AEM 7500 PS 4 Tianli Xia

1.

$$
\begin{aligned}
u_i(a,s,\varepsilon_i;\theta) &= u_{oi}(a_i,a_{-i},s_i;\theta)+\varepsilon_i(a_i) \\
&= \begin{cases} -\beta_0 a_i-\beta_{1i}a_i a_{-i}-\beta_2(2-a_i)s_i+\varepsilon_i(a_i) \text{ if } a_i>0 \\ -\beta_{3i}a_{-i}+\varepsilon_i(a_i) \text{ if } a_i=0 \end{cases}
\end{aligned}
$$

$\theta=(\beta_0,\beta_{1i},\beta_2,\beta_{3i})_{i=\{R,W\}}$ are the parameters. In this setup, each country's utility is a function of action sets $a=(a_i,a_{-i})=(a_r,a_w)$, state variables s, random shocks ε_i and parameters to be determined θ.

The utility function consists of two parts: the deterministic component u_{0i} and the stochastic component $\varepsilon_i(a_i)$ – one shock each for each action. Here the actions are to choose the pollution levels. In the utility function, countries are modelled to have disutility over pollution both from its own emission and also contigent on the other country's decision/actional emission a_{-i}.

(a) β_0 denotes country i's utility loss from per unit of pollution from its own; $(\beta_0>0)$

(b) β_{1i} denotes country i's utility loss from per unit of pollution from the other country when it pollutes. $(\beta_1>0)$

(c) β_2 denotes the utility loss for country i from bearing the abatement cost. $(\beta_2>0)$

(d) β_{3i} denotes country i's utility loss from per unit of pollution from the other country when it doesn't pollute. $(\beta_3>0)$

   Hence the trade-off for each firm is that producing more pollution hurts itself but also reducing the abatement costs. The larger the state variable s_i the larger the abatement cost.

   Also when $a_i=0$ the country is not affected by the other country's pollution. This is odd but needed in order to identify parameter β_{1i} since otherwise the $\beta_{1i}a_{-i}$ is the same for any a_i so we are unable to identify it from other parameters.

2. It is a game because a player's payoff depends interactively on the other player's actions. It is static in the sense that the decision today won't affect future outcomes and players don't take dynamic consequences into account .It can also be seen from the fact that the data is cross-sectional.

3. We can estimate the ex-ante choice probability non-parametrically.

$$\sigma_i(a_i=k|\vec{s}=\vec{m})=\frac{\#\text{ times action=k and state=m}}{\#\text{ times state=m}}.$$

To make it simpler, we can assume that countries account for expected payoff so the probability above becomes,

$$\hat{\sigma}_i(a_i=k|\vec{s}=\vec{m})=\frac{\exp\left(\sum_{a_{-i}}\Pr(a_{-i})\,u_i(k,a_{-i}|s)\right)}{\sum_{a_i}\exp\left(\sum_{a_{-i}}\Pr(a_{-i})\,u_i(a_i,a_{-i}|s)\right)}$$

which simplifies the problem a little since the linearality of a_{-i} allows us to replace a_{-i} with $\sum_{a_{-i}}a_{-i}\sigma_{-i}(a_{-i}|s)$.

GMM

Table 1: Ex-ante choice probability estimates for red country

a_r	s_r	s_w	Prob
0	0	0	1.000
1	0	0	0.000
2	0	0	0.000
0	1	0	1.000
1	1	0	0.000
2	1	0	0.000
0	2	0	0.348
1	2	0	0.652
2	2	0	0.000
0	0	1	0.000
1	0	1	0.853
2	0	1	0.147
0	1	1	0.000
1	1	1	0.658
2	1	1	0.342
0	2	1	0.000
1	2	1	0.474
2	2	1	0.526
0	0	2	0.121
1	0	2	0.515
2	0	2	0.364
0	1	2	0.000
1	1	2	1.000
2	1	2	0.000
0	2	2	0.000
1	2	2	0.000
2	2	2	1.000

Table 2: Ex-ante choice probability estimates for white country

a_w	s_r	s_w	Prob
0	0	0	0.033
1	0	0	0.967
2	0	0	0.000
0	1	0	0.000
1	1	0	0.275
2	1	0	0.725
0	2	0	0.522
1	2	0	0.391
2	2	0	0.087
0	0	1	0.000
1	0	1	0.882
2	0	1	0.118
0	1	1	0.000
1	1	1	0.013
2	1	1	0.987
0	2	1	0.316
1	2	1	0.474
2	2	1	0.211
0	0	2	0.061
1	0	2	0.939
2	0	2	0.000
0	1	2	0.000
1	1	2	0.778
2	1	2	0.222
0	2	2	0.000
1	2	2	0.400
2	2	2	0.600

