IO II- Pset 2

Tom Hierons & Alexander Petrov

February 27, 2021

1

1.1

Letting $\varepsilon_t := (\varepsilon_{1t}, \varepsilon_{2t})$ we can bound this as in Ciliberto and Tamer by (2009):

$$\int_{\mathcal{R}_1} f(\varepsilon_t) d\varepsilon_t \leq \mathbb{P}(y_{1t} = 1, y_{2t} = 0 | x_t) \leq \int_{\mathcal{R}_1} f(\varepsilon_t) d\varepsilon_t + \int_{\mathcal{R}_2} f(\varepsilon_t) d\varepsilon_t$$

Where:

$$\mathcal{R}_2 := [-x_t\alpha, \delta - x_t\alpha]^2$$
$$\mathcal{R}_1 := \Big([-x_t\alpha, \infty) \times (-\infty, \delta - x_t\alpha]\Big) \setminus \mathcal{R}_2$$

So that $\mathcal{R}_1$ is the unique outcome area and $\mathcal{R}_2$ is the multiple outcome area.

1.2

1.

The price is endogenous and correlated with the structural error which may be known to firms. There may also be selection, leading z_{it} to be correlated with the error in the subsample of firms that enter the market.

2.

It depends what is meant by valid instrument and the definition of the errors. If an instrument w_{it} is exogenous and relevant in equation (2) for the subset of firms that have entered and p_{it} is the only endogenous variable for this subset then, by definition, it solves the endogeneity problem through linear IV. However, we only observe equation (2) for entering firms. So it is possible that w_{it} is correlated with p_{it} and uncorrelated with ξ_{it} *in the population as a whole* but that $\mathbb{E}[w_{it}\varepsilon_{it}] \neq 0$ so that in the subset of firms that we observe it returns and inconsistent estimator when used as an IV. z_{it} may also be endogenous given the selection concerns given above.

3.

Define:

$$\zeta_t := (\varepsilon_{1t}, \varepsilon_{2t}, \xi_{1t}, \xi_{2t})$$
$$\tilde{\mathcal{R}}_1 := \mathcal{R}_1 \times (-\infty, e] \times \mathbb{R} \subseteq \mathbb{R}^4$$
$$\tilde{\mathcal{R}}_2 := \mathcal{R}_2 \times \mathbb{R}^2 \subseteq \mathbb{R}^4$$

Then given the pdf $f_{\varepsilon\xi}$ from the question we have:

$$\int_{\tilde{\mathcal{R}}_1} f_{\varepsilon\xi}(\zeta_t) d\zeta_t \leq \mathbb{P}(\xi_{1t} | y_{1t} = 1, y_{2t} = 0, x_t, z_{1t}, z_{2t}) \leq \int_{\tilde{\mathcal{R}}_1} f_{\varepsilon\xi}(\zeta_t) d\zeta_t + \int_{\tilde{\mathcal{R}}_2} f_{\varepsilon\xi}(\zeta_t) d\zeta_t$$

1.3

1.

Let subscripts denote that the expectation is being taken with respect to that firm's information set and beliefs. Then firm i's expected utility maximising strategy given their beliefs is to follow:

$$y_{it} = \mathbb{1}\Big\{\alpha x_t - \delta \mathbb{E}_i[y_{-i,t}] + \varepsilon_{1t} \geq 0\Big\} \tag{1}$$

As illustrated above- each firm's action is a random variable from the perspective of the other firm: namely it maps realisations of their shock to entry/exit decisions. Computing the expectation we get:

$$\mathbb{E}_i[y_{-i,t}] = 1 - F_\varepsilon\big(- \alpha x_t + \delta \mathbb{E}_{-i}[y_{it}]\big)$$

Using the logistic assumption a BNE must therefore satisfy the following system of equations in terms of $p_{2t} := \mathbb{E}_1[y_{2t}]$ and $p_{1t} := \mathbb{E}_2[y_{1t}]$:

$$p_{2t} = \frac{\exp\big(\alpha x_t - \delta p_{1t}\big)}{1 + \exp\big(\alpha x_t - \delta p_{1t}\big)}$$

$$p_{1t} = \frac{\exp\big(\alpha x_t - \delta p_{2t}\big)}{1 + \exp\big(\alpha x_t - \delta p_{2t}\big)}$$

With the strategies following realisation given by equation (1).

