Rapport du projet Econometrics of Competition

Sujet de Fusion 1

Enguérand Acquarone
Charlotte Pasquier
Benjamin Phan
Jean-Michel Roufosse

Janvier 2018

Question 1

On considère deux entreprises vendant des boissons sucrées et se faisant concurrence dans 300 villes. Nous supposons que chacune de ces villes est habitée par 1000 habitants pouvant consommer de l'eau sucrée produite par les deux entreprises ou de l'eau du robinet non sucrée (*outside option*).

Le bien-être des consommateurs dépend non seulement du taux de sucre présent dans la boisson consommée mais également de son prix. Par conséquent, la fonction d'utilité du consommateur i dans la ville v, par rapport à la consommation de la boisson j peut être modélisée sous la forme suivante :

$u_{ivj} = \alpha + \beta_1 sucre_{vj} - \beta_2 sucre_{vj}^2 - \gamma prix_{vj} + \epsilon_{ivj} + \xi_{vj}$ avec $i \in \{1, ..., 1000\}$, $j \in \{0,1,2\}$, $v \in \{0, ..., 300\}$

On suppose que l'utilité du consommateur augmente avec le taux de sucre contenu dans la boisson consommée. Nous ajoutons par ailleurs la variable explicative "sucre au carré" afin d'introduire un effet de satieté. Ainsi, plus le taux de sucre est élevé et plus le gain marginal d'une augmentation du taux de sucre est faible et peut éventuellement entraîner une diminution du bien-être du consommateur suivant les paramètres.

Deux termes d'erreurs sont introduits. Le premier ϵ_{ivj} représente les préférences individuelles (non capturées par le taux de sucre) de chaque consommateur pour le produit. Le second terme d'erreur ξ_{vj} dépend uniquement de la ville et du produit - il pourrait par exemple s'agir d'une préférence de l'ensemble des individus d'une ville pour l'un des produit en raison de la réputation de l'entreprise dans cette zone.

Avant de procéder au calcul des prix d'équilibre et parts de marché prévalant dans chacune des villes il est nécessaire de simuler les données.

Les taux de sucre des boissons produites par les deux entreprises dans chacune des 300 villes sont simulées selon une loi Beta. La loi Beta, permet d'obtenir un taux de sucre compris entre 0 et 1. De plus, la forme de la densité nous permet d'approcher ce que serait le comportement rationnel d'une entreprise - à savoir allouer un taux de sucre plébiscité par le consommateur moyen sur une grande majorité des marchés tout en permettant une adaptation du taux de sucre par marché.

Par soucis de simplification des calculs, les termes inobservés sont quant à eux simulés à l'aide d'une loi Generalized Extreme Value (GEV) de paramètres $(0, 1, 0)$.

Enfin, nous simulons également les coûts marginaux des deux entreprises sur chaque marché. Le coût marginal de production de l'entreprise j sur le marché v dépend positivement du taux de sucre et est par conséquent égal à : $C_{vj} = a.sucre_{vj} + \eta_{vj}$ avec η_{vj} un terme d'erreur simulé à l'aide d'une loi Gamma. La loi Gamma étant positive on est assuré d'avoir des coûts positifs.

Les paramètres utilisés pour la simulation des données sont les suivants : $\alpha = 1, \beta_1 = 2, \beta_2 = 1$, $\gamma = 1$ et $a = 0,5$.

Afin de calculer les prix d'équilibre et parts de marché nous devons dans un premier temps déterminer la demande sur chaque marché.

Les marchés pouvant être considérés comme distincts, la demande pour une boisson dans chaque ville correspond au nombre d'individus dont l'utilité est maximisée par la consommation du produit. Ainsi $\forall j' \neq j$, la demande du produit j sur le marché v est égale à :

$$
\begin{aligned}
D_{vj} &= \sum_i \mathbb{1}_{u_{ivj} \geq u_{ivj'}} \\
&= \sum_i \mathbb{1}_{\alpha + \beta_1 sucre_{vj} - \beta_2 sucre_{vj}^2 - \gamma prix_{vj} + \epsilon_{ivj} + \xi_{vj} \geq \alpha + \beta_1 sucre_{vj'} - \beta_2 sucre_{vj'}^2 - \gamma prix_{vj'} + \epsilon_{ivj'} + \xi_{vj'}} \\
&= \sum_i \mathbb{1}_{\beta_1 (sucre_{vj} - sucre_{vj'}) - \beta_2 (sucre_{vj}^2 - sucre_{vj'}^2) - \gamma (prix_{vj} - prix_{vj'}) + \epsilon_{ivj} - \epsilon_{ivj'} + \xi_{vj} - \xi_{vj'} \geq 0}
\end{aligned}
$$