1. The moment conditions that we use is

$$E\left[\vec{y}_t - \vec{\sigma}\left(s_t, \hat{\phi}^e, \theta\right)\right] = 0$$

when $\vec{y}_t$ is a 6x1 vector denoting the choices (3) made by 2 players. But if the observed probability of choosing $a = \{0, 1\}$ is matched by our models:

$$\begin{aligned} E\left[\vec{y}_{i0} - \vec{\sigma}_{i0}\left(s_t, \hat{\phi}^e, \theta\right)\right] &= 0 \\ E\left[\vec{y}_{i1} - \vec{\sigma}_{i1}\left(s_t, \hat{\phi}^e, \theta\right)\right] &= 0 \end{aligned}$$

then mechanically the probability of choosing $a = \{2\}$ is exactly matched because the probability sums to 1!

$$E\left[\vec{y}_{i2} - \vec{\sigma}_{i2}\left(s_t, \hat{\phi}^e, \theta\right)\right] = 0$$

So effectively we have only four moments but six parameters. This makes the model not identified.

In order to make the model identifiable, we have to interact the error terms with state variables s_t under the assumption that s_t is not related to the prediction error. The sample analog of our estimate is

$$\begin{aligned} m(\theta) &= \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T}\left(\vec{y}_t - \vec{\sigma}\left(s_t, \hat{\phi}^e, \theta\right)\right) \\ \frac{1}{T}\sum_{t=1}^{T}\left(\vec{y}_t - \vec{\sigma}\left(s_t, \hat{\phi}^e, \theta\right)\right) s_t \end{bmatrix} \\ \hat{\theta} &= \arg\min m(\theta)' W m(\theta) \end{aligned}$$

2. The code consists three part.

File	
Ps4_gmm.m	Clean the data
Estimation.m	Call fminsearch
GMM.m	Objective function to minimize

Table 3: **Estimation results:** The first stage GMM uses identity matrix as the weighting matrix, the effective GMM uses inversion of the variance-covariance matrix of the moments as the weighting matrix. The results are in general similar, but the standard errors are reduced under efficient GMM. The standard errors are bootstrap. I also bootstrap the 95% confidence interval non-parametrically.

	First stage GMM				**Effective GMM**			
	Coef.	Std.	CI_lower	CI_upper	Coef.	Std.	CI_lower	CI_upper
beta0	-0.981	0.164	-1.232	-0.694	-1.000	0.141	-1.219	-0.767
beta1r	1.391	0.228	1.038	1.780	1.466	0.217	1.101	1.801
beta2	0.081	0.155	-0.198	0.317	0.072	0.156	-0.179	0.333
beta3r	0.739	0.240	0.332	1.140	0.785	0.239	0.361	1.204
beta1w	1.361	0.314	0.818	1.905	1.421	0.317	0.919	1.930
beta3w	2.739	0.612	1.943	3.993	2.825	0.595	1.982	3.918
Obj. function	6.931E-05				4.621E-06			

3. Bootstrap: now we draw 100 times with replacement from the original data, generate new observed choice for each time, and estimate the model to match the observed choice probabilities each time. Then we can form confidence intervals from these 100 estimates.

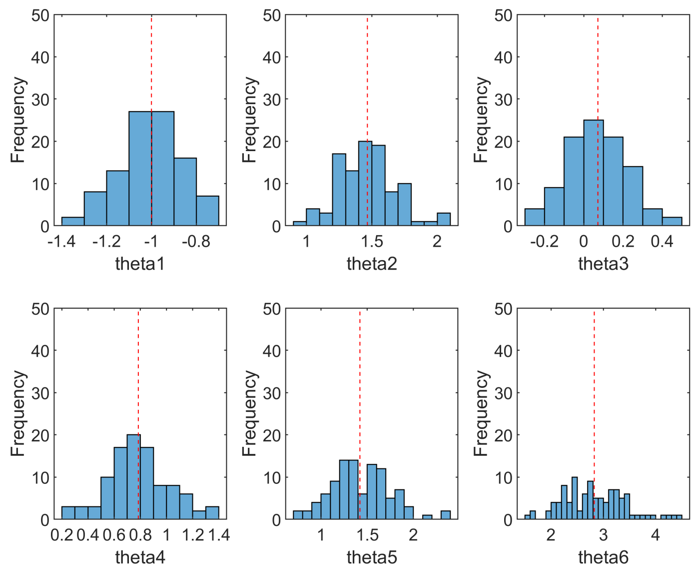

Figure 1: Bootstrap results for efficient GMM

4. Interpret: β_{3i} is precisely estimated, probably because that the probability of choosing $a_i = 0$ is characterized by β_{3i} only. $\beta_0, \beta_2, \beta_3$ are negative while β_1 is positive, this says that is when the other country is not emitting pollutants, my own pollution is always good. But when the other country pollutes, my own pollutant hurts.

