ARE 261 PSet 1

Cassandra Turk

October 2023

1

1.1

1.1.1

Done in Stata

1.1.2

Done in Stata

1.1.3

Done in Stata

1.1.4

Done in Stata

1.2

1.2.1

Because the omitted bin is the below 0 one, in Table 1 the coefficient for the C32+ bin indicates that relative to an additional day below 0 degrees an additional day above 32 degrees average temperature results in a 20 percent increase in log employment. We see that the standard error for this is very small, indicating a significant difference from zero in the coefficient.

1.2.2

We see the coefficients plotted in Figure 2. We then used these coefficients to predict log income per capita, and then graphed that against average temperature to get Figure 1. This graph seems to predict much lower income per capita for high temperatures, which seems to disagree with the notion that employment is higher with more hot days.

1.2.3

Following the suggestion, we interacted number of days in each bin with the number of days above 32 degrees, which allows us to see how cooler regions (regions with fewer 32+ days) respond to changes in temperature relative to warmer regions.

We see in this regression in Table 1 that not all interaction term coefficients are significant; the interaction with 0 to 4 is not significantly different than 0. However, several of the interaction term coefficients are significant at the .05 level, suggesting that there is probably treatment effect heterogeneity in this dataset.

2

2.1

We see in 3 that when controlling for other variables, the change in particulate level has no significant effect on the change in log home prices, even though with no covariates, there is a significant positive effect. This suggests an issue of omitted variable bias under the simple regression specification.

To check whether there is OVB, we want to see how measures of economic shocks effect measures of particulates. It seems reasonable to expect that economic shocks will have an effect on total particulate emissions; a steel or cement boom in a given county could greatly increase TSP, while a shift to less manufacturing-heavy jobs would likely reduce the amount of particulates. We estimate this by regressing the change in particulates on a few economic shock variables, found in Table 4. This shows us that all three chosen shock factors (income, unemployment, and manufacturing) all significantly predict total particulates, in what seem like reasonable directions; increased manufacturing leads to increased TSP, and more unemployment leads to less TSP. If these factors explain some of the variation in both housing prices and particulates, they should be included in our main regression.

2.2

To be a valid instrument, the selected variable must adhere to the exclusion (is the instrument predictive of the errors or other covariates?) and relevance (does the instrument predict the treatment variable?) principles. Relevance is much easier to test, so we do this first in Table 5, by regressing change in particulates on the presence of regulation. Regulation of air quality turns out to be a good predictor of change in particulate levels; the coefficient is significant (p¡.01), and very negative (a treated county is predicted to have a nearly nine-unit reduction in pollution), suggesting that, if exclusion holds, this regulation is likely to be a good instrument.

We then attempt to "test" the exclusion restriction. In practice, this is more or less impossible, but we get at the idea of it in Table 6 where we check the predictive value of the covariates for regulation. Unfortunately, for several variables, it seems that regulation can be predicted by a change in the covariate. Note that we have not–and cannot–test whether the regulation is correlated with the error terms. We will have to rely on economic intuition to argue that the error terms will be uncorrelated with the regulation, despite it seeming likely that there are unobserveables that correlate with regulation; for example, having a neighboring county with high emissions, or having been selected to have an air quality monitor. However, we will proceed onwards.

2.3

We check the first stage and reduced form estimates in Table 7. With no covariates, the existence of regulation predicts a reduction of 8.774 in the geometric mean of TSP, while with covariates the prediction is reduced to 8; both are significant. In the reduced form, we see that the existence of the regulation predicts a 4.8 percent increase in housing prices without covariates, and a 3.9 percent increase with. Again, both are significant at the .05 level.

We can use these estimates to get a coefficient estimate for the 2SLS predictor; we simply divide the reduced-form coefficient by the respective first stage coefficient, which gives us our 2SLS coefficient, as seen in Table 8. The result is that, using an instrument, the no-control specification predicts a 5.5 percent decrease in home price for a one-unit reduction in the geometric mean of TSP, and the specification with controls predicts a decrease of 4.9 percent. Both are significant.