Calculons l'espérance de la demande :

$$\mathbb{E}(D_{vj}(sucre, prix, \xi_{vj}, \xi_{vj'})) = \mathbb{E}\left(\sum_i \mathbb{1}_{\beta_1(sucre_{vj}-sucre_{vj'})-\beta_2(sucre^2_{vj}-sucre^2_{vj'})-\gamma(prix_{vj}-prix_{vj'})+\epsilon_{ivj}-\epsilon_{ivj'}+\xi_{vj}-\xi_{vj'}\geq 0|\forall j'\neq j}\right)$$

$$= \sum_i \mathbb{E}\left(\mathbb{1}_{\epsilon_{ivj'}\leq \underbrace{\beta_1(sucre_{vj}-sucre_{vj'})-\beta_2(sucre^2_{vj}-sucre^2_{vj'})-\gamma(prix_{vj}-prix_{vj'})+\xi_{vj}-\xi_{vj'}}_{A_{vjj'}}+\epsilon_{ivj}|\forall j'\neq j}\right)$$

$$= \sum_i \mathbb{P}(\epsilon_{ivj'} \leq A_{vjj'} + \epsilon_{ivj}|\forall j' \neq j)$$

$$= \sum_i \int_{\epsilon_{ivj}} \mathbb{P}(\epsilon_{ivj'} \leq A_{vjj'} + \epsilon_{ivj}|\forall j' \neq j) f(\epsilon_{ivj}) d\epsilon_{ivj}$$

$$= \sum_i \int_{\epsilon_{ivj}} f(\epsilon_{ivj}) \prod_{j'\neq j} F(A_{vjj'} + \epsilon_{ivj}) d\epsilon_{ivj}$$

$$= \sum_i \int_{\epsilon_{ivj}} e^{-\epsilon_{ivj}} e^{-e^{-\epsilon_{ivj}}} \prod_{j'\neq j} e^{-e^{-A_{vjj'}-\epsilon_{ivj}}} d\epsilon_{ivj}$$

$$= \sum_i \int_{\epsilon_{ivj}} e^{-\epsilon_{ivj}} e^{-e^{-\epsilon_{ivj}}(1+\sum_{j'\neq j} e^{-A_{vjj'}})} d\epsilon_{ivj}$$

$$= \sum_i \frac{1}{B_{vj}} \int_{\epsilon_{ivj}} e^{-\epsilon_{ivj}} B_{vj} e^{-e^{-B\epsilon_{ivj}}} d\epsilon_{ivj}$$

$$= \sum_i \frac{1}{B_{vj}} = \frac{n}{B_{vj}}$$

Avec :

$$A_{vjj'} = \beta_1(sucre_{vj} - sucre_{vj'}) - \beta_2(sucre^2_{vj} - sucre^2_{vj'}) - \gamma(prix_{vj} - prix_{vj'}) + \xi_{vj} - \xi_{vj'}$$

et

$$B_{vj} = 1 + \sum_{j'\neq j} e^{-A_{vjj'}}$$

On pose $\delta_{vj} = \alpha + \beta_1 sucre_{vj} - \beta_2 sucre^2_{vj} - \gamma prix_{vj} + \xi_{vj}$ et on a donc $A_{vjj'} = \delta_{vj} - \delta_{vj'}$ et en particulier $\delta_{v0} = \alpha + \xi_{v0}$ qu'on normalise à 0.

On a alors :

$$\begin{aligned}\mathbb{E}(D_{vj}(sucre, prix, \xi_{vj}, \xi_{vj'})) &= \frac{n}{1 + \sum_{j'\neq j} e^{-\delta_{vj}-\delta_{vj'}}} \\ &= \frac{n}{1 + e^{-\delta_{vj}} \sum_{j'\neq j} e^{\delta_{vj'}}} \\ &= \frac{ne^{\delta_{vj}}}{\sum_{j'} e^{\delta_{vj'}}}\end{aligned}$$

L'entreprise j maximise son espérance de profit Π_{vj} dans chaque ville séparément à l'aide des prix p_{vj}. L'espérance de profit s'écrit de la manière suivante :

$$\Pi_{vj} = \mathbb{E}(D_{vj}(sucre, prix, \xi_{vj}, \xi_{vj'}))(p_{vj} - (a.sucre_{vj} + \eta_{vj}))$$