MLE

1. Again, since we have calculated the ex-ante predicted choice probabilities for each action a_i conditional on the states, instead of matching the population level moments, we can ask what is the probability of our oberved dataset and search over parameter space to maximize the likelihood function:

$$\max_{\theta} L(\theta) = \sum_i \sum_t \log\left(\sigma_i\left(y_{it}|s_t, \hat{\phi}^e, \theta\right)\right)$$

2. The code consists three part.

File	
Ps4_mle.m	Clean the data
Estimation.m	Call fminsearch
MLE.m	Objective function to minimize

3. Bootstrap: now we draw 100 times with replacement from the original data, generate new observed choice for each time, and estimate the model to match the observed choice probabilities each time. Then we can form confidence intervals from these 100 estimates.

4. Interpretation: The MLE results are quite different from the MLE results. I think the MLE here is a pseudo-MLE since the ex-ante belief on the other player's choice probability is estimated, hence MLE may magnify the sample errors. A paper by Berry, Ostrovsky, and Pakes (Rand 2007) discussed this issue. Here the pollution becomes beneficial itself for the country. I tend to believe the GMM more due to its less restrictive assumptions.

Table 4: **Estimation results:** The psuedo-MLE estimation results. The results are quite different from what we get in GMM. The standard errors are bootstrap. I also bootstrap the 95% confidence interval non-parametrically.

	Psuedo MLE			
	Coef.	Std.	CI$_l ower$	CI$_u pper$
beta0	-0.603	0.125	-0.821	-0.413
beta1r	0.480	0.174	0.193	0.777
beta2	-0.723	0.110	-0.899	-0.542
beta3r	-0.403	0.194	-0.743	-0.111
beta1w	-0.147	0.220	-0.533	0.170
beta3w	0.135	0.402	-0.478	0.782

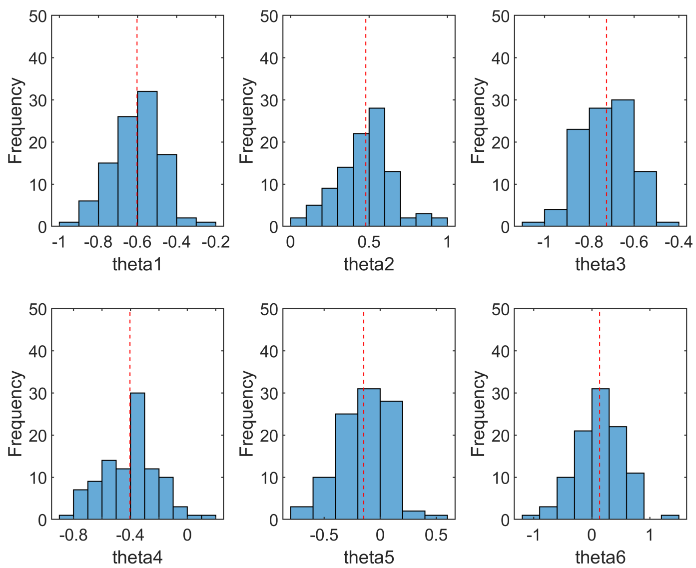

Figure 2: Bootstrap results for efficient GMM

What questions does this structural model help you answer?

The structural model answers the question about how each model primitives affects country's utility functions, in particular, the structural assumption allows us to look at the externality (peer effects) imposed by the other country. Furthermore, with a structural model, I can conduct welfare analysis (for example, how current situation deviates from the social optimal level) and counterfactual analysis (for example, how global agreement will affect the utility level of each country). These excercises are hard to do in the reduced form model.

Additional state variables:

It may be interesting to make all the parameters country-specific so we can say more about the heterogeneity of two countries. Currently, β_0 and β_2 are assumed to be homogeneous across countries. It might be worthwhile to see how each country is affected by its own pollution and abatement costs differently.

Also we make function form assumptions in the utility function (a_{-i} affects i only through the interaction channel $a_i a_{-i}$. It might be interesting to see how it gets affected by a_{-i} seperately.)