Spatial Economics – Assignment 3

Elia Di Gregorio (h12039313@s.wu.ac.at) Max Heinze (h11742049@s.wu.ac.at)
Sophia Ludescher (h11801892@s.wu.ac.at)

May 3, 2024

Contents

The executable code that was used in compiling the assignment is available on GitHub at https://github.com/maxmheinze/spatial.

Task A

Naïve Panel Model and Specification Search

We begin by fitting the spatially naïve panel model using the `plm` function. The output is printed below.

```
cm1 <- plm(logc ~ logp + logy, data = cigs, effect = "twoway", model = "within",
    index = c("state", "year"))
```

	Dependent variable:
	logc
logp	-1.035^{***}
	(0.042)
logy	0.529^{***}
	(0.047)
Observations	1,380
R^2	0.394
Adjusted R^2	0.359
F Statistic	424.344^{***} (df = 2; 1303)
Note:	*p<0.1; **p<0.05; ***p<0.01

As the true spatial econometricians we are, we now ignore everything that is reminiscent of economic theory and venture on a standard specification testing path. We begin by performing a Lagrange Multiplier test for spatial lags in both the error and the dependent. The results are printed below.

```
slmtest(cm1, cigm, test = "lme")
```

```
##
##  LM test for spatial error dependence
##
## data:  formula (within transformation)
## LM = 54.655, df = 1, p-value = 1.437e-13
## alternative hypothesis: spatial error dependence
```

```
slmtest(cm1, cigm, test = "lml")
```

```
##
##  LM test for spatial lag dependence
##
## data:  formula (within transformation)
## LM = 46.901, df = 1, p-value = 7.468e-12
## alternative hypothesis: spatial lag dependence
```

As we can see, both tests return significant p-values, meaning that we reject the null hypotheses of spatial dependence being present in the errors or the dependent, respectively. We thus continue by performing robust LM tests for spatial dependence in both the errors and the dependent which account for the respectively other type of spatial dependence potentially being present. The results of these two tests are printed below.

```
slmtest(cm1, cigm, test = "rlme")
```

```
##
##  Locally robust LM test for spatial error dependence sub spatial lag
##
## data:  formula (within transformation)
## LM = 8.9106, df = 1, p-value = 0.002835
## alternative hypothesis: spatial error dependence
```

```
slmtest(cm1, cigm, test = "rlml")
```

```
##
##  Locally robust LM test for spatial lag dependence sub spatial error
```

```
##
## data:  formula (within transformation)
## LM = 1.1563, df = 1, p-value = 0.2822
## alternative hypothesis: spatial lag dependence
```

Since this time, testing for spatial dependence in the error given spatial dependence in the dependent returns a significant result, and testing for spatial dependence in the dependent given spatial dependence in the errors does not, we conclude that based on these test results, a **SEM model** is the most appropriate specification.

Estimating a SEM Model and Comparing it to an SLX Model

We use the `spml` function from the `splm` package to estimate a spatial panel model using Maximum Likelihood. Consistent with our previous answer, we estimate a **SEM model**, and we do this by setting `lag = FALSE` as well as `spatial.error = "b"`.

```
cm2 <- spml(logc ~ logp + logy, data = cigs, listw = cigm, effect = "twoways", model =
"within",
    index = c("state", "year"), lag = FALSE, spatial.error = "b")
```

The results of this estimation are printed in the following.

	Dependent variable:
	logc
logp	-1.004^{***}
	(0.040)
logy	0.554^{***}
	(0.049)
Spatial error rho	0.240^{***}
	(0.033)
Observations	1,380
Note:	*p<0.1; **p<0.05; ***p<0.01

Next, we estimate an **SLX model** using the `plm` function in combination with the `slag` function to create spatial lags. Since there neither is a lag of the dependent nor are there spatial errors, there is no need to use a function for a dedicatedly spatial panel model.

```
cm3 <- plm(logc ~ logp + logy + slag(logp, listw = cigm) + slag(logy, listw = cigm),
    data = cigs, effect = "twoways", model = "within", index = c("state", "year"))
```

Again, the results are printed below.

