



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

Arnaud Blanc, Nikita Gusarov, Sasha Picon

Tuteur :
Adélaïde Fadhuile

Niveau d'études :
Master 2

Parcours :
Chargé d'études économiques et statistique

Université Grenoble Alpes
Faculté d'économie et gestion

2019 - 2020

Contents

Introduction	1
1. Les pesticides	1
2. Le marché du vin français	2
3. Le cadre théorique	4
4. Les données	8
5. L'étude statistique	9
6. Modélisation	12
7. Résultats des estimations	15
9. Conclusions	20
Annexes	21
References	36

Introduction

Aujourd'hui, l'utilisation des pesticides est un problème majeur de l'agriculture. Celle-ci utilise la majeure partie des pesticides en France. Il s'agit d'un enjeu majeur du développement durable car ils ont un impact important sur les risques environnementaux et sanitaires.

Les pesticides sont utilisés dans l'agriculture pour protéger la production. Il est supposé que les pesticides servent à protéger les rendements. En effet, les aléas climatiques influencent sur le développement de champignons ou de maladie. Ainsi, les pesticides permettent de protéger les cultures contre les aléas climatiques et de ne pas perdre de production.

Dans ce travail nous cherchons à comprendre et estimer les effets des pesticides sur le marché des vins simples. De cette façon nous chercherons à étudier l'équilibre sur le marché des vins simples ce qui est censé de nous donner des résultats plus précis et fiables.

1. Les pesticides

Mettre des sources partout !

Pour lutter contre l'utilisation des pesticides l'Etat Français et l'union européenne ont mis en place des mesures. Ainsi, l'Etat Français lors du Grenelle de l'environnement de 2006 a fixé ces objectifs. Ainsi, le plan ECOPHYTO 2018 visait à réduire de 50% l'utilisation des pesticides de synthèse. Le deuxième objectif est le passage en agriculture biologique à 6% de la surface agricole utilisée en 2010 et vise 20% en 2020.

En 2008, les 30 produits les plus toxiques les plus toxiques sont interdits. Une taxe sur les phytosanitaires a aussi été mise en place. Cette taxe est croissante avec leur niveau de toxicité. Cette taxe devait augmenter au fil des années et l'octroi de crédits d'impôt en faveur de l'agriculture biologique.

Malgré tous ces efforts, l'utilisation des pesticides perdurent. La France enregistre une hausse des ventes de produits phytosanitaires de 10000 tonnes, entre 2011 et 2016. En 2008, le nombre de doses unitaires a été créé pour enregistrer l'évolution de la demande de pesticide. On remarque que les doses utilisées augmentent de 12% en 2014-2016 par rapport à 2009-2011.

Etat actuel

Contrairement aux attentes des autorités, on ne remarque aucune baisse de l'utilisation de pesticides. Le Nodu a connu une hausse de 23% entre 2008 et 2017. Certaines critiques ont été faites sur l'utilisation du Nodu. Il est possible d'utiliser le nombre de substances actives utilisées. Mais, cet indicateur connaît lui aussi une hausse de 15% entre 2011 et 2017.

Néanmoins, les politiques ont quand même eu quelques effets positifs, puisque l'achat des produits les plus dangereux baisse de 6% en 2017.

Les grandes cultures sont les premières utilisatrices de pesticides. Elles représentent 67,4% de l'utilisation de pesticides. La deuxième culture est celle de la vigne ce qui représente 14,4% des pesticides utilisés.

Comment baisser l'utilisation de pesticides

Afin de baisser l'utilisation des pesticides, des méthodes de cultures ont été développées pour baisser l'utilisation des pesticides. Il est possible d'utiliser différents mode de culture. On peut en retenir trois principaux.

Le premier est l'agriculture intensive. Elle ne limite pas le recours aux pesticides.

Le deuxième est l'agriculture raisonnée. Elle limite le recours aux pesticides en fonction de seuils.

Le troisième niveau est l'agriculture biologique. Elle supprime les traitements avec des produits phytosanitaires de synthèse.

Les professionnels proposent de commencer par utiliser l'agriculture raisonnée qui permet de réduire les doses de pesticides légales. Ensuite l'agriculture doit se déplacer vers l'agriculture biologique qui n'utilise aucun produit phytosanitaire de synthèse.

2. 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.

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, Caracciolo, and Pomarici 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 (2004), 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, Caracciolo, and Pomarici (2014).

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 également 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. En même temps, 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, Caracciolo, and Pomarici 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.

Ajouter des articles proches par la méthodologie à notre

- le cas du modèle simple,
- le cas du marché liée,
- le cas des clusters.

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 et elle est particulièrement 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.

3. Le cadre théorique

Les hypothèses théoriques

Ajouter les références ...

Comme proposé dans la littérature, notre étude sur les vins non coûteux (non IGP) est effectuée au niveau du pays Cembalo, Caracciolo, and Pomarici (2014) pour deux raisons. D'abord, 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 (Anderson, Nelgen, and others 2011). De plus, la perception des produits de consommation varie d'un pays à l'autre (MÄKELÄ et al. 2006).

Rachat du vin par les enseignes (grand surfaces) ... KREMER and VIOT (2004)

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.

