

Etude des effets des pesticides dans la production des vins de table

Analyse empirique des marchés

A. Blanc, N. Gusarov, S. Picon

19/12/2019

1 Introduction

Quel est l'effet de l'utilisation des pesticides sur le marché des vins simples ?

Dans cette étude, nous chercherons à étudier l'équilibre sur le marché du vin.

2 Les pesticides

- Présentation du problème des pesticides
- Etat actuel
- Comment baisser l'utilisation de pesticides

2.1 Présentation du problème des pesticides

- Source de nombreux débats sur la santé et l'environnement.
- Le rôle actuel :
 - Moyen de protection contre les aléas climatiques ;
 - Outil pour la réservation du rendement.
- Plusieurs mesures mises en places pour réduire leurs usages :
 - des interdiction des produits les plus toxiques ;
 - l'instauration d'une taxe, payée par les agriculteurs (Butault et al, 2011).
- Malgres les efforts l'utilisation perdure :
 - Hausse des ventes de produits phytosanitaires ;
 - Augmentation des doses utilisés (+12% en 2014-2016) ;

2.2 Etat actuel

Contrairement aux attentes des autorités aucune baisse de l'utilisation de pesticides :

- Le nombre de doses unité augmente de 23% entre 2008 et 2017 ;
- Le nombre de substances actives utilisées a augmenté de 15% entre 2011 et 2017 ;
- Une baisse des produits les plus dangereux de 6%, en 2017 (Moghaddam et al, 2019) ;
- Les grandes cultures (blés, etc...) sont les premières utilisatrices de pesticides 67.4% ;
- Les vignes sont les deuxièmes 14.4% (Butault et al, 2011).

2.3 Comment baisser l'utilisation de pesticides

Les méthodes contemporaines visant à baisser l'utilisation des pesticides sont :

- Le changement de mode de culture :

- agriculture biologique ;
- agriculture raisonnée ;
- La diversification des cultures, ce qui est impossible pour la vigne (Moghaddam et al, 2019).

3 Le marché du vin français

La France est l'un des principaux producteurs et vendeurs de vin dans le monde. En effet, la France représente 10% de la surface de vigne dans le monde. La surface de vigne française se répartit dans 65 des 95 départements de la métropole. En France, il y a plus de 750000 hectares de vignes qui sont exploitées en 2018. Ainsi, en France, une exploitation agricole sur cinq est une exploitation viticole. Cela représente 85000 exploitations. La production de vins en France, représentait 4,6 milliard de litres. Cela représentait plus de 17% de la production totale de vin. En volume de production la France se place donc en deuxième position derrière le volume de production de l'Italie. 3% de la surface agricole est consacrée à la production de vin. Néanmoins, le vin représente 15% de la production agricole en valeur. Du côté du consommateur, la France est le deuxième pays consommateur de vin derrière les Etats Unis. En effet, la consommation de vin en France représentait plus de 3,5 milliards de bouteille, en 2018. Néanmoins, on remarque une baisse de la consommation Française depuis une trentaine d'année.

3.1 Le problème d'hétérogénéité

Il existe une forte hétérogénéité entre les différents labels mais aussi à l'intérieur de ces labels.

Dans le commerce du vin, il est courant de diviser les vins en deux grandes classes en fonction de leurs prix (Cembalo et al., 2014) :

- les vins de qualité inférieure, les moins chers avec les caractéristiques de qualité de base ;
- les vins de qualité supérieure plus chers, dotés de caractéristiques qualitatives complexes et d'une image de grande valeur.

De plus, pour les vins français, selon Steiner (2014), le système européen de classification des "*vins de qualité produits dans certaines régions*" (VQPRD) contient à la fois des vins AOC et des "*vins de haute qualité provenant d'un vignoble régional agréé*" (VDQS). Les vins de cépage appartiennent à la catégorie des vins autres que VQPRD, qui comprend les **vins de table** et les **vins de pays**.

En tenant compte des spécificités du marché du vin français, nous utilisons la méthodologie du ministère d'agriculture et divisons le marché en deux parties :

- La gamme haute (les vins IGP, vendus dans des magasins spécifiques) ;
- La gamme basse (les vins non IGP, vendus en grands surfaces).

La première partie est soumise à des règlements spécifiques : limitations des quantités produites, origine contrôlé, un caractère de la demande spécifique. La deuxième, c'est-à-dire le marché des vins moins chers, est aussi complexe. Les produits classés dans cette catégorie sont susceptibles d'avoir un certain degré d'hétérogénéité, comme cela a été montré par Cembalo et al. (2014).

3.2 Les vins de table

Ces vins sans indication géographique (sans IG) ont vu leurs transactions augmenter en volume pour toutes les couleurs. Ainsi, on remarque que pour les vins rouges les transactions ont augmenté

de 10%, pour les rosées la hausse représentait 52%, pour les vins blancs les volumes de transactions ont presque été doublé. Néanmoins on remarque une baisse des cours des vins sans indication géographique. En effet, on remarque que les prix moyens pour les vins rouges et rosées sans indication géographique baisse de 3%. Le prix moyen des vins blancs baisse quand à eux de 12%, pour la campagne 2019/2020. Sur les deux mois de campagne, les échanges de Vin sans indication géographique est de 142 milliers d’hectolitres. Cela correspond à une hausse de 39% par rapport à la campagne précédente. Les ventes représentent 92 milliers d’hectolitres. La tendance sur le marché des vins sans indication géographique s’explique par une forte hausse des vins blancs. En effet, ceux-ci connaissent une hausse de près de 28 milliers d’hectolitres, soit une hausse de 232% vis-à-vis de la campagne de 2018-2019. Les vins rosés connaissent également une hausse. Néanmoins, celle-ci reste modeste puisque les ventes augmentaient de 61% par rapport à la campagne 2018/2019. Néanmoins, les ventes de vins rouges ont légèrement baissé. Le cours des Vins sans indication géographique baisse par rapport à la campagne précédente.

Lors de la campagne 2018/2019, les ventes de vins en grande distribution sont en baisse. Cela peut s’expliquer par une hausse des prix moyens. Les ventes de vins représentent 8,7 millions d’hectolitres et un chiffre d’affaires de 4,1 milliards d’euros avec un prix moyen de 4,73 euro/litre. La baisse de la consommation de vins rouges s’aggrave avec une baisse de 8% par rapport à la campagne de 2017/2018. Les vins blancs connaissent aussi une faible baisse de 1,2% en volume par rapport à la consommation de la campagne précédente. Pour finir, les ventes de vins rosés ont baissé lors de la campagne 2018/2019. En effet, on enregistre une baisse de 3,9% en volume par rapport à la campagne 2017/2018. La consommation de vin sans indication géographique est de 6% en volume contre 3% en valeur. Les ventes de vins sans indications géographiques sont en légère hausse dans la campagne 2018/2019 par rapport à la campagne 2017/2018.

Dans notre étude, nous traitons uniquement les vins simples (non IGP). La situation sur ce marché est sensée influencer l’utilisation des pesticides, car les volumes de productions sont plus significatives que pour le marché des vins IGP.

Suivant le raisonnement des chercheurs Cembalo et al. (2014), dans une catégorie de vin avec une fourchette de prix étroite, il existe une homogénéité presque parfaite due à des vins ayant des attributs intrinsèques simples, une complexité de qualité médiocre et donc une différenciation peu marquée.

Cela nous permet d’analyser le marché par département est non par des marques/produits.

3.3 Utilisation des pesticides dans la viticulture

Les phytosanitaires sont très utilisés dans les cultures comme la viticulture. Il s’agit donc d’un intrant important pour la production de vin. Ainsi, la viticulture utilisait 15% de produit phytosanitaire. La pression sanitaire varie selon les productions. Elle est forte en viticulture. De la même façon, la pression phytosanitaire varie selon les régions. Ainsi, pour la vigne l’IFT varie de 7 en Provence à 22 en Champagne.

4 Le Modèle théorique

4.1 Les hypothèses théoriques

Comme proposé dans la littérature, notre étude sur les vins non coûteux (non IGP) est effectué au niveau du pays Cembalo et al. (2014) pour deux raisons :

- Les prix de vente moyens des marchés sont différents en raison des droits de douane à l'importation et des taxes à la consommation différents
- La perception des produits de consommation varie d'un pays à l'autre

La plupart des bouteilles achetées sont achetées dans la grande distribution. Néanmoins, dans un souci de simplicité nous estimerons que les consommateurs achètent leurs bouteilles directement auprès du viticulteur. Donc nous supprimerons tous les intermédiaires entre le producteur et le marché final.

Quand aux exportations et les importations, n'ayant pas la possibilité contrôler le montant des vins non IGP exportés/importés, nous laissons ces effets au terme d'erreur. Nous ignorons les interactions internationales complètement.

Pour conclure, nos suppositions au niveau du marché des vins sont les suivantes :

- La demande pour les vins simples est unique pour toute la France. On n'observe pas les quantités consommées par départements, mais pour tout le pays, avec un prix unique.
- La production du vin varie par département, suite à des différences climatologiques.
- On n'observe que l'équilibre sur le marché au niveau du pays (la quantité demandée est égale à la quantité offerte par l'ensemble des régions).

En ce qui concerne les pesticides, nous supposons que :

- La demande des pesticides est inélastique au prix, ce qui nous permet d'exclure la partie de l'offre des pesticides de notre analyse. La quantité de pesticides utilisée demande seulement des intentions et des besoins des agriculteurs.

