Econ 244 - Homework #4

1) (CRE vs Fixed Effects)

(a) Prove that the LSDV estimator of this model is consistent for β:

The LSDV estimator is given estimating

$$Y_{it} = \sum_{j=1}^{N} \gamma_j D_{it}^j + \beta X_{it} + \epsilon_{it}$$

By the Frisch-Waugh-Lovell theorem, the estimated coefficient β is equivalent to regressing $\tilde{Y}_{it}$ on $\tilde{X}_{it}$, $\tilde{Y}_{it}$ and $\tilde{X}_{it}$ are the regressions of Y and X on the dummy variables. This is equivalent to demeaning the variables by individual i, so $\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$ and $\tilde{X}_{it} = X_{it} - \bar{X}_i$. This is consistent for β if $E[\tilde{\epsilon}_{it}\tilde{X}_{it}] = 0$:

$$\begin{aligned} E[\tilde{\epsilon}_{it}\tilde{X}_{it}] &= E[(\epsilon_{it} - \bar{\epsilon}_i)(X_{it} - \bar{X}_i)] \\ &= E[(\epsilon_{it} - \bar{\epsilon}_i)(X_{it} - \bar{X}_i)] \end{aligned}$$

which holds if there is strict exogeneity of ϵ_{it}, because $\bar{\epsilon}_i$ is a function of future and lag values of ϵ_{it} as well; that is, if X_{it} is uncorrelated with not only ϵ_{it} but also future and lag values of ϵ_{it}.

(b) Show that estimating $Y_{it} = \alpha_0 + \alpha_1 \bar{X}_i + \beta X_{it} + u_{it}$ yields the same estimate as the LSDV estimator:

By using the F-W-L theorem, estimating the above is equivalent to estimating

$$Y_{it} - \bar{Y} = \alpha_1(\bar{X}_i - \bar{X}) + \beta(X_{it} - \bar{X}) + (u_{it} - \bar{u})$$

That is, regressing the population-demeaned variables on each other without a constant. Applying F-W-L again, note that the annihilator matrix for $\bar{X}_i - \bar{X}$ is

$$M = I - \frac{1}{T\sum_j(\bar{X}_j - \bar{X})^2}\left[[(\bar{X}_i - \bar{X})(\bar{X}_j - \bar{X})]\right]$$

that is, every element of the last term is $(X_i - \bar{X})(X_j - \bar{X})$. Applying the annihilator matrix to the LHS, the ith element of the LHS vector is

$$\begin{aligned}
(MY)_i &= Y_{it} - \bar{Y} - \frac{(\bar{X}_i - \bar{X}) \sum_{j,t} \left(\bar{X}_j - \bar{X}\right) \left(Y_{jt} - \bar{Y}\right)}{T \sum_j (\bar{X}_j - \bar{X})^2} \\
&= Y_{it} - \bar{Y} - \frac{(\bar{X}_i - \bar{X}) \sum_j \left(\bar{X}_j - \bar{X}\right) \sum_t \left(Y_{jt} - \bar{Y}\right)}{T \sum_j (\bar{X}_j - \bar{X})^2} \\
&= Y_{it} - \bar{Y} - \frac{(\bar{X}_i - \bar{X}) \sum_j \left(\bar{X}_j - \bar{X}\right) \left(\bar{Y}_j - \bar{Y}\right)}{\sum_j (\bar{X}_j - \bar{X})^2} \\
&= Y_{it} - \bar{Y} - \frac{(\bar{X}_i - \bar{X}) \sum_j \left(\bar{X}_j - \bar{X}\right) \left([a_0 + a_1\bar{X}_j + \beta\bar{X}_j + \bar{u}_j] - [a_0 + a_1\bar{X} + \beta\bar{X} + \bar{u}]\right)}{\sum_j (\bar{X}_j - \bar{X})^2} \\
&= Y_{it} - \bar{Y} - \frac{(\bar{X}_i - \bar{X}) \sum_j \left(\bar{X}_j - \bar{X}\right) \left((a_1 + \beta)(\bar{X}_j - \bar{X}) + \bar{u}_j - \bar{u}\right)}{\sum_j (\bar{X}_j - \bar{X})^2} \\
&= Y_{it} - \bar{Y} - (a_1 + \beta)(\bar{X}_i - \bar{X}) - \frac{(\bar{X}_i - \bar{X}) \sum_j \left(\bar{X}_j - \bar{X}\right) (\bar{u}_j - \bar{u})}{\sum_j (\bar{X}_j - \bar{X})^2} \\
&= Y_{it} - \bar{Y}_i
\end{aligned}$$

and similar algebra applied to regressing $X_{it} - \bar{X}$ on $\bar{X}_i - \bar{X}$ shows that the full regression is equivalent to regressing $Y_{it} - \bar{Y}_i$ on $X_{it} - \bar{X}_i$, which is the same as LSDV.