On suppose donc que l'entreprise connaît les caractéristiques de tous les produits i.e. les taux de sucres ainsi que les caractéristiques propres des boissons dans chaque ville : $\xi_{vj'} \forall j'$ ou qu'elle arrive à trouver pour chaque ville le prix qui maximise son profit. Afin de trouver le prix d'équilibre, on fixe le prix de l'une des entreprise et on cherche la meilleure réponse de l'autre entreprise. On dérive donc le log de Π_{vj}/n par rapport au prix p_{vj} (on pose $C_{vj} = c_{vj} sucre_{vj} + \eta_{vj}$) :

$$\begin{aligned}\frac{\partial log(\Pi_{vj}/n)}{\partial p_{vj}} &= \frac{\partial log(e^{\delta_{vj}})}{\partial p_{vj}} - \frac{\partial log(\sum_{j'} e^{\delta_{vj'}})}{\partial p_{vj}} + \frac{\partial log(p_{vj} - C_{vj})}{\partial p_{vj}} \\ &= -\gamma + \frac{\gamma e^{\delta_{vj}}}{\sum_{j'} e^{\delta_{vj'}}} + \frac{1}{p_{vj} - C_{vj}}\end{aligned}$$

Les conditions de premier ordre donnent $\frac{\partial log(\Pi_{v1}/n)}{\partial p_{v1}} = 0$ et $\frac{\partial log(\Pi_{v2}/n)}{\partial p_{v2}} = 0$

$$1 - \gamma(p_{v1} - C_{v1}) + \gamma(p_{v1} - C_{v1}) \frac{e^{\delta_{v1}}}{\sum_{j'} e^{\delta_{vj'}}} = 0 \tag{1}$$

De même,

$$1 - \gamma(p_{v2} - C_{v2}) + \gamma(p_{v2} - C_{v2}) \frac{e^{\delta_{v2}}}{\sum_{j'} e^{\delta_{vj'}}} = 0 \tag{2}$$

Ces équations définissent des conditions nécessaires pour un équilibre de Nash dans la ville v. On décide de résoudre numériquement ce système de deux équations à deux inconnues (les prix intervenant dans les δ correspondants, le système ne peut se résoudre analytiquement). Il est à noter que dans un premier temps l'optimisation numérique des prix renvoyait des prix très négatifs pour un petit nombre de ville. Après investigation nous avons trouvé que le package d'optimisation que nous avions utilisé estimait de manière trop approximative le jacobien de la matrice représentant le système linéaire à résoudre. Nous avons donc implémenté le jacobien à la main afin d'arriver à des prix d'équilibre cohérents.

Afin d'obtenir les parts de marché nous calculons l'utilité des individus lors de la consommation de chaque type de boisson (aux prix d'équilibre obtenus précédemment). Le produit consommé par un individu est celui qui lui procure la plus grande satisfaction, à savoir la plus grande utilité. Dès lors la part de marché s_{vj} est égale au nombre d'individus dont l'utilité est maximisée par la consommation du produit j, et ce divisé par la population de chaque ville.

Question 2 : Statistiques descriptives

Les résultats étant similaires quel que soit le nombre de villes, nous nous contenterons de présenter les résultats obtenus pour 300 villes. Commençons par analyser les prix d'équilibre calculés lors de l'optimisation numérique.

D'après le tableau Figure 1, la boisson sucrée produite par l'entreprise 1 est vendue en moyenne à un prix de 2 euros. Les prix de vente varient par ailleurs relativement peu d'une ville à l'autre. 75% des boissons sont vendues à un prix inférieur à 2,22 euros, le prix de vente minimal etant de 1,09 euros contre 4,76 euros pour le prix de vente maximal. Les paramètres définis precedemment nous permettent donc d'obtenir des résultats cohérents et semblables aux prix pratiqués pour des produits comparables.

	prix_1
Nombre	300
Moyenne	1.99
Écart-type	0.65
Minimum	1.09
Q1	1.55
Médiane	1.83
Q3	2.22
Maximum	4.76

FIGURE 1 – Statistiques descriptives sur le prix des boissons sucrées pour 300 villes de l'entreprise 1

Interessons nous desormais aux coût de production des boissons sucrées dans chacune des villes.

	couts1
Nombre	300
Moyenne	0.32
Écart-type	0.48
Minimum	0.01
Q1	0.14
Médiane	0.20
Q3	2.22
Maximum	3.70

