Problem set 4

Natalia Serna

1 Setup

1. Results from the OLS, IV and full model are reported in table (**??**). These estimates are obtained with the code presented in the appendix. This code is written in R and replicates the MATLAB code provided. However, when analyzing the MATLAB code, I noticed several things that could potentially affect the estimates obtained with it, which haven't been mentioned nor addressed before:

   - When performing the IV regression of the mean utility level, product characteristics which are assumed to be exogenous, are not included within the set of IVs for the first stage regression. They only enter the second stage. Although the argument is that brand fixed effects capture this, there is variation within a brand of sugar content and mushiness, which wouldn't be accounted for.
   - Following the argument that product characteristics are exogenous as well as the constant term, they should be included in the matrix of instruments when computing the criterion function. Nevertheless, the MATLAB code does not include them.
   - The estimator for the weight matrix in the criterion function is not the optimal weight matrix. According to Nevo's Practitioner's Guide, A should be a consistent estimator for $E[Z'\omega\omega'Z]$, where Z is the matrix of instruments and ω are the residuals from the mean utility regression. This suggests the appropriate estimator is $\hat{A} = (J)^{-1}(Z'\hat{\omega}\hat{\omega}'Z)$, but the MATLAB code implements $\hat{A} = Z'Z$.
   - The problem set mentions that the MATLAB code implements the results in Nevo's Practitioner's Guide, but it does not include market fixed effects, which are used in the paper.

- Finally, there are three sources of variance in the model, two of which are not accounted in the estimation of the variance covariance matrix: first there is the intrinsic variance due to ξ. Second, we have an error induced by simulation of individuals -that in this application seems particularly high given that we only simulate 20 of them- and third, we have variance coming from the fact that we don't observe the population shares but just a sample. In my code I account for the three sources of variance, and show that they are not negligible.

Table (1) shows some summary statistics pooled across markets. The average price is 0.125, the proportion of cereals that are mushy is 33%, the average sugar content is 8.62, and the average market share of a product in a market is 1.98%.

Table 1: Pooled summary statistics

	Mean	Median	sd
Price	0.1257	0.1240	0.0290
Shares	0.0198	0.0112	0.0256
Sugar	8.6250	8.5000	5.7879
Mushy	0.3333	0.0000	0.4715

Table (**??**) shows the estimation results for the OLS Logit model with and without brand fixed effects, the IV Logit model with and without brand fixed effects, and the random coefficients logit. The first thing worth noticing is that not accounting for price endogeneity severely underestimates demand elasticity, which is inconsistent with profit maximization. Underestimation of demand elasticity also explains why predicted markups are higher for OLS models compared to the IV and Mixed logit models. Using brand fixed effects in the OLS and IV specifications tends to bias the estimate for sugar content in the utility function, indicating that higher sugar content is less preferred. However after controlling for interactions with number of children in the full model, we can see that sugar content increases the mean utility level. In all specifications, mushy cereals are also preferred. In particular, in the full model we find that price significantly increase the dispersion of the distribution of utility levels, but not the sugar content nor the mushiness. Results also show that high income individuals are less price sensitive.

Table 2: Demand estimation

	OLS		IV		Full model					
					Mean	St. dev.	Interactions			
							Income	Income2	Age	Child
Constant	-2.9945	-2.7391	-2.8468	-2.6011	-5.5670	0.3124	0.4693	2.6104	0.5629	0.2814
	(0.1117)	(0.0888)	(0.1119)	(0.0883)	(0.8436)	(0.6067)	(0.3395)	(0.7295)	(0.2209)	(0.4502)
Price	-10.1059	-28.9474	-11.3601	-29.3589	-31.7361	1.8295	3.0993	13.2909	-0.4852	0.3313
	(0.8800)	(0.9854)	(0.8813)	(0.9592)	(0.1664)	(0.3230)	(0.5872)	(0.5815)	(0.0750)	(0.4518)
Sugar	0.0461	-0.0159	0.0477	-0.0157	0.0451	-0.0129	-0.1407	-0.5710	-0.0357	-0.0005
	(0.0044)	(0.0033)	(0.0044)	(0.0033)	(0.5844)	(0.5052)	(0.5456)	(0.7750)	(0.5332)	(0.2800)
Mushy	0.0523	0.4869	0.0417	0.4960	0.7308	0.1711	0.5290	1.3563	0.3747	-0.4702
	(0.0520)	(0.0414)	(0.0516)	(0.0409)	(1.1436)	(0.9091)	(0.3679)	(1.3314)	(0.0905)	(0.4431)
Fixed effects										
Brand	No	Yes	No	Yes	Yes					
R-squared	0.0790	0.1127	0.0921	0.1147	0.2732					
GMM Obj					18.7408					