Facteurs de production ... Laporte and PICHERY (1996)

Les coûts des viticulteurs ... Laporte and PICHERY (1996)

Avant de conclure, nous proposons au lecteur une liste exhaustive des suppositions sur le comportement du marché des vins simples. Premièrement, nous supposons que chaque département à une fonction de production unique déterminé par des spécificités historiques, les traditions, la législation, le terroir, ainsi que des conditions météorologiques et géographiques. Les effets sont fixes au niveau départemental et peuvent être isolés par des transformations spécifiques des données (ex : une transformation Within). Deuxièmement, la quantité vendue sur le marché départemental est consommée au sein du même département. C'est une hypothèse trop restrictive, qui nous éloigne de la réalité, mais nous devrions l'adopter si nous voulons intégrer les relations entre l'offre et la demande dans notre modèle. Afin de vérifier cette hypothèse nous allons construire deux modèles différents. Finalement, les effets qu'on vise à estimer sont des effets moyens au niveau départemental. C'est à dire nous allons obtenir un estimateur des effets moyens pour l'ensemble des départements inclus dans notre analyse, ou des effets moyens au sein des groupes des départements, si nous révélons des différences significatives entre les départements. Un autre modèle nous permettra de vérifier et justifier cette hypothèse.

En ce qui concerne les pesticides, nous supposons d'abord, que l'utilisation des pesticides par les viticulteurs est entrélinéaire avec la demande sur le vin et les préférences des consommateurs. De plus, nous posons, que la demande des pesticides est inélastique au prix, ce qui nous permet d'exclure les interactions entre les fournisseurs des pesticides et les agriculteurs de notre analyse. C'est-à-dire, la quantité de pesticides utilisés par les agriculteurs correspond seulement à leurs intentions et besoins.

Pour sommeriser cette partie, on résume que ce travail va porter sur les effets des pesticides sur l'offre des vins simples. Nous allons tester certaines hypothèses sur le comportement et l'organisation des relations sur le marché des vins simples en comparant des différents modèles. Puis, nous pourrions choisir entre ces modèles différents un le plus vraisemblable, qui nous servira pour répondre à la question de recherche.

Formalisation

En formalisant notre modèle théorique de base, nous posons, que l'offre agrégée pour toute la France est donnée identiquement par l'équation suivante :

$$Q_o = \sum_{i=1}^N q_{oi} \quad (1)$$

Avec la quantité offerte déterminé par des contraintes de production et le prix sur le marché :

$$q_{oi} = a_i + b_i P_{oi} + c_i X_i \quad (2)$$

Où X est 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 et la surface disponible, alors l'effet c_{i1} : $c_i = (c_{i1}, c_{i2})$ représente l'effet d'utilisation des pesticides dans la production du vin sur l'offre du dernier.

Cette équation déjà en soit permet d'estimer les effets d'utilisation des pesticides sur le marché du vin. Appelons cette modèle théorique M1 pour le référencier en futur, nous permettant de distinguer le cas sans interactions simultanées entre l'offre et la demande.

Il faut tenir compte que de cette façon nous ignorons plusieurs effets pervers, tels que :

- La structure du marché interne de la France ;
- La mobilité des produits finis entre des différents départements ;
- L'exportation et l'importation du vin.

Toutefois, ces résultats ne seront valables que dans la situation où la quantité du vin simple offerte sur le marché est déterminée seulement par le producteur et n'est pas entréliée avec la demande. Comme nous avons vu dans la section précédente, la demande peut influencer les décisions des viticulteurs (ex: le choix de la procédure technique à suivre, d'utiliser ou non les pesticides, etc). Dans le cas pareil, nous devrions prendre en compte les interactions entre l'offre et la demande. A ce but nous introduisons également la demande dans notre analyse.

La demande agrégée du vin en France peut s'écrire sous la forme suivante :

$$Qd = \sum_{i=1}^N qd_i$$

Où $i \in \{1, \dots, N\}$ sont des départements, chacun ayant sa propre fonction de la demande unique :

$$qd_i = \alpha_i + \beta_i Pd_i + \gamma_i Z_i$$

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).

Pour intégrer cette information dans notre *framework* analytique, nous devons construire une système d'équations. Il y existe plusieurs façons de le faire.

Dans le premier cas, nous pouvons essayer de capter les effets au niveau national. Pour le faire nous réécrivons les deux équation (de la demande et de l'offre respectivement) sous la forme suivante :

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

$$Qd = \sum_{i=1}^N (\alpha_i + \beta_i Pd_i + \gamma_i Z_i) = \sum_{i=1}^N \alpha_i + \sum_{i=1}^N \beta_i Pd_i + \sum_{i=1}^N \gamma_i Z_i$$

Ce qui nous produira un système des deux équations, avec $Qd = Qo$ dans la situation d'équilibre :

$$Qd = \sum_{i=1}^N \alpha_i + \sum_{i=1}^N \beta_i Pd_i + \sum_{i=1}^N \gamma_i Z_i$$

$$Qo = \sum_{i=1}^N a_i + \sum_{i=1}^N b_i Po_i + \sum_{i=1}^N c_i X$$

Néanmoins, ce cas se réleve d'être très complexe. D'abord, les effets peuvent être différents pour tous les départements, ce qui nous conduira à une augmentation dans le nombre des paramètres à estimer significative. De plus, même si tous les effets sont identiques pour l'ensemble des départements, des contraintes au niveau des données peuvent se révéler d'être trop restrictives en réduisant au néant la puissance statistique de notre estimateur (ex : le nombre des observation par années très faible). Dans le deux cas nous faisons face à un impasse.

