Statistiques descriptives des variables clé

A. Blanc, N. Gusarov, S. Picon

Question de recherche.

Dans cette étude, nous nous intéressons à l'effet de la quantité de pesticides utilisé sur l'équilibre du marché des vins de table.

Modèle économétrique.

Dans cette étude, nous nous intéressons à l'effet de la quantité de pesticides utilisé sur l'équilibre du marché des vins de table ? Le modèle économique : Formalisant notre modèle théorique, nous posons, que la demande agrégé de vin a la forme suivante :

$$Qd_t = \alpha_d + \beta_d Pd_t + \gamma_d Z_t \tag{1}$$

Avec Z étant l'ensemble des variables ayant l'influence sur la demande du vin, dans le cas le plus simple nous n'utilisons que les revenus (c'est une des variables les plus utilisées dans des études empiriques sur le marché du vin).

L'offre agrégé pour toute la France est donnée par l'équation suivante :

$$Qo_t = \sum_{i=1}^{N} q_{i,t} \tag{2}$$

Avec :
— Qd : la quantité demandée de vin en hectolitre
— Pd : Le prix du vin moyen en euros/hectolitres
— Z : Le revenu disponible brut déflaté
Ou $i \in \{1, ..., N\}$ sont des régions, chacun ayant sa propre fonction de production et d'offre unique :

$$q_{i,t} = a_i + b_i Po_t + c_i X_{i,t} \tag{3}$$

Avec X étant un vecteur des variables explicatives influençant la production (dans le cas le plus simple nous ne prenons en compte que les quantités des pesticides utilisées). Plus précisément :
— q_i : la quantité de vin en hectolitre dans chaque département
— Po : prix moyen en hectolitre
— X : la quantité de pesticide
— Y : La superficie en hectare
Nous pouvons réécrire l'équation de l'offre sous la forme :

$$Qo_{i,t} = \sum_{i=1}^{N}(a_i + b_i Po_t + c_i X_{i,t}) = \sum_{i=1}^{N} a_i + \sum_{i=1}^{N} b_i Po_t + \sum_{i=1}^{N}(c_i X_{i,t}) \tag{4}$$

Nous obtenons enfin un système de $N+2$ équations :

$$\begin{aligned}
Qd_t &= \alpha_d + \beta_d Pd_t + \gamma_d Z_t \\
Qo_t &= \sum_{i=1}^{N} q_{i,t} \\
q_1 &= a_1 + b_1 Po_t + c_1 X_{1,t} \\
&\vdots \\
q_N &= a_N + b_N Po_t + c_N X_{N,t}
\end{aligned}$$

A l'équilibre nous ayons $Po_t = Pd_t = P_t$ et $Qo_t = Qd_t = Q_t$.

Modèle économétrique.

N'ayant les valeurs que pour l'équilibre, nous pouvons réécrire notre modèle comme :

$$\begin{aligned}
Q_t &= \alpha_d + \beta_d P_t + \gamma_d Z_t + \epsilon_t \\
Q_t &= \sum_{i=1}^{N} q_{i,t} \\
q_1 &= a_1 + b_1 P_t + c_1 X_{1,t} + u_{1,t} \\
&\vdots \\
q_N &= a_N + b_N P_t + c_N X_{N,t} + u_{N,t}
\end{aligned}$$

Ce qui nous donne :

$$\alpha_d + \beta_d P_t + \gamma_d Z_t + \epsilon_t = \sum_{i=1}^{N} a_i + \sum_{i=1}^{N} b_i P_t + \sum_{i=1}^{N} c_i X_{i,t} + \sum_{i=1}^{N} u_{i,t} \tag{5}$$

Le problème apparaisse au niveau du terme $\sum_{i=1}^{N}(c_i X_{i,t})$. Si $cor(c_i X_{i,t}) \neq 0$ on a autant des termes c_i dans notre équation de départ que le nombre des départements étudié N. C'est à nous obtiendrons une équation structurelle du type :

$$P_t = \frac{\sum_{i=1}^{N} a_i - \alpha_d}{\beta_d - \sum_{i=1}^{N} b_i} + \frac{\sum_{i=1}^{N} c_i X_{i,t}}{\beta_d - \sum_{i=1}^{N} b_i} + \frac{-\gamma_d}{\beta_d - \sum_{i=1}^{N} b_i} Z_t + \frac{\sum_{i=1}^{N} u_{i,t} - \epsilon_t}{\beta_d - \sum_{i=1}^{N} b_i} \tag{6}$$