2. Table (3) shows the implied markups and marginal costs from the five specifications. As I mentioned before, marginal costs are underestimated in models where price endogeneity is not accounted for and high markups are inconsistent with a relatively less elastic demand in those models. In the full model, the average markup is 33.2% of the average price, and the average marginal cost is 0.084. While in the OLS model without brand fixed effects these are 93% and 0.008, respectively. To compute the markups I use the following equation:

$$\hat{b} = \hat{\Omega}^{-1}\hat{s}_{jt}$$

where

$$\hat{\Omega} = \begin{cases} -\frac{\partial s_{jt}}{\partial p_{rt}} & \text{if} \quad \text{firm } f \text{ produces } j \text{ and } r \\ 0 & \text{otherwise} \end{cases}$$

So for instance if we have three firms in one market each producing 3,2, and 1 products, $\hat{\Omega}$ will look like:

$$\begin{pmatrix} \partial s_1/\partial p_1 & -\partial s_1/\partial p_2 & -\partial s_1/\partial p_3 & 0 & 0 & 0 \\ -\partial s_2/\partial p_1 & \partial s_2/\partial p_2 & -\partial s_2/\partial p_3 & 0 & 0 & 0 \\ -\partial s_3/\partial p_1 & -\partial s_3/\partial p_2 & \partial s_3/\partial p_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \partial s_4/\partial p_4 & -\partial s_4/\partial p_5 & 0 \\ 0 & 0 & 0 & -\partial s_5/\partial p_4 & \partial s_5/\partial p_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \partial s_6/\partial p_6 \end{pmatrix}$$

Table 3: Markups and marginal cost in monetary units

	Markups		
	Mean	Median	St. dev.
OLS w/o brand f.e	0.1172	0.1150	0.0145
OLS w/ brand f.e	0.0409	0.0402	0.0050
IV w/o brand f.e	0.1043	0.1023	0.0129
IV w/ brand f.e	0.0403	0.0396	0.0050
Full model	0.0418	0.0396	0.0095
	Marginal cost		
	Mean	Median	St. dev.
OLS w/o brand f.e	0.0085	0.0075	0.0339
OLS w/ brand f.e	0.0848	0.0830	0.0301
IV w/o brand f.e	0.0215	0.0201	0.0331
IV w/ brand f.e	0.0854	0.0836	0.0300
Full model	0.0839	0.0819	0.0324

3. Table (4) shows the equilibrium prices and market share following the Post-Nabisco merger and table (5) shows the equilibrium after the GM-Quaker merger for all of the specifications. To compute the equilibrium prices and quantities after a merger, we want to find the vector p^* that solve the following equation :

$$p^* = \hat{mc} + \hat{\Omega}_{post}^{-1}(p^*)s(p^*)$$

This implies finding the fixed point of such contraction mapping. To do so I follow these steps in every market t:

(a) Fix a vector of prices p^0

(b) Using the estimates for $\hat{\bar{\alpha}}, \hat{\bar{\beta}}, \hat{\xi}_{jt}$, compute $\hat{\delta}_{jt}$, as:

$$\hat{\delta}_{jt} = \hat{\bar{\alpha}}p^0 + \mathbf{x}_{jt}\hat{\bar{\beta}} + \hat{\xi}_{jt}$$

