Paul Le Tran
plt377
30/09/21
ECO388E

Problem Set 1

1a). Estimating the linear model using OLS, we obtain the following parameter values and respective standard error (SE) estimates:

$\left(\widehat{\theta_1}, \widehat{\theta_2}\right) = (0.0055, 1.0105);$
$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right)\right) = (0.0225, 0.0227).$

1b). Estimating the linear model using maximum likelihood estimation (MLE), we obtain the following parameter values and respective SE estimates (excluding one for σ):

$\left(\widehat{\theta_1}, \widehat{\theta_2}, \hat{\sigma}\right) = (0.0055, 1.0105, 0.4998);$
$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right)\right) = (0.0226, 0.0230).$

1c). Estimating the linear model using generalised method of moment (GMM), we obtain the following parameter values and respective SE estimates:

$\left(\widehat{\theta_1}, \widehat{\theta_2}\right) = (0.0055, 1.0105);$
$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right)\right) = (0.0229, 0.0231).$

σ is no longer a parameter that needs to be identified in GMM for this linear model because we have invertibility. This allows us to not have to make assumptions about the full distribution of ϵ_i up to parameters.

Recall that the asymptotic variance for the GMM estimator is

$$(\Gamma' A\Gamma)^{-1}(\Gamma' AVA\Gamma)(\Gamma' A\Gamma)^{-1},$$

where *Γ, A,* and *V* are matrices described in the problem set. Recall that in our model, we have two parameters and two moments. Therefore, we can see that *Γ* and *V* are both 2x2 matrices. From algebra, we thus can get:

$$\begin{aligned}(\Gamma' A\Gamma)^{-1}(\Gamma' AVA\Gamma)(\Gamma' A\Gamma)^{-1} &= (\Gamma^{-1}A^{-1}\Gamma'^{-1})(\Gamma' AVA\Gamma)(\Gamma^{-1}A^{-1}\Gamma'^{-1}) \\ &= (\Gamma^{-1}A^{-1})(AVA)(A^{-1}\Gamma'^{-1}) \\ &= \Gamma^{-1}V\Gamma'^{-1},\end{aligned}$$

which is our desired result.

2a). We estimate the non-linear, invertible model using MLE. We obtain the following parameter values and respective SE estimates (excluding one for σ):

$$\left(\widehat{\theta_1}, \widehat{\theta_2}, \hat{\sigma}\right) = (3.0659, 1.0272, 0.9633);$$
$$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right)\right) = (0.1953, 0.2529).$$

2b). We estimate the non-linear, invertible model using GMM. We obtain the following parameter values and respective SE estimates:

$$\left(\widehat{\theta_1}, \widehat{\theta_2}\right) = (3.0574, 1.0394);$$
$$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right)\right) = (0.1861, 0.2158).$$

- *Which of* **2a).** and **2b).** *is robust to heteroscedasticity of* ε_i *with respect to* x_i*? Why?*
  Recall that for MLE, one needs to assume the full distribution of the error term up to parameters. This includes making a stance on what the variance of the error terms are. In contrast, one only needs to assume mean independence, $E[\epsilon_i|x_i] = 0$, in invertible GMM. This means that invertible GMM isn't making a stance on the variance of the error term, and is therefore robust to heteroscedasticity.

- *Explain why this GMM estimator is not efficient (given the assumption that* $E[\epsilon_i|x_i] = 0$*)?*
  In the context of this problem, recall that the model is exactly identified because the number of moments is equal to the number of parameters estimated. As a result, the choice of weight matrix *A* should not matter. We know that when it comes to non-invertible GMM estimation, the estimator is not efficient because it does not contain all possible higher-order moments of ϵ_i's distribution, which capture more information about the variance). This is even the case when we make an assumption on the distribution of ϵ_i up to parameters.