FIGURE 2 – Statistiques descriptives sur le coût de production des boissons sucrées pour 300 villes de l'entreprise 1

En moyenne, une boisson sucrée coûte $0,32$ euros à produire, le coût de production minimal étant de 0,01 euros contre 3,70 euros pour le coût de production maximal. Les distributions des coûts de production ainsi que des prix de vente d'équilibre sont représentées sur le graphique suivant :

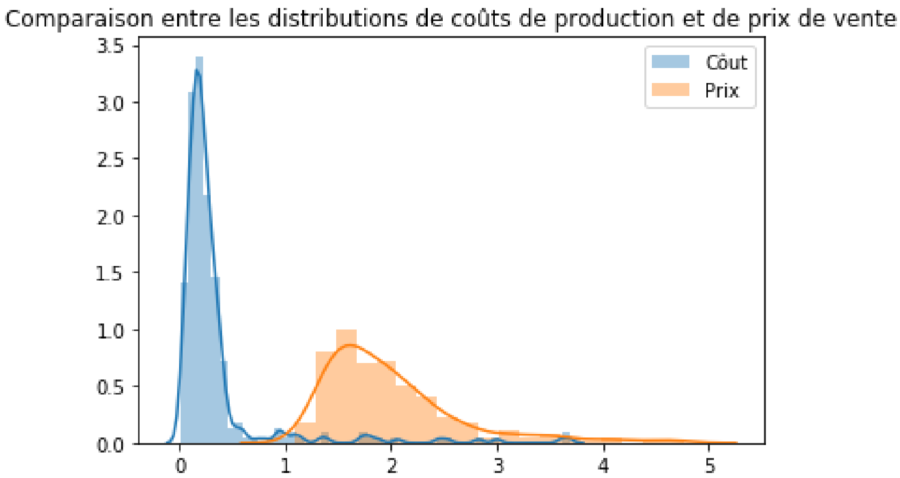

FIGURE 3 – Histogramme des distributions du coût de production d'une boisson sucrée (bleu) et de son prix de vente (orange) pour l'entreprise 1

Les prix de vente d'équilibre varient bien plus que les coûts de production d'une ville à l'autre. En effet, les coûts de production dépendent uniquement du taux de sucre contenu dans la boisson (et d'un terme d'erreur) alors que les prix d'équilibre résultent d'un choix stratégique de la part de l'entreprise qui maximise son profit sur chacun des marchés.

Concentrons nous desormais sur les marges réalisées par l'entreprise 1 dans chaque ville . D'après la figure 2, l'entreprise 1 réalise en moyenne une marge de 85% sur les boissons sucrées vendues dans toutes les villes, et, il en est de même pour l'entreprise 2.

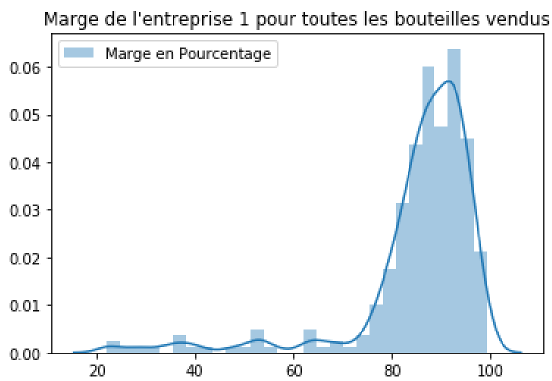

FIGURE 4 – Histogramme de la marge réalisée en pourcentage

Enfin, nous concluerons par une analyse des corrélations de l'ensemble des variables.