(c) Compute choice probabilities $\hat{s}_{ijt}$ using the estimates $\hat{\Pi}, \hat{\Sigma}, \hat{\delta}$ as:

$$\hat{s}_{ijt} = \frac{exp(\hat{\delta}_{jt} + \mu_{ijt})}{\sum_{k=0}^{J} exp(\hat{\delta}_{kt} + \mu_{itt})}$$

where

$$\mu_{ijt} = [\hat{\Pi}D_i + \hat{\Sigma}v_i][p^0 \quad \mathbf{x}_{jt}]$$

(d) Compute market shares $\hat{s}_{jt}$ as:

$$\hat{s}_{jt} = \frac{1}{n}\sum_{i=1}^{20} \hat{s}_{ijt}$$

(e) Compute derivatives of demand as:

$$\frac{\partial s_{jt}}{\partial p_{kt}} = \begin{cases} -\frac{1}{n}\sum_{i=1}^{20} \hat{\alpha}_i \hat{s}_{ijt}(1-\hat{s}_{ijt}) & \text{if} \quad j = k \\ \frac{1}{n}\sum_{i=1}^{20} \hat{\alpha}_i \hat{s}_{ijt}\hat{s}_{ikt} & \text{if} \quad j \neq k \end{cases}$$

(f) Predict markups using:

$$\hat{b} = \hat{\Omega}^{-1}\hat{s}_{jt}$$

where $\hat{\Omega}_{post}$ is as defined previously. The difference being that if we assume firms 1 and 2 merge, then $\hat{\Omega}_{post}$ will look like:

$$\begin{pmatrix} \partial s_1/\partial p_1 & -\partial s_1/\partial p_2 & -\partial s_1/\partial p_3 & -\partial s_1/\partial p_4 & -\partial s_1/\partial p_5 & 0 \\ -\partial s_2/\partial p_1 & \partial s_2/\partial p_2 & -\partial s_2/\partial p_3 & -\partial s_2/\partial p_4 & -\partial s_2/\partial p_5 & 0 \\ -\partial s_3/\partial p_1 & -\partial s_3/\partial p_2 & \partial s_3/\partial p_3 & -\partial s_3/\partial p_4 & -\partial s_3/\partial p_5 & 0 \\ -\partial s_4/\partial p_1 & -\partial s_4/\partial p_2 & -\partial s_4/\partial p_3 & \partial s_4/\partial p_4 & -\partial s_4/\partial p_5 & 0 \\ -\partial s_5/\partial p_1 & -\partial s_5/\partial p_2 & -\partial s_5/\partial p_4 & -\partial s_5/\partial p_4 & \partial s_5/\partial p_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & \partial s_6/\partial p_6 \end{pmatrix}$$

(g) Calculate new vector of prices

$$p^1 = \hat{mc} + \hat{b}$$

(h) Repeat (a)-(g) with updated price vector until $||p^{h+1} - p^h|| < tol$, where h is the iteration.

Table 4: Post-Nabisco merger

	Prices		
	Mean	Median	St. dev.
OLS w/o brand f.e	0.1262	0.1240	0.0292
OLS w/ brand f.e	0.1287	0.1272	0.0294
IV w/o brand f.e	0.1259	0.1240	0.0292
IV w/ brand f.e	0.1285	0.1270	0.0294
Full model	0.1318	0.1296	0.0311
	Market shares		
	Mean	Median	St. dev.
OLS w/o brand f.e	0.0199	0.0111	0.0261
OLS w/ brand f.e	0.0245	0.0077	0.0510
IV w/o brand f.e	0.0198	0.0112	0.0251
IV w/ brand f.e	0.0243	0.0076	0.0502
Full model	0.0063	0.0008	0.0195