2.

For $x_t = 1$ we have that when $(\alpha, \delta) = (1, 1)$ the unique probabilities are $(p_{1t}, p_{2t}) \in \{(0.599, 0.599)\}$. For $(\alpha, \delta) = (3, 6)$ we get $(p_{1t}, p_{2t}) \in \{(0.929, 0.071), (0.5, 0.5), (0.071, 0.929)\}$.

For $x_t = 2$ we have that $(\alpha, \delta) = (1, 1)$ the unique probabilities are $(p_{1t}, p_{2t}) \in \{(0.773, 0.773)\}$. For $(\alpha, \delta) = (3, 6)$ we get $(p_{1t}, p_{2t}) \in \{(0.711, 0.849), (0.785, 0.785), (0.849, 0.711)\}$.

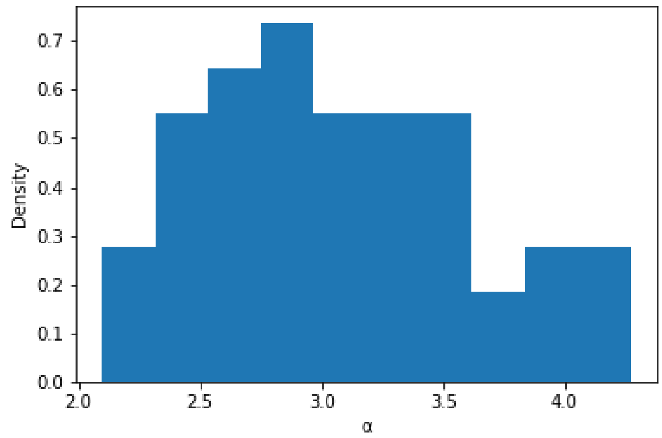
In all cases the corresponding strategies are given by equation (1).

3.

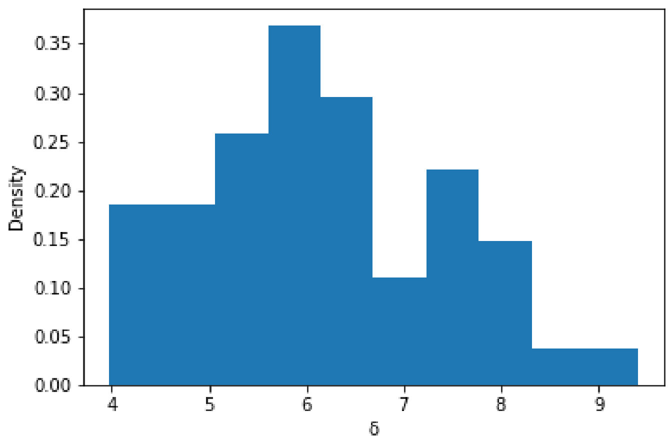
The MLE estimator is implemented using a fixed point approach. Given proposed values for α, δ we solve for p_t according to the equation for a symmetric BNE in each market (given the realisation of x_t). Given these values for p_t, we then update α, δ to maximise the conditional log-likelihood given by:

$$\begin{aligned}\mathcal{L}(\alpha, \delta; \{y_{it}, x_t, p_t\}) &= \sum_{t=1}^{T} \sum_{i=1}^{2} (y_{it} \log \mathbb{P}(y_{it} = 1 | x_t, p_t; \alpha, \delta) + (1 - y_{it}) \log(1 - \mathbb{P}(y_{it} = 1 | x_t, p_t; \alpha, \delta))) \\ &= \sum_{t=1}^{T} \sum_{i=1}^{2} (y_{it} \log(1 - F_\varepsilon(\delta p_t - \alpha x_t) + (1 - y_{it}) \log(F_\varepsilon(\delta p_t - \alpha x_t))\end{aligned}$$

Where $F_\varepsilon(.)$ is the logistic CDF. Further details are given in the attached code. As suggested we run $S = 50$ Monte Carlo repetitions for samples with $T = 1000$ markets. The histogram of the results is presented in the figure below. The estimates are around the truth.