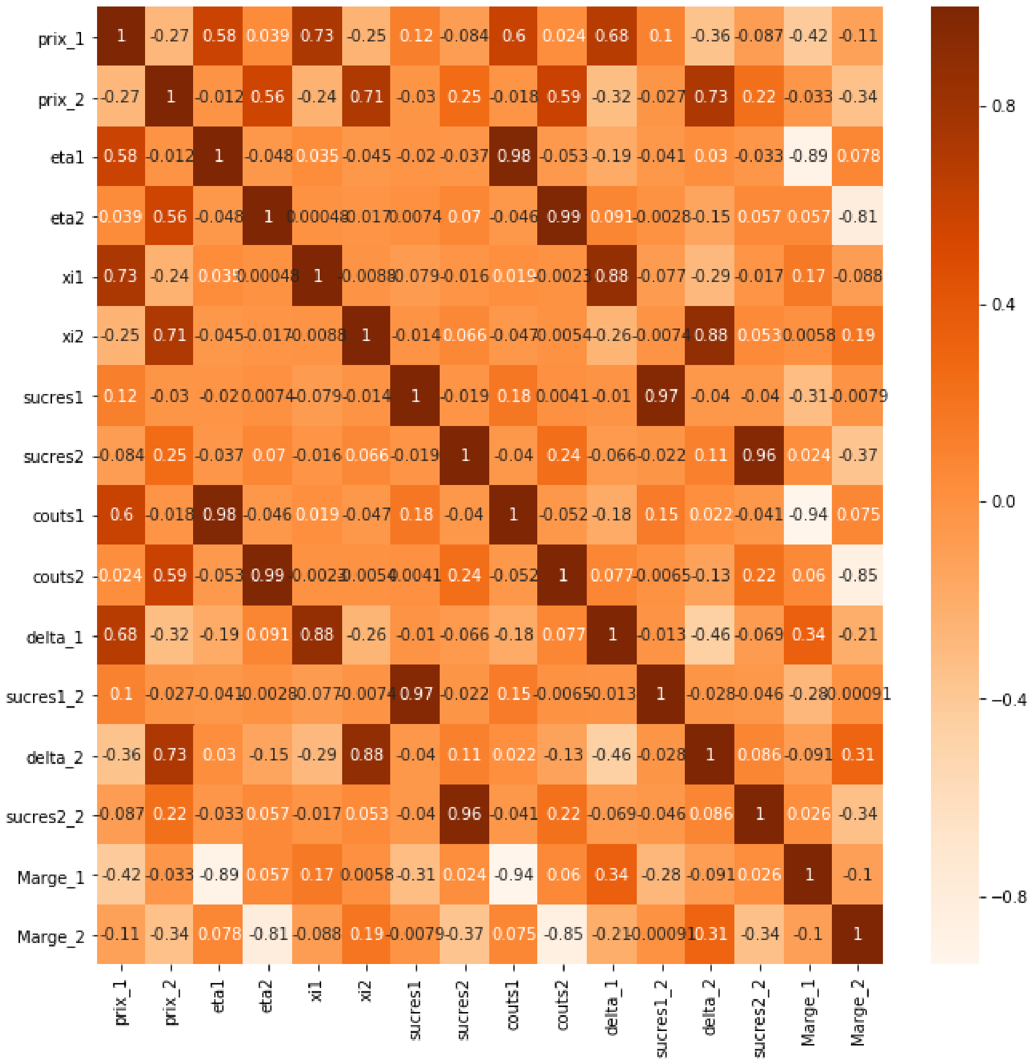

FIGURE 5 – Matrice des Corrélations de toutes les variables de notre base de données

On se concentrera en particulier sur les variables corrélées avec la marge réalisée par l'entreprise lors de la vente de boissons sucrées. Marge et prix de vente sont corrélés negativement, le coefficient de corrélation étant égal à -0,42. Ainsi, alors que nous nous attendions à ce que la fixation d'un prix de vente élevé entraine une augmentation de la marge, il semblerait que les plus grandes marges soient réalisées sur des boissons peu chères. Cette relation provient très certainement d'une forte corrélation entre prix et coûts. La production de boissons très sucrées étant couteuse, mais, ne permettant pas la mise en place de prix suffisamment élevés pour à la fois attirer sufisamment de consommateurs et, réliser une marge interessante.

Le taux de marge est ainsi également corrélé au taux de sucre. Cette corrélation est de -0,31 (quasiment 0 avec le terme de degré 2), ce qui signifie qu'un taux de sucre elevé est associé à une marge plus faible qu'une boisson dont le taux de sucre est modéré.

Les prix de vente de l'entreprise 1 et 2 sont à quant à eux corrélés négativement, ce qui laisse supposer qu'une augmentation des prix de la part d'une des entreprises entraîne une baisse des prix de l'entreprise concurrente.