  When considering invertible GMM estimation, we only make assume $E[\epsilon_i|x_i] = 0$), meaning we do not know anything else about ϵ_i's distribution. Therefore, even with optimal weight matrix *A* and the model exactly identified, we are unable to add any high-order moments to the moment condition due to not knowing anything else about ϵ_i's distribution. It is this lack of higher-order moments that makes both invertible GMM estimators (this problem) and non-invertible GMM estimators (problem 3) less efficient than ML estimators.

Paul Frame

21/9/21

ECO388E

plt377

ECO388E Problem Set 1, Problem 3a

3a). Inverse demand curve: $p_i = 100 - q_i$ $\Big\}$ $p_i = 100 - (q_1 + q_2)$

Total quantity: $q_i = q_1 + q_2$.

$MC_1 = \exp(\theta_1 + \theta_2 x_i^1 + \sigma \varepsilon_i^1)$

$MC_2 = \exp(\theta_1 + \theta_2 x_i^2 + \sigma \varepsilon_i^2)$

$TR_1 = p_i q_1 = (100 - q_1 - q_2) q_1 = 100 q_1 - q_1^2 - q_1 q_2$

$TR_2 = p_i q_2 = (100 - q_1 - q_2) q_2 = 100 q_2 - q_2^2 - q_1 q_2$

$\therefore MR_1 = \frac{\partial TR_1}{\partial q_1} = 100 - 2q_1 - q_2$

$MR_2 = \frac{\partial TR_2}{\partial q_2} = 100 - 2q_2 - q_1$

Profit maximisation of firm 1:

$MR_1 = MC_1 \Rightarrow 100 - 2q_1 - q_2 = \exp(\theta_1 + \theta_2 x_i^1 + \sigma \varepsilon_i^1)$

$\Rightarrow \boxed{q_1 = \frac{100 - q_2 - \exp(\theta_1 + \theta_2 x_i^1 + \sigma \varepsilon_i^1)}{2}}$

Profit maximisation of firm 2:

$MR_2 = MC_2 \Rightarrow 100 - 2q_2 - q_1 = \exp(\theta_1 + \theta_2 x_i^2 + \sigma \varepsilon_i^2)$

$\Rightarrow \boxed{q_2 = \frac{100 - q_1 - \exp(\theta_1 + \theta_2 x_i^2 + \sigma \varepsilon_i^2)}{2}}$

(2) Observe that the 2 firms are identical. Therefore, we know that $q_1^* = q_2^*$.

$\therefore q_1^* = \frac{100 - q_1^* - \exp(\theta_1 + \theta_2 x_i^1 + \sigma \varepsilon_i^1)}{2}$

$q_1^* + \frac{q_1^*}{2} = \frac{100 - \exp(\theta_1 + \theta_2 x_i^1 + \sigma \varepsilon_i^1)}{2}$

$\frac{3}{2} q_1^* = \frac{100 - \exp(\theta_1 + \theta_2 x_i^1 + \sigma \varepsilon_i^1)}{2}$

$$\boxed{\begin{aligned} \therefore q_1^* &= \frac{100 - \exp(\theta_1 + \theta_2 x_i^1 + \sigma \varepsilon_i^1)}{3}; \\ q_2^* &= \frac{100 - \exp(\theta_1 + \theta_2 x_i^2 + \sigma \varepsilon_i^2)}{3}; \\ p_i^* &= 100 - q_1^* - q_2^*. \end{aligned}}$$

3b). We estimate the non-linear, non-invertible model using MLE. We obtain the following parameter values and respective SE estimates (excluding one for σ):

$$\left(\widehat{\theta_1}, \widehat{\theta_2}, \hat{\sigma}\right) = (1.0039, 1.0529, 0.0955);$$
$$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right)\right) = (0.0091, 0.1281).$$

3c). Attempting to estimate this model with only a mean independence assumption would be difficult due to what the integrating-out process requires (both complete and partial). In the context of choosing to integrate out ϵ_i^1, in order to approximate the likelihood contribution of *i* (after integrating out ϵ_i^1), one would need to know what the density of ϵ_i^1 is up to parameters. If one does not, the one cannot take simulated draws of the error term being integrated out. It is because of this reason that when estimating a non-linear, non-invertible model with MLE, one needs to assume the distribution of the error terms up to parameters.