2.4

Note: exclusion has not been "tested" again. It seems unlikely that 1974 particulates would perform better in an exclusion setting than the regulator indicator.

We begin in Table 9 by estimating the first stage and reduced form regressions; for every unit increase in the geometric mean of 1974 particulates, we observe a unit reduction in the geometric mean of particulates of .275 (.243 with controls). The reduced form shows every unit increase in the geometric mean of 1974 particulates predicting a 1 percent (1.2 percent with controls) increase in house prices. All mentioned coefficients are significant at the .05 level.

Once again, we can divide each reduced form estimate by its first stage counterpart to get our 2SLS estimate; when we re-estimate 2SLS in Table 10 we see a significant reduction in prices, at the magnitude of 4 percent with no controls, and 5 percent with controls.

2.5

These graphs can be found at Figure 3 and Figure 4. We can exploit this discontinuity in treatment to estimate a regression discontinuity at the 1974 particulates = 75 level, but this relationship is only valid if there is continuity in the outcome variable at the treatment discontinuity.

We estimate the LATE by graphing the first stage and reduced form expected means as in 2SLS; in fact, we see that the jump at the discontinuity is positive and less than .05 for the reduced form estimate. This suggests that regulation has a positive effect on housing prices, which is the same effect we see in the reduced-form estimate. The jump in the first stage figure also seems similar in magnitude to the corresponding regression, and has the same negative sign.

2.6

The graph of this can be found at Figure 5. We see the predicted curve (generated from the results of the regression without particulate matter) appears quite smooth near the discontinuity, with very little variation in price. This bodes well for our continuity in outcome requirement!

2.7

We view these graphs in Figure 6 and Figure 7.

First, we discuss the graph depicting change in particulates by regulation. We see that the decrease in particulate matter gets larger in regulated counties as they get closer to the right-hand cutoff; this makes sense, as they would likely need to reduce more than the non-regulation counties. The counties to the lefthand side of the graph likely needed to reduce their emmissions less, as they had a lower average overall, and possibly as few as two bad days.

In the graph depicting log house prices, we see that towards the cutoff, prices in the regulated counties diverge noticeably upwards. One can interpret this as the change in log housing price due to change in particulate matter due to being regulated.

2.8

Measuring the ATE relies on no unmeasured confounding variables, and randomized assignment; in this case, it's hard to assume random assignment of the treatment, so the best we can do is find the local average treatment effect. This is well-identified using 2SLS, because we can measure the change in housing prices as a function of only the change in particulates that were due to the effect of regulation.

2.9

While studying direct-effect regressions, we failed to find an effect of change in particulate matter on change in log house price. After some consideration, we decided to exploit a regulation policy to predict change in pollution; this allowed us to run an instrumental variable regression that gave us a negative effect of pollution on housing prices. This is more in line with what we know about peoples' preferences, as in general we expect that people prefer less pollution. We took this instrument a step further, and predicted change in pollution using the 1974 average pollution, which is a criteria for the regulation. This gave us a similar negative effect to what we saw in the last IV. However, IV relies partially on instrument exogeneity from the control variables, which we see a violation of in Table 6. Additionally, there should be no correlation between the instrument (regulation or 1974 pollution) and the error terms of the regression, which is impossible to test, and in this case, may also be hard to justify, given the many unobserveables that tend to be correlated with pollution.