4.2 Formalisation

En formalisant notre modèle théorique, nous posons, que la demande de vin a la forme suivante :

$$Q_d = \alpha_d + \beta_d P_d + \gamma_d Z \quad (1)$$

Avec Z étant l'ensemble des variables ayant une 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 totale pour toute la France est donnée par l'équation suivante :

$$Q_o = \sum_{i=1}^N q_i \quad (2)$$

Où $i \in \{1, \dots, N\}$ sont des départements, chacun ayant sa propre fonction de production et d'offre unique :

$$q_i = a_i + b_i P_o + c_i X \quad (3)$$

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). Nous pouvons réécrire l'équation de l'offre sous la forme :

$$Q_o = \sum_{i=1}^N (a_i + b_i P_o + c_i X) = \sum_{i=1}^N a_i + \sum_{i=1}^N b_i P_o + \sum_{i=1}^N c_i X \quad (4)$$

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

$$\begin{aligned}
Q_d &= \alpha_d + \beta_d P_d + \gamma_d Z \\
Q_o &= \sum_{i=1}^N q_i \\
q_{1,t} &= a_1 + b_1 P_o + c_1 X \\
&\vdots \\
q_{N,t} &= a_N + b_N P_o + c_N X
\end{aligned}$$

Quand même, parce que nous pouvons supposer une présence des contraintes au niveau des données, nous devrions prévoir des modifications possibles pour notre modèle. Les contraintes principales sont au niveau du manque des données au niveau des années, c'est-à-dire que nous risquons d'avoir une très faibles variation intra-annuelle des prix et de revenus pour pouvoir identifier les coefficient associés par un passage à l'équation structurelle.

Une de ces modification possibles est l'introduction d'une contrainte supplémentaire au niveau de la demande sur le vin de table. Afin de pouvoir identifier les effets de toutes les variables par un système AIDS, nous pouvons supposer, que tout le vin produit dans un département est consommé dans le même département. C'est une hypothèse forte, qui nous éloigne de la réalité, parce que de cette façon nous ignorons plusieurs effets pervers, tels que :

- La structure du marché interne de la France ;
- La mobilité de la production entre les différents départements ;
- L'export du vin ;
- La consommation des vins importés.

Nous pouvons tout de même ignorer ces effets, car nous visons à estimer les effets moyens pour tous les départements. De cette façon lors d'aggregation des effets au niveau national nous allons mitiger les biais possibles.

Alors, nous pouvons réécrire notre système d'équations sous la forme suivante :

$$\begin{aligned}
qd_1 &= \alpha_1 + \beta P_{1,d} + \gamma_1 Z_1 \\
&\vdots \\
qd_N &= \alpha_N + \beta P_{N,d} + \gamma_N Z_N \\
qo_1 &= a_1 + b P_{1,o} + c_1 X_1 \\
&\vdots \\
qo_N &= a_N + b P_{N,o} + c_N X_N
\end{aligned}$$

Il faut spécifier, que nous supposons les effets de prix sont identiques pour tous les département en moyenne, tandis que nous laissons quand même les effets des autres variables dépendantes (ex : le revenu et les pesticides) de varier par département.

5 Les données

5.1 Sources des données :

Nous avons utilisé les bases des données suivantes pour notre analyse :

- Les données de ventes de pesticides par département (INERIS)
- Les données sur les prix du vin (France Agrimer)
- Les données sur la population (INSEE)
- Les données sur la production de vin (SSM Finances Publiques)

5.2 Les variables utilisées pour notre modèle

Dans notre étude nous faisons face à un problème avec deux variables endogènes et trois variables exogènes.

- Variables endogènes :
 - la quantité totale produite de vin rouge et blanc non IG par département (en hectolitres, en log),
 - le prix moyen des vins rouges-blancs (indice, en log).
- Variables exogènes :
 - le revenu médian par département (en euros par personne par année, en log),
 - la surface agricole destinée aux vins de table (en hectares, en log),
 - la quantité des pesticides utilisés sur la vigne (indice, en log).

Au niveau des pesticides, on va s'intéresser plus particulièrement aux quantités de produits vendus par département entre 2009 et 2017 utilisés principalement sur les cultures viticoles. Il faut faire preuve de vigilance sur le conditionnement des produits qui n'est pas exprimé dans la même unité au sein de cette base : en litres ou en kilos. Dans notre étude nous allons étudier l'impact de la masse totale des pesticides utilisés. Pour pouvoir le faire, nous créons un indice qui permet de prendre en compte les évolutions des différents types des produits à la fois. Nous créons un indice simple :

$$P = \frac{\sum_j p_{j,t} q_{j,t}}{\sum_j p_{j,0} q_{j,0}}$$

Avec j désignant le produit j , et p étant un coefficient de pondération (dans le cas le plus simple $p = 1$).

En ce qui concerne les données sur le prix du vin, on s'intéresse principalement au prix moyen des vins rouge- rosés et blancs sans IG (Indication Géographique) sur la période 2009-2017. Ces prix sont déflatés par l'indice des prix à la consommation (base 100 en 2014). On ne considère ici que le prix moyen déflaté au niveau national. Dans le deuxième modèle nous avons besoin de créer artificiellement un estimateur qui va varier par département. Dans ce but nous créons l'indice de prix du vin de table départementale, calculé de façon suivante :

$$P = \frac{p_{rouge,t} q_{rouge,t} + p_{blanc,t} q_{blanc,t}}{p_{rouge,0} q_{rouge,0} + p_{blanc,0} q_{blanc,0}}$$

Avec t étant l'année au temps t .

Au niveau des données sur la population, la variable qui nous intéresse ici est relative au niveau de revenu, exprimée au niveau départemental (laquelle, si besoin nous pourrions facilement agréger

au niveau national). Plus précisément, on va utiliser le revenu médian par département. Il est aussi déflatée de l'indice des prix à la consommation (base 100 en 2014).

Enfin, nous avons exploité les variables suivantes au niveau de la production de vin : la surface totale de culture viticole en hectares, la surface utilisée pour les vins non IG, la quantité produite de vins rouges-rosés et blancs sans IG en hectolitres, pour chaque département et sur la période 2009-2017. Afin d'obtenir un estimateur de la quantité de vin totale on prend la somme totale des tous types du vin sans IG.

5.3 Dictionnaire des variables

Table 1: Dictionnaire des variables

Variable	Description
année	année
ndep	numéro de département
si	superficie de vigne sans indication géographique en hectare en log
qi	quantité de vins produits en hectolitre en log
ipi	indice des prix du vin sans IG déflatés en log
ri	revenu disponible brut des ménages français déflatés en log
iki	indice de quantité de pesticides achetés en log
t	la tendance temporelle

Les propriétés de ces données sont suivantes :

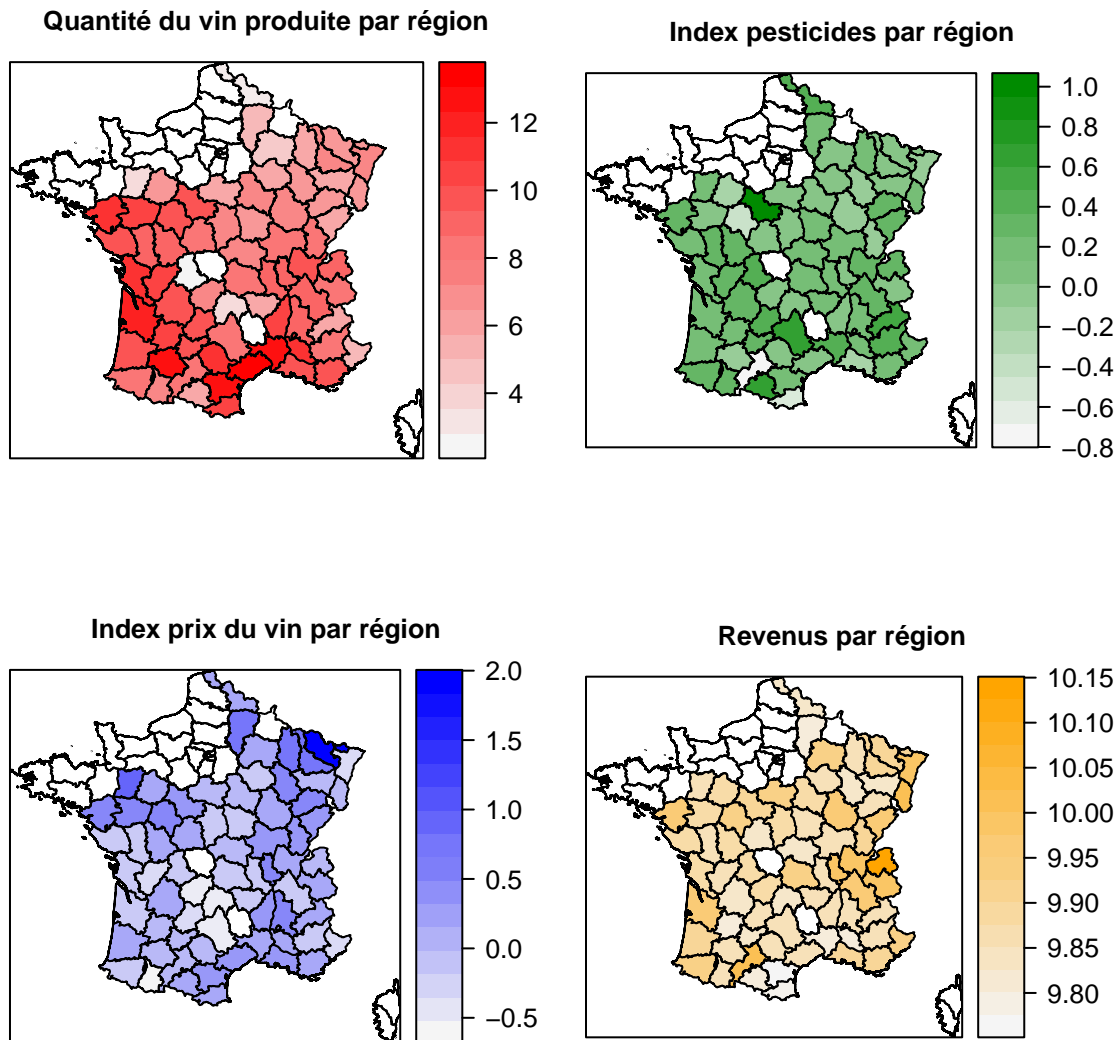
- Toutes les variables varient par département et par année.
- Le période temporelle comprise dans notre échantillon est de 2012 à 2016.
- Nous ne considérons que les régions produisant du vin.
- Nous éliminons les effets fixes pour en substrayant les moyennes départementales.
- Données en panel "cylindrées".
- Nombre des individus large (69 départements, qui produisent le vin et utilisent des pesticides) et le nombre des périodes pauvre (5 périodes).