(a) α



(b) δ

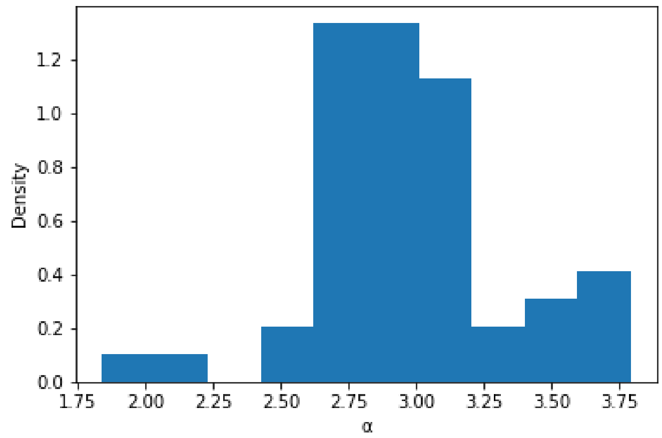
Figure 1: Histograms

4.

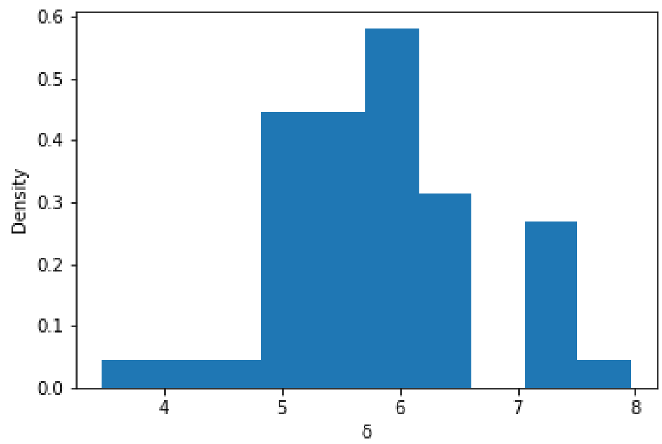
Again the MLE estimator is implemented using a fixed point approach but with an added complication. Given proposed values for α, δ we solve for p_{it}^k according to the equations for a BNE. We then order them according to the given rule and calculate the probability of selecting each of these equilibria is $\lambda_{it}^k(x_t)$ as given in the problem statement. Given these values for p_{it}^k and $\lambda_{it}^k(x_t)$, we then update α, δ to maximise the conditional log-likelihood given by:

$$\mathcal{L}(\alpha, \delta; \{y_{it}, x_t, p_{it}^k, \lambda_{it}^k(x_t)\}) = \sum_{t=1}^{T} \sum_{i=1}^{2} \Bigg(y_{it} \log \left[\sum_{k=1}^{K} \lambda_{it}^k(x_t) \big(1 - F_\epsilon(\delta p_{-i,t} - \alpha x_t)\big) \right] + (1 - y_{it}) \log \left[\sum_{k=1}^{K} \lambda_{it}^k(x_t) F_\epsilon(\delta p_{-it} - \alpha x_t) \right] \Bigg)$$

As in part 3 I present histograms of the estimates.



(a) α



(b) δ

Figure 2: Histograms

Again the histograms show that the estimates cluster around the truth across simulations.

5.

Note that we have:

$$\lambda_{it}^k(x_t, 1) = \frac{\exp\left(\frac{1+k}{2}\right)}{\sum_{j=1}^K \exp\left(\frac{1+j}{2}\right)} = \frac{\exp\left(\frac{k}{2}\right)}{\sum_{j=1}^K \exp\left(\frac{j}{2}\right)}$$

And also:

$$\lambda_{it}^k(x_t, 0) = \frac{\exp\left(\frac{k}{2}\right)}{\sum_{j=1}^K \exp\left(\frac{j}{2}\right)}$$