3d). We estimate the non-linear, non-invertible model using method of simulated moments (MSM). We obtain the following parameter values and respective SE estimates in the <u>first stage</u>:

$$\left(\widehat{\theta_1}, \widehat{\theta_2}, \hat{\sigma}\right) = (1.0056, 0.9269, 0.0970);$$
$$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right), SE(\hat{\sigma})\right) = (0.0095, 0.1175, 0.0077).$$

When doing the <u>second stage</u> of MSM, we obtain the following:

$$\left(\widehat{\theta_1}, \widehat{\theta_2}, \hat{\sigma}\right) = (1.0064, 0.9376, 0.0967);$$
$$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right), SE(\hat{\sigma})\right) = (0.0095, 0.1178, 0.0077).$$

When comparing the results in both stages, my computational results show that the change in the SE estimates for the parameters between stages is essentially negligible. This is potentially due to the sensitivity of MSM to initial values chosen for numerical optimisation. Taking my results as is though, they seem to indicate that the second stage process is more about finding the "correct" parameter estimates, and less about decreasing the SE estimates of said parameters.

3e). When estimating the same model using MSM but with only the first three moments, we obtain the following parameter estimates and respective SE estimates in the <u>first stage</u>:

$$\left(\widehat{\theta_1}, \widehat{\theta_2}, \hat{\sigma}\right) = (1.0069, 0.9989, 2.6644e - 07);$$
$$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right), SE(\hat{\sigma})\right) = (0.0141, 0.3097, 0.4508).$$

When doing the <u>second stage</u> of MSM, we obtain the following:

$$\left(\widehat{\theta_1}, \widehat{\theta_2}, \hat{\sigma}\right) = (1.0075, 1.0066, 2.7348e - 07);$$
$$\left(SE\left(\widehat{\theta_1}\right), SE\left(\widehat{\theta_2}\right), SE(\hat{\sigma})\right) = (0.0140, 0.3048, 0.4445).$$

If we are to compare the second stage results to those of part **3d).**, we see that the parameter estimates are essentially identical. However, it's immediately noticeable that the SE estimates have increased when omitting the fourth moment (the moment involving p_i^2). In the context of this moment, these results seem to indicate that the moment involvingp_i^2 is more sensitive to variance. In other words, this moment captures a decent amount of the characteristics of the variance of p_i. Omitting the moment would therefore result in the MSM estimates to be less efficient.

3f). The main change we see done to the model is that there is now an additional error term, η_i, to the price equation, p_i. In terms of intuitive changes, because η_i is observable to neither firm, they will essentially set the new error term to equal zero in their FOCs. This means when it comes to either firm's profit maximisation, the optimal quantities will not change from those of the original model. However, optimal price p_i^* is no longer pinned down by only the optimal quantities of both firms as before. This is due to the existence of η_i. Therefore, this error term can be interpreted as the optimal price being subject to potential demand shocks not accounted for by either firm. Note that this interpretation occurs when η_i is large. However, if η_i is close to zero, the model results are essentially identical to those of the original model.

In terms of changes to estimation procedures, having η_i as an additional error term would allow for partially integrating out both original error terms $\epsilon_i^1, \epsilon_i^2$. This means we would need to take simulated draws for said error terms when doing MLE (two sets of simulated draws instead of just one as in the original model when doing MLE). An additional change is that we now have four parameters to estimate: $(\theta_1, \theta_2, \sigma, \phi)$. When it comes to procedural changes to GMM/MSM estimation, observe that the new model has four parameters and four moments (if we use the four moments from the original model). Therefore, our model will be exactly identified. Additionally, we will need to do three sets of simulated draws instead of two because the model now has three error terms.