Une des modification possibles dans ce cas sera 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 d'équations, nous pouvons supposer, que tout le vin produit dans un département est consommé dans le même département. Dans ce cas nous pourrions obtenir des estimateurs pour les effets moyens au niveau départemental. Toutefois, c'est une supposition forte, laquelle nous éloigne de la réalité.

Théoriquement, 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 en estimant le coefficient moyen unique pour tous les départements nous allons mitiger les biais possibles.

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

$$\begin{aligned} qd_i &= \alpha_i + \beta Pd_{i,d} + \gamma Z_i \\ qo_i &= a_i + bPo_{i,o} + cX_i \end{aligned}$$

Où $qd_i = qo_i$ et $Pd_i = Po_i$, ce qui permet de rélier les équations au niveau départemental. Les coefficients b , c , β et γ sont supposé fixes pour tous les départements nous donnent un estimateur des effets moyens au niveau de la France. L'effet des pesticides dans la production du vin seront captés par le terme c_1 : $c = (c_1, c_2)$ dans ce cas.

Néanmoins, nous nous posons la question, comment réagir dans le cas où les effets sont différents pour des différents départements suite à des spécificité des marchés locaux, géographiques ou autres ? On peut supposer, qu'il existent au moins quelques groupes majeurs ayant des caractéristiques et comportements similaires. Dans ce cas nous pourrions construire des clusters, qui regroupent des départements ayant des caractéristiques identiques. Cela nous permettra de modéliser les effets moyens par cluster en réduisant les biais éventuels.

Ce système peut être formalisé par K systèmes d'équations suivantes :

$$\begin{aligned} qd_{i_{c=const}} &= \alpha_{i_{c=const}} + \beta_{c=const} Pd_{i_{c=const},d} + \gamma_{c=const} Z_{i_{c=const}} \\ qo_{i_{c=const}} &= a_{i_{c=const}} + b_{c=const} Po_{i_{c=const},o} + c_{c=const} X_{i_{c=const}} \end{aligned}$$

Où c décrit l'appartenance des départements à un des groupes (clusters).

4. Les données

Avant de passer à la discussion des modèles économétriques il nous faut prendre connaissance de la nature des données en notre disposition. Dans cette partie de notre travail nous allons présenter la base des données utilisé lors de cette étude. Nous commencerons par une présentation des sources et des types des données extraits de ces sources. Puis, nous procéderons avec la description des méthodes et techniques utilisées pour transformer ces données et les rendre traitables. Finalement, nous présenterons un dictionnaire des variables pour nos bases des données.

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)

Les variables utilisées pour notre modèle

Révérier tous les sources et la nature des données ...

Expliciter la procédure de création des variables

Preciser les effets attendus des variables

Discuter les externalités (ou c'est mieux de l'inclure dans la partie théorique ? ou contextualisation ? A VOIR)

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 période 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).

Toutes les variables subissent une transformation logarithmique, ce qui nous permet d'interpréter les effets estimés plus facilement. Pour un modèle logarithmique nous pourrions traiter les estimateurs obtenus comme l'élasticité de la demande/l'offre par rapport à des facteurs différents. Ainsi, nous cherchons particulièrement l'élasticité de quantité offerte sur le marché par rapport à la quantité des pesticides utilisés.

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 simple et qui utilisent des pesticides) et le nombre des périodes pauvre (5 périodes).

5. 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. Puis, nous allons finir avec étude des données alternées par une transformation **within**.

Visualisation au niveau de la France

Pour le premier analyse il peut être intéressant de voir la situation du point de vue géographique. Nous visualisons les valeurs moyens par département des différentes variables (une partie de la représentation se trouve dans l'annexe X).

D'abord nous étudions le comportement de la variable dépendante de notre système. La quantité du vin sans IG produit par département semble pouvoir être corréliée à partir de la figure suivante.

Quantité du vin produite par département

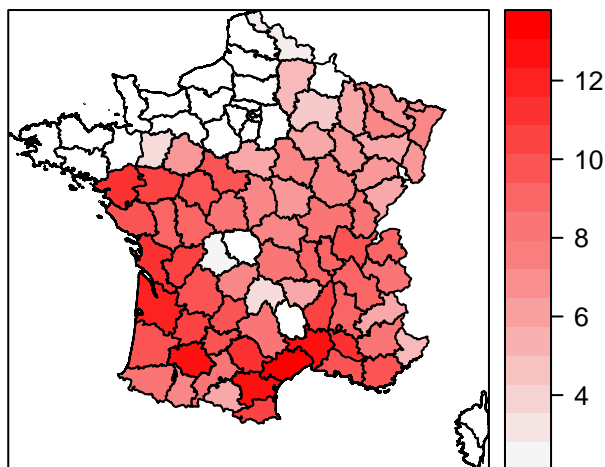


Figure 1: Les quantités du vin non-IG moyennes par département

Puis, nous observons le comportement du reste des variables (les représentations graphiques sont groupées dans l'annexe X). L'indice des prix se comporte pratiquement comme la quantité du vin produite, car cet indice fut construit par biais de cette variable. Les autres moyennes ne semblent pas avoir des structures corrélées dans l'espace au niveau de la France. Dans notre analyse nous nous laissons liberté d'ignorer les effets possibles d'autocorrélation spatiale dans nos données, parce que au moment de constructions de notre base des données nous avons ignoré les départements ne produisant pas le vin simple, mais qui peuvent quand même jouer son rôle si nous étions à prendre en compte la structure spatiale de nos données.