6 L'étude statistique

Dans cette partie de travail nous allons explorer les données collectées.

De l'étude de la variance pour les données en panel avec des statistiques générales, nous passerons vers l'étude des interdependances des variables. Pui, nous allons finir avec étude des donnée alternées par une transformation **Within**.

6.1 Visualisation au niveau de la France



6.2 Etude de la variance

Le tableau suivant regroupe les statistiques descriptives essentielles :

- Moyennes
- Variance sur l'échantillon complet
- Variance *between*
- Variance *within*

Il est facile à remarquer que la variance *between* est plus significative que la variance *within*. Cela nous amène à l'idée qu'il faut utiliser un modèle qui permettra d'estimer et corriger ces inégalités entre les individus, car nous sommes plus intéressés par des effets individuels moyens (les effets moyens pour tous les individus).

Table 2: Etude de la variance

	Mean	Overall	Between	Within
Index prix	0.175	0.568	0.368	0.434
Index pesticides	0.170	0.333	0.239	0.234
Surface	4.892	1.986	1.955	0.410
Revenus	9.891	0.061	0.061	0.011
Temps	3	1.416	0	1.416

De plus, il est interessant d’observer les résultats obtenus pour le test de Chow comparant le modèle complet (*pooled model*) contre les modèles au effet fixes et randomes.

Sauf le cas de la surface nous ne pouvons pas rejeter l’hypothese nulle, specifiant que les individus ont des effets identiques pour toute la population.

Table 3: Les p-valeurs de pooling-test de Chow

	Random	Fixed
Index prix	0.535	0.533
Index pesticides	0.485	0.451
Surface	0	0.0001
Revenus	0.297	0.247

6.3 Visualisatoin des interdependances

Maintenant passons à l’étude des graphiques bivariés pour notre variable dependante principale (la quantité du vin) et les regresseurs.

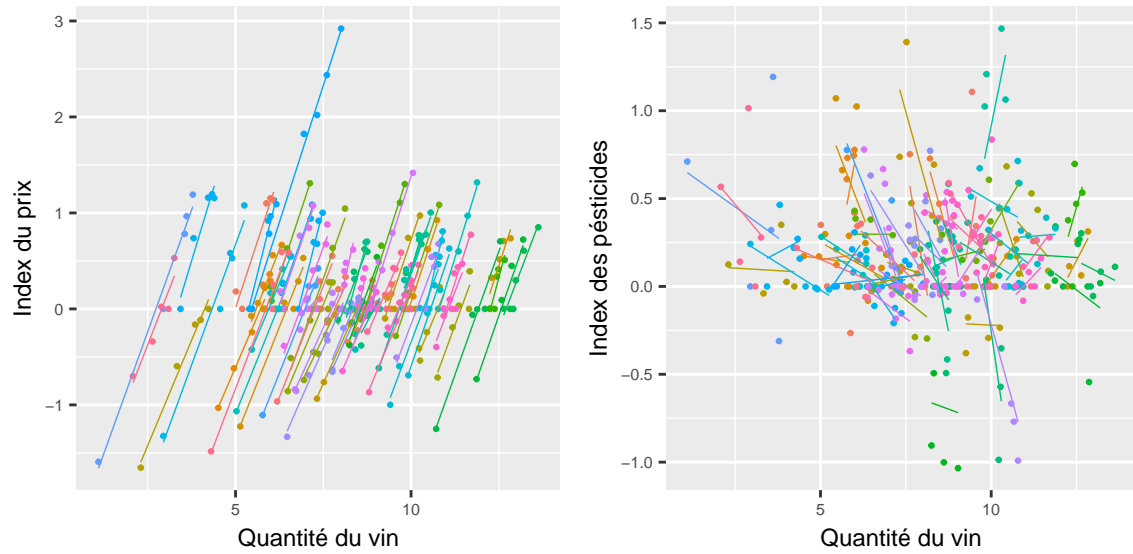


Figure 1: L'étude bivarié, partie 1

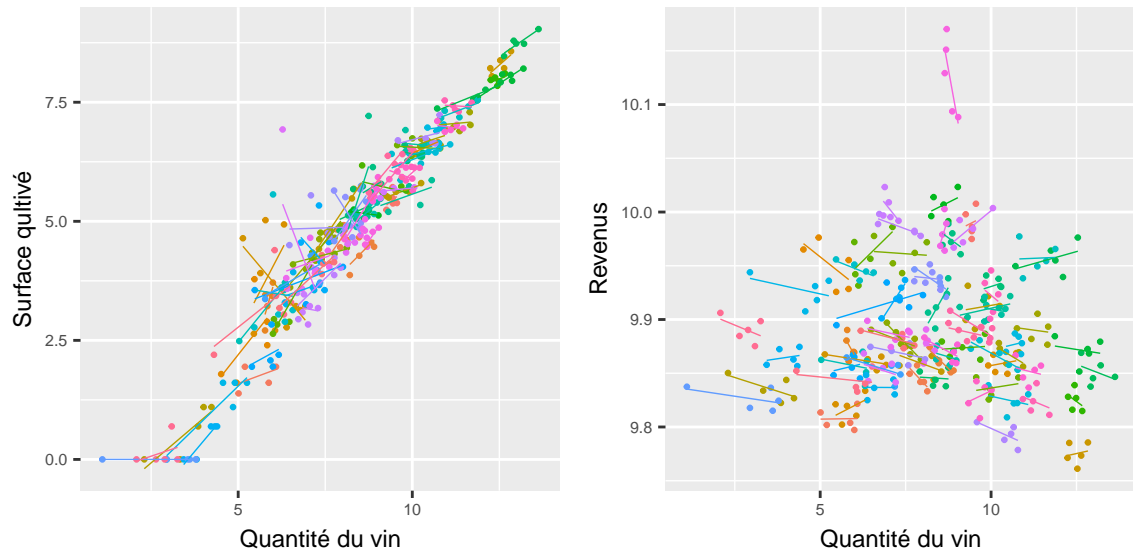


Figure 2: L'étude bivarié, partie 2

6.4 L'étude des types d'effets

Nous avons déjà vu, qu'il est fortement probable que nous faisons face à un modèle aux effets fixes individuelles. Il faut quand même le justifier. Pour faire cela, nous allons effectuer le test de multiplicateur de Lagrange sur la nature des effets fixes. Selon les résultats il est évidente que nous avons des effets fixes au niveau individuel pour toutes les variables. Pour la variable de surface nous avons des effets à la fois individuels et temporels.

Table 4: p-valeurs de Lagrange multiplier test

	Individual	Time	Twoways
Index prix	0	0.169	0
Index pesticides	0	0.222	0
Surface	0	0.030	0
Revenus	0	0.248	0

6.5 L'analyse de la corrélation

Maintenant nous allons comparer deux tableaux de corrélation. Le premier tableau comprend les résultats pour les données telles-quelles, le deuxième par contre intègre les résultats pour les données sous la transformation *within*.

	Quantité du vin	IP	Surface	Revenus	Index pesticides	Temps
Quantité du vin	1.0000000	0.1538882	0.9558586	-0.0265558	-0.0783682	-0.0359595
IP	0.1538882	1.0000000	0.0451726	-0.0373917	-0.1267030	0.0430476
Surface	0.9558586	0.0451726	1.0000000	-0.0567015	-0.0604548	-0.0640070
Revenus	-0.0265558	-0.0373917	-0.0567015	1.0000000	-0.0515319	0.1188471
Index pesticides	-0.0783682	-0.1267030	-0.0604548	-0.0515319	1.0000000	0.2908564
Temps	-0.0359595	0.0430476	-0.0640070	0.1188471	0.2908564	1.0000000

Les relations entre les variables mieux ressortent pour les données transformées.

	Quantité du vin	IP	Surface	Revenus	Index pesticides	Temps
Quantité du vin	1.0000000	0.9608421	0.3655017	-0.1601237	-0.2275316	-0.1994410
IP	0.9608421	1.0000000	0.2892935	-0.0085803	-0.1265659	0.0562534
Surface	0.3655017	0.2892935	1.0000000	-0.1656929	-0.1908549	-0.3103175
Revenus	-0.1601237	-0.0085803	-0.1656929	1.0000000	0.2280372	0.6521815
Index pesticides	-0.2275316	-0.1265659	-0.1908549	0.2280372	1.0000000	0.4139250
Temps	-0.1994410	0.0562534	-0.3103175	0.6521815	0.4139250	1.0000000

6.6 La transformation within

Avant de terminer cette partie d'étude statistique de notre échantillon il faut encore présenter les relations bivariées pour les données transformées (sous la transformation *within*).