Cela nous risque de poser des problèmes lors d'estimation et dérivation des coefficients des équation de départ. Quand même, si nous posons que c_i n'est pas corrélé avec $X_{i,t}$ et $cor(c_i X_{i,t}) = 0$, nous pouvons supposer que :

$$\sum_{i=1}^{N} c_i X_{i,t} = \frac{1}{N} \sum_{i=1}^{N} c_i \frac{1}{N} \sum_{i=1}^{N} X_{i,t} \tag{7}$$

Ce qui revienne de l'idée que $E(XY) = E(X)E(Y)$ si $cor(X,Y) \neq 0$. Dans ce cas, l'équation structurelle s'écrit comme :

$$P_t = \frac{\sum_{i=1}^{N} a_i - \alpha_d}{\beta_d - \sum_{i=1}^{N} b_i} + \frac{\sum_{i=1}^{N} c_i}{\beta_d - \sum_{i=1}^{N} b_i} \frac{1}{N} \sum_{i=1}^{N} X_{i,t} + \frac{-\gamma_d}{\beta_d - \sum_{i=1}^{N} b_i} Z_t + \frac{\sum_{i=1}^{N} u_{i,t} - \epsilon_t}{\beta_d - \sum_{i=1}^{N} b_i} \tag{8}$$

Ce qu'on peut réécrire comme :

$$P_t = \pi_1 + \pi_2 \frac{1}{N} \sum_{i=1}^{N} X_{i,t} + \pi_3 Z_t + v_t \tag{9}$$

Respectivement on peut dériver équation structurelle pour Q :

$$\begin{aligned}
Q_t = (\alpha_d + \beta_d \frac{\sum_{i=1}^{N} a_i - \alpha_d}{\beta_d - \sum_{i=1}^{N} b_i}) + (\beta_d \frac{\sum_{i=1}^{N} c_i}{\beta_d - \sum_{i=1}^{N} b_i}) \frac{1}{N} \sum_{i=1}^{N} X_{i,t} + \\
(\gamma_d + \beta_d \frac{-\gamma_d}{\beta_d - \sum_{i=1}^{N} b_i}) Z_t + (\epsilon_t + \beta_d \frac{\sum_{i=1}^{N} u_{i,t} - \epsilon_t}{\beta_d - \sum_{i=1}^{N} b_i})
\end{aligned} \tag{10}$$

Ce qui se réécrit sous forme :

$$Q_t = \theta_1 + \theta_2 \frac{1}{N} \sum_{i=1}^{N} X_{i,t} + \theta_3 Z_t + w_t \tag{11}$$

Le reste est estimé comme :

$$\begin{multline} q_{i,t} = (a_i + b_i \frac{\sum_{i=1}^{N} a_i - \alpha_d}{\beta_d - \sum_{i=1}^{N} b_i}) + (c_i + b_i \frac{\sum_{i=1}^{N} c_i}{\beta_d - \sum_{i=1}^{N} b_i}) X_{i,t} + \\ (b_i \frac{-\gamma_d}{\beta_d - \sum_{i=1}^{N} b_i}) Z_t + (u_{i,t} + b_i \frac{\sum_{i=1}^{N} u_{i,t} - \epsilon_t}{\beta_d - \sum_{i=1}^{N} b_i}) \end{multline} \tag{12}$$

En simplifiant on le réécrit :

$$q_{i,t} = \psi_{i,1} + \psi_{i,2} X_{i,t} + \psi_{i,3} Z_t + e_{i,t} \tag{13}$$

Ce qui avec i le numéro de département, nous donne suffisamment des différences entre les coefficients pour identifier les paramètres des équations de départ.