	Dependent variable:
	logc
logp	-1.017^{***}
	(0.042)
logy	0.608^{***}
	(0.060)
slag(logp, listw = cigm)	-0.220^{***}
	(0.077)
slag(logy, listw = cigm)	-0.219^{***}
	(0.080)
Observations	1,380
R^2	0.400
Adjusted R^2	0.364
F Statistic	217.105^{***} (df = 4; 1301)
Note:	*p<0.1; **p<0.05; ***p<0.01

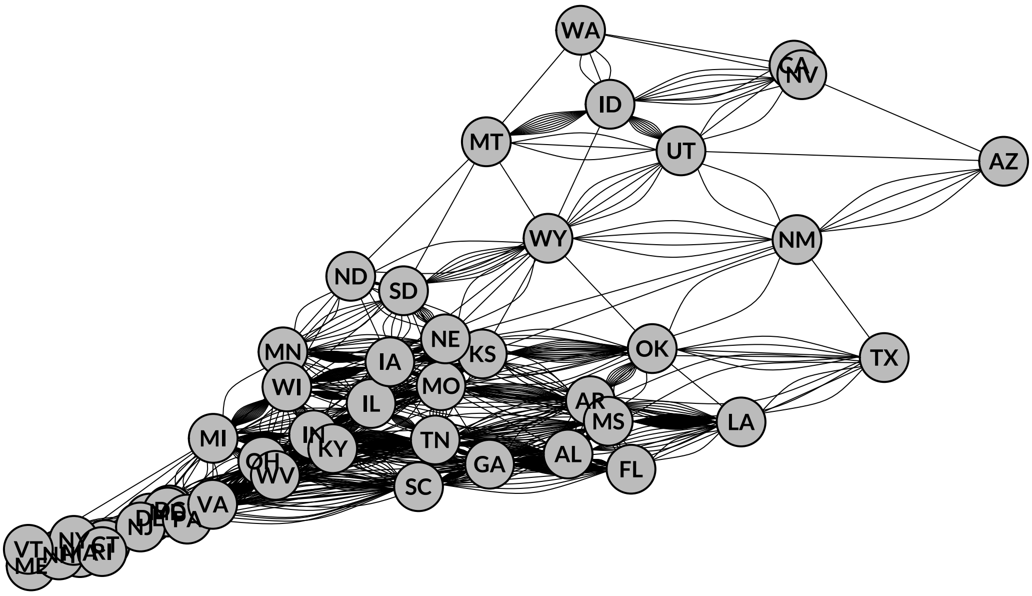
Distance Decay SLX

```
cigd <- read_excel("./assignment3/data/cigarettes/cigar_states.xls") %>%
    select(longitude, latitude) %>%
    as.matrix() %>%
    `rownames<-`(cign)

cigi <- distm(cigd, fun = distVincentyEllipsoid)

cigj <- (cigi/1e+06)^(-3) %>%
    `diag<-`(0) %>%
    mat2listw(style = "B")
```

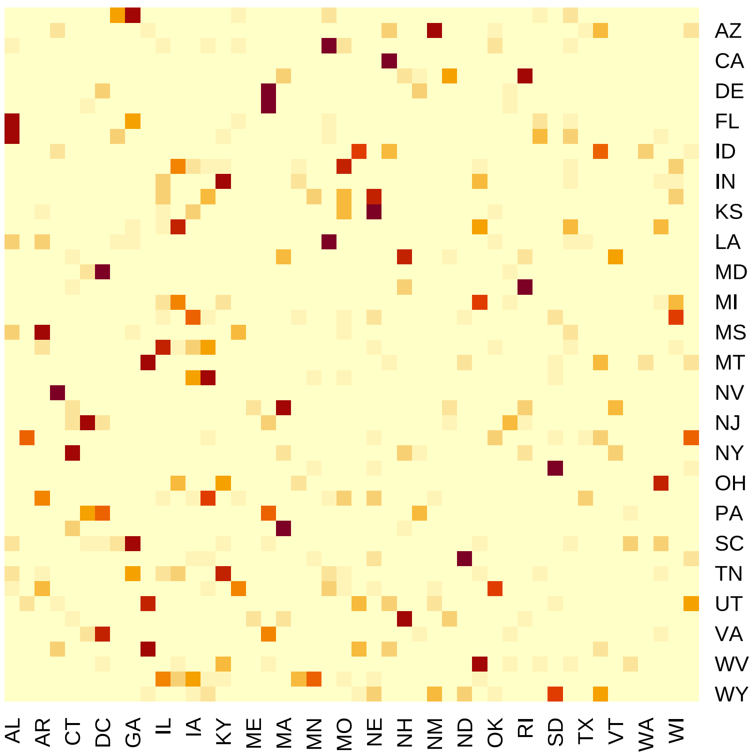
```
cm4 <- plm(
  logc ~ logp + logy + slag(logp, listw = cigj) + slag(logy, listw = cigj), #SLX w/ distance
decay
  data = cigs,
  effect = "twoways",
  model = "within",
  index = c("state", "year")
)
```