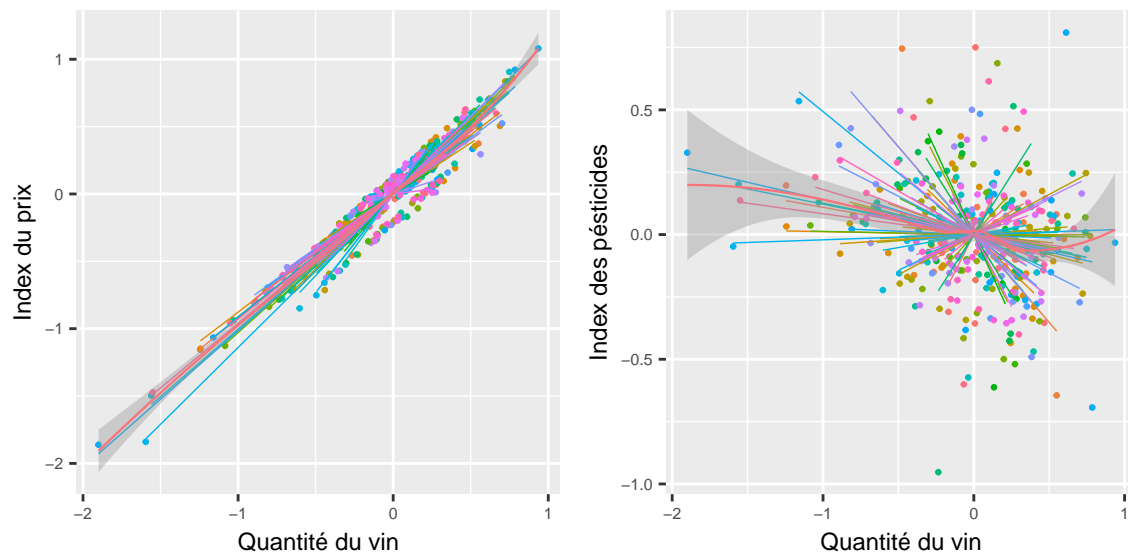


Figure 3: Relations bivariés dans le cas de transformation within, partie 1

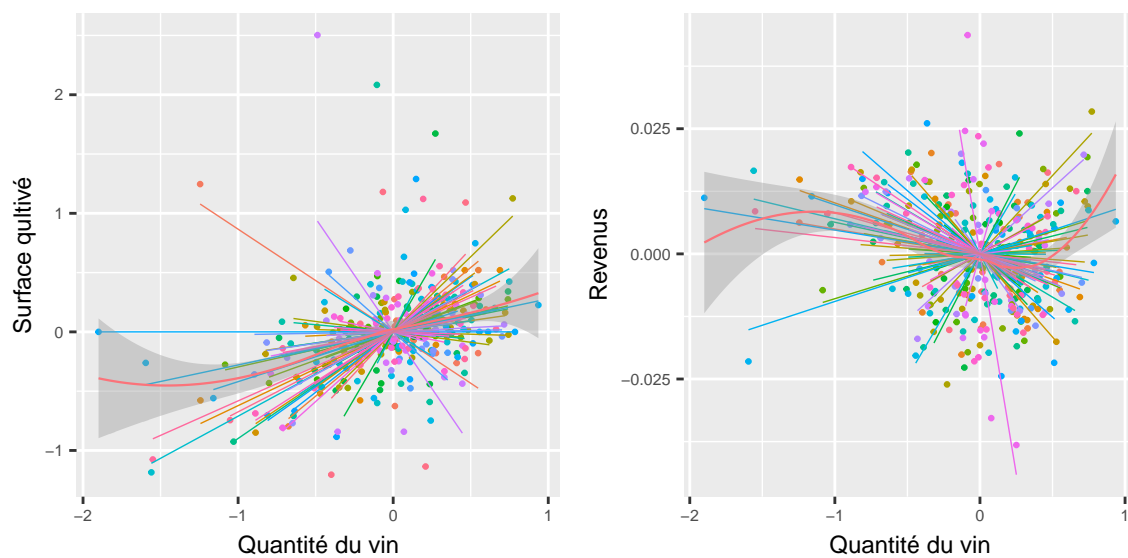


Figure 4: Relations bivariés dans le cas de transformation within, partie 2

7 Modélisation

- Présentation de la méthode
- Les estimations
 - OLS, WLS et SUR
 - 2SLS, W2SLS, 3SLS et i3SLS

7.1 Présentation de la méthode

L'AIDS et les autres modèles de demande cités dans la littérature ont de nombreuses lacunes qui les rendent impropres pour l'estimation du marché du vin, selon Cembalo et al. (2014). Quand même, dans notre étude nous allons utiliser ce modèle là, sous des suppositions restrictives.

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

7.2 Modèle économétrique

Par ce modèle nous visons à estimer les effets moyens pour tous les départements. De cette façon lors d'aggregation des effets au niveau national nous allons mitiger les biais possibles, liés à la misspecification du modèle.

Nous pouvons réécrire notre système d'équations sous la forme suivante :

$$\begin{aligned}
 qd_{1,t} &= \alpha_1 + \beta Pd_{1,t} + \gamma_1 Z_{1,t} + \epsilon_{1,t} \\
 &\vdots \\
 qd_{N,t} &= \alpha_N + \beta Pd_{N,t} + \gamma_N Z_{N,t} + \epsilon_{1,t} \\
 qo_{1,t} &= a_1 + bPo_{1,t} + c_1 X_{1,t} + u_{1,t} \\
 &\vdots \\
 qo_{N,t} &= a_N + bPo_{N,t} + c_N X_{N,t} + u_{N,t}
 \end{aligned}$$

Nous posons que l'offre et la demande sont égaux au niveau de département. L'offre de département vise à satisfaire la demande interne du même département. En termes d'aggregation ex-post des effets estimés, nous sommes sensé de tomber sur l'équilibre au niveau du marché national. C'est-à-dire :

$$\begin{aligned}
 qd_{1,t} &= qo_{1,t} \\
 &\vdots \\
 qd_{N,t} &= qo_{N,t}
 \end{aligned}$$

Au pint d'équilibre nous avons également l'égalité des prix :

$$Po_{1,t} = Pd_{1,t}$$

De cette façon nous obtenons un système des systèmes des équations :

$$\begin{aligned}
 q_{1,t} &= \alpha_1 + \beta P_{1,t} + \gamma_1 Z_{1,t} + \epsilon_{1,t} \\
 &\vdots \\
 q_{N,t} &= \alpha_N + \beta P_{N,t} + \gamma_N Z_{N,t} + \epsilon_{1,t} \\
 q_{1,t} &= a_1 + bP_{1,t} + c_1 X_{1,t} + u_{1,t} \\
 &\vdots \\
 q_{N,t} &= a_N + bP_{N,t} + c_N X_{N,t} + u_{N,t}
 \end{aligned}$$

En simplifiant l'écriture nous pouvons la représenter sous la forme suivante :

$$\begin{aligned}q_{i,t} &= \alpha_i + \beta P_{i,t} + \gamma_i Z_{i,t} + \epsilon_{i,t} \\q_{i,t} &= a_i + b P_{i,t} + c_i X_{i,t} + u_{i,t}\end{aligned}$$

D'ici nous avons à notre disposition deux outils d'identification des effets étudiés :

- Résolution du système par le passage aux équations structurelles.
- Doubles moindres carrés, lesquels on peut utiliser car on s'intéresse principalement au rôle des pesticides dans la production du vin ;

7.3 La stratégie d'identification

Nous attendons à ce que l'estimateur de 3SLS, qui permet de capter tous les effets de corrélations entre les équations en présence de plusieurs variables exogènes nous permettra d'obtenir des estimations fiables.

- Traitement des systèmes d'équations liées (simultaneity bias)
- L'estimateur de 3SLS donne des résultats similaires à l'estimateur de ILS
- L'estimateur est consistant
- La distribution pour les estimateurs suit une loi normale seulement dans des grands échantillons
- L'estimateur est efficient (asymptotiquement)

Quand même dès le début nous envisageons des problèmes possibles avec les résultats obtenus :

- Faible représentation des effets hétérogènes entre les régions (nous estimons seulement les effets moyens)
- Les interférences induites par l'hétérogénéité

8 Résultats

Dans cette section nous allons présenter les premiers résultats économétriques. Lors de ces premières estimations nous supposons que les effets sont identiques pour tous les départements (on fait la correction pour les effets fixes au niveau départemental quand même).

Nous allons présenter en particulier :

- Les coefficients estimés avec leurs variances
- L'efficacité et comparaison des estimateurs
- Étude des erreurs
 - La distribution des erreurs
 - * La normalité
 - * Centrage sur 0
 - * Indépendance des variables explicatives
 - L'autocorrélation des résidus
 - L'hétéroscédasticité

Nous estimons un ensemble des différents modèles afin de pouvoir choisir la méthode la plus raisonnable. Les modèles suivants sont traités séparément :

- Modèles simples :

- OLS, WLS (ou nous concentrons notre attention sur l'équation d'offre du vin)
- SUR (ce qui n'est pas justifié, mais on l'inclus quand même)
- Modèles des équations simultanées avec des variables endogènes :
 - 2SLS, W2SLS
 - 3SLS et 3SLS itérés (le dernier modèle étant similaire à FIML)

8.1 Les résultats OLS, WLS et SUR

	OLS	WLS	SUR
Demande: ipi	0.93*** (0.01)	0.93*** (0.01)	0.93*** (0.01)
Demande: ri	-5.75*** (0.47)	-5.75*** (0.47)	-2.00*** (0.33)
Offre: ipi	0.90*** (0.01)	0.90*** (0.01)	0.92*** (0.01)
Offre: si	0.08*** (0.01)	0.08*** (0.01)	0.02* (0.01)
Offre: iki	-0.17*** (0.02)	-0.17*** (0.02)	-0.05** (0.02)
Demande: R ²	0.95	0.95	0.94
Offre: R ²	0.94	0.94	0.93
Demande: Adj. R ²	0.95	0.95	0.94
Offre: Adj. R ²	0.94	0.94	0.93
Num. obs. (total)	690	690	690

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 7: Statistical models

8.2 Indépendance des résidus

	OLS D	OLS O	WLS D	WLS O	SUR D	SUR O
Vin	0.2317615	0.2444187	0.2317615	0.2444187	0.2728127	0.2746202
IP	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000
Surface	0.2707361	0.0000000	0.2707361	0.0000000	0.3131039	0.2364897
Revenus	0.0000000	-0.4799169	0.0000000	-0.4799169	-0.3931430	-0.5421354
Pesticides	-0.3082912	0.0000000	-0.3082912	0.0000000	-0.3726957	-0.2814999

8.3 L'autocorrelation et l'hétéroscédacité

Table 9: Les statistiques test de Durbin-Watson

	OLS	WLS	SUR
Equation de demande	1.112	1.112	1.569
Equation d'offre	1.062	1.062	1.552