The next design we estimated was a regression discontinuity; we exploited the design of the regulation to estimate the effect of the regulation on pollution, and the effect of pollution on housing prices. However, this design relies on continuity of the endogenous variable across the discontinuity–which we showed to be a reasonable assumption by predicting housing prices with the controls, and showing that around the discontinuity point, this prediction seemed fairly smooth. Using the overlap of the policy and the area under the cutoff, we saw that the regulation does seem to have an effect on pollution levels. One consideration to take into account when estimating a regression discontinuity is whether groups are able to manipulate their status around the threshold; for example, if certain counties knew they would be able to reduce emissions cheaply, and were able to get financial assistance from the government as a result of being in nonattainment, counties may have chosen to emit slightly more than they otherwise would have, and be classified as under regulation. There may be other reasons why various counties would have manipulated their status around the edges of the cutoffs, and this would effect the legitimacy of the design.

In short, there are benefits and drawbacks to every regression design, and we will never be able to say with full certainty that we are absolutely correct. However, it seems reasonable to take away from our analysis that a reduction in pollution is predictive of an increase in housing prices.

3 Tables

Table 1: Real Farm Proprietors Income on Cubic Spline of Mean Temperature

	(1) OLS	(2) FE
(mean) splineC1	-0.000665***	-0.000396***
	(0.0000)	(0.0000)
(mean) splineC2	0.00539***	0.00350***
	(0.0002)	(0.0003)
(mean) splineC3	-0.0141***	-0.0114***
	(0.0009)	(0.0009)
(mean) splineC4	0.00863***	0.00927***
	(0.0014)	(0.0015)
Constant	8.041***	9.110***
	(0.0430)	(0.1079)
Observations	86555	86554

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Farm Employment on Temperature Bins

	(1) OLS	(2) FE	(3) interacted	(4) interacted FE
(mean) temp0to4	-0.00167***	0.000264**	-0.00164***	0.000303**
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
(mean) temp4to8	-0.00381***	0.000329***	-0.00351***	0.000279**
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
(mean) temp8to12	-0.000232	0.000395***	-0.000393	0.000354**
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
(mean) temp12to16	-0.00222***	0.000819***	-0.00260***	0.000806***
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
(mean) temp16to20	0.0000885	0.00128***	0.0000434	0.00128***
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
(mean) temp20to24	0.00610***	0.00130***	0.00614***	0.00130***
	(0.0002)	(0.0001)	(0.0002)	(0.0001)
(mean) temp24to28	-0.00310***	0.00157***	-0.00315***	0.00157***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(mean) temp28to32	0.000993***	0.00109***	0.00131***	0.00105***
	(0.0002)	(0.0001)	(0.0002)	(0.0002)
(mean) tempA32	0.0217***	0.00286***	0.408***	-0.0646***
	(0.0009)	(0.0005)	(0.0780)	(0.0196)
temp_interact_32_temp0to4			-0.00228***	0.0000414
			(0.0005)	(0.0001)
temp_interact_32_temp4to8			-0.00173***	0.000380***
			(0.0002)	(0.0001)
temp_interact_32_temp8to12			-0.000443	0.000208***
			(0.0002)	(0.0001)
temp_interact_32_temp12to16			-0.000604**	0.000228***
			(0.0002)	(0.0001)
temp_interact_32_temp16to20			-0.000353	0.000174**
			(0.0002)	(0.0001)
temp_interact_32_temp20to24			-0.00111***	0.000176**
			(0.0002)	(0.0001)
temp_interact_32_temp24to28			-0.00114***	0.000224***
			(0.0002)	(0.0001)
temp_interact_32_temp28to32			-0.00168***	0.000144**
			(0.0002)	(0.0001)
temp_interact_32_tempA32			-0.00176***	0.000129*
			(0.0002)	(0.0001)
Constant	6.758***	6.358***	6.768***	6.360***
	(0.0281)	(0.0283)	(0.0282)	(0.0283)
Observations	126619	126619	126619	126619

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: House Price on Particulates