Etude de la variance

Passons maintenant à l'étude de la variance. Nous allons décortiquer la variance par type (between et within) afin d'obtenir une idée sur le choix préférable de la dimension d'aggrégation de nos données, car il peut se révéler que la théorie ne correspond pas à la réalité (ex: nous faisons face aux effets fixes par année et non par département).

Le tableau suivant regroupe les statistiques descriptives essentielles :

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

Table 1: Etude de la variance

	Mean	Overall	Between	Within
Index prix	1.431	1.339	1.012	0.883
Index pesticides	1.257	0.483	0.335	0.350
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

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 départements). Ce qui est complètement conforme à notre hypothèse qu'on a exprimé lors de la formalisation du modèle économique théorique.

De plus, il est intéressant d'observer les résultats obtenus pour le test de Chow comparant le modèle complet (*pooled model*) contre les modèles à effet fixes et aléatoires. Le tableau suivant regroupe les p-valeurs de ce test pour les modèles univariés différents.

Table 2: Les p-valeurs de pooling-test de Chow

	Random	Fixed
Index prix	0	0
Index pesticides	0.354	0.294
Surface	0	0.0001
Revenus	0.297	0.247

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

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 (individuels, temporels ou en double dimension). Selon les résultats des tests il est difficile de choisir arbitrairement un type d'effets. Il est évident que nous avons des effets fixes au niveau individuel ou des fixes en double dimension pour toutes les variables.

Table 3: p-valeurs de Lagrange multiplier test

	Individual	Time	Twoways
Index prix	0	0.256	0
Index pesticides	0	0.229	0
Surface	0	0.030	0
Revenus	0	0.248	0

Selon les résultats obtenus, ainsi que les evidences théoriques des études antérieures nous décidons de ne garder que les effets fixes au niveau individuel afin de faciliter l'analyse.

L'analyse de la correlation

Dans le tableau ci-dessous nous presentons les correlations des variables après la correction pour les effets fixes individuels (nous effectuons la transformation *within* sur nos données en substrayant les moyennes individuelles pour l'ensemble des variables). Dans les annexes nous proposons également un tableau de correlation pour les données non-transformées, ce qui permet d'observer les inégalités et une pauvre représentativité des liens entre les variables pour les données initiales.

Particulièrement nous pouvons remarquer une forte correlation entre la quantité offerte et le prix d'équilibre. Egalement ...

6. Modélisation

Séparer les modèles (OLS, 3SLS avec justification par 2SLS et la comparaison avec i3SLS, clusters en OLS et 3SLS).

Justifier le choix des modèles par 3 cas théoriques. Discuter les avantages et les inconveniences

Ajouter des liens avec des études méthodologiques précédents.

Pour le modèle 2SLS préciser la forme, tester les instruments

Arbitrage du choix de 2SLS vs 3SLS

Cette partie du travail abordera la formulation économétrique de notre problème. Nous allons débiter par la présentation des notions théoriques implémentés dans ce travail, suivis par la formalisation économétrique du modèle théorique que nous avons spécifié dans la section 5. Après, nous expliquerons la stratégie d'identification utilisée.

Présentation de la méthodologie

L'AIDS (*almost ideal demand system*) 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, Caracciolo, and Pomarici (2014). Quand même, dans notre étude nous allons utiliser un approche similaire à ce modèle là, sous des suppositions restrictives.

Ce modèle nous permettra de simuler l'équilibre sur le marché du vin, prenant ainsi en compte la plupart des facteurs incitant les producteurs du vin d'utiliser les pesticides.

Modèle économétrique

Dans cette section nous allons presenter une par une nos modèles économétriques correspondant chacune à un des trois cadres théoriques possibles. Tous les modèles visent à estimer les effets moyens pour tous les départements sous des hypothèses différentes du fonctionnement du marché. Dans tous les cas, l'aggregation des effets au niveau national (ou au niveau des groupes) nous permet de mitiger les biais eventuels, liés à la misspecification du modèle.

Pour le cadre où nous n'observons pas des interactions entre la demande et l'offre sur le marché (M1), nous ésimons un modèle simple. Nous écrivons notre modèle sous la forme suivante :

$$qo_{i,t} = a_1 + bPo_{i,t} + cX_{i,t} + u_{i,t}$$

A ce point nous avons un choix : soit nous supposons que les agriculteurs sont des preneurs des prix, ce qui nous permet de traiter le prix comme une variable exogène; soit nous devrions construire un estimateur IV afin de traiter l'endogénéité eventuelle de l'index des prix. Evidement le premier cas est le plus simple, mais pour justifier l'implementation de cette méthode nous devrions effectuer des tests d'énogénéité des prix. Le deuxième cas est beaucoup plus réaliste, puisque les viticulteurs sont rarement preneurs des prix et l'offre aussi joue son rôle sur l'équilibre du marché.

Dans la dernière situation nous utilisons les idées de MacKay and Miller (2018), supposant que les variables déterminant la demande sont des instruments fiables pour la prédiction des variables endogènes dans l'équation d'offre (bien que dans notre cas nous ignorons les effets des interactions entre l'offre et la demande). Particulieremnt ici nous pourrions utiliser les données sur les revenus afin d'instrumenter le niveau des prix (l'indice des prix du vin).