8.4 Le comportement des résidus

8.5 Les PDF des résidus

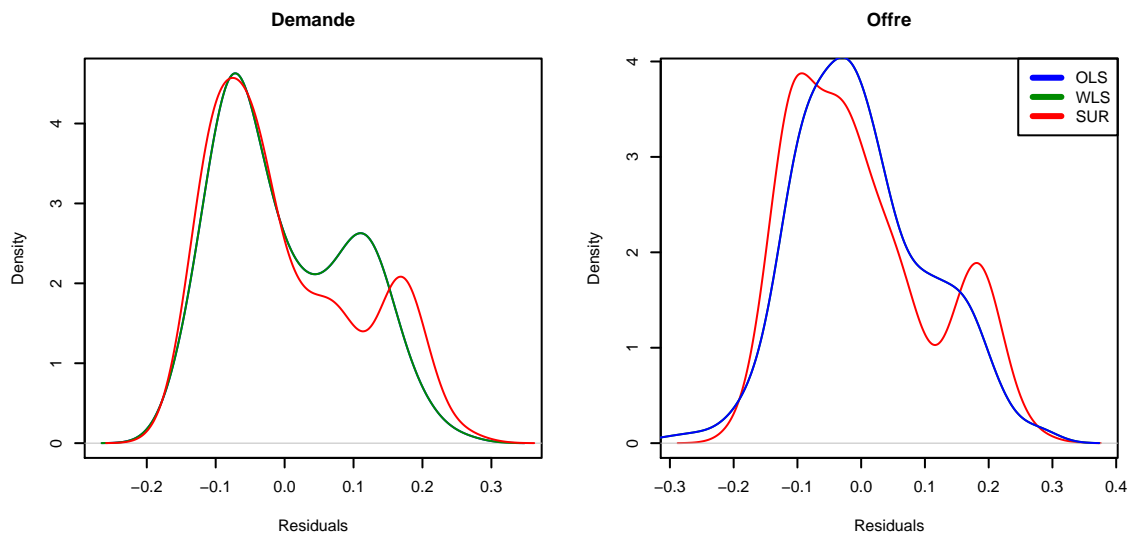


Figure 5: Les PDF des résidus

8.6 Les résidus contre la variable prédite

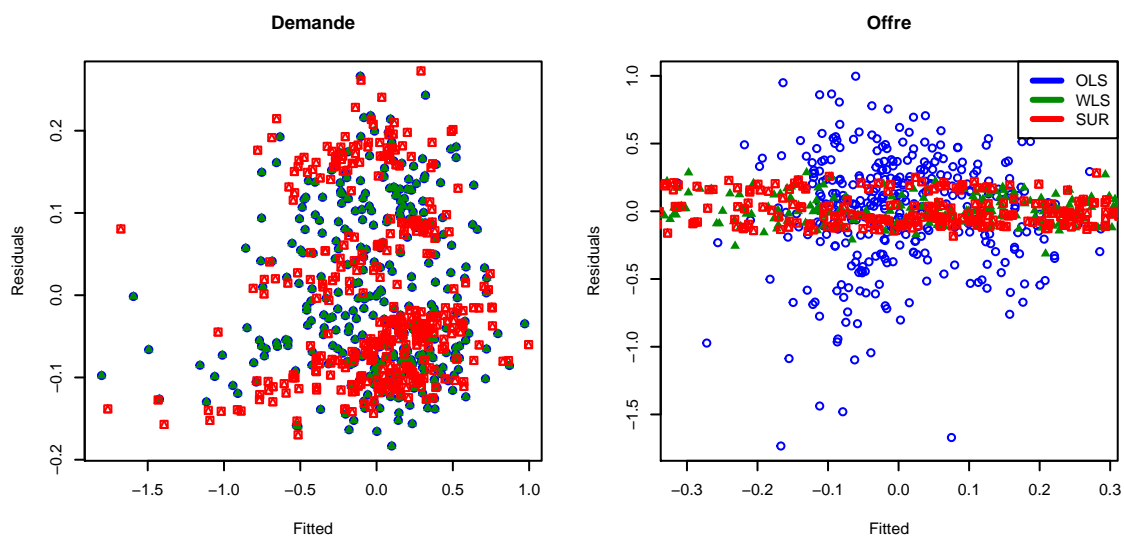


Figure 6: Les résidus contre la variable prédite

8.7 Les résultats 2SLS, W2SLS, 3SLS et i3SLS

	2SLS	W2SLS	3SLS	i3SLS
Demande: ipi	1.19*** (0.06)	1.19*** (0.06)	1.19*** (0.06)	1.19*** (0.06)
Demande: ri	-5.67*** (0.71)	-5.67*** (0.71)	-5.67*** (0.71)	-5.67*** (0.71)
Offre: ipi	-1.22 (1.97)	-1.22 (1.97)	-0.71 (1.96)	-0.81 (1.60)
Offre: si	0.70 (0.59)	0.70 (0.59)	0.46 (0.58)	0.51 (0.47)
Offre: iki	-0.46 (0.34)	-0.46 (0.34)	-0.73* (0.32)	-0.68* (0.26)
Demande: R ²	0.88	0.88	0.88	0.88
Offre: R ²	-3.37	-3.37	-1.60	-1.90
Demande: Adj. R ²	0.88	0.88	0.88	0.88
Offre: Adj. R ²	-3.40	-3.40	-1.62	-1.92
Num. obs. (total)	690	690	690	690

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 10: Statistical models

8.8 Comparaison des modèles

Table 11: Hausman 3SLS consistency test

	Test	Resultats
1	2SLS contre 3SLS	0.350
2	2SLS contre i3SLS	0.735

8.9 Le comportement des résidus

Table 12: Shapiro-Wilk test de normalité

	2SLS	3SLS	i3SLS
Equation de demande	0.00003	0.00003	0.00003
Equation d'offre	0.00000	0.00000	0.00000

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: lun., déc. 16, 2019 - 13:26:38

Table 13: Test de Bartlett sur l'hétéroscédasticité

	2SLS	3SLS	i3SLS
Equation de demande	0.975	0.975	0.975
Equation d'offre	0.0003	0.002	0.001