Question 3

On estime les paramètres du modèle à l'aide des parts de marché des entreprises j dans les villes v s_{vj}. On a montré que :

$$\mathbb{E}(D_{vj}(sucre, prix, \xi_{vj}, \xi_{vj'})) = \frac{ne^{\delta_{vj}}}{\sum_{j'} e^{\delta_{vj'}}}$$

On en déduit la part de marché de l'entreprise j dans la ville v :

$$s_{vj} = \frac{e^{\delta_{vj}}}{\sum_{j'} e^{\delta_{vj'}}}$$

Donc,

$$\log\left(\frac{s_{vj}}{s_{v0}}\right) = \delta_{vj}$$

On observe les parts de marché simulées donc on peut estimer les paramètres α, β_1, β_2 et γ. Comme le prix est endogène, on instrumente le prix de l'entreprise 1 par le taux de sucre et le carré du taux de sucre de la boisson de l'entreprise 2 et vice versa. En effet le prix de l'entreprise j est corrélé aux caractéristiques inobservées de l'entreprise j : si l'entreprise j a une boisson de qualité exceptionnellement bonne dans une ville v, elle est incitée à pratiquer un prix élevé dans cette ville car beaucoup de consommateurs seront intéressés par cette boisson malgré le prix. En revanche, par construction de nos données, les taux de sucres sont indépendants des caractéristiques inobservées, ce qui semble une modélisation raisonable. De plus, le prix de la première entreprise dépend du taux de sucre de l'autre entreprise car ce taux influe sur la qualité de la boisson de l'autre entreprise et donc la première entreprise choisit son prix en fonction pour s'adapter à la concurrence. Le taux de sucre du concurrent semble donc être un bon candidat pour servir de variable instrumentale et on applique le même raisonnement pour le carré de ce taux.

On implémente ainsi, la régression linéaire suivante : $log(\frac{s_{vj}}{sv_0}) = \omega_0 + \omega_1 sucres_{vj} + \omega_2 sucres_{vj}^2 + \omega_3 \widehat{prix_{vj}} + erreur$ avec $j \in \{1, 2\}$ où $\widehat{prix_{vj}}$ est l'estimation du prix du produit j dans la ville v obtenue grâce à la regression du prix sur les variables instrumentales retenues(taux de sucre et taux de sucre au carré du produit de l'autre entreprise) ainsi que sur les variables exogènes. Notons, que la base de données utilisée pour la regression est composée de 600 individus, chaque ligne correspondant au produit vendu par l'une des deux entreprise dans une des 300 villes.

Les résultats de la régression sont présentés en Annexe (voir table A), le R^2 ajusté est égal à 0,927 ; ainsi, le taux de sucre, ainsi que le taux de sucre au carré et le prix expliquent plus de 90 % de la variation de la variable dépendante. Nous obtenons les estimations suivantes $\widehat{\omega_1} = 2,08$; $\widehat{\omega_2} = -0,97$ et $\widehat{\omega_3} = -1,1$. La regression IV permet d'obtenir de très bonnes approximations des paramètres utilisés pour la simulation des données : β_1=2, β_2=1, γ=1. Notons que, l'utilité présentée page 1 dépend par construction négativement du taux de sucre au carré ainsi que du prix, ceci explique les différences de signes entre paramètres et paramètres estimés.
On note que l'estimateur de α vaut $0,15$ alors que le vrai paramètre a été fixé à 1. Cependant, une fonction d'utilité est définie à une constante près alors il n'est pas gênant de ne pas retrouver le vrai paramètre. Il est même normal qu'on ne retrouve pas ce paramètre car on aurait pu le fixer à une autre valeur et on aurait obtenu les mêmes simulations. On remarque d'ailleurs que cet estimateur a une grande p-valeur et un grand intervalle de confiance ce qui illustre l'incapacité du modèle à calculer le paramètre.

Question 4

On recommence l'estimation des paramètres avec 600 et 1000 villes. Les estimations obtenues sont les suivantes :
- R^2_{ajuste}=0,935 ; $\widehat{\omega_1} = 2,20$; $\widehat{\omega_2} = -1,07$ et $\widehat{\omega_3} = -1,20$ lorsque nous simulons les données pour 600 villes (voir table B en Annexe)
-R^2_{ajuste}=0,952 ; $\widehat{\omega_1} = 2,02$; $\widehat{\omega_2} = -0,9$ et $\widehat{\omega_3} = -1,16$ lorsque nous simulons les données pour 1000 villes (voir table C en Annexe)

L'augmentation du nombre de villes, et par conséquent du nombre d'observations utilisées dans la régression doit en raison de la loi forte des grands nombres entrainer une meilleure estimation des paramètres. Si les qualités de prédiction et d'estimation des paramètres semblent s'améliorer , nous n'observons pas pour autant de nette convergence des paramètres estimés vers leur vraie valeur. Tous les paramètres estimés étant très proches de leur vraie valeur, et ce quel que soit le nombre de villes incluses dans la régression, il est fort probable que la convergence ait lieu pour un nombre de ville inférieur à 300, les différences d'estimation entre les regressions pour 300,600 et 1000 villes resultants alors uniquement des bruits simulés.