Table 5: GM-Quaker merger

	Prices		
	Mean	Median	St. dev.
OLS w/o brand f.e	0.1304	0.1284	0.0297
OLS w/ brand f.e	0.1323	0.1303	0.0297
IV w/o brand f.e	0.1297	0.1277	0.0296
IV w/ brand f.e	0.1321	0.1301	0.0296
Full model	0.1384	0.1343	0.0349
	Market shares		
	Mean	Median	St. dev.
OLS w/o brand f.e	0.0197	0.0110	0.0261
OLS w/ brand f.e	0.0222	0.0081	0.0478
IV w/o brand f.e	0.0196	0.0112	0.0251
IV w/ brand f.e	0.0220	0.0081	0.0469
Full model	0.0067	0.0009	0.0205

4. In computing the equilibrium prices and quantities following a merger there are several implicit assumptions: one is that marginal costs are not changing from the observed scenario to the counterfactual. This means that we are not allowing the merged firm to become more cost efficient or experience economies of scale after the merger. The second is that we are assuming that preferences remain constant from the observed

scenario to the counterfactual, which means that any advertising efforts that the firms can engage in will not affect the marginal utility for each of the product characteristics. Another assumption is that product characteristics or product portfolios are also the same, this can be problematic in the case the merged firm decides to eliminate products in order take advantage of complementarities or substitution between the remaining ones it produces. From the observed to the counterfactual, we are also assuming that the form of competition between firms is not changing, that is, we assume firms compete in prices. However, one could expect that, following a merger, the merged firm becomes a leader and the rest of firms, followers. Finally, we are assuming that all firms have perfect information regarding what the merged firm will do after the merger, but one could also argue against this assumption. All of these problems are related to the underlying assumptions of the model. To address the first, one can jointly estimate the demand and supply side of the market to be able to predict new marginal costs in the counterfactual. To address the second and third criticisms one would have to allow for endogenous product characteristics and product choices.

5. There are several differences between the equilibrium prices and quantities following a merger using the different demand specifications. The most notorious one from tables (4) and (5) is that the OLS and IV models underestimate the new equilibrium price. These models show that after the Post-Nabisco merger, prices were around 0.003 monetary units higher than before the merger without any significant changes in equilibrium market shares. However with the full model we are able to predict that post equilibrium prices are 0.0061 monetary units higher and the average market share falls significantly. Differences in equilibrium prices between the two sets of models are less striking in the case of the GM-Quaker merger. The full model predicts prices were 0.0127 monetary units higher than before the merger. But differences in market shares seems to be important with the OLS and IV predicting almost no change in market shares while the full model predicting a significant decrease relative to the benchmark scenario.