8.10 Les PDF des résidus

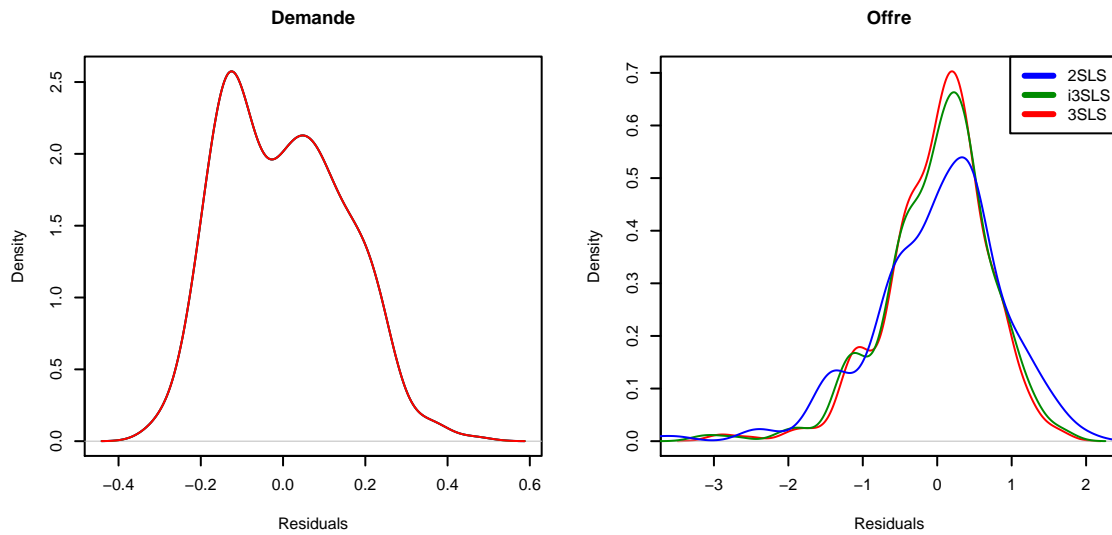


Figure 7: Les PDF des résidus

8.11 Les résidus contre la variable prédite

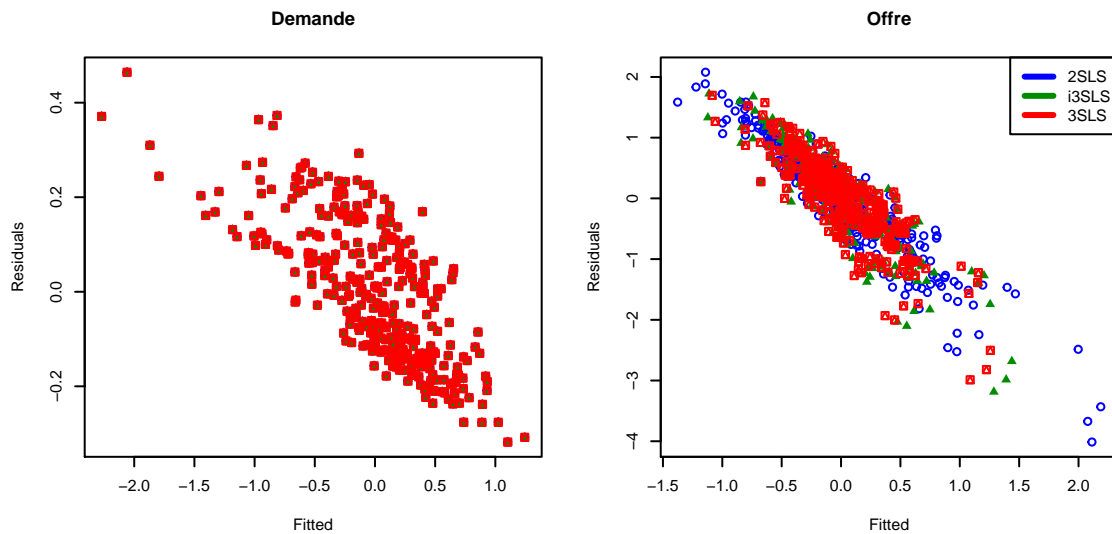


Figure 8: Les résidus contre la variable prédite

8.12 Les résidus et les prédictions pour i3SLS

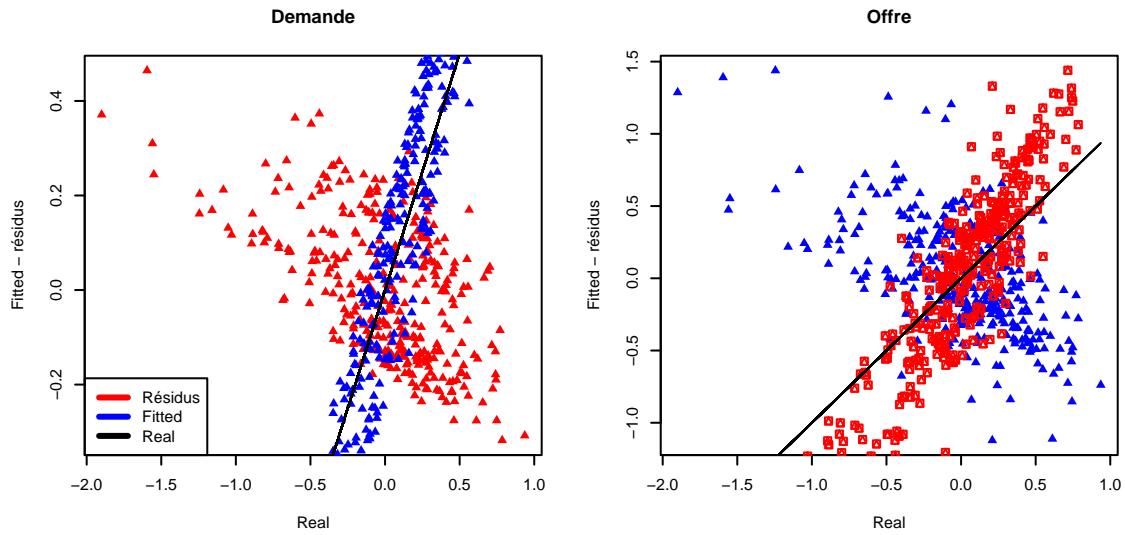


Figure 9: Les résidus et les prédictions

8.13 L'autocorrelation

8.14 L'autocorrelation sur 2 dimensions pour i3SLS

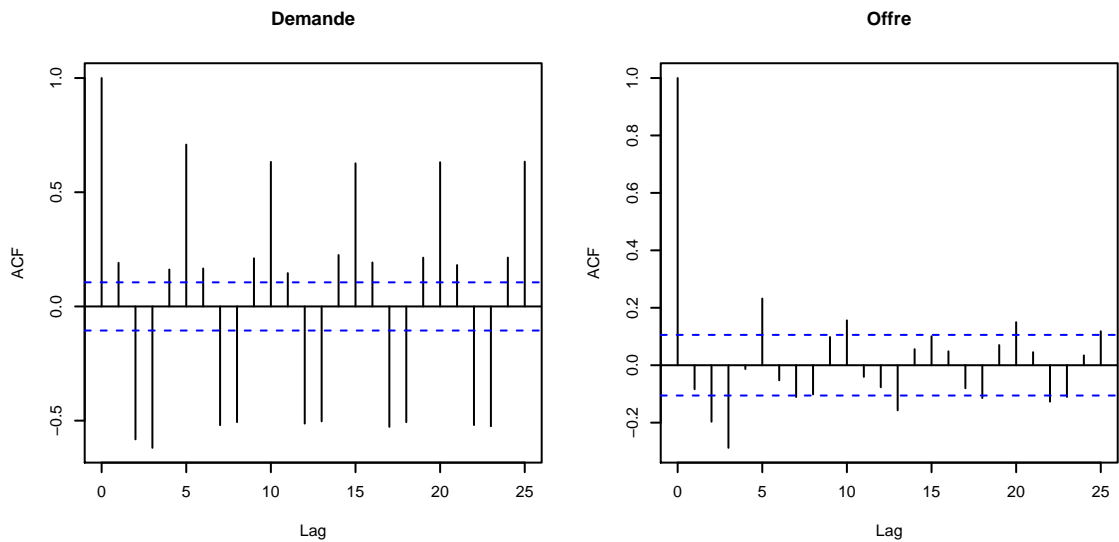


Figure 10: L'autocorrelation pour le modèle i3SLS

8.15 L'indépendance des résidus

9 Clustering et la séparation des groupes (Between transformation approach)

Nous avons vu dans le comportement des résidus une nature non-aléatoire groupée. Cela nous amène à l'idée de construire k-clusters pour modéliser les relations par groupe.

Nous supposons que les départements ayant des valeurs moyennes interannuelles proches (transformation Between) ont le comportement identique. La clusterisation est effectuée sur les données Between pour les départements.

Nous pouvons supposer que le nombre des clusters optimal est entre 3 et 5. Prenant en compte les graphiques des résidus vus lors d'analyse des modèles nous allons supposer qu'il n'y a que 3 clusters principaux.

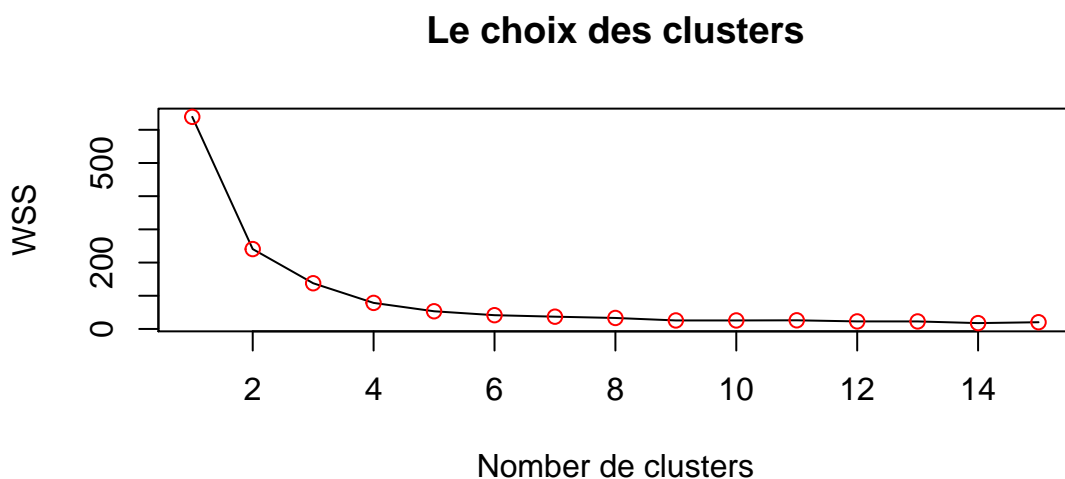


Figure 11: Le choix des clusters

Les groupes sont définies par des caractéristiques suivantes :

Table 14: Les centres des clusters

	qi	ipi	si	ri	iki	.1
1	5.349	0.190	2.311	9.880	0.216	17
2	8.207	0.154	4.879	9.914	0.136	31
3	10.940	0.192	7.000	9.867	0.184	21

9.1 Modélisation

Nous évaluons le système en introduisant les variables de groupe (dummy variables) sous l'hypothèse des résidus joints.

	OLS	2SLS	3SLS
Demande: ipi	0.93*** (0.01)	1.14*** (0.05)	1.13*** (0.05)
Demande: ri1	-5.42*** (0.94)	-4.20** (1.30)	-5.64*** (1.21)
Demande: ri2	-6.58*** (0.68)	-7.21*** (0.93)	-7.69*** (0.91)
Demande: ri3	-4.54*** (0.92)	-4.30*** (1.25)	-4.21*** (1.19)
Offre: ipi	0.89*** (0.01)	0.41* (0.20)	0.41* (0.20)
Offre: si1	0.05** (0.02)	0.19** (0.07)	0.18* (0.07)
Offre: si2	0.09*** (0.02)	0.20** (0.06)	0.21** (0.06)
Offre: si3	0.30*** (0.06)	0.74** (0.23)	0.72** (0.22)
Offre: iki1	-0.22*** (0.05)	-0.26* (0.11)	-0.36*** (0.10)
Offre: iki2	-0.18*** (0.04)	-0.34** (0.10)	-0.34*** (0.10)
Offre: iki3	-0.09* (0.04)	0.03 (0.11)	-0.06 (0.10)
Demande: R ²	0.95	0.90	0.90
Offre: R ²	0.94	0.72	0.73
Demande: Adj. R ²	0.95	0.90	0.90
Offre: Adj. R ²	0.94	0.72	0.72
Num. obs. (total)	690	690	690

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 15: Statistical models

9.2 Tests

Etudions la validité du modèle 3SLS :

Table 16: Hausman 3SLS consistency test

Test		Resultats
1	2SLS contre 3SLS	0

La normalité des résidus :