Avec les deux premières équations structurelles on obtient les coefficients pour la première équation de départ, qui décrit la demande agrégé :

$$\alpha_d = \theta_1 - \frac{\pi_1 \theta_2}{\pi_2} \tag{14}$$

$$\beta_d = \frac{\theta_2}{\pi_2} \tag{15}$$

$$\gamma_d = \frac{\theta_2 \pi_3}{\pi_2} - \theta_3 \tag{16}$$

Ainsi bien que pour celle, qui décrit l'offre agrégé :

$$\sum_{i=1}^{N} a_i = \theta_1 - \frac{\pi_1 \theta_3}{\pi_3} \tag{17}$$

$$\sum_{i=1}^{N} b_i = \frac{\theta_3}{\pi_3} \tag{18}$$

$$\sum_{i=1}^{N} c_i = \theta_2 - \frac{\theta_3 \pi_2}{\pi_3} \tag{19}$$

Les coefficients uniques pour les régions a_i, b_i et c_i sont a identifier séparément avec les estimateurs du reste des équations. On les obtient d'une manière suivante :

$$a_i = \psi_{i,1} - \frac{\psi_{i,3} \pi_1}{\pi_3} \tag{20}$$

$$b_i = \frac{\psi_{i,3}}{\pi_3} \tag{21}$$

$$c_i = \psi_{i,2} - \frac{\psi_{i,3} \pi_2}{\pi_3} \tag{22}$$

En ce qui concerne la variance des estimateurs obtenus, il reste encore à vérifier.

Estimations simples

On commence par construire le modèle simple, afin de voir les relations de base sur le marché du vin :

	Dependent variable :
	qi
p	0.541*
	(0.278)
s	1.074***
	(0.019)
qki	0.141***
	(0.028)
Constant	−1.001
	(1.212)
Observations	414
R^2	0.915
Adjusted R^2	0.915
Residual Std. Error	0.652 (df = 410)
F Statistic	1,477.087*** (df = 3 ; 410)
Note :	*p<0.1 ; **p<0.05 ; ***p<0.01

Les estimations pour le modèle sous supposition que $cor(c_i, X_i) \neq 0$:

	Dependent variable :
	Y
X[,]p	-0.046
	(0.068)
X[,]r	0.038
	(0.111)
X[,]V3	-0.180
	(0.911)
X[,]V4	0.149
	(0.380)
X[,]V5	-1.458
	(0.911)
X[,]V6	0.393
	(0.380)
X[,]V7	0.363
	(0.911)
X[,]V8	-0.211
	(0.380)
X[,]V9	0.230
	(0.911)
X[,]V10	-0.093
	(0.380)
X[,]V11	-0.158
	(0.911)
X[,]V12	-0.151
	(0.380)
X[,]V13	0.643
	(0.911)
X[,]V14	-0.534
	(0.380)
X[,]V15	-0.725
	(0.797)
X[,]V16	0.216^{*}
	(0.113)
X[,]V17	0.417
	(0.797)
X[,]V18	-0.255^{**}
	(0.113)
X[,]V19	0.793
	(0.797)
X[,]V20	-0.248^{**}
	(0.113)

	Dependent variable :
	Y
X[,]V21	-0.190
	(0.797)
X[,]V22	0.316^{***}
	(0.113)
X[,]V23	2.227^{***}
	(0.797)
X[,]V24	-0.338^{***}
	(0.113)
X[,]V25	0.355
	(0.797)
X[,]V26	0.031
	(0.113)
X[,]V27	0.540
	(0.335)
X[,]V28	-0.040
	(0.119)
X[,]V29	0.104
	(0.335)
X[,]V30	0.192
	(0.119)
X[,]V31	0.158
	(0.335)
X[,]V32	-0.149
	(0.119)
X[,]V33	0.392
	(0.335)
X[,]V34	-0.241^{**}
	(0.119)
X[,]V35	-0.258
	(0.335)
X[,]V36	0.010
	(0.119)
X[,]V37	0.065
	(0.335)
X[,]V38	0.094
	(0.119)
X[,]V39	-0.994^{*}
	(0.557)
X[,]V40	0.174
	(0.204)

	Dependent variable :
	Y
X[,]V41	−0.753
	(0.557)
X[,]V42	0.322
	(0.204)
X[,]V43	−0.048
	(0.557)
X[,]V44	−0.014
	(0.204)
X[,]V45	−0.402
	(0.557)
X[,]V46	0.249
	(0.204)
X[,]V47	0.117
	(0.557)
X[,]V48	0.281
	(0.204)
X[,]V49	−0.483
	(0.557)
X[,]V50	0.177
	(0.204)
X[,]V51	−0.531
	(0.371)
X[,]V52	0.502***
	(0.121)
X[,]V53	0.337
	(0.371)
X[,]V54	0.176
	(0.121)
X[,]V55	0.408
	(0.371)
X[,]V56	0.272**
	(0.121)
X[,]V57	0.114
	(0.371)
X[,]V58	−0.041
	(0.121)
X[,]V59	−0.244
	(0.371)
X[,]V60	0.205*
	(0.121)