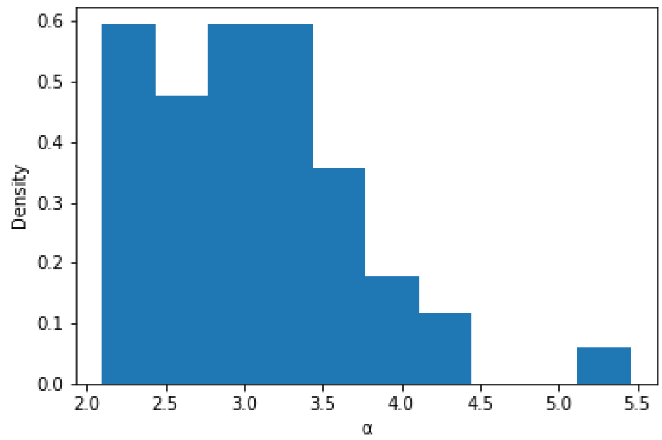
So combining we find that:

$$\begin{aligned}
\mathbb{P}(\text{equilibrium } k \text{ selected}|x_t) &= \mathbb{P}(u_t = 1|x_t)\lambda_{it}^k(x_t, 1) + \big(1 - \mathbb{P}(u_t = 1|x_t)\big)\lambda_{it}^k(x_t, 0) \\
&= \frac{\exp\left(\frac{k}{2}\right)}{\sum_{j=1}^K \exp\left(\frac{j}{2}\right)} \\
&= \lambda_{it}^k(x_t)
\end{aligned}$$

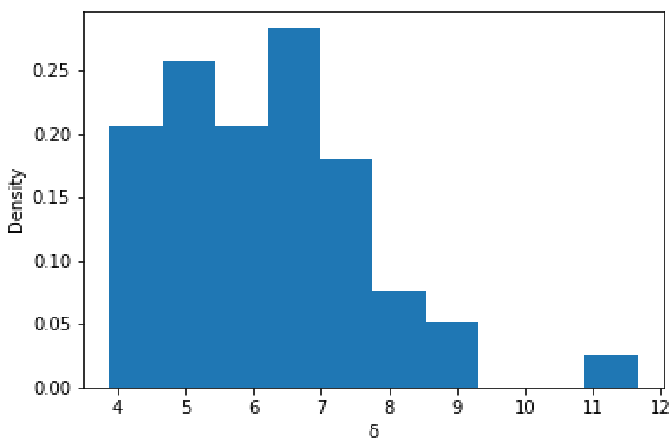
So that the equilibrium selection mechanism turns out to be the same as in part 4. The method from part 4 will therefore return a consistent estimate. In general if the equilibrium selection mechanism is misspecified then this will not be the case.

6. Bonus discussion

As suggested in the hint- we could also think about running the algorithm from part 3 on data generated according to part 4 or the algorithm from part 4 on data generated according to part 3. In both cases the equilibrium selection mechanism is misspecified and we would expect the estimators to be inconsistent. This intuition is supported by the following plots.

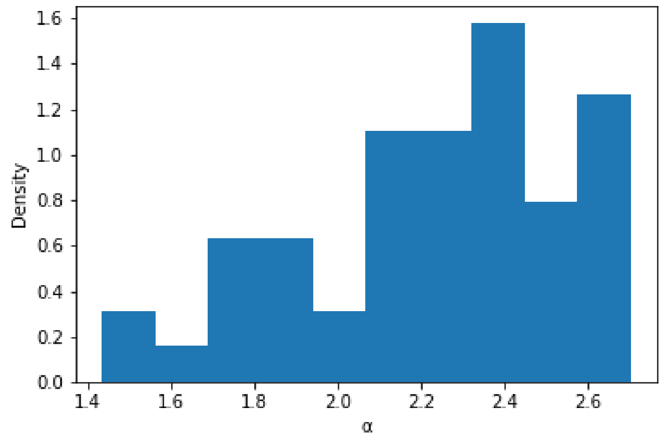


(a) α

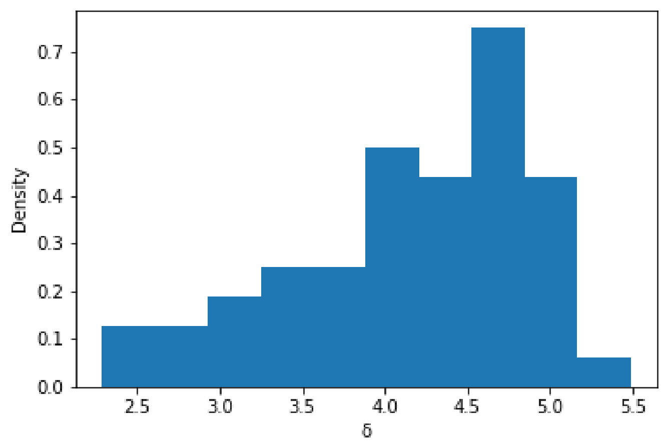


(b) δ

Figure 3: Histogram of data with data as in part 4 and algorithm from part 3



(a) α



(b) δ

Figure 4: Histograms with data as in part 3 but algorithm from part 4

In both cases we see that the algorithms perform poorly and are inconsistent.