(c) Describe a procedure for testing for zero correlation between the person effects and X_{it}: one could regress T_{it} on $\bar{X}_{it}$ and X_{it}, as before, and t-test the hypothesis that the coefficient on $\bar{X}_{it}$ is 0. If one rejects the hypothesis, then that is evidence for correlation between the person effects and X_{it}.

(d) Propose estimators for $Var(v_i)$ and $Var(\alpha_i)$:

For $Var(v_i)$, an estimator is the saple variance, given by $\frac{1}{N-1} \sum_i (\hat{\alpha}_i - \hat{a}_0 - \hat{a}_1 \bar{X}_i)^2$, where $\hat{\alpha}$ is taken from the fixed-effects regression and α_0 and α_1 is taken from the regression from part (b).

For $Var(\alpha_i)$, since $\alpha_i = a_0 + a_1\bar{X}_i + v_i$, the estimator for α_i is $Var(\bar{X}_i) + \hat{Var}(v_i)$.

2) (Event Study)

(a) See do file.

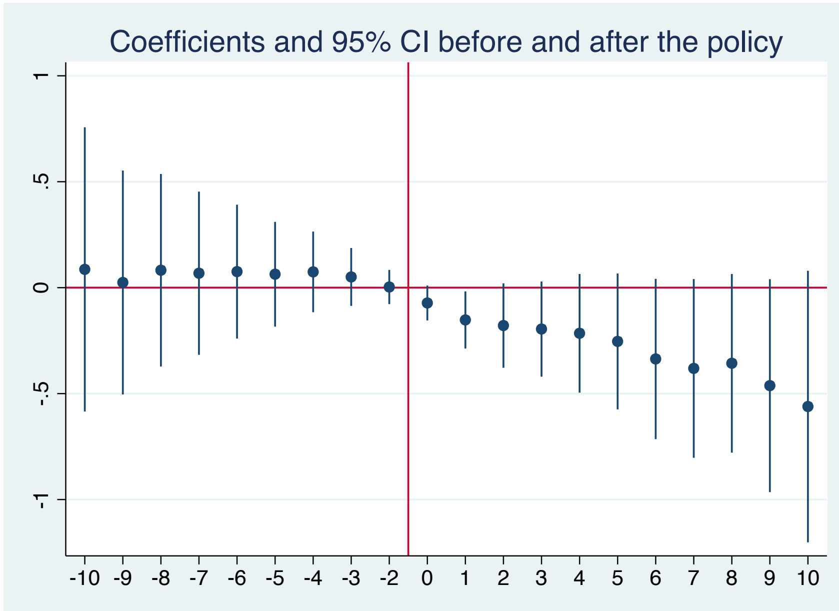
(b) See do file and table 1.

(c) It appears arrests go down in the years following the policy relative just before the curfew was put in place. In the years prior to the policy, arrest rates were fairly flat. However, the estimates are fairly imprecise outside of a couple year band around the policy. Note that data from 10 years before the policy and 10 years after were binned into the -10 and 10 dummies.