Passons maintenant au modèle plus complexe (M2), basé sur l'hypothèse que la demande influence l'offre, affectant également le mode d'utilisation des pesticides par les agriculteurs. Nous pouvons réécrire notre système d'équations dans ce cas sous la forme suivante :

$$\begin{aligned} qo_{i,t} &= a_1 + bPo_{i,t} + cX_{i,t} + u_{i,t} \\ qd_{i,t} &= \alpha_i + \beta Pd_{i,t} + \gamma Z_{i,t} + \epsilon_{i,t} \end{aligned}$$

Nous posons que l'offre et la demande sont égaux au niveau de département : $qd_{i,t} = qo_{i,t}$. C'est à dire l'offre interne du 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. En d'autre mots, le système (qui implique : $Qd = Qo$) :

$$qd_{i,t} = qo_{i,t}$$

Au point d'équilibre nous rencontrons également l'égalité des prix :

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

De cette façon nous obtenons un système des systèmes des équations. En simplifiant l'écriture nous pouvons la représenter sous la forme suivante :

$$\begin{aligned} q_{i,t} &= \alpha_i + \beta P_{i,t} + \gamma Z_{i,t} + \epsilon_{i,t} \\ q_{i,t} &= a_i + bP_{i,t} + cX_{i,t} + u_{i,t} \end{aligned}$$

Et finalement, nous pouvons estimer les deux modèles (M1 et M2) en regroupant les départements par leurs caractéristiques. Appelons ces modèles M3.1 et M3.2 respectivement.

Le premier prenant la forme :

$$qo_{i,t} = a_1 + bPo_{i,t} + cX_{i,t} + u_{i,t}$$

Tandis que le dernier :

$$\begin{aligned} q_{i_c,t} &= \alpha_{i_c} + \beta P_{i_c,t} + \gamma Z_{i_c,t} + \epsilon_{i_c,t} \\ q_{i_c,t} &= a_{i_c} + bP_{i_c,t} + cX_{i_c,t} + u_{i_c,t} \end{aligned}$$

Avec c décrivant l'appartenance de département à un des clusters.

Pour finir cette partie, résumons que nous avons à notre disposition plusieurs chemins différents à traiter ce modèle du point de vue économétrique. Le plus simple est d'estimer les effets des pesticides sur l'offre du vin en ignorant les impacts du comportement des consommateurs sur les producteurs. Cette méthode implique une estimation par OLS simples (ou IV-OLS, lesquels introduisent la notion d'endogénéité des prix). D'autre côté, nous pouvons implémenter les triples moindres carrés (nous devrions comparer les résultats obtenus avec un système d'équations non-réligées, estimé par 2SLS afin de traiter l'endogénéité), qui nous permettront d'obtenir des résultats identiques aux résultats d'estimations des équations structurelles sous l'hypothèse des interactions entre l'offre et la demande. Cette méthode offre la possibilité d'estimer le système d'équations avec plusieurs variables endogènes en prenant en compte les deux côtés du marché à la fois. Finalement, si on trouve qu'il y existe une hétérogénéité entre les départements en termes d'équilibre interne, nous pourrions réestimer les modèles en clusterisant nos *individus* (départements) par des différentes classes selon leurs attributs, pour après estimer les équations par cluster.

Hypothèses sur les résultats

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

les plus fiables. Cette méthode nous permet à dépasser le biais de simultanéité qui apparaisse dans le cas d'estimation des systèmes d'équations liées (dans notre cas nous étudions les effets des pesticides sur l'offre et production du vin simple sous hypothèse de présence des effets du marché). L'estimateur pareil donne des résultats similaires à l'estimateur de ILS (*indirect least squares*). De plus, sa version itérée (qui converge à des résultats similaires à ceux obtenus par l'estimation avec maximum de vraisemblance) donne des résultats avec la moindre variance.

Les propriétés de cet estimateurs sont :

- Consistence ;
- Efficience (asymptotique) ;
- La distributions pour les estimateurs suit une loi normale suelement dans des grands échantillons.

Quand même dès le debut nous envisageons que cet estimateur ne refletera pas la nature du marché. C'est pourquoi nous, dans ce travail, testons plusieurs modèles.

Parmis les inconveniences eventuelles on a également la faible representation des effets hetérogenes entre les départements par le modèle. Nous estimons seulement les effets moyens et ainsi ignorons les differences des élasticités pour des départements differents. Hereusement ce problème peut être rémédiee par l'introduction des clusters, regroupant des département ayant le comportement similaire.

Finalement, il existe des effets qu'on ignore completement, mais qui risquent d'intervenir. Par exemple, nous ignorons la présence d'autocorrelation spatiale et/ou temporelle dans notre modèle. Egalement, un nombre probablement insuffisant des facteurs est utilisé dans ce modèle, ce qui risau d'apporter le biais des variables omises dans nos estimations.

7. Résultats des estimations

Dans cette séction nous allons presenter les résultats économétriques pour des differents modèles ainsi que les comparer.

Nous estimons un enseble des differents modèles possibles afin de pouvoir choisir la méthode la plus raisonnable. Les modèles suivantes sont traitées séparément :

- M1 : modèle simple sans interactions entre l'offre et la demande ;
- M2 : modèle complexe visant à integrer les interactions entre l'offre et la demande en presence des variables endogènes ;
- M3 : les modèles sur les données clustérisés (M3.1 et M3.2 respectivement pour les deux cas précédents).