Question 5

On estime dans un premier temps les paramètres. On trouve alors α, β_1, β_2 et γ ainsi que les erreurs empiriques ξ. Ensuite il faut estimer les coûts. D'après les équations (1) et (2) définissant l'équilibre de Nash dans la ville v on a :

$$\begin{aligned}
1-\gamma(p_{v1}-C_{v1})+\gamma(p_{v1}-C_{v1})\frac{e^{\delta_{v1}}}{\sum_{j'}e^{\delta_{vj'}}} &= 0\\
1-\gamma(p_{v1}-C_{v1})+\gamma(p_{v1}-C_{v1})s_{v1} &= 0\\
1-\gamma p_{v1}+\gamma p_{v1}s_{vj} &= \gamma C_{v1}(s_{v1}-1)\\
C_{v1} &= p_{v1}-\frac{1}{\gamma(1-s_{v1})}
\end{aligned}$$

On peut donc estimer les coûts à l'aide de l'estimation de γ.
Par ailleurs, on trouve aisément la boisson conservée car il s'agit de celle dont le taux de sucre est le plus faible.
On ajoute un "'" aux variables pour indiquer les variables correspondant à l'entreprise fusionnée.
Les calculs de demande faits dans la partie 1. sont toujours valables, avec une entreprise de moins. Ainsi, on sait que l'espérance de demande de l'entreprise fusionnée est :

$$\mathbb{E}(D'_v(sucre', prix', \xi'_v)) = \quad = \frac{ne^{\delta'_v}}{1+e^{\delta'_v}}$$

On maximise alors le profit de l'entreprise fusionnée, ou plutôt le log du profit divisé par n ce qui revient au même.

$$\begin{aligned}
\frac{\partial log(\Pi'_v/n)}{\partial p'_v} &= \frac{\partial log(e^{\delta'_v})}{\partial p'_v}-\frac{\partial log(1+e^{\delta'_v})}{\partial p'_v}+\frac{\partial log(p'_v-C'_v)}{\partial p'_v}\\
&= -\gamma+\frac{\gamma e^{\delta'_v}}{1+e^{\delta'_v}}+\frac{1}{p'_v-C'_v}
\end{aligned}$$

La condition de premier ordre donne alors

$$\begin{aligned}
0 &= -\gamma+\frac{\gamma e^{\delta'_v}}{1+e^{\delta'_v}}+\frac{1}{p'_v-C'_v}\\
0 &= 1-\gamma(p'_v-C'_v)+\gamma(p'_v-C'_v)\frac{e^{\delta'_v}}{1+e^{\delta'_v}}
\end{aligned}$$

On résout cette équation à une inconnue numériquement ce qui donne un prix d'équilibre pour chaque ville.
On compare ensuite ce prix au prix de la boisson la moins sucrée avant la fusion. Le prix moyen avant fusion de la boisson la moins sucrée est de $1,30$ euros et le prix après fusion est de $1,88$ euros, soit une hausse de 45%.
Afin de déterminer si le consommateur est perdant ou non dans la fusion, on calcule d'abord l'utilité moyenne estimée des consommateurs avant la fusion à l'aide des paramètres estimés. Sans prendre en compte les termes d'erreurs individuels ϵ_{ivj} qui sont inconnus on aurait une utilité de l'eau constante égale à 0. Or, cela conduit à des incohérences car on observe des consommateurs qui consomment autre chose que de l'eau du robinet, ce qui leur apporte une utilité estimée inférieure à 0. Similairement, on observe des consommateurs buvant de l'eau du robinet alors que l'utilité estimée à boire une autre boisson est supérieure à 0. C'est pourquoi on simule une erreur conditionnelle pour l'eau du robinet. Quand le consommateur boit de l'eau du robinet, on simule une erreur ϵ telle que $\epsilon \sim GEV(0,1,0)|\epsilon > max(\hat{u_1},\hat{u_2})$ et lorsqu'il ne boit pas de l'eau on simule une erreur ϵ telle que $\epsilon \sim GEV(0,1,0)|\epsilon < max(\hat{u_1},\hat{u_2})$ avec $\hat{u_j}$ l'utilité estimée à consommer le bien j. Pour faire cette simulation, on simule suivant les cas une loi uniforme sur $[0, F(max(\hat{u_1},\hat{u_2}))]$ ou $[F(max(\hat{u_1},\hat{u_2})),1]$ et on applique ensuite l'inverse de la fonction de répartition de la loi extreme value.