2 Dynamics

2.3.0.1

The engine replacement is characterized by a decrease in the mileage. That is, we assign value 1 to periods immediately after which there is a drop in the mileage.

2.3.0.2

The conditional independence assumption allows to disentangle the transition probabilities for the state variable (mileage) and the shock ε:

$$p(x_{t+1}, \varepsilon_{t+1} | x_t, \varepsilon_t, d) = q(\varepsilon_{t+1} | x_{t+1}) p(x_{t+1} | x_t, d) \tag{2}$$

2.3.0.3

See the attached code.

2.3.0.4

Starting with the value function $V_\theta(x, \varepsilon)$:

$$\begin{aligned} V_\theta(x, \varepsilon) &= \max_d \left[u(x, d) + \varepsilon_d + \beta \int_y \int_{\varepsilon'} V_\theta(y, \varepsilon') p(d\varepsilon) p(dy | x, d) \right] \\ &= \max_d \left[u(x, d) + \varepsilon_d + \beta \int_y v(y) p(dy | x, d) \right] \end{aligned}$$

Integrating both sides over ε and using the property of T1EV:

$$\begin{aligned} v(x) &= \int_{\varepsilon} \max_{d} \left[u(x,d) + \varepsilon_d + \beta \int_y v(y) p(dy|x,d) \right] p(d\varepsilon) \\ &= \log \left[\sum_d \exp \left(u(x,d) + \beta \int_y v(y) p(dy|x,d) \right) \right] + k \end{aligned}$$

Integrating again over the next period mileage variable y:

$$EV(x,d) = \int_y \log \left[\sum_d \exp\left(u(y,d) + \beta EV(y,d) \right) \right] p(dy|x,d) \tag{3}$$

2.3.0.5

$$\begin{aligned} P(d_t = 0|x_t) &= P\left(EV(x_t, d_t = 0) > EV(x_t, d_t = 1) \right) && (4) \\ &= \frac{\exp\left(u(x_t, d_t = 0) + \beta EV(x_t, d_1 = 0)\right)}{\exp\left(u(x_t, d_t = 0) + \beta EV(x_t, d_1 = 0)\right) + \exp\left(u(x_t, d_t = 1) + \beta EV(0, d_t = 1)\right)} && (5) \\ &= \frac{1}{1 + \exp\left(u(x_t, d_t = 1) + \beta EV(0, d_t = 1) - u(x_t, d_t = 0) - \beta EV(x_t, d_t = 0)\right)} && (6) \end{aligned}$$

Nested Fixed Point

2.3.1.1

$EV(x,1)$ and $EV(0,0)$ are equal, therefore we can effectively get rid of d in calculations of the value function.

2.3.1.2

EV is a $K \times 1$ vector; u_0 and u_1 are also $K \times 1$ vectors; EV_0 is the first element of EV.

$$EV = \Pi \log \left[\exp\left(u_0 + \beta EV\right) + \exp \left(u_1 + \beta \begin{bmatrix} EV_0 \\ \vdots \\ EV_0 \end{bmatrix} \right) \right] \tag{7}$$

2.3.1.3 - 2.3.1.5

See the attached code.

2.3.1.6

The estimated parameters are in Table 1. We encountered computational difficulties and cannot fully trust the obtained results. In particular, the likelihood couldn't be estimated during the search for the optimum (at most steps the yielded value of the likelihood was `nan`). Moreover, the obtained parameter values do not make much sense, as θ_1 and θ_2 are unlikely to be both negative.
Also, we again encountered difficulties with computing gradient using `Jax`. It seems not to tolerate the fact that while computing the fixed point EV, exponents are taken of extremely low values. We couldn't find a way to alleviate this and make `Jax` compute the derivative.

	Estimate
θ_1	-1034.13
θ_2	-9.65
θ_3	11.29

Table 1: Results of parameter estimation