3g). With the introduction of parameter ρ and error term α in each firm's marginal cost functions, this means both FOCs and thus profit maximisations will be different from those of the original model (note that both firms face the same values of the new error term). As a result, the general optimal quantities chosen by each firm will be different from those of the original model. This is because the firms are now basically subject to potential supply shocks which affect their marginal costs. Such a scenario would occur when ρ is large. If ρ were small, the model would behave closely or identically to the original model.

In terms of estimation changes, having α as an additional error term would allow for partially integrating out both original error terms $\epsilon_i^1, \epsilon_i^2$. This means we would need to take simulated draws for both error terms when doing MLE (two sets of simulated draws instead of just one). An additional change is that we now have four parameters to estimate: $(\theta_1, \theta_2, \sigma, \rho)$. When doing GMM/MSM estimation, observe that the new model has four parameters to estimate and four moments (if we use the four moments from the original model). Therefore, our model will be exactly identified. Additionally, we would need to do three sets of simulated draws instead of two because the model now has three error terms.

3h). The following analyses are done referring to the model in part **3f).** Suppose we observe the data $(p_i, x_i^1, x_i^2, q_i^1, q_i^2)$. Assume we observe that q_i^1, q_i^2 are similar across markets, x_i^1, x_i^2 are similar across markets, and p_i's is similar across markets. The former two observations tell us that σ should be small. The last observation in tandem would also imply that ρ is small as well. If we observe that q_i^1, q_i^2 are different across markets, x_i^1, x_i^2 are different across markets, and p_i's is similar across markets, we would be able to say that σ is large, but ρ is small. These various observation scenarios tell us overall that such an observable dataset allows us to pin down the value/magnitude of both parameters ρ, σ.

Now assume our observable dataset is (p_i, x_i^1, x_i^2, q_i). Assume we see q_i's is similar across markets, x_i^1, x_i^2 are similar across markets, and p_i's is similar across markets. Because we are only able to observe total quantity, it is difficult to say from the data alone if σ is small. However, we are able to say that ρ is small from our third observation. Overall, this means that this observable dataset allows us to pin down the value/magnitude of ρ, but not that of σ.

Finally, assume our observable dataset is the original dataset of the problem set: (p_i, x_i^1, x_i^2). Assume we see x_i^1, x_i^2 are similar across markets and p_i's is similar across markets. We unfortunately can't infer what the value of σ is due to how it's possible for the individual quantities to differ by a lot across markets under our observed data. Furthermore, even if we see similar p_i's across markets, we aren't able to pin down if this is due to being in a world with small σ only, small ρ only, or both.

3i). The following analyses are done referring to the model in part **3g).** Suppose our observable dataset is $(p_i, x_i^1, x_i^2, q_i^1, q_i^2)$. Recall that our model has both marginal cost equations sharing the same error term α. Assume we observe that q_i^1, q_i^2 are similar across markets, x_i^1, x_i^2 are similar across markets, and p_i's is similar across markets. This means we can say that σ is small. However, we are unable to pin down ρ due to the associated error term not only being in the marginal cost equations, but is shared in both equations as well. Now assume we observe that q_i^1, q_i^2 are quite different across markets and x_i^1, x_i^2 are different across markets. We are able to pin down σ as being large, but ρ still remains uncertain from the data alone.

Suppose our observable dataset is now (p_i, x_i^1, x_i^2, q_i). Assume we see q_i's is similar across markets, x_i^1, x_i^2 are similar across markets, and p_i's is similar across markets. Because we only observe total quantity and not individual quantities, even with similar x_i^1, x_i^2 across markets, we are unable to pin down the value of σ. We also run into the issue of α being in both marginal cost equations. As a result, observations in the data alone cannot allow us to pin down the value of ρ as well.