"SOLUTIONS"
Problem Set 2: Static Entry Games

Matt Grennan

January 29, 2008

These are my attempt at the second problem set for the second year Ph.D. IO course at NYU with Heski Bar-Isaac and Allan Collard-Wexler in Fall 2007. They are offered as suggested "solutions", and they benefit from discussions with the instructors. Any errors are my own.

In this problem set we consider estimating static entry games from data on automobile service centers in isolated US towns.

1 Examining the data

(a) Summary statistics

Table 1 displays summary statistics for population, housing, and autoroute (whether or not the town is on a major travel route) variables in our data

Table 1. Summary Statistics

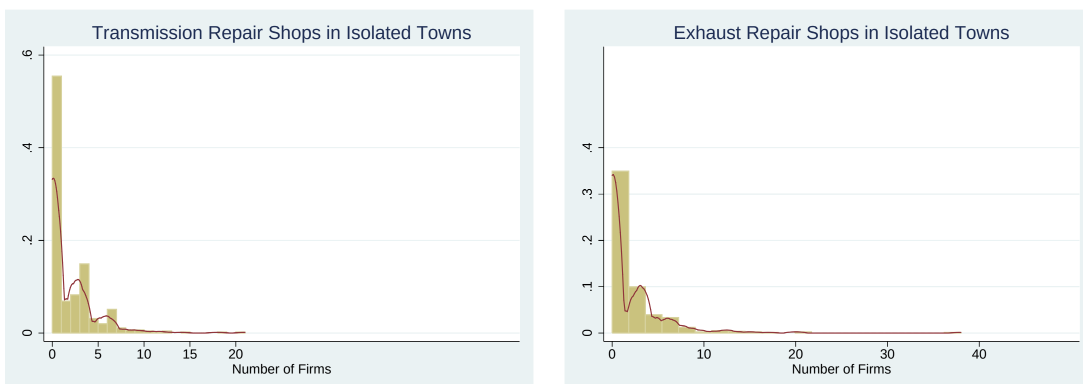
Variable	Observations	Mean	Std. Dev.	Min	Max
population	449	12,031	14,269	4,019	176,576
housing	449	22,153	19,910	989	132,640
autoroute	449	0.28062	0.44981	0	1

Both the population and housing statistics indicate the large amount of variation in the data set of sizes of these isolated towns. We might wonder if the market structure and competition between establishments isn't fundamentally different in towns of 176,576 people versus towns of 4,019 people, but we will hope that our models are "good enough" for the purposes of this problem set. The other statistic of note is that 28% of the towns are located on major auto routes, which is something we might think matters for the demand for automobile repairs shops (though we won't include it in our models in this problem set).

(b) Distribution of market sizes

Figure 1 shows the distribution of the number of firms per market for the transmission (811112) and exhaust (811113) sectors.

Figure 1: Distribution of the number of firms per market



Here we notice that the distributions take approximately the same shape, with transmission repair shops having relatively more likelihood of zero entry and less likelihood of greater than 20 firms. Both distributions seem to indicate that monopoly is rare relatively to oligopoly.

(c) Unobservable market characteristics

Regressing the number of transmission and exhaust establishments, respectively, on the population we find that both are positively and significantly correlated with the size of the town, with the results shown in Table 2.

Table 2. Regressions of the number of establishments on population

	Transmission	Exhaust
constant	0.7684	0.4311
	(0.1492)	(0.1881)
population	7.29e-5	12.58e-5
	(0.8e-5)	(1.0e-5)

The residuals from these regressions contain other market characteristics that are correlated with the number of each type of establishment. We can see if these unobserved market factors are similar across the two types of establishments by examining the correlation between the residuals, which we find is $\rho = 0.5331$. This tells us that there are very likely unobserved market-specific characteristics that affect entry of both types of establishments in similar ways (thinking of our summary statistics, autoroute comes to mind as one possible such characteristic).

The correlation between the residuals also tells us something about the econometrics of estimating these regressions, namely that we are in a "seemingly unrelated regressions" (SUR) situation. Thus we might get more efficient estimates by estimating the two regressions simultaneously via GLS, however, since in our case the independent variables are the same, there is no such efficiency gain (see Greene 5th Ed., p.341-343 for a discussion).