Table 17: Shapiro-Wilk normality test

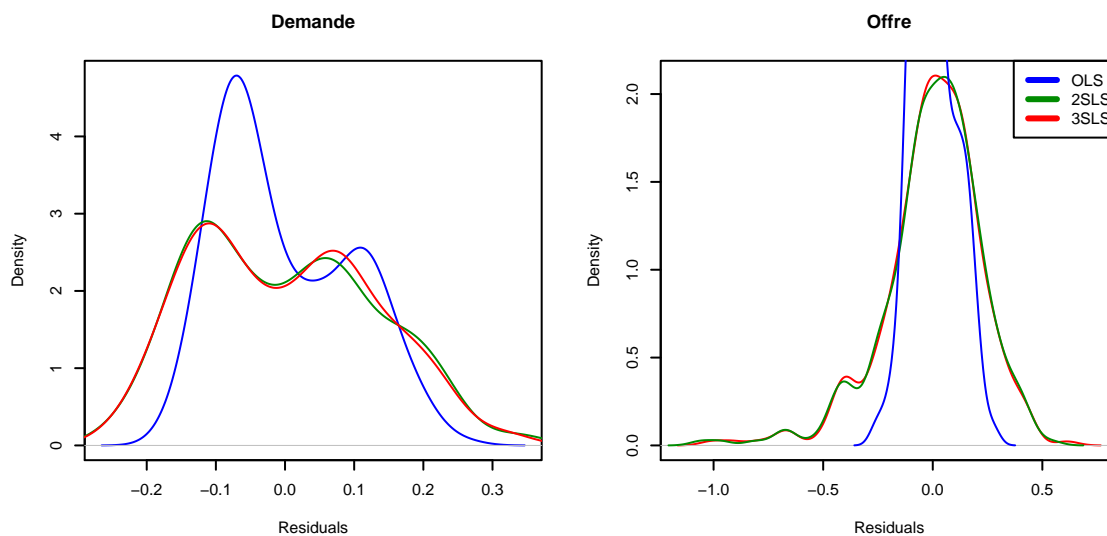
	OLS	2SLS	3SLS
Equation de demande	0	0.00001	0.00001
Equation d'offre	0.001	0	0

L'heteroscedacité :

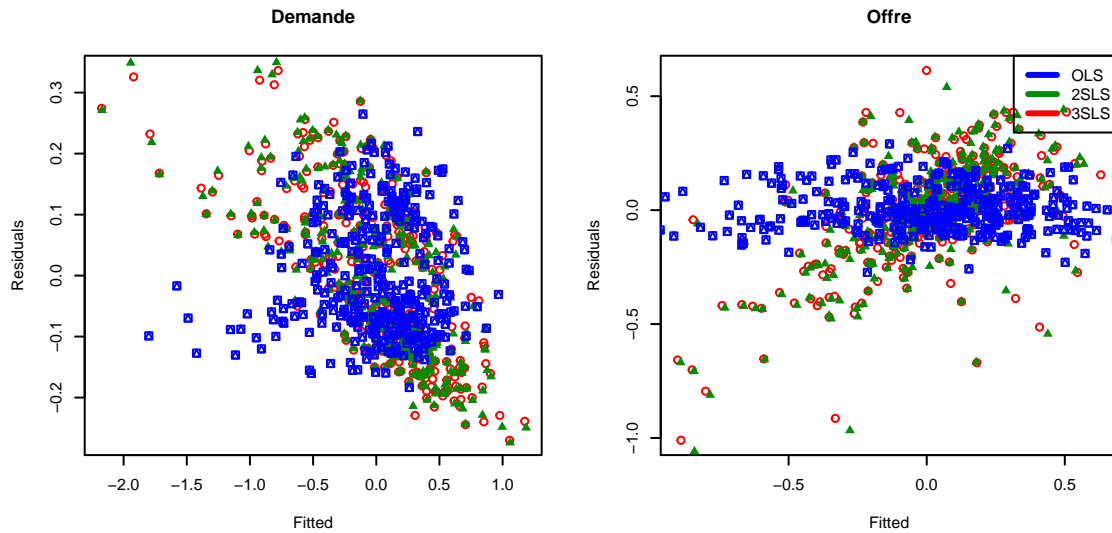
Table 18: Bartlett heteroscedasticity test

	OLS	2SLS	3SLS
Equation de demande	0.977	0.987	0.974
Equation d'offre	0.989	0.00000	0.00000

Les PDF des résidus :



Les résidus contre les variables prédites :



Conclusion : les problèmes restent non résolus.

10 Clustering et la séparation des groupes (WIDE data approach)

Nous avons vu dans le comportement des résidus une nature non-aléatoire groupé. Cela nous amène à l'idée de construire k-clusters pour modéliser les relations par groupe.

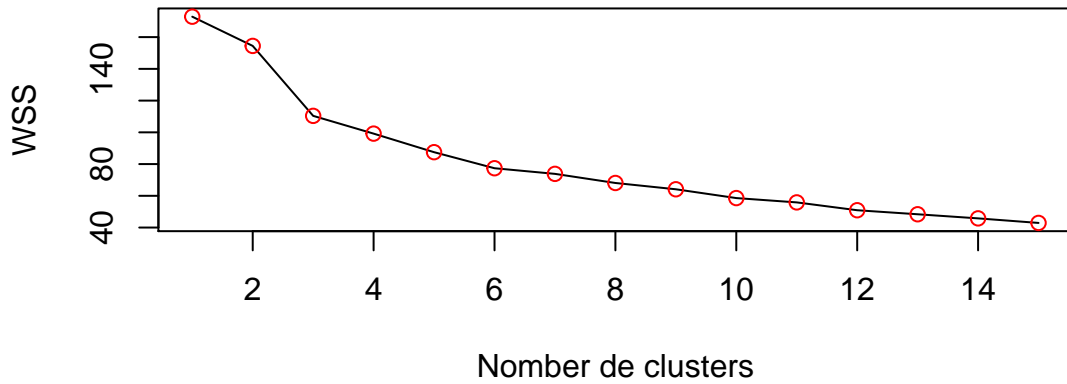
D'abord on compare le comportement des cluster pour les données à l'information complete et les données Within.

10.1 Transformation Within

Comme nous pouvons voir dans les résultats le nombre des cluster optimaux est trop large pour les séparer dans l'analyse :

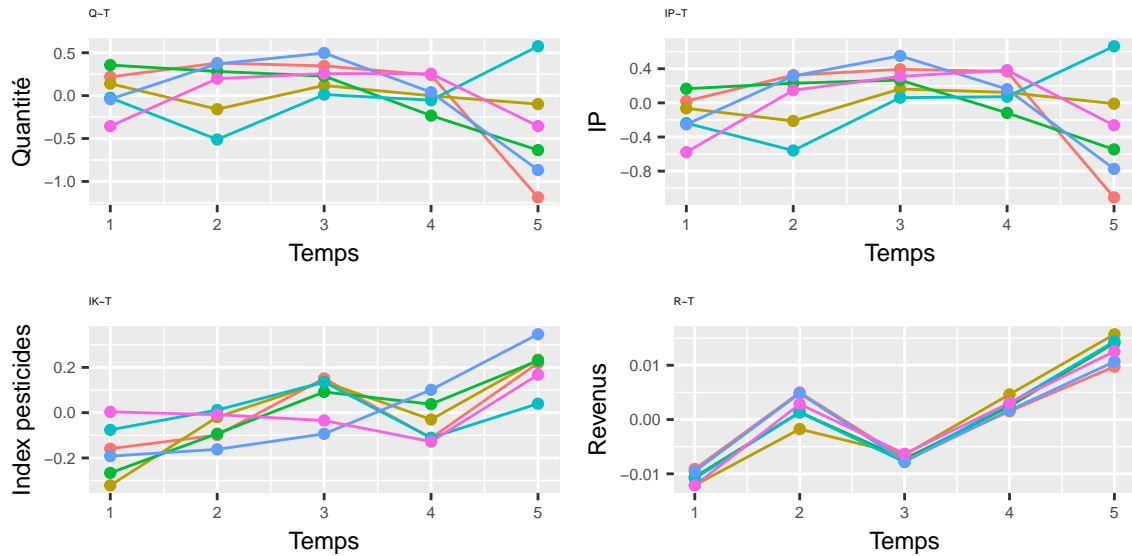
Nous pouvons supposer que le nombre des clusters optimal est entre 6 et 15.

Le choix des clusters



Les groupes sont définies par des caractéristiques suivantes :

Representation graphique :

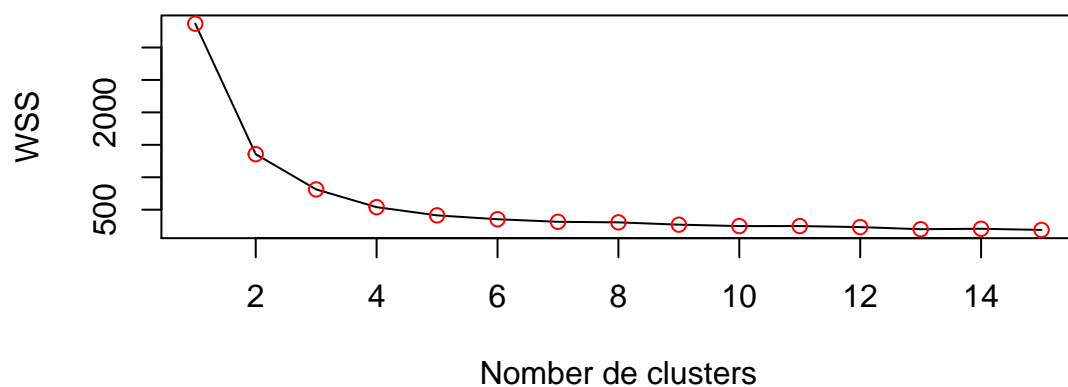


10.2 Information complete

Dans le cas d'information complete on a :

Nous pouvons supposer que le nombre des clusters optimal est entre 3 et 5. Prenant en compte les graphiques des résidus vus lors d'analyse des modèles nous allons supposer qu'il n'y a que 3 clusters principaux.