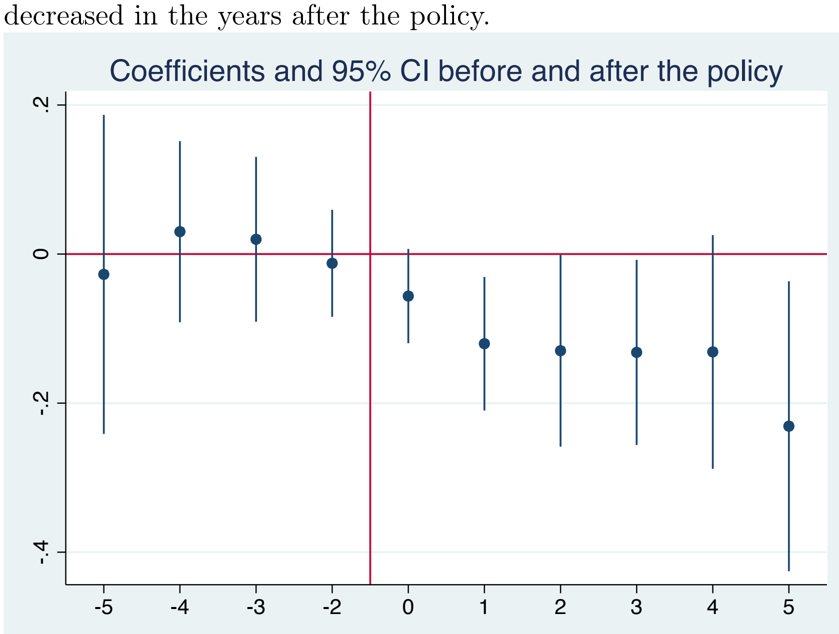
Table 1: Question 2b and 2e

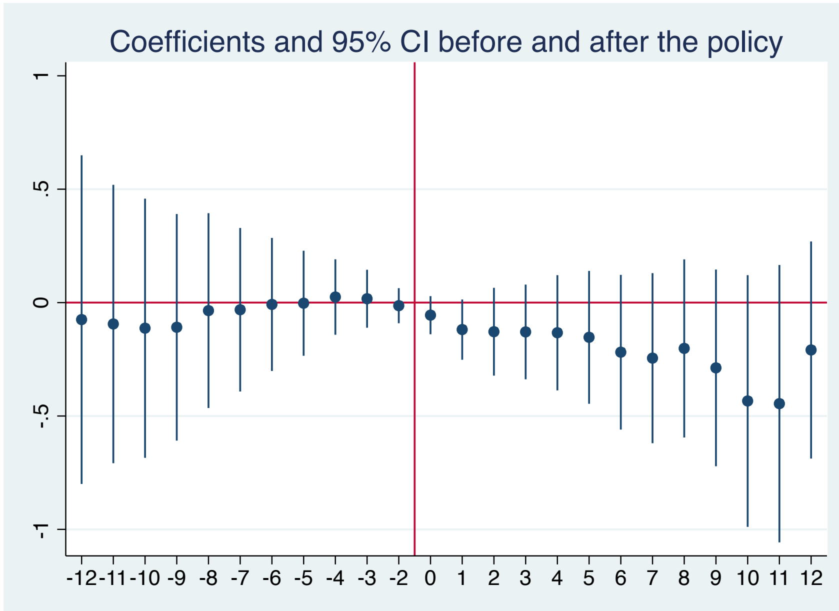
	(1) Log Arrests	(2) Log Arrests
0	-0.0724*	-0.0447
	(0.04097)	(0.03410)
1	-0.153**	-0.0994*
	(0.06718)	(0.05405)
2	-0.179*	-0.103
	(0.09928)	(0.08530)
3	-0.196*	-0.0969
	(0.11195)	(0.09814)
4	-0.216	-0.0984
	(0.13955)	(0.13022)
5	-0.254	-0.127
	(0.15976)	(0.15272)
6	-0.337*	-0.188
	(0.18844)	(0.18051)
7	-0.381*	-0.225
	(0.21022)	(0.21005)
8	-0.357*	-0.206
	(0.21014)	(0.21727)
9	-0.463*	-0.293
	(0.25040)	(0.25997)
10	-0.561*	-0.423
	(0.31938)	(0.37817)
-2	0.00306	-0.0266
	(0.04029)	(0.03919)
-3	0.0505	-0.0118
	(0.06806)	(0.06462)
-4	0.0744	-0.0217
	(0.09482)	(0.07838)
-5	0.0634	-0.0682
	(0.12321)	(0.09156)
-6	0.0756	-0.116
	(0.15743)	(0.11055)
-7	0.0681	-0.163
	(0.19197)	(0.13063)
-8	0.0822	-0.174
	(0.22642)	(0.15078)
-9	0.0243	-0.277
	(0.26328)	(0.17508)
-10	0.0862	-0.342
	(0.33427)	(0.20880)
Constant	6.545***	6.556***
	(0.26179)	(0.10881)
Year Fixed Effects	X	X
City Fixed Effects	X	X
City-Specific Time Trend		X
R^2	0.804	0.870
Observations	1297	1297