2 Bresnahan-Reiss model

The original Bresnahan and Reiss (1991) model assumes symmetry among firms in each market as well as two further behavioral assumptions (which are necessary conditions for a Nash Equilibrium in pure strategies for their entry game):

A1 Firms that enter make positive profits

$$\pi(N_m, X_m) + \epsilon_m > 0 \qquad \text{whenever } N_m > 0$$

A2 Further entrants would earn negative profits

$$\pi(N_m + 1, X_m) + \epsilon_m < 0$$

These assumptions combine to predict a unique number of firms in each market in equilibrium. We complete the model by specifying a profit function

$$\pi(N_m, X_m) = \alpha + X_m\beta + [\mathbf{1}_{\{N_m \geq 2\}}, \mathbf{1}_{\{N_m \geq 3\}}, \max(0, N_m - 3)]\gamma$$

where N_m and X_m are the number of firms and log of the population, respectively, in market m; ϵ_m are i.i.d. $\mathcal{N}(0,1)$; and $\theta := (\alpha, \beta, \gamma)$ are parameters to be estimated. Note this is essentially a Probit model on cross sectional data, so we need to assume the i.i.d. and additively separable unobservables as well as normalize scale (variance equals 1) and position (profits for not entering equal 0) for identification.

(a) Constructing the likelihood

We can then estimate the parameters via maximum likelihood, where the likelihood is

$$\begin{aligned}
\mathcal{L}(\theta|N, X, \pi(\cdot)) &= \prod_{m=1}^{M} Pr(N_m = n|X_m, \theta, \pi(\cdot)) \\
&= \prod_{m=1}^{M} Pr\left([\pi(N_m, X_m; \theta) + \epsilon_m > 0]\mathbf{1}_{\{n>0\}}, [\pi(N_m + 1, X_m; \theta) + \epsilon_m < 0]\right) \\
&= \prod_{m=1}^{M} \Phi[-\pi(n + 1, X_m; \theta)] - \mathbf{1}_{\{n>0\}}\Phi[-\pi(n, X_m; \theta)]
\end{aligned}$$

(b) Estimation

Maximizing the corresponding log-likelihood function gives us the parameter estimates in Table 3[1] below.

[1]Standard errors are obtained via the inverse of the outer product of the gradients of the log likelihood w.r.t. the parameters (often called the Bernt-Hall-Hall-Hausman estimator for the asymptotic covariance matrix).

Table 3. MLE for Bresnahan-Reiss model

constant	-6.6483
	(0.5309)
log population	0.7115
	(0.0064)
$1_{\{N \geq 2\}}$	-0.2087
	(0.0013)
$1_{\{N \geq 3\}}$	-0.2754
	(0.0019)
$\max(0, N-3)$	-0.2566
	(0.0001)
log Likelihood = -701.41	
N = 449	

Implementing this estimation required some tweaks on a simple maximizing of the likelihood. First, I put in a penalty function to restrict the impact of competition to be non-positive. While the model is technically identified without this restriction, this data does not seem to have enough independent variation in X_m and N_m to do so. Thus the non-positivity constraint is a "sensible", theory-based assumption that helps identify the model with this data. Second, much of the parameter space gives vectors that conflict with the combination of our model and data in that $Pr(N_m = n | X_m, \theta, \pi(\cdot))$ can be non-positive (for our profit function in this model this occurs exactly when the impact of competition is non-negative). Such a case automatically makes the likelihood zero, so again I put in a penalty function to prevent it. Note that in this specific model, the above two cases are exactly linked, so only one of the "fixes" is necessary. Finally, using *fminsearch* for this log likelihood is not robust to starting values, so I tried several values and chose the parameter vector with the largest log likelihood value.

Looking at the estimation results, we see that all the coefficients are highly statistically significant. The constant is negative, indicating there may be some fixed costs of entry. Population has a positive marginal impact, as we might expect. A somewhat surprising result is that the impact of additional competitors does not appear to decrease as the number of firms in the market increases.