Le choix des clusters

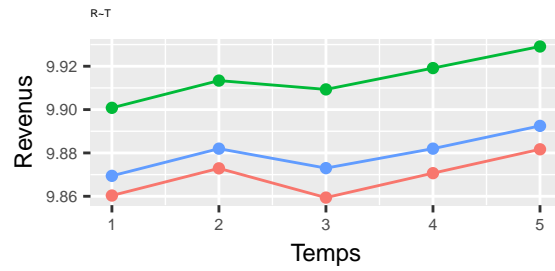
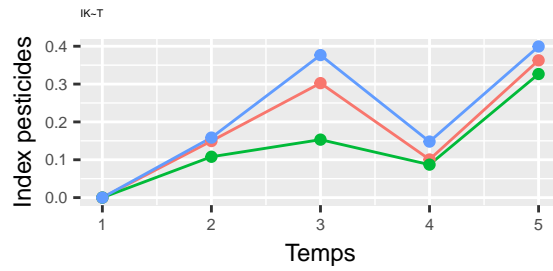
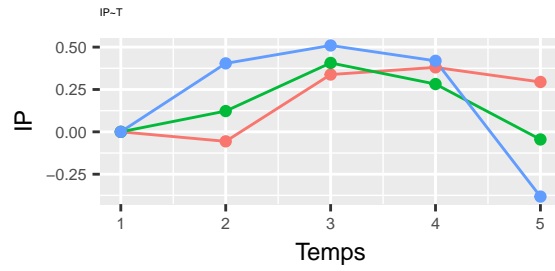
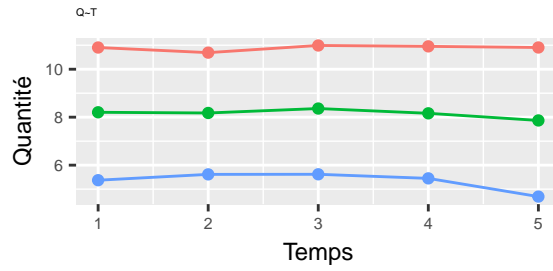


Les groupes sont définies par des caractéristiques suivantes :

Table 19: Les centres des clusters

	qi	ipi	si	ri	iki	n	k	t
1	10.905002	0	7.048328	9.860348	0	22	1	1
2	10.691657	-0.056005	6.945538	9.872895	0.149462	22	1	2
3	10.990015	0.338205	6.984754	9.859395	0.302499	22	1	3
4	10.952753	0.380444	6.855669	9.87067	0.100718	22	1	4
5	10.906341	0.294342	6.830466	9.881679	0.36276	22	1	5
6	8.204277	0	4.987838	9.900816	0	30	2	1
7	8.175461	0.122615	5.007068	9.913384	0.10792	30	2	2
8	8.360791	0.406721	4.908536	9.909318	0.153133	30	2	3
9	8.161618	0.281943	4.746188	9.919147	0.087094	30	2	4
10	7.864193	-0.044913	4.637459	9.929119	0.326554	30	2	5
11	5.37234	0	2.473289	9.869431	0	17	3	1
12	5.61569	0.404064	2.55531	9.881971	0.158519	17	3	2
13	5.619106	0.509682	2.302302	9.872995	0.376623	17	3	3
14	5.448956	0.419102	2.217837	9.881977	0.147986	17	3	4
15	4.689674	-0.381999	2.008663	9.892485	0.399327	17	3	5

Representation graphique :



10.3 Modélisation

Nous évaluons le système en introduisant les variables de groupe (dummy variables) sous l'hypothèse des résidus joints.

10.4 Tests

	OLS	2SLS	3SLS
Demande: ipi	0.93*** (0.01)	1.14*** (0.05)	1.14*** (0.05)
Demande: ri1	-4.63*** (0.91)	-4.64*** (1.24)	-4.49*** (1.19)
Demande: ri2	-6.56*** (0.69)	-7.07*** (0.94)	-7.66*** (0.91)
Demande: ri3	-5.43*** (0.94)	-4.18** (1.31)	-5.65*** (1.22)
Offre: ipi	0.89*** (0.01)	0.42* (0.20)	0.43* (0.20)
Offre: si1	0.30*** (0.06)	0.74** (0.23)	0.72** (0.22)
Offre: si2	0.09*** (0.02)	0.19** (0.06)	0.19** (0.06)
Offre: si3	0.05** (0.02)	0.19** (0.07)	0.17* (0.07)
Offre: iki1	-0.09* (0.04)	0.02 (0.11)	-0.07 (0.10)
Offre: iki2	-0.18*** (0.04)	-0.33** (0.10)	-0.33** (0.10)
Offre: iki3	-0.22*** (0.05)	-0.26* (0.11)	-0.35*** (0.10)
Demande: R ²	0.95	0.90	0.90
Offre: R ²	0.94	0.73	0.74
Demande: Adj. R ²	0.95	0.90	0.90
Offre: Adj. R ²	0.94	0.73	0.74
Num. obs. (total)	690	690	690

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 20: Statistical models

Etudions la validité du modèle 3SLS :

Table 21: Hausman 3SLS consistency test

	Test	Resultats
1	2SLS contre 3SLS	0.0003

La normalité des résidus :

Table 22: Shapiro-Wilk normality test

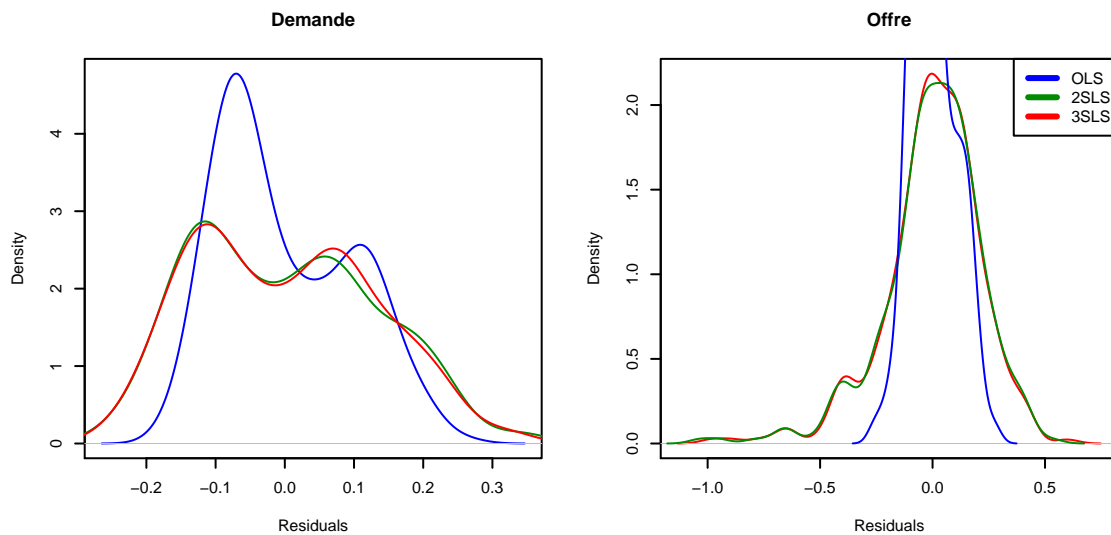
	OLS	2SLS	3SLS
Equation de demande	0	0.00001	0.00001
Equation d'offre	0.001	0	0

L'heteroscedacité :

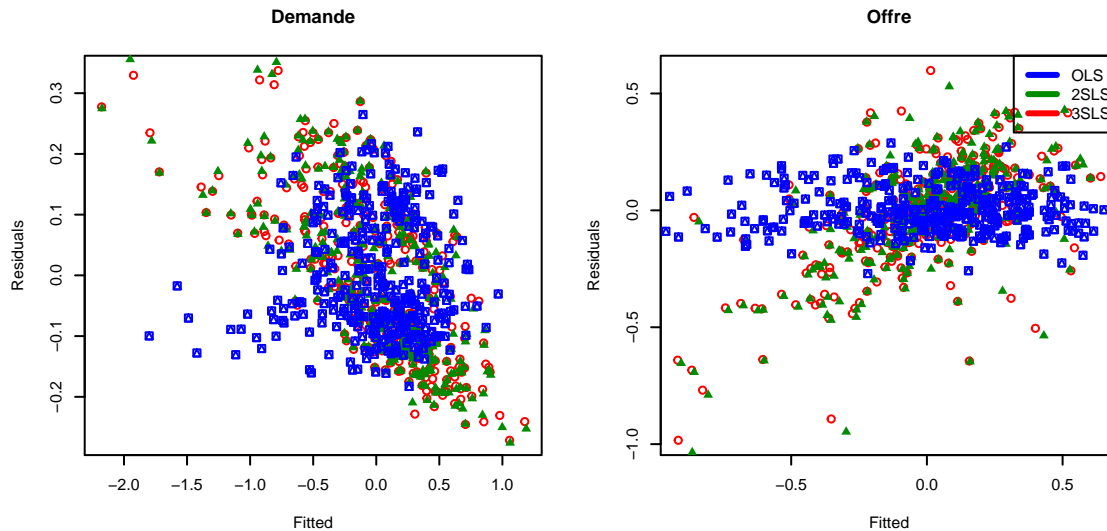
Table 23: Bartlett heteroscedasticity test

	OLS	2SLS	3SLS
Equation de demande	0.975	0.985	0.969
Equation d'offre	0.989	0.00000	0.00000

Les PDF des résidus :



Les résidus contre les variables prédites :



Conclusion : les problèmes restent non résolus.

11 Conclusions

11.1 Conclusions

- Le marché du vin
- Le rôle des pesticides
- Validité

11.2 Le marché du vin

- Un comportement inattendus
 - Les effets de substitution contre les produits de la haute gamme
 - Les effets négatives du revenu
 -

11.3 Le rôle des pesticides

- Confirmation des résultats des études précédentes
 - Utilisés pour réduire les pertes

11.4 Validité

- Faible validité du modèle économétrique
 - Variables ommisses
- Cembalo L., Caracciolo F., & Pomarici E. (2014). “Drinking cheaply : the demand for basic wine in Italy.” *Australian Journal of Agricultural and Resource Economics*, 58(3). 374-391.
- Butault J-P., Delame N., Jacquet F. & Zardet G. (2011). “L’utilisation des pesticides en France: état des lieux et perspectives de réduction.” *Notes et études socio-économiques*, 35. 7-26

- Pujol J. (2017). “Apports des produits phytosanitaires en viticulture et climat : une analyse à partir des enquêtes pratiques culturelles.” *Agreste Les Dossiers*. 39. 3-25

Cembalo, Caracciolo, and Pomarici (2014)

References

Cembalo, Luigi, Francesco Caracciolo, and Eugenio Pomarici. 2014. “Drinking Cheaply: The Demand for Basic Wine in Italy.” *Australian Journal of Agricultural and Resource Economics* 58 (3): 374–91.