2 Appendix: Code

```
blp <- function(theta_nlin, x, brand=F, iv=T, supply=T, n, A=NULL){

  #Arrange the non linear coefficients
  sig <- theta_nlin[1:4]
  sig <- diag(sig)
  pi <- cbind(theta_nlin[5:8], theta_nlin[9:12], theta_nlin[13:16], theta_nlin[17:20])

  #Organize demographic variables
  dj1 <- seq(1, 20, 1)
  dj1 <- paste("v", dj1, sep="")
```

```
dj2 <- seq(21, 40, 1)
dj2 <- paste("v", dj2, sep="")
dj3 <- seq(41, 60, 1)
dj3 <- paste("v", dj3, sep="")
dj4 <- seq(61, 80, 1)
dj4 <- paste("v", dj4, sep="")

xt <- split(x, as.factor(x$t))

#Fixed point algorithm to compute mean utilities
deltas <- NULL
va <- NULL
sij_a <- NULL

for(i in 1:length(xt)){
  xtt <- xt[[i]]
  xi <- data.matrix(xtt[,c("constant", "price", "sugar", "mushy")])
  d <- data.matrix(xtt[,c(dj1, dj2, dj3, dj4)])
  sjo <- c(xtt$share, 1-sum(xtt$share))
  v <- t(rnorm(n))
  va <- rbind(va,v)
  v <- matrix(rep(t(v),ncol(xi)), ncol=n, byrow = TRUE)
  del <- matrix(1, nrow=nrow(xi), ncol=1)
  sij <- matrix(0, ncol=n, nrow=nrow(xi)+1)
  mu <- matrix(0, ncol=n, nrow=nrow(xi))
  for(j in 1:nrow(d)){
    dj <- rbind(d[j,dj1], d[j,dj2], d[j,dj3], d[j,dj4])
    muj <-  xi%*%sig%*%v + xi%*%pi%*%dj
    mu <- mu+(muj/(nrow(d)))
    u <- del%*%rep(1, n) + muj
    exp_u <- exp(rbind(u, rep(0, ncol(u))))
    sijt <- sweep(exp_u, 2, colSums(exp_u),'/')
    sij <- sij + sijt
  }
  sj <- rowMeans(sij)
  delp <-  log(sjo) - log(sj) + c(del,1)
  dels <- rbind(c(c(del,1)), c(delp))
  tol <-  dist(dels)
  tol <- as.numeric(tol)

  repeat{
    del <- delp
    u <- t(t(del[1:(length(del)-1)]))%*%rep(1, n) + mu
    exp_u <- exp(rbind(u, rep(0, ncol(u))))
    sij <- sweep(exp_u, 2, colSums(exp_u),'/')
    sj <- rowMeans(sij)
    delp <-  log(sjo) - log(sj) + del
    dels <- rbind(c(del), c(delp))
    tol <-  dist(dels)
    tol <- as.numeric(tol)
    if (tol<1e-14) break
  }
  deltas <- rbind(deltas, t(t(delp[1:(length(delp)-1)])))
  u <- t(t(delp[1:(length(delp)-1)]))%*%rep(1, n) + mu
  exp_u <- exp(rbind(u, rep(0, ncol(u))))
  sij_a[[i]] <- sweep(exp_u, 2, colSums(exp_u),'/')
}
va <<- va
```

```r
#Mean utility regression
if(iv==T){
  vars <- seq(1, 20, 1)
  vars <- paste("z", vars, sep="")
  vars <- paste(vars, collapse="+")
  vars <- as.name(vars)
  form <- paste("price", vars, sep="~")
  p_h <- predict(lm(as.formula(form), data=x))
  if(brand==T) vars <- c("-1", "p_h", "brand") else vars <- c("p_h", "sugar", "mushy")
  vars <- paste(vars, collapse="+")
  vars <- as.name(vars)
  form <- paste("deltas", vars, sep="~")
}else{
  if(brand==T) vars <- c("-1", "price", "brand") else vars <- c("price", "sugar", "mushy")
  vars <- paste(vars, collapse="+")
  vars <- as.name(vars)
  form <- paste("deltas", vars, sep="~")
}
mean_u <<- lm(as.formula(form), data=x)
nam <- names(coefficients(mean_u))
nam <- which(nam=="price"|nam=="p_h")

#Minimum distance estimates for brand dummy f.e
if(brand==T){
  ymd <- coefficients(mean_u)[2:length(coefficients(mean_u))]
  hvcov <- vcov(mean_u)[2:length(coefficients(mean_u)),2:length(coefficients(mean_u))]
