

Empirical IO

PS1: BLP structural estimation

Hans Martinez

Mar 03, 2021

Homogeneous consumers model

I used Julia for this problem set. This solution and the code can be found in my github repo. My true parameters are $\{\beta, \alpha, \sigma, \eta\} = \{2.5, 4.0, 2.0, 1.5\}$.

- a. Description of the simulated data is summarized in table 1. Likewise, the correlations between the unobserved and observed variables are described in table 2.

Table 1: Summary statistics of simulated data

	Obs	Mean	Std. Dev.	Min	Max
x	100	0.022	1.058	-2.131	3.574
ξ	100	-0.060	1.029	-2.667	2.686
w	100	0.081	0.930	-2.603	1.888
ω	100	0.172	0.937	-2.247	2.524
p	100	0.226	2.274	-4.663	5.118
δ	100	-0.909	7.084	-17.930	12.869

Table 2: Correlations between observed and unobserved variables

	x	w	p
p	0.6972	0.4050	1.000
ξ	0.1267	-0.0056	0.4971
ω	-0.1890	-0.0759	0.2851

- b. Table 3 displays the results for the estimates of α and β , ignoring the endogeneity of prices in columns (1) and (2).
- c. Columns (3)-(8) display the results for different variations of instrumental variables estimation.

Discussion: The results that are closer to the true parameters are the ones from the model (5). Model (5) includes the intercept in both stages. Not surprisingly, OLS is biased and performs poorly in comparison to IV when prices are endogenous.

Table 3: Results from estimations

	OLS		IV Reduced Form		IV2S			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(Intercept)	-12.672*** (0.083)		-12.662*** (0.071)		-12.602*** (0.495)			-13.054*** (0.486)
x	1.888*** (0.106)	2.922* (1.645)	1.693*** (0.097)	2.589 (1.758)	2.418*** (0.881)	6.036** (2.399)	3.811 (2.524)	2.368*** (0.876)
p	-3.623*** (0.049)	-4.502*** (0.762)	-3.452*** (0.052)	-4.208*** (0.931)				
w			-0.548*** (0.093)	-0.932 (1.687)				
z					-3.986*** (0.515)	-6.629*** (1.390)	-5.109*** (1.466)	-3.946*** (0.509)
N	100	100	100	100	100	100	100	100
R^2	0.987	-2.200	0.990	-2.190	0.542	-2.523	-2.862	0.542

Columns (7) and (8) do not include an intercept in the first stage of the IV estimation.

Heterogeneous consumers model (BLP)

For these simulations, I keep the values for the true parameters. Namely, $\{\beta, \alpha, \sigma, \eta\} = \{2.5, 4.0, 2.0, 1.5\}$.

- a. Table 4 describes the simulated data with 10,000 heterogeneous agents. Table 5 describes the correlation between the observed and unobserved variables for this new draw.

Table 4: Summary statistics from simulated data with heterogeneous consumers

	Obs	Mean	Std. Dev.	Min	Max
x	100	0.008	1.100	-2.808	2.998
ξ	100	0.005	1.033	-2.239	2.762
w	100	0.013	0.963	-2.663	2.217
ω	100	-0.217	0.919	-2.686	1.819
p	100	-0.188	2.253	-6.412	5.821
δ	100	0.775	6.941	-16.212	21.801
ν	10000	-0.001	1.009	-4.121	3.717

Table 5: Correlations between observed and unobserved variables

	x	w	p
p	0.6800	0.4127	1
ξ	0.01017	-0.0229	0.4312
ω	-0.1120	0.0372	0.3136

- b. The estimates of the OLS regression assuming homogeneous taste parameters are shown in table 6. We can see that simple OLS is biased when we do not account for the heterogeneity of the taste parameters. In particular, β is being pushed up, whereas α is pushed down. In contrast to problem 1, using the cost shifters as instruments and doing a simple IV in reduced form does not help solve the endogeneity problem.