Les résultats en absence des interactions

	OLS	WLS	SUR
Demande: ipi	0.33*** (0.02)	0.33*** (0.02)	0.32*** (0.02)
Demande: ri	-9.00*** (1.46)	-9.00*** (1.46)	-1.10 (0.66)
Offre: ipi	0.30*** (0.02)	0.30*** (0.02)	0.32*** (0.02)
Offre: si	0.23*** (0.04)	0.23*** (0.04)	0.03 (0.02)
Offre: iki	-0.16*** (0.05)	-0.16*** (0.05)	-0.02 (0.02)
Demande: R ²	0.50	0.50	0.46
Offre: R ²	0.52	0.52	0.46
Demande: Adj. R ²	0.50	0.50	0.45
Offre: Adj. R ²	0.52	0.52	0.46
Num. obs. (total)	690	690	690

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

Table 4: Statistical models

Les résultats 2SLS, W2SLS, 3SLS et i3SLS

	2SLS	W2SLS	3SLS	i3SLS
Demande: ipi	0.79*** (0.15)	0.79*** (0.15)	0.79*** (0.15)	0.79*** (0.15)
Demande: ri	-13.07*** (2.76)	-13.07*** (2.76)	-13.07*** (2.76)	-13.07*** (2.76)
Offre: ipi	-0.28 (0.25)	-0.28 (0.25)	-0.25 (0.25)	-0.25 (0.24)
Offre: si	0.47*** (0.13)	0.47*** (0.13)	0.45*** (0.13)	0.45*** (0.12)
Offre: iki	-0.11 (0.09)	-0.11 (0.09)	-0.17* (0.08)	-0.17* (0.08)
Demande: R ²	-0.41	-0.41	-0.41	-0.41
Offre: R ²	-0.87	-0.87	-0.74	-0.75
Demande: Adj. R ²	-0.42	-0.42	-0.42	-0.42
Offre: Adj. R ²	-0.89	-0.89	-0.75	-0.76
Num. obs. (total)	690	690	690	690

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

Table 5: Statistical models

Clusterisation et modélisation par groupe

Between

Nous avons vus 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.

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.

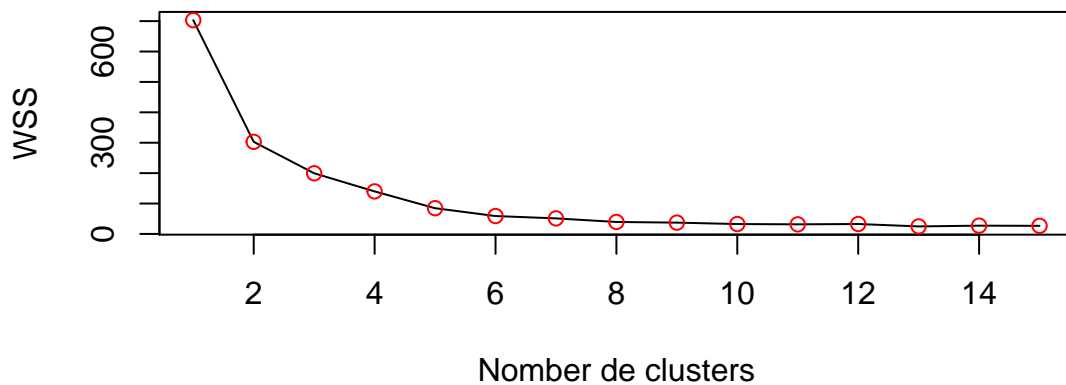


Figure 2: Le choix des clusters

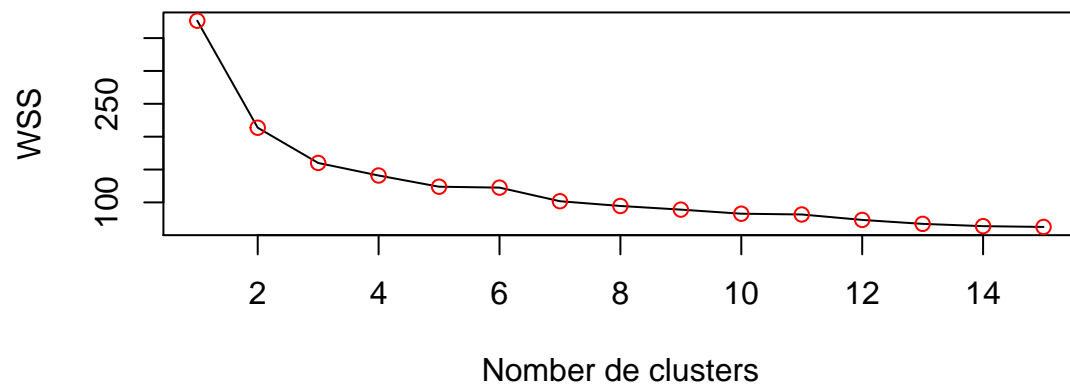
Within

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.