  ymd <- matrix(c(as.numeric(na.omit(ymd))),nrow=nrow(hvcov), ncol=1)
  xmd <- xt[[1]]
  xmd <- data.matrix(xmd[,c("constant", "sugar", "mushy")])
  hdmd <- solve(t(xmd)%*%solve(hvcov)%*%xmd)%*%t(xmd)%*%solve(hvcov)%*%
  matrix(c(ymd),nrow=nrow(hvcov), ncol=1)
  resmd <- ymd-xmd%*%hdmd
  semd  <-  sqrt(diag(solve(t(xmd)%*%solve(hvcov)%*%xmd)))
  coefs <<- c(hdmd[1], coefficients(mean_u)[nam], hdmd[2:3])
  se <<- c(semd[1], sqrt(vcov(mean_u)[nam,nam]), semd[2:3])
  Rsq <<- 1-(t(resmd-mean(resmd))%*%(resmd-mean(resmd)))/(t(ymd-mean(ymd))%*%
  (ymd-mean(ymd)))
  Rsq_G <<- 1-(t(resmd)%*%solve(hvcov)%*%resmd)/(t(ymd-mean(ymd))%*%
  solve(hvcov)%*%(ymd-mean(ymd)))
}

#Marginal cost estimates
elasticities <- NULL
mc <- NULL
b <- NULL
for(i in 1:length(xt)){
  xtt <- xt[[i]]
  d <- data.matrix(xtt[,c(dj1, dj2, dj3, dj4)])
  aij <- matrix(0, ncol=n, nrow=nrow(xtt))
  for(j in 1:nrow(d)){
    dj <- rbind(d[j,dj1], d[j,dj2], d[j,dj3], d[j,dj4])
    a <-  coefficients(mean_u)[nam] + t(t(xtt$price))%*%sig[2,2]%*%va[i,] +
    t(t(xtt$price))%*%pi[2,]%*%dj
    aij <- aij+a
  }
  aij <- aij/nrow(d)
  sijt <- sij_a[[i]]
  sijt <- sijt[1:(nrow(sijt)-1),]
  sj <- rowMeans(sijt)
```

```
    ep <- rowMeans(aij*sijt*(1-sijt))
    ec <- ((-aij)*sijt)%*%t(sijt)/n
    diag(ec) <- ep
    elasticities[[i]] <- ec*(xtt$price/xtt$share)
    su <- summary(xtt$firm)
    frm <- NULL
    for(k in 1:length(su)) {frm[[k]] <- matrix(1, nrow=su[k], ncol=su[k])}
    om <-  data.matrix(bdiag(frm))
    om <- (-1)*om*ec
    bt <- solve(om)%*%sj
    b <- rbind(b, bt)
    mct <- xtt$price-bt
    mc <- rbind(mc, mct)
  }
  mc <<- mc
  b <<- b

  #Estimation of pricing equation
  if(supply==T){
    if(brand==T) w <<- lm(mc ~ sugar + mushy + brand + t, data=x)
    else w <<- lm(mc ~ sugar + mushy + t, data=x)
    g <- matrix(c(residuals(mean_u),residuals(w)),
    nrow=length(c(residuals(mean_u),residuals(w))), ncol=1)
    ge <<- g
    if(iv==T){
      vars <- seq(1, 20, 1)
      vars <- paste("z", vars, sep="")
      if(brand==T) Z <- data.matrix(rbind(cbind(x[,vars], dummy.data.frame(as.data.frame(x$brand))),
      cbind(x[,vars],dummy.data.frame(as.data.frame(x$brand)))))
      else Z <- data.matrix(rbind(x[,vars],x[,vars]))
    }else{
      Z <- data.matrix(rbind(dummy.data.frame(as.data.frame(x$brand)),
      dummy.data.frame(as.data.frame(x$brand))))
    }
  }else{
    g <- residuals(mean_u)
    ge <<- residuals(mean_u)
    if(iv==T){
      vars <- seq(1, 20, 1)
      vars <- paste("z", vars, sep="")
      if(brand==T) Z <- data.matrix(cbind(x[,vars], dummy.data.frame(as.data.frame(x$brand))))
      else Z <- data.matrix(x[,vars])
    }else{
      if(brand==T) Z <- data.matrix(dummy.data.frame(as.data.frame(x$brand)))
    }
  }

  #Criterion function
  if(brand == T | iv == T) gmm <-  ((t(g)%*%Z)%*%solve(A)%*%(t(Z)%*%g))/nrow(xt[[1]])
  if(brand == F & iv == F) gmm <-  (t(g)%*%(g))/nrow(xt[[1]])
  if(length(gmm)==0) gmm <- 1e8
  return(gmm)
}
```