	Dependent variable :
	Y
X[,]V61	−0.007
	(0.371)
X[,]V62	0.026
	(0.121)
X[,]V63	−0.069
	(0.173)
X[,]V64	0.008
	(0.075)
X[,]V65	−0.379**
	(0.173)
X[,]V66	0.211***
	(0.075)
X[,]V67	0.012
	(0.173)
X[,]V68	0.149**
	(0.075)
X[,]V69	−0.027
	(0.173)
X[,]V70	0.023
	(0.075)
X[,]V71	0.475***
	(0.173)
X[,]V72	−0.353***
	(0.075)
X[,]V73	−0.042
	(0.173)
X[,]V74	0.280***
	(0.075)
X[,]V75	−0.528
	(0.829)
X[,]V76	0.082
	(0.434)
X[,]V77	1.643**
	(0.829)
X[,]V78	−0.746*
	(0.434)
X[,]V79	0.061
	(0.829)
X[,]V80	0.138
	(0.434)

	Dependent variable :
	Y
X[,]V81	−0.196
	(0.829)
X[,]V82	0.287
	(0.434)
X[,]V83	−1.013
	(0.829)
X[,]V84	0.394
	(0.434)
X[,]V85	−1.047
	(0.829)
X[,]V86	0.192
	(0.434)
X[,]V87	−0.166
	(0.156)
X[,]V88	−0.064
	(0.078)
X[,]V89	−0.270*
	(0.156)
X[,]V90	−0.147*
	(0.078)
X[,]V91	−0.426***
	(0.156)
X[,]V92	0.013
	(0.078)
X[,]V93	0.076
	(0.156)
X[,]V94	−0.166**
	(0.078)
X[,]V95	−0.181
	(0.156)
X[,]V96	−0.431***
	(0.078)
X[,]V97	0.183
	(0.156)
X[,]V98	−0.080
	(0.078)
X[,]V99	−0.049
	(0.248)
X[,]V100	−0.008
	(0.076)

	Dependent variable :
	Y
X[,]V101	0.049
	(0.248)
X[,]V102	−0.101
	(0.076)
X[,]V103	−0.186
	(0.248)
X[,]V104	0.172**
	(0.076)
X[,]V105	−0.075
	(0.248)
X[,]V106	−0.064
	(0.076)
X[,]V107	0.033
	(0.248)
X[,]V108	−0.112
	(0.076)
X[,]V109	0.208
	(0.248)
X[,]V110	0.018
	(0.076)
X[,]V111	1.021
	(1.385)
X[,]V112	−0.698
	(0.786)
X[,]V113	−0.577
	(1.385)
X[,]V114	0.304
	(0.786)
X[,]V115	0.959
	(1.385)
X[,]V116	−0.631
	(0.786)
X[,]V117	−0.326
	(1.385)
X[,]V118	0.190
	(0.786)
X[,]V119	0.485
	(1.385)
X[,]V120	−0.260
	(0.786)

	Dependent variable :
	Y
X[,]V121	1.749
	(1.385)
X[,]V122	−0.978
	(0.786)
X[,]V123	−0.139
	(0.246)
X[,]V124	0.275**
	(0.118)
X[,]V125	0.074
	(0.246)
X[,]V126	0.060
	(0.118)
X[,]V127	0.021
	(0.246)
X[,]V128	0.114
	(0.118)
X[,]V129	0.153
	(0.246)
X[,]V130	0.156
	(0.118)
X[,]V131	0.307
	(0.246)
X[,]V132	−0.036
	(0.118)
X[,]V133	0.182
	(0.246)
X[,]V134	−0.017
	(0.118)
X[,]V135	1.493
	(1.184)
X[,]V136	−1.063**
	(0.499)
X[,]V137	0.932
	(1.184)
X[,]V138	−0.590
	(0.499)
X[,]V139	0.297
	(1.184)
X[,]V140	−0.218
	(0.499)
Constant	8.385***
	(0.026)

	Dependent variable :
	Y
Observations	420
R^2	0.965
Adjusted R^2	0.947
Residual Std. Error	0.508 (df = 279)
F Statistic	54.908*** (df = 140 ; 279)
Note :	*p<0.1 ; **p<0.05 ; ***p<0.01