Table 6: OLS results assuming homogeneous taste parameters

	(1)	(2)	(3)	(4)
(Intercept)	1.350*** (0.201)	1.419*** (0.191)		
x	5.066*** (0.249)	4.640*** (0.265)	5.211*** (0.298)	4.884*** (0.328)
p	-3.592*** (0.121)	-3.300*** (0.142)	-3.690*** (0.145)	-3.466*** (0.175)
w		-0.852*** (0.244)		-0.665** (0.303)
N	100	100	100	100
R^2	0.900	0.911	0.854	0.861

- c. I started fixing $\sigma = 1.98$ for the BLP fixed-point algorithm. I set the outside good share greater than 20% and used these values as the ‘observed’ shares, s_j . So, for initial values, I set $\delta_0 = \log(s_j) - \log(s_0)$. Figure 1 shows the distribution of the recovered mean utilities ($\hat{\delta}$).

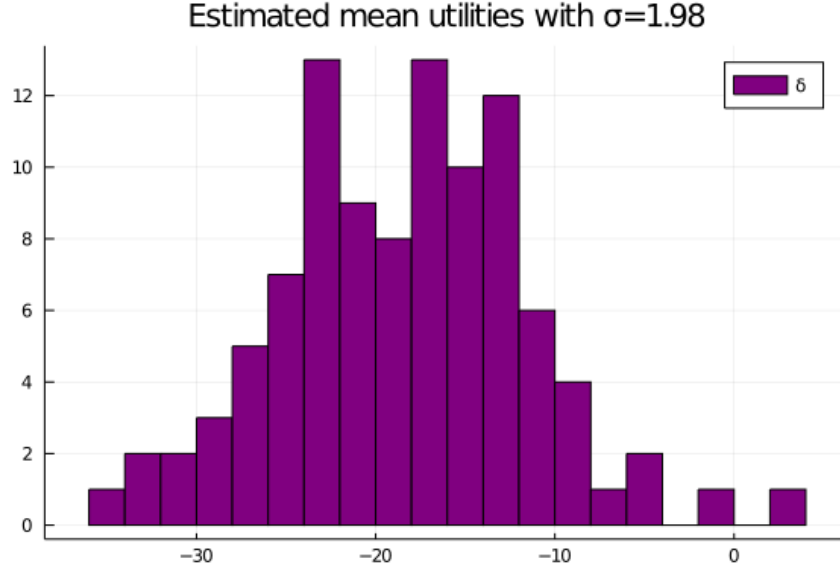


Figure 1: Estimated mean utilities from the BLP contraction mapping algorithm.

- d. Table 7 shows estimates for $\hat{\alpha}$ and $\hat{\beta}$. We can observe that the best model is the IV2S (5).
- e. Therefore, using estimates of table 7, column (5), the implied product attributes are graphically displayed in figure 2.
- f. Following BLP, I defined the matrix of instruments as $Z = \left[\mathbf{1} \sum_{r \neq j \notin \mathcal{F}_f} \mathbf{x}_{rk} \sum_{r \neq j \notin \mathcal{F}_f} \mathbf{w}_{rk} \right]$. That is, the vector of ones, the sum of all the characteristics of the other products except its own, and the sum of all of the cost shifters of the other products except its own. This implicitly assumes that each product is produced by one firm. Therefore, my moment condition is $G = E[Z\hat{\xi}]$, and the criterion function becomes:

Table 7: Using estimated mean utilities to recover α & β

	OLS		IV Reduced Form		IV2S			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(Intercept)	-19.235*** (0.086)		-19.186*** (0.073)		-19.312*** (0.479)		-18.458*** (0.469)	
x	2.034*** (0.107)	-0.030 (2.397)	1.735*** (0.101)	-1.569 (2.680)	2.575*** (0.801)	-3.038 (3.306)	2.587*** (0.802)	2.840 (3.285)
p	-3.637*** (0.052)	-2.244* (1.166)	-3.432*** (0.054)	-1.189 (1.429)				
w			-0.598*** (0.093)	-3.133 (2.470)				
z					-4.026*** (0.486)	-0.081 (1.995)	-4.038*** (0.487)	-4.306** (1.996)
N	100	100	100	100	100	100	100	100
R^2	0.985	-6.697	0.990	-6.571	0.551	-6.988	0.551	-6.626

Columns (7) and (8) do not include an intercept in the first stage of the IV estimation.

$$\|G\| = Z\hat{\xi}'WZ'\hat{\xi}$$

where W is the identity matrix.

g. Then, I repeat the process for the grid σ_i for $i \in [1.5, 2.5]$ increasing in steps of 0.01. So, a total of 100 points. Figure 3 displays the log of the value of the criterion function for all the grid. The minimum value is reached at $\sigma = 2.06$, which is closest to the true value 2. Finally, table 8 displays the corresponding values for the minimum value of the criterion function. Estimates are very close to the true parameters.

Table 8: Using estimated mean utilities to recover α & β

	<i>Coeff</i>	<i>SE</i>
<i>Const</i>	-19.37	0.4796
$\hat{\beta}$	2.478	0.8018
$\hat{\alpha}$	-4.021	0.4863
$\hat{\sigma}$	2.04	

As an additional exercise, Berry mentions that second and higher order basis functions are highly collinear with instruments that are similar to ours. Therefore, we ran the whole exercise with $Z = \left[\mathbf{1} (\sum_{r \neq j \notin \mathcal{F}_f} \mathbf{x})^2 (\sum_{r \neq j \notin \mathcal{F}_f} \mathbf{w}_{rk})^2 \right]$. The value of the criterion function with these instruments for the whole grid is displayed in figure 4. The corresponding estimates to the lowest value are shown in table 9.

Table 9: Recovered α & β with different set of instruments

	<i>Coeff</i>	<i>SE</i>
<i>Const</i>	-19	0.4767
$\hat{\beta}$	3.145	0.797
$\hat{\alpha}$	-4.055	0.4834
$\hat{\sigma}$	1.62	

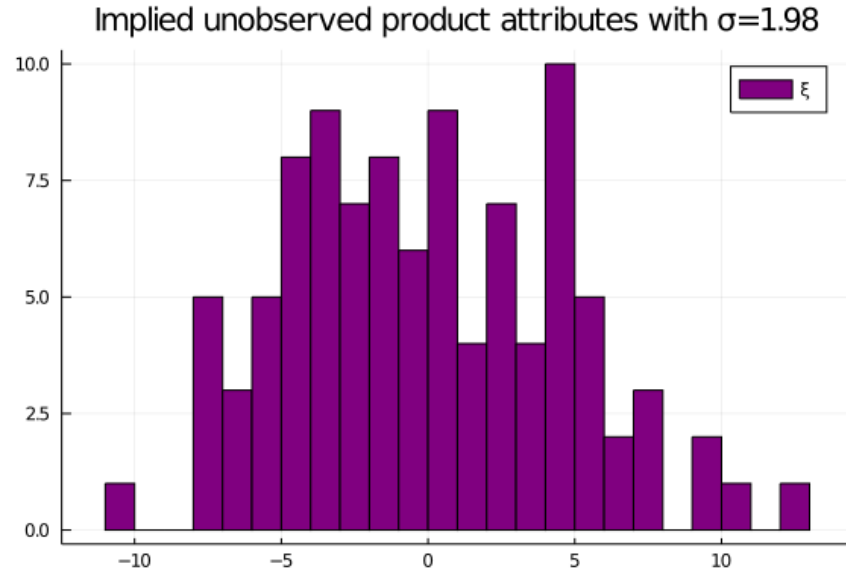


Figure 2: Implied unobserved product attributes after estimating mean utilities from the BLP contraction mapping algorithm.