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

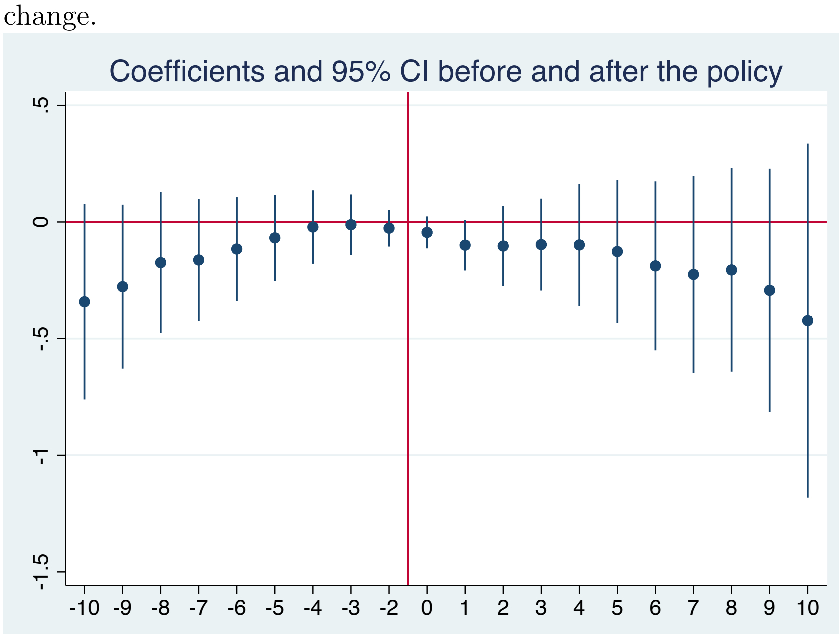


(d) We can restrict our analysis to just look 5 years before and after the policy or 12 years before and after. We can see a slight change in the coefficients but the general pattern is similar–that arrest rates were fairly flat beforehand and decreased in the years after the policy.





(e) When we add a linear city-specific trend, we get a slightly different story from the data. It appears that arrests were slightly increasing prior to the policy change and then decreased afterwards. Perhaps explaining the impetus for the policy change.



3) (Dynamic Panel)

4) (CRE Event Study)

(a) See do file.

(b) Estimating the model with “mvreg”.

(c) To ease notation, let $H \equiv \{85, 87, 88, 89, 90\}$ be the set of event dates.

	Log of Arrests Made in						
	1984	1985	1986	1987	1988	1989	1990
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Enacted in 1985	-0.50	-0.76	-0.71	-0.60	-0.59	-0.65	-0.85
	(0.66)	(0.64)	(0.72)	(0.76)	(0.72)	(0.71)	(0.68)
Enacted in 1987	0.70	0.49	0.47	0.20	0.03	0.06	0.03
	(0.47)	(0.46)	(0.52)	(0.54)	(0.52)	(0.51)	(0.49)
Enacted in 1988	1.01**	0.92**	1.01*	1.08*	1.18**	1.16**	1.18**
	(0.47)	(0.46)	(0.52)	(0.54)	(0.52)	(0.51)	(0.49)
Enacted in 1989	-0.62	-0.67*	-0.69	-0.76*	-0.56	-0.74*	-0.87**
	(0.39)	(0.38)	(0.43)	(0.45)	(0.43)	(0.42)	(0.40)
Enacted in 1990	-0.06	-0.15	-0.19	-0.22	-0.28	-0.26	-0.26
	(0.34)	(0.33)	(0.37)	(0.39)	(0.37)	(0.37)	(0.35)
Year Intercept	6.80***	6.91***	6.90***	6.86***	6.87***	6.97***	7.00***
	(0.10)	(0.10)	(0.11)	(0.12)	(0.11)	(0.11)	(0.11)
N	53						

First, observe that we can write $\pi_{k,t}$ as including the dynamic causal effects as well as the endogenous component. That is, we can write:

$$\pi_{k,t} = \delta_{t-k}\mathbf{1}(t \geq k) + \eta_k$$

As is stated in the original problem, this directly implies the first set of linear restrictions:

$$\pi_{k,t} = \pi_{k,t'} \ \forall (t,t') < k$$

Next, observe that for all h such that $k, k' \in H$ and for all t and s, we have

$$\begin{aligned}\pi_{k',k'+t} - \pi_{k,k+t} &= \eta_{k'} - \eta_k \\ &= \pi_{k',k'+s} - \pi_{k,k+s}\end{aligned}$$

Similarly, we also have for $k, k' \in H$ and time periods t and s:

$$\begin{aligned}\pi_{k,k+t+s} - \pi_{k,k+t} &= \delta_{t+s} - \delta_t \\ &= \pi_{k',k'+t+s} - \pi_{k',k'+t}\end{aligned}$$