3 Mazzeo model using moment inequalities

We want to look at competitive interactions between transmission and exhaust repair shops using Mazzeo's model of entry where type choice is endogenous. This model makes the following three behavioral assumptions (which are again necessary conditions for a pure strategy Nash equilibrium in Mazzeo's game):

A1 Firms that enter make positive profits (for both types h and l)

$$\pi^h(X_m, N_m^h, N_m^l) + \epsilon_m^h > 0$$

$$\pi^l(X_m, N_m^h, N_m^l) + \epsilon_m^l > 0$$

A2 Further entrants would earn negative profits (for either type)

$$\pi^h(X_m, N_m^h + 1, N_m^l) + \epsilon_m^h < 0$$

$$\pi^l(X_m, N_m^h, N_m^l + 1) + \epsilon_m^l < 0$$

A3 No entrant wants to switch type

$$\pi^h(X_m, N_m^h, N_m^l) + \epsilon_m^h > \pi^l(X_m, N_m^h - 1, N_m^l + 1) + \epsilon_m^l$$

$$\pi^l(X_m, N_m^h, N_m^l) + \epsilon_m^l > \pi^h(X_m, N_m^h + 1, N_m^l - 1) + \epsilon_m^h$$

The above are as listed in the problem set, but you may recall from our work in problem 2 that the requirement that firms that enter make positive profits should be conditional on $N_m > 0$ since profits need not be positive (and probably aren't) if $N_m = 0$. This has consequences for how to construct v_1 and v_2 in what follows. Since no one was expected to notice this (we didn't at first), I will do the problem twice. First, I will proceed as suggested in the problem set—but keep in mind this is wrong (and thus I won't interpret parameter estimates in this section)! After that, I will go through a way to do it that fixes the mistake.

3.1 As suggested in the original problem . . .

We can form moment inequalities from the identifying assumption that each of these six inequalities must hold in expectation. To construct our objective function, we use the sample analog of these expectations, i.e. we let

$$\begin{aligned}
v_1 &= \left(-\sum_m \pi^h(X_m, N_m^h, N_m^l) Z_m\right)^+ \\
v_2 &= \left(-\sum_m \pi^l(X_m, N_m^h, N_m^l) Z_m\right)^+ \\
v_3 &= \left(\sum_m \pi^h(X_m, N_m^h + 1, N_m^l) Z_m\right)^+ \\
v_4 &= \left(\sum_m \pi^l(X_m, N_m^h, N_m^l + 1) Z_m\right)^+ \\
v_5 &= \left(-\sum_m [\pi^h(X_m, N_m^h, N_m^l) - \pi^l(X_m, N_m^h - 1, N_m^l + 1)] Z_m\right)^+ \\
v_6 &= \left(-\sum_m [\pi^l(X_m, N_m^h, N_m^l) - \pi^h(X_m, N_m^h + 1, N_m^l - 1)] Z_m\right)^+
\end{aligned}$$

where $Z = [1\ X]$ are instruments, giving us 12 moment inequalities to estimate 6 parameters. We will combine these moments in a way that weights them all equally, using the objective function

$$Q(\theta) := \sum_{j=1}^{6} v_j v_j' \tag{1}$$

and again we complete the model by specifying the profit functions

$$\begin{aligned}\pi^h(X, N^l, N^h) &= \alpha^h + X\beta^h + \theta^{hh}\log(N^h+1) + \theta^{hl}\log(N_l+1)\\ \pi^l(X, N^l, N^h) &= \alpha^l + X\beta^l + \theta^{lh}\log(N^h+1) + \theta^{ll}\log(N_l+1)\end{aligned}$$

where here we normalize $\beta^h = \beta^l = 1$ for scale (though we could technically normalize one and estimate the other since the difference is identified from our "no switching type" moment).

(a) Partial identification

Minimizing the criterion function, we find that $Q(\theta) = 0$ for some vector θ. Thus the model is partially identified, and there is potentially a set Θ of parameter vectors that satisfy $Q(\theta) = 0$, $\forall\theta \in \Theta$.