	Dependent variable:
	logc
logp	-1.017^{***}
	(0.042)
logy	0.608^{***}
	(0.060)
slag(logp, listw = cigm)	-0.220^{***}
	(0.077)
slag(logy, listw = cigm)	-0.219^{***}
	(0.080)
Observations	1,380
R^2	0.400
Adjusted R^2	0.364
F Statistic	217.105^{***} (df = 4; 1301)
Note:	*p<0.1; **p<0.05; ***p<0.01

```r
plot(cigg, vertex.size = 10, vertex.color = "#BBBBBB", vertex.label.cex = 0.7,
vertex.label.font = 2,
    vertex.label.family = "Lato", vertex.label.color = "black", edge.color = "black",
    edge.width = 0.5, asp = 0)
```




```r
heatmap(cigk, Rowv = NA, Colv = "Rowv")
```




```r
rowSums(cigk) %>%
    sort(decreasing = TRUE) %>%
    knitr::kable(col.names = "Row Sum")
```

	Row Sum
MD	11026.67198
DC	9945.46534
RI	4756.35048
MA	4728.87454
DE	2663.12139
CT	1913.64955
NH	1878.45853
PA	1082.00099
NJ	983.52610
NY	968.47870
VT	746.96390
VA	615.20596
ME	386.75814
KY	333.66481
OH	308.21679
IN	303.99804
WV	271.93286
NV	253.46415
CA	248.41830
KS	215.91668
NE	209.03832
IL	207.81238
GA	201.84045
AL	193.97210
MO	191.43001
IA	185.93943
TN	176.79008
MS	160.89378
AR	140.08363
WI	127.73114
MI	123.99357

	Row Sum
SC	100.11099
FL	98.10254
SD	90.44541
MN	85.93133
ND	71.63010
LA	70.08663
OK	60.44297
WY	38.15529
ID	35.44483
UT	32.25951
TX	25.43502
MT	25.15952
NM	24.47703
AZ	14.19491
WA	10.75984

	Eigenvector Centrality
MD	1.0000000
DC	0.9871728
DE	0.2112788
PA	0.0682453
VA	0.0486044
NJ	0.0310696
CT	0.0033553
WV	0.0030710
RI	0.0028691
NY	0.0025882
MA	0.0023642
OH	0.0013980
NH	0.0013860
VT	0.0008475
SC	0.0007672
KY	0.0006332
IN	0.0003951
MI	0.0003824
ME	0.0003685
GA	0.0002656
TN	0.0002394
AL	0.0000095
IL	0.0000043
WI	0.0000021
FL	0.0000017
MO	0.0000013
MS	0.0000012
IA	0.0000005
AR	0.0000005
MN	0.0000003
KS	0.0000002
LA	0.0000001
NE	0.0000001
OK	0.0000000
SD	0.0000000
TX	0.0000000
ND	0.0000000
WY	0.0000000
NM	0.0000000
MT	0.0000000
UT	0.0000000

	Eigenvector Centrality
AZ	0.0000000
ID	0.0000000
NV	0.0000000
CA	0.0000000
WA	0.0000000

Task B

Task B.1 - Unit of Observation

Describe the unit of observation: * subnational "cells" –> grid squares * used to conduct a geographically disaggregated analysis of civil conflict and weather shocks

What are their areas? * cells have dimensions of 110 km by 110 km (square)

Why and to which (relative) extent do they differ? * spatial resolution?

Task B.2 - Un-normalized weights matrix