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

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

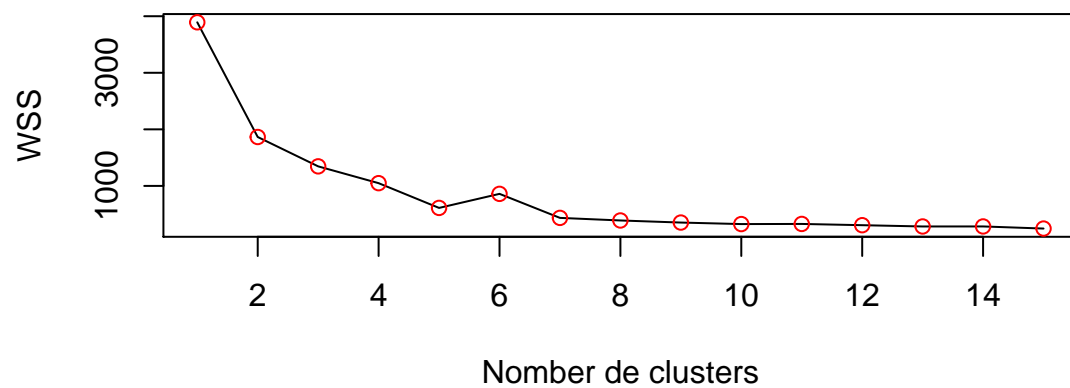
Nous pouvons supposer que le nombre des clusters optimal est entre 6 et 15.



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.



Modélisation

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

Les résultats obtenus sont suivants :

	OLS	2SLS	3SLS
Demande: ipi	0.33*** (0.02)	0.15*** (0.03)	0.16*** (0.03)
Demande: ri1	-1.69 (1.92)	0.99 (2.17)	0.48 (1.19)
Demande: ri2	-17.14*** (2.09)	-18.38*** (2.35)	-3.30** (1.26)
Demande: ri3	-22.12* (10.47)	13.13 (12.23)	13.93 (12.17)
Offre: ipi	0.37*** (0.02)	0.16*** (0.03)	0.19*** (0.03)
Offre: si1	0.14 (0.08)	0.21* (0.09)	0.03 (0.05)
Offre: si2	0.26*** (0.04)	0.31*** (0.05)	0.06* (0.03)
Offre: si3	-2.58*** (0.37)	-0.68 (0.47)	-0.58 (0.46)
Offre: iki1	-0.06 (0.08)	0.00 (0.09)	-0.01 (0.05)
Offre: iki2	-0.18*** (0.05)	-0.20*** (0.06)	-0.04 (0.03)
Offre: iki3	13.81*** (3.21)	8.01* (3.74)	4.03 (2.99)
Demande: R ²	0.54	0.42	0.37
Offre: R ²	0.59	0.46	0.42
Demande: Adj. R ²	0.54	0.42	0.36
Offre: Adj. R ²	0.59	0.45	0.41
Num. obs. (total)	690	690	690