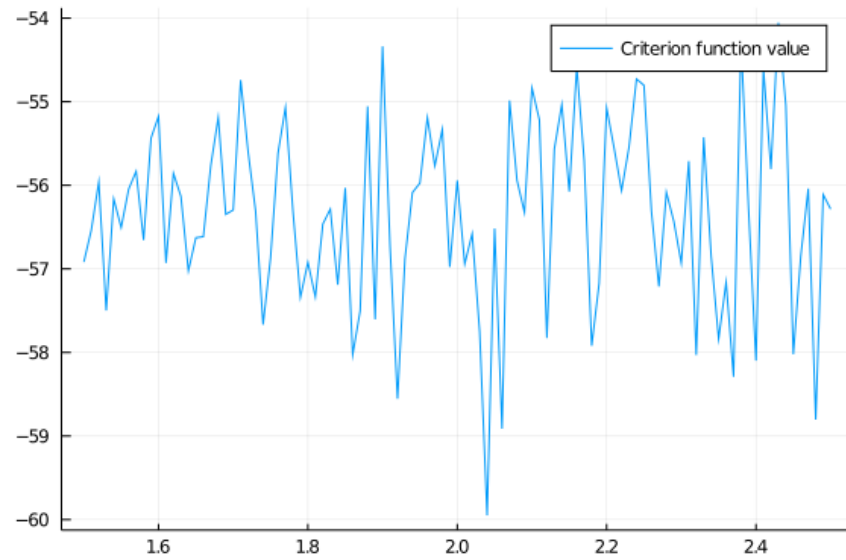


Figure 3: Value of the criterion function at different values of σ .

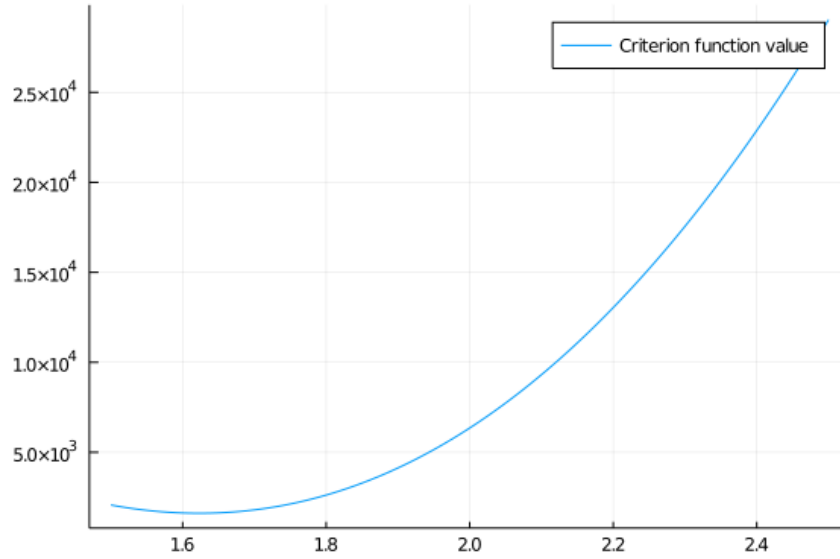


Figure 4: Value of the criterion function at different values of σ for different sets of instruments.

- h. From first problem we can observe that OLS is biased because of endogeneity of prices and unobserved product characteristics. In that case, IV could help us to identify the parameters using the cost shifter as instrument. However, when we have heterogeneous taste parameters, IV is not longer enough. The BLP algorithm helped us recover the parameters of interest nicely.

Discussion: Even tough BLP estimates were not very precise, the true parameters were inside the confidence intervals of the estimates. This was true even when we changed the instruments. In the case of σ , we didn't estimated the s.e. because it was outside of the scope of the problem set, we could expect that both reported estimates were not different from the true parameter and from each other.