** (0.098)
Controls	No	Yes	Yes
R^2	0.008	0.556	0.859
F	8.183	39.260	.
N	975	970	970

Equation (3) includes quadratics, cubics, and interactions of the controls.

Changes in TSPs instrumented for with 1975 TSPs nonattainment status.

Huber-White standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

may be determining outcomes at the threshold. Lastly, plotting the empirical distribution of the metrics used to determine regulation status would allow one to verify whether there is endogenous sorting over those metrics at the threshold.

- (e) The housing market can be viewed as an implicit market for houses as well as for the amenities (and disamenities) offered by different housing locations. If people have different WTP for clean air, then individuals with low WTP for clean air will sort into high-pollution areas and individuals with high WTP for clean air will sort into low-pollution areas. If there are no mobility constraints then each housing sale will reflect a match between preferences and amenities, and there will be no opportunities for Pareto improvements on those market outcomes. To the extent that there is any sorting, pollution reductions in high pollution areas will raise housing prices less than equivalent pollution reductions in low pollution areas, all else equal.

Consider the following statistical model:

$$\Delta y_{it} = \bar{\Theta} \Delta T_{it} + (\Theta_i - \bar{\Theta}) \Delta T_{it} + \Delta u_{it} \quad (1)$$

$$\Delta T_{it} = \Pi Z_i + \Delta v_{it} \quad (2)$$

where i indexes counties and t indexes time, Δy_{it} is the change in housing prices, ΔT_{it} is the change in pollution, Z_i is an instrument for changes in pollution (in this case, a regulation status indicator), and $\bar{\Theta}$ is the Average Treatment Effect (ATE). Two-stage least-squares will identify the ATE if the standard maintained hypotheses hold and $E[\Theta_i \cdot \Delta T_{it}] = 0$, viz., there is no bias introduced by the endogenous sorting of homeowners across counties.

In the context of hedonic theory, the ATE represents homeowners' average WTP for clean air. If the ATE cannot be identified because $E[\Theta_i \cdot \Delta T_{it}] > 0$, but the standard maintained hypotheses still hold, then two-stage least-squares can identify the Local Average Treatment Effect (LATE), which in this case is the average WTP for clean air among homeowners in high-pollution areas. Given the previously discussed sorting behavior, one would expect this to be smaller than the WTP for the entire population.

- (f) Two-stage estimates using a “Garen-type” control function can be seen in Table 7. These estimates are similar to those of Table 5, which suggests that the single control function implicit in two-stage least-squares does a good job of “absorbing” any bias. The coefficient on the first control function is large and significant in Columns 1 and 2, which implies that there is a lot of omitted variables bias in the first-difference estimates, even after regression adjustment. Also, the coefficient on the second control function is positive and significant in Column 1, which suggests that there is sorting based on tastes. The magnitude of this estimate drops by an order of magnitude after regression adjustment, however, and is insignificant in Column 3. This suggests that sorting can largely be explained by observable differences across counties.

Using a “Garen-type” control function identifies the ATE as long as the omitted variables bias and

selectivity bias are linear in the treatment and the instrument:

$$E[\Delta u_{it} | \Delta T_{it}, Z_i] = \lambda_T \Delta T_{it} + \lambda_Z Z_i \quad (3)$$

$$E[\Theta_i | \Delta T_{it}, Z_i] = \psi_T \Delta T_{it} + \psi_Z Z_i \quad (4)$$

Table 7: Control Function Estimates of the Capitalization of 1970–80 Changes in TSPs Pollution, with Correction for Selectivity Bias Due to Random Coefficients

	(1)	(2)	(3)
Mean TSPs ($\times 1/100$)	−0.328** (0.162)	−0.208** (0.105)	−0.201 (0.498)
v_i (first-stage residual) ($\times 1/100$)	0.504*** (0.168)	0.266** (0.106)	0.217 (0.499)
$v_i \times \text{mean TSPs}$ ($\times 1/10000$)	0.116*** (0.043)	0.048* (0.026)	0.024 (0.063)
Controls	No	Yes	Yes
R^2	0.049	0.560	0.856
F			
N	1000	995	995

Equation (3) includes quadratics, cubics, and interactions of the controls.

Changes in TSPs instrumented for with 1975/76 TSPs nonattainment status.

Bootstrap standard errors in parentheses (1000 reps.).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- (g) This study used pollution reductions induced by the Clean Air Act to estimate the average WTP for clean air, as it is capitalized in housing values. By using regulation status as an instrument for pollution reduction, the omitted variables bias that confounds conventional first-difference estimates was controlled for. By adding a “Garen-type” control function, (linear) selectivity bias was also controlled for—though the evidence suggests that the bias introduced by selectivity is small. The estimated causal effect of pollution reduction on housing values was robust to all IV specifications, at roughly 0.2–0.4 % per $1\mu\text{g}/\text{m}^3$ reduction in TSP. This robustness and the consistency of the results with economic intuition lend additional credibility to this study.