	(1) OLS	(2) OLS 2	(3) OLS weighted	(4) OLS 2 weighted
dgtsp	0.00104***	0.000371*	0.00299***	0.000354
	(0.0002)	(0.0002)	(0.0003)	(0.0002)
dwhite		-0.782***		-1.287***
		(0.1021)		(0.1032)
ddens		0.0000346*		0.0000332**
		(0.0000)		(0.0000)
dfeml		-0.898		-1.701*
		(0.5340)		(0.7495)
dage65		-3.041***		-2.427***
		(0.3744)		(0.4821)
dhs		0.799***		-0.385
		(0.1516)		(0.1990)
dcoll		-0.323		-0.438
		(0.2132)		(0.2835)
durban		-0.0155		-0.0667
		(0.0530)		(0.0852)
dunemp		-1.138***		-4.125***
		(0.2436)		(0.3146)
dpoverty		-0.888***		-1.087***
		(0.1757)		(0.2716)
dplumb		-0.0387		-0.0218
		(0.1983)		(0.3534)
dincome		0.0000925***		0.0000976***
		(0.0000)		(0.0000)
dvacant		0.364*		1.269***
		(0.1435)		(0.1953)
downer		-0.190		0.257
		(0.1088)		(0.1438)
drevenue		-0.0000289		0.000150***
		(0.0000)		(0.0000)
dtaxprop		0.00000487		-0.000352***
		(0.0001)		(0.0001)
depend		0.0000272		-0.000106*
		(0.0000)		(0.0000)
dmnfcg		-0.0761		-0.831***
		(0.1188)		(0.1845)
Constant	1.083***	0.886***	1.072***	1.096***
	(0.0058)	(0.0288)	(0.0081)	(0.0370)
Observations	1001	994	1001	994

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Particulates on Shock Factors

	(1)
	OLS
dincome	0.00815***
	(0.0009)
dunemp	-42.06
	(34.6223)
dmnfcg	105.6***
	(21.1793)
Constant	-18.01***
	(1.3124)
Observations	1001

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Particulates on Regulation

	(1) Relevance	(2) Relevance 2
tsp7576	-8.774***	-8.009***
	(1.1841)	(1.1852)
dwhite		-0.538
		(14.0247)
ddens		0.00368*
		(0.0015)
dfeml		208.9*
		(101.0707)
dage65		-83.15
		(64.7605)
dhs		22.84
		(26.7747)
dcoll		58.37
		(38.1499)
durban		-1.471
		(11.4528)
dunemp		-58.64
		(42.2579)
dpoverty		29.27
		(36.7578)
dplumb		14.41
		(47.5451)
dincome		0.00527***
		(0.0014)
dvacant		28.40
		(26.2721)
downer		-17.20
		(19.3559)
drevenue		0.00715*
		(0.0030)
dtaxprop		-0.0298***
		(0.0089)
depend		-0.00592
		(0.0061)
dmnfcg		104.0***
		(24.5866)
Constant	-11.89***	-13.87**
	(0.8121)	(5.0020)
Observations	1001	994

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Regulation on Covariates

	(1)
	Exclusion
dwhite	-1.793***
	(0.3744)
ddens	-0.0000497
	(0.0000)
dfeml	-10.16***
	(2.7102)
dage65	0.205
	(1.7490)
dhs	-1.138
	(0.7222)
dcoll	-2.580*
	(1.0270)
durban	-0.182
	(0.3093)
dunemp	0.361
	(1.1412)
dpoverty	3.596***
	(0.9860)
dplumb	-1.655
	(1.2830)
dincome	0.000105**
	(0.0000)
dvacant	0.829
	(0.7090)
downer	1.133*
	(0.5215)
drevenue	0.000237**
	(0.0001)
dtaxprop	-0.000985***
	(0.0002)
depend	0.0000677
	(0.0002)
dmnfcg	-0.914
	(0.6634)
Constant	0.648***
	(0.1335)
Observations	994

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Particulates on Regulation and House Prices on Regulation