(b) Estimating the identified set

Finding the exact set Θ in the parameter space is not a simple thing to do. We will actually find the hyper-cube that contains the set. We will do this by solving constrained optimization problems, each of which will give one of the bounds in one of the dimensions of the parameter space. For example, to find the lower bound for α^h, we will solve the problem

$$\begin{aligned}\min_{\theta} \quad & \alpha^h \\ s.t. \quad & Q(\theta) = 0\end{aligned} \tag{2}$$

which we can modify to find the lower bounds for other parameters by substituting them for α_h in (2) as the objective to be minimized; and upper bounds can be found by changing min to max. We can implement this in Matlab using the constrained minimization function *fmincon*. One could either specify our six inequality constraints or use $Q(\theta) = 0$ as a nonlinear equality constraint (I did the latter). The results of this procedure give the hypercube containing the identified set described in Table 4[2] below.

[2]Inference on the appropriate confidence interval for the identified set is an active area of research. The state-of-the-art to my knowledge is a subsampling procedure which is in the spirit of bootstrapping (but slightly different). I would recommend interested parties to Professor Stoye.

Table 4. Bounds on the identified set of parameters in Mazzeo model

Variable	Lower bound	Upper bound
α^h, constant for transmission	-6.3588	-2.4213
β^h, log population for transmission	1.0000	1.0000
θ^{hh}, log(transmission firms + 1) for transmission	-12.2635	-4.9190
θ^{hl}, log(exhaust firms + 1) for transmission	-15.0958	-0.6883
α^l, constant for exhaust	-7.1446	-3.9515
β^l, log population for exhaust	1.0000	1.0000
θ^{lh}, log(transmission firms + 1) for exhaust	-16.7776	-1.6889
θ^{ll}, log(exhaust firms + 1) for exhaust	-3.1053	-1.8872

Here I restricted the impact of competition to be negative, and I found the values of the bounds to be sensitive to the starting parameters. More discussion in the next section.

3.2 A better way . . .

As stated previously, the problem comes from the fact that the first behavioral assumption is really

A1 Firms that enter make positive profits

$$\pi(N_m, X_m) + \epsilon_m > 0 \qquad \text{whenever } N_m > 0$$

and thus our identifying moment assumption is

$$\begin{aligned} 0 &< E[\pi(N_m, X_m) + \epsilon_m | N_m > 0] \\ &= \pi(N_m, X_m) + E[\epsilon_m | N_m > 0] \end{aligned}$$

where since we are taking a conditional expectation, we have $E[\epsilon_m|N_m > 0] > 0$ (according to our model of profit maximizing entry), and this means that our original construction of v_1 and v_2 uses bounds that are too tight. There is, however, a clever way (due to Allan) to construct moments that will allow us to drop these unobservables from our moment condition.

We can create an inequality that is an implication of A1 and A2, i.e.

A1* Firms that enter in market i make more profits than a further entrant in market $j \neq i$

$$[\pi(N_i, X_i) + \epsilon_i] - [\pi(N_j + 1, X_j) + \epsilon_j] > 0 \qquad \text{whenever } N_i, N_j > 0$$

and from this inequality we can construct the following moment

$$\begin{aligned} 0 &< E\left[[\pi(N_i, X_i) + \epsilon_i] - [\pi(N_j + 1, X_j) + \epsilon_j] | N_i, N_j > 0\right] \\ &= \pi(N_i, X_i) - \pi(N_j + 1, X_j) + E[\epsilon_i - \epsilon_j | N_i, N_j > 0] \\ &= \pi(N_i, X_i) - \pi(N_j + 1, X_j) + E[\epsilon_i | N_i > 0] - E[\epsilon_j | N_j > 0] \\ &= \pi(N_i, X_i) - \pi(N_j + 1, X_j) \end{aligned}$$

where the unobservables cancel out. We can then construct the sample analog of this by either taking all $i \neq j$ pairs or a random sample of these pairs (I chose a single random order of j to match with the i). Thus we follow the procedures in the previous section, but with the modified conditions

$$\begin{aligned} v_1^* &= \left(- \sum_{N_i, N_j > 0} \left[\pi^h(X_i, N_i^h, N_i^l) - \pi^h(X_j, N_j^h + 1, N_j^l)\right] Z_i \right)^+ \\ v_2^* &= \left(- \sum_{N_i, N_j > 0} \left[\pi^l(X_i, N_i^h, N_i^l) - \pi^l(X_j, N_j^h + 1, N_j^l)\right] Z_i \right)^+ \end{aligned}$$

in place of v_1 and v_2 in our criterion function (1).