On calcule ensuite l'utilité moyenne après la fusion à l'aide de l'estimation de la fonction d'utilité, des nouveaux prix et de ces erreurs calculées pour l'eau du robinet. On compare ces deux utilités moyennes pour déterminer si le consommateur est perdant. On trouve ainsi une utilité moyenne de $0,11$ avant la fusion et de $-0,31$ après la fusion ce qui est défavorable au consommateur. On recommence ainsi avec différents facteurs de réduction du coût pour trouver la réduction nécessaire pour que le consommateur ne soit pas perdant : on recalcule le prix d'équilibre et les boissons consommées et enfin l'utilité des consommateurs. On trouve finalement qu'une réduction de 10% rend la fusion profitable aux consommateurs avec une utilité moyenne de $0,47$.

Appendices

A Régression pour 300 villes

```
IV-2SLS Estimation Summary
        ==============================================================================
          Dep. Variable:                      l1   R-squared:                      0.9272
          Estimator:                     IV-2SLS   Adj. R-squared:                 0.9269
          No. Observations:                  600   F-statistic:                    206.64
          P-value (F-stat)                0.0000   Distribution:                  chi2(3)
          Cov. Estimator:                 robust

                                   Parameter Estimates

          ==================================================================================
                         Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
          ----------------------------------------------------------------------------------
          Intercept         0.1524     0.4061     0.3753     0.7074     -0.6436      0.9484
          sucres1           2.0809     0.4425     4.7031     0.0000      1.2137      2.9481
          sucres1\_carre    -0.9675     0.3251    -2.9763     0.0029     -1.6046     -0.3304
          prix\_1           -1.1052     0.2589    -4.2694     0.0000     -1.6126     -0.5978
          ==================================================================================

          Endogenous: prix\_1
          Instruments: sucres2, sucres2\_carre
          Robust Covariance (Heteroskedastic)
          Debiased: False
```

B Régression pour 600 villes

```
IV-2SLS Estimation Summary
          ==============================================================================
          Dep. Variable:                     l1   R-squared:                      0.9347
          Estimator:                    IV-2SLS   Adj. R-squared:                 0.9345
          No. Observations:                1200   F-statistic:                    497.96
          P-value (F-stat)               0.0000   Distribution:                  chi2(3)
          Cov. Estimator:                robust


                                      Parameter Estimates
          =================================================================================
                         Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
          ---------------------------------------------------------------------------------
          Intercept         0.3401     0.3497     0.9724     0.3309     -0.3454      1.0256
          sucres1           2.2012     0.2114     10.413     0.0000      1.7869      2.6155
          sucres1\_carre   -1.0738     0.1165    -9.2159     0.0000     -1.3022     -0.8455
          prix\_1          -1.2047     0.2108    -5.7147     0.0000     -1.6179     -0.7915
          =================================================================================

          Endogenous: prix\_1
          Instruments: sucres2, sucres2\_carre
          Robust Covariance (Heteroskedastic)
          Debiased: False
          IVResults, id: 0x2234ddcbeb8
```

C Régression pour 1000 villes

```
IV-2SLS Estimation Summary
          ==============================================================================
          Dep. Variable:                     l1   R-squared:                      0.9520
          Estimator:                    IV-2SLS   Adj. R-squared:                 0.9520
          No. Observations:                2000   F-statistic:                    771.35
          Distribution:                 chi2(3)   P-value (F-stat)                0.0000
          Cov. Estimator:                robust

                                Parameter Estimates
          ==================================================================================
                         Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
          ----------------------------------------------------------------------------------
          Intercept         0.3002     0.3859     0.7781     0.4365     -0.4561      1.0565
          sucres1           2.0271     0.2093     9.6830     0.0000      1.6168      2.4374
          sucres1\_carre    -0.9090     0.0978    -9.2992     0.0000     -1.1006     -0.7174
          prix\_1           -1.1689     0.2299    -5.0853     0.0000     -1.6194     -0.7184
          ==================================================================================

          Endogenous: prix\_1
          Instruments: sucres2, sucres2\_carre
          Robust Covariance (Heteroskedastic)
          Debiased: False
          IVResults, id: 0x2235f0e9550
```