	(1) First Stage	(2) First Stage 2	(3) Reduced Form	(4) Reduced Form 2
tsp7576	-8.774***	-8.009***	0.0486***	0.0391***
	(1.1841)	(1.1852)	(0.0128)	(0.0087)
dwhite		-0.538		-1.212***
		(14.0247)		(0.1034)
ddens		0.00368*		0.0000366**
		(0.0015)		(0.0000)
dfeml		208.9*		-1.201
		(101.0707)		(0.7451)
dage65		-83.15		-2.465***
		(64.7605)		(0.4774)
dhs		22.84		-0.329
		(26.7747)		(0.1974)
dcoll		58.37		-0.310
		(38.1499)		(0.2812)
durban		-1.471		-0.0596
		(11.4528)		(0.0844)
dunemp		-58.64		-4.161***
		(42.2579)		(0.3115)
dpoverty		29.27		-1.228***
		(36.7578)		(0.2710)
dplumb		14.41		0.0527
		(47.5451)		(0.3505)
dincome		0.00527***		0.0000950***
		(0.0014)		(0.0000)
dvacant		28.40		1.244***
		(26.2721)		(0.1937)
downer		-17.20		0.204
		(19.3559)		(0.1427)
drevenue		0.00715*		0.000143***
		(0.0030)		(0.0000)
dtaxprop		-0.0298***		-0.000321***
		(0.0089)		(0.0001)
depend		-0.00592		-0.000111*
		(0.0061)		(0.0000)
dmnfcg		104.0***		-0.756***
		(24.5866)		(0.1812)
Constant	-11.89***	-13.87**	1.001***	1.063***
	(0.8121)	(5.0020)	(0.0088)	(0.0369)
Observations	1001	994	1001	994

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: 2SLS House Prices on Particulates

	(1) 2SLS	(2) 2SLS 2
dgtsp	-0.00553**	-0.00488***
	(0.0018)	(0.0013)
dwhite		-1.215***
		(0.1272)
ddens		0.0000546***
		(0.0000)
dfeml		-0.181
		(0.9910)
dage65		-2.871***
		(0.5989)
dhs		-0.218
		(0.2465)
dcoll		-0.0246
		(0.3612)
durban		-0.0668
		(0.1040)
dunemp		-4.448***
		(0.3924)
dpoverty		-1.085**
		(0.3315)
dplumb		0.123
		(0.4329)
dincome		0.000121***
		(0.0000)
dvacant		1.383***
		(0.2401)
downer		0.120
		(0.1788)
drevenue		0.000178***
		(0.0000)
dtaxprop		-0.000467***
		(0.0001)
depend		-0.000140*
		(0.0001)
dmnfcg		-0.248
		(0.2685)
Constant	0.935***	0.996***
	(0.0305)	(0.0517)
Observations	1001	994

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Particulates on 1974 Particulates and House Prices on 1974 Particulates

	(1) First Stage	(2) First Stage 2	(3) Reduced Form	(4) Reduced Form 2
mtspgm74	-0.275***	-0.243***	0.00109***	0.00124***
	(0.0297)	(0.0326)	(0.0003)	(0.0002)
dwhite		-22.18		-1.101***
		(14.6954)		(0.1094)
ddens		0.00300		0.0000400***
		(0.0015)		(0.0000)
dfeml		139.8		-0.742
		(102.5767)		(0.7636)
dage65		-95.90		-2.375***
		(64.8420)		(0.4827)
dhs		20.10		-0.304
		(26.8896)		(0.2002)
dcoll		35.64		-0.199
		(38.5781)		(0.2872)
durban		-8.553		-0.0409
		(11.7851)		(0.0877)
dunemp		-41.45		-4.318***
		(42.5680)		(0.3169)
dpoverty		3.316		-1.107***
		(36.5913)		(0.2724)
dplumb		-17.99		0.203
		(48.1232)		(0.3583)
dincome		0.00555***		0.0000932***
		(0.0014)		(0.0000)
dvacant		40.40		1.195***
		(26.5167)		(0.1974)
downer		-3.263		0.133
		(19.6586)		(0.1464)
drevenue		0.00587*		0.000148***
		(0.0030)		(0.0000)
dtaxprop		-0.0334***		-0.000305***
		(0.0089)		(0.0001)
depend		-0.00623		-0.000106*
		(0.0061)		(0.0000)
dmnfcg		89.56***		-0.692***
		(24.7984)		(0.1846)
Constant	1.692	-0.0271	0.953***	0.991***
	(2.0026)	(5.5959)	(0.0222)	(0.0417)
Observations	974	967	974	967