Again we find that there exists many parameter vectors such that $Q^*(\theta) = 0$, so we are partially identified. One such vector is $\theta = [\alpha^h = -3.04 \quad \beta^h =$

1.00 $\theta^{hh} = -5.29$ $\theta^{hl} = -0.96$ $\alpha^l = -5.04$ $\beta^l = 1.00$ $\theta^{lh} = -3.47$ $\theta^{ll} = -2.92$]. Though this vector is just one of many that could be true given our data and model, it does give us an idea of the relative values of the different coefficients. The negative constants again suggest fixed costs of entry. The impact of competition appears to be negative and perhaps stronger within types than between types, though in this case there is an asymmetric impact of competition in that entry of transmission firms appears to hurt exhaust firms more than exhaust firms hurt transmission firms.

Proceeding with estimation of the hypercube containing the identified set Θ, we get the results in Table 4* below.

Table 4*. Bounds on the identified set of parameters in Mazzeo model

	Lower bound	Upper bound
α^h, constant for transmission	-6.0874	-2.7736
β^h, log population for transmission	1.0000	1.0000
θ^{hh}, log(transmission firms + 1) for transmission	-10.0261	-5.0534
θ^{hl}, log(exhaust firms + 1) for transmission	-11.0090	-0.4852
α^l, constant for exhaust	-6.4283	-3.5635
β^l, log population for exhaust	1.0000	1.0000
θ^{lh}, log(transmission firms + 1) for exhaust	-17.4398	-1.1287
θ^{ll}, log(exhaust firms + 1) for exhaust	-100.0000	-1.6317

Again here it is worth making some comments about the estimation details before discussing the results. As in problem 2, I constrained the effect of competition to be negative, but in this case this assumption has extra bite because it doesn't have the direct mapping to the objective function effect that it does in the maximum likelihood approach. Also, I found the bounds were not robust to different starting values (the results shown are for the particular θ given above as the starting value). This is a result of the fact that *fmincon* is a local optimization routine and our $Q(\theta) = 0$ constraint is highly nonlinear. Thus even this approach of finding the hypercube containing the identified set suffers from the potentially non-smooth shape of the set's boundaries. I have not attempted to try enough starting values to obtain more robust bounds (though obviously the hypercube would grow if I did so). The discussion of the results should be interpreted with this caveat.

Turning to the results, it is hard to draw many decisive conclusions by looking at these bounds. The negativity of the constants, and thus the presence of some fixed costs of entry, are one definitive conclusion. It also appears that all of the competitive effects are bounded away from zero.

There are many ways in which such large bounds on the hypercube can be misleading, though, since the coefficients are correlated through our model. One way to dig deeper is to look at many specific $\theta \in \Theta$, as we did above with one vector. We could also presumably improve our results by getting more data or by modifying our model (perhaps incorporating the autoroute variable and/or some interactions that would make the profit function more flexible).

3.3 MLE vs. Moment Inequalities approaches

It is worth thinking about the pluses and minuses of examining Mazzeo's model with the moment inequalities approach we use here versus his MLE approach. The big differences are summarized in Figure 2 below.

Figure 2: MLE versus Moment Inequalities

MLE	Moment Inequalities
assumes $\epsilon_m \sim \mathcal{N}(0,1) \quad \forall m$	assumes $E[\epsilon_m = 0] \quad \forall m$
imposes selection among multiple equilibria	consistent with necessary conditions for all equilibria
point identified	set identified

Thus the moment inequalities approach has the advantage of eliminating what might be ad-hoc assumptions on the unobservables and equilibrium selection with the disadvantage that the parameters will only be set identified. As we have seen in this problem set, this disadvantage can be very large in that the bounds on the identified set might be difficult to calculate robustly and/or the bounds may be so large that one is left with little to conclude. However, a large identified set in itself is a result in that it emphasizes exactly how much identification is coming from the assumptions on the unobservables and equilibrium selection in the maximum likelihood approach.