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: 2SLS House Prices on Particulates

	(1) 2SLS	(2) 2SLS 2
dgtsp	-0.00399**	-0.00512***
	(0.0014)	(0.0013)
dwhite		-1.215***
		(0.1306)
ddens		0.0000553***
		(0.0000)
dfeml		-0.0265
		(1.0081)
dage65		-2.866***
		(0.6146)
dhs		-0.201
		(0.2535)
dcoll		-0.0170
		(0.3697)
durban		-0.0847
		(0.1100)
dunemp		-4.530***
		(0.4046)
dpoverty		-1.090**
		(0.3414)
dplumb		0.111
		(0.4467)
dincome		0.000122***
		(0.0000)
dvacant		1.402***
		(0.2486)
downer		0.117
		(0.1842)
drevenue		0.000178***
		(0.0000)
dtaxprop		-0.000475***
		(0.0001)
depend		-0.000138*
		(0.0001)
dmnfcg		-0.234
		(0.2675)
Constant	0.960***	0.991***
	(0.0237)	(0.0522)
Observations	974	967

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4 Graphs

Figure 1: 1.2.2

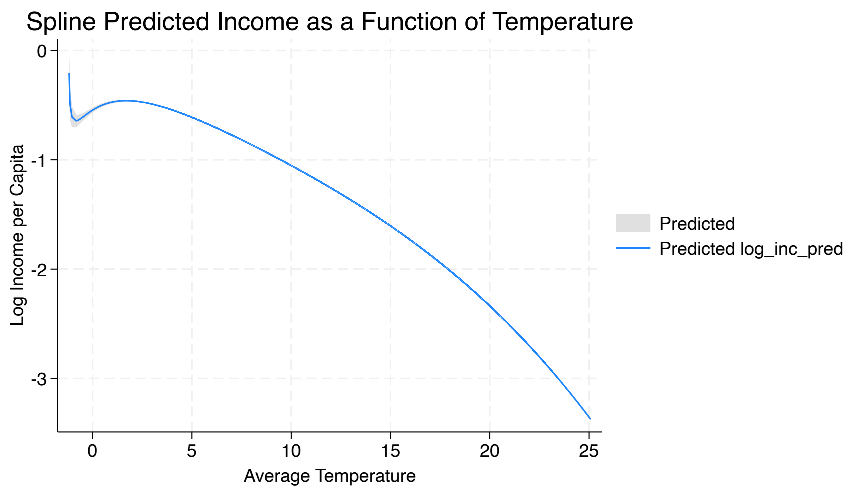


Figure 2: 1.2.2

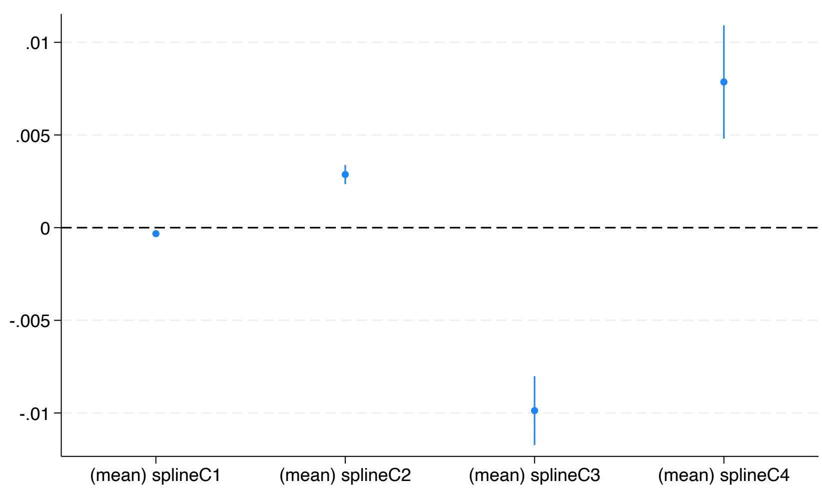


Figure 3: 2.5.1

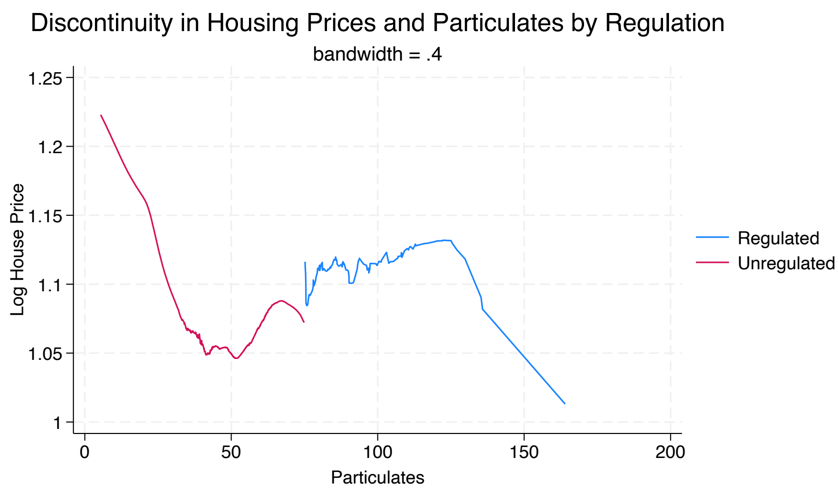


Figure 4: 2.5.2

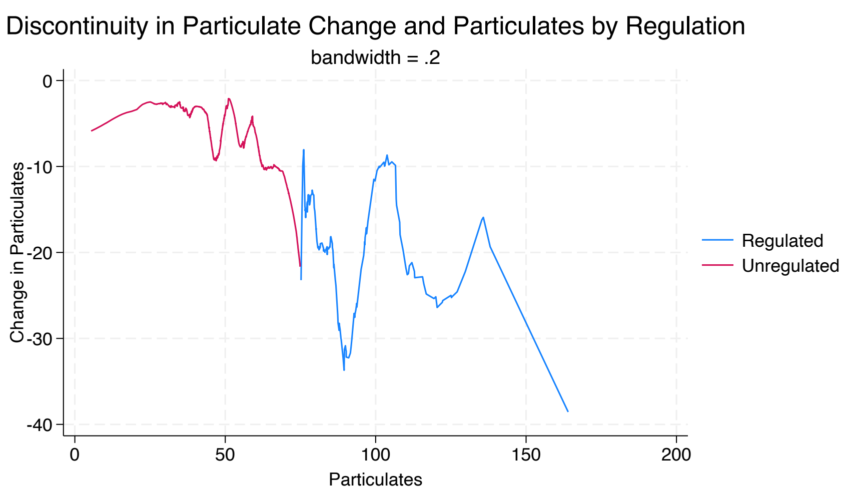


Figure 5: 2.6.1

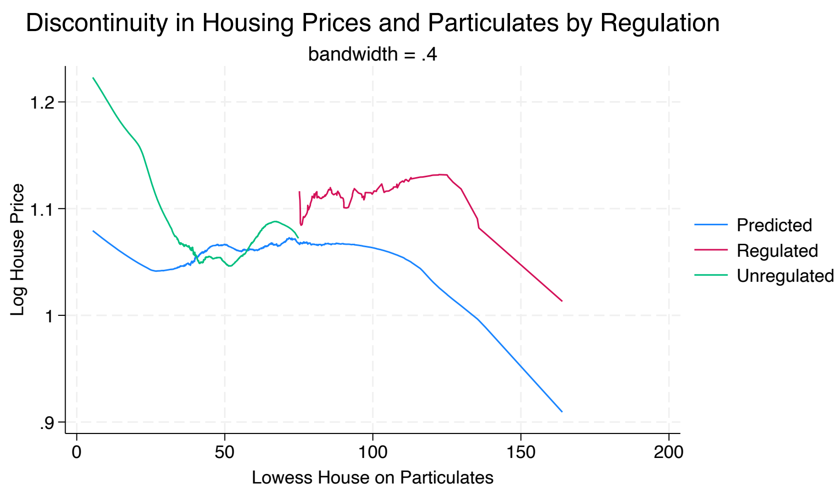


Figure 6: 2.6.2

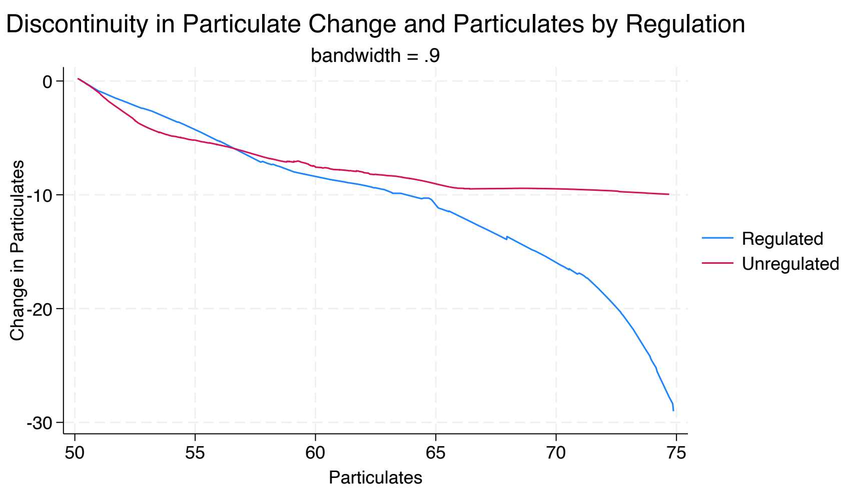
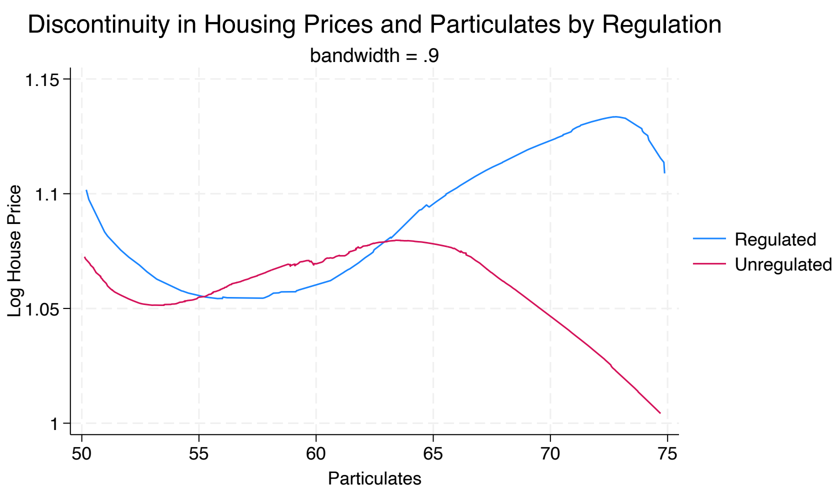


Figure 7: 2.7



Discontinuity in Housing Prices and Particulates by Regulation
bandwidth = .9
1.15
1.1
1.05
1
Log House Price
50
55
60
65
70
75
Particulates
Regulated
Unregulated