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

Table 6: Statistical models

9. Conclusions

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

Le marché du vin

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

Le rôle des pesticides

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

Validité

- Faible validité du modèle économétrique
 - Variables omises

Annexes

A Les statistiques descriptives

A1 Les moyennes par département

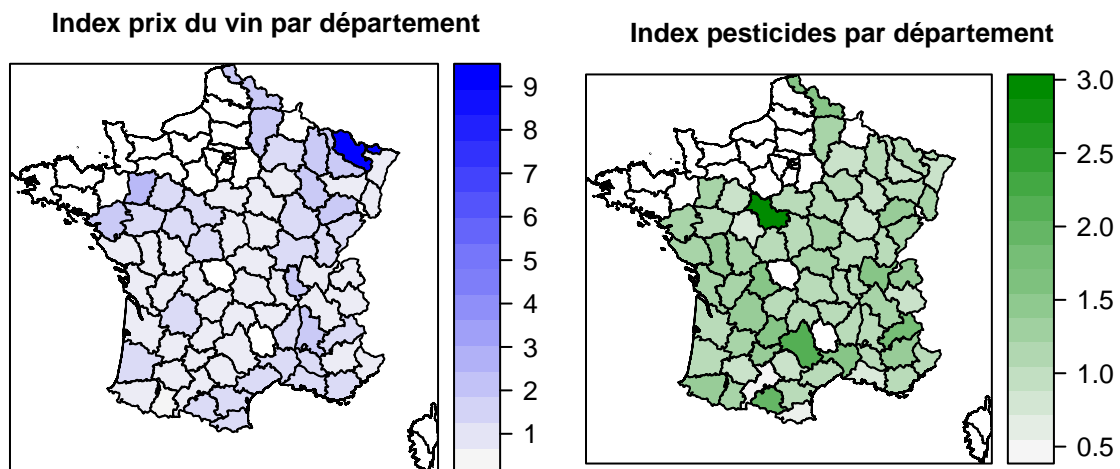


Figure 3: Les valeurs moyennes par département, partie 1

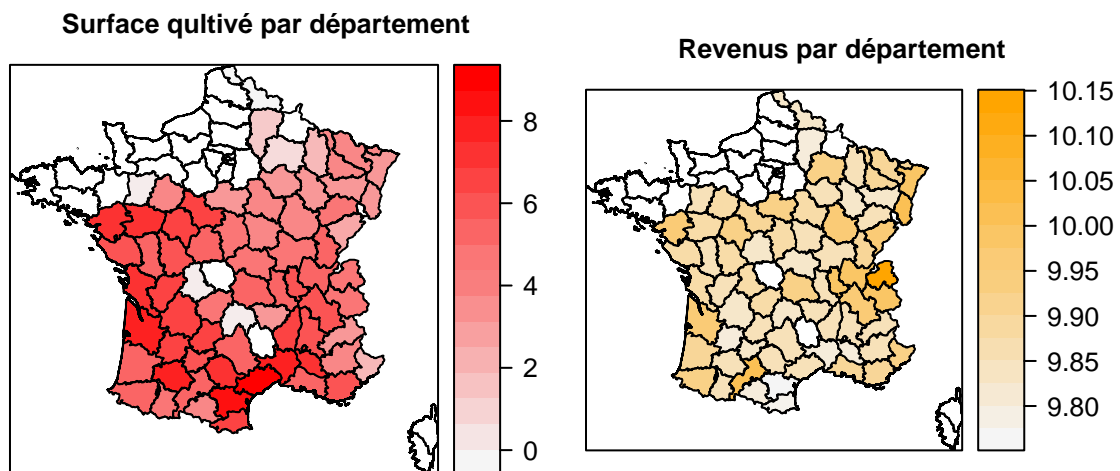


Figure 4: Les valeurs moyennes par département, partie 2

A2 Les graphiques bivariés

Cas général

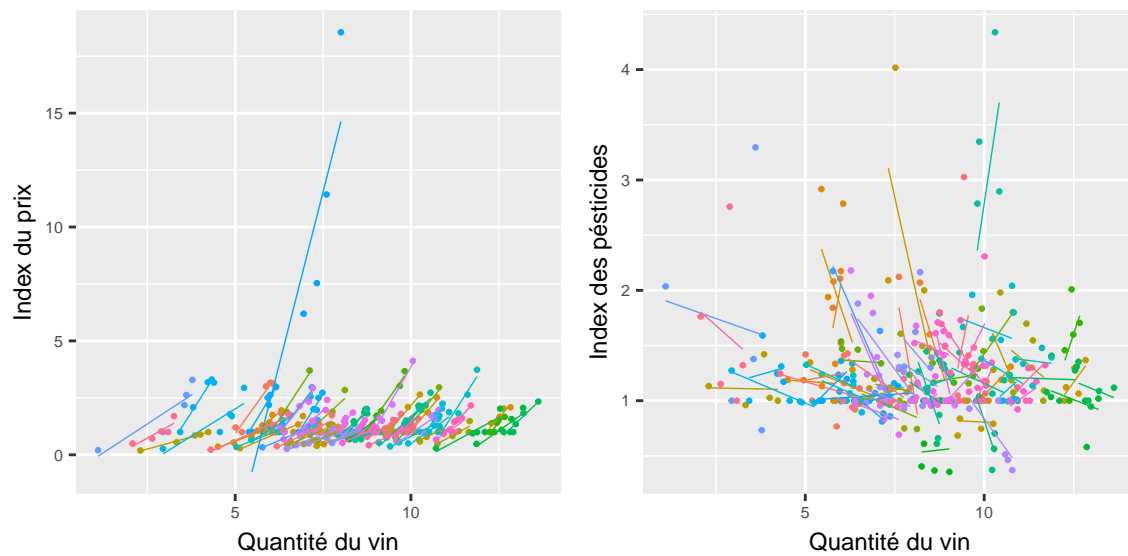


Figure 5: L'étude bivarié, partie 1

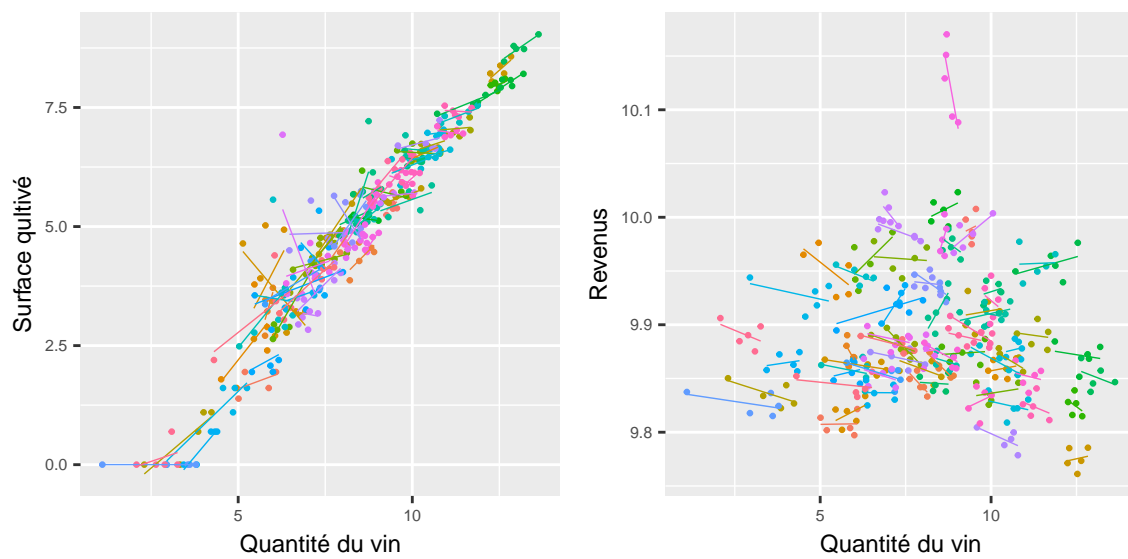


Figure 6: L'étude bivarié, partie 2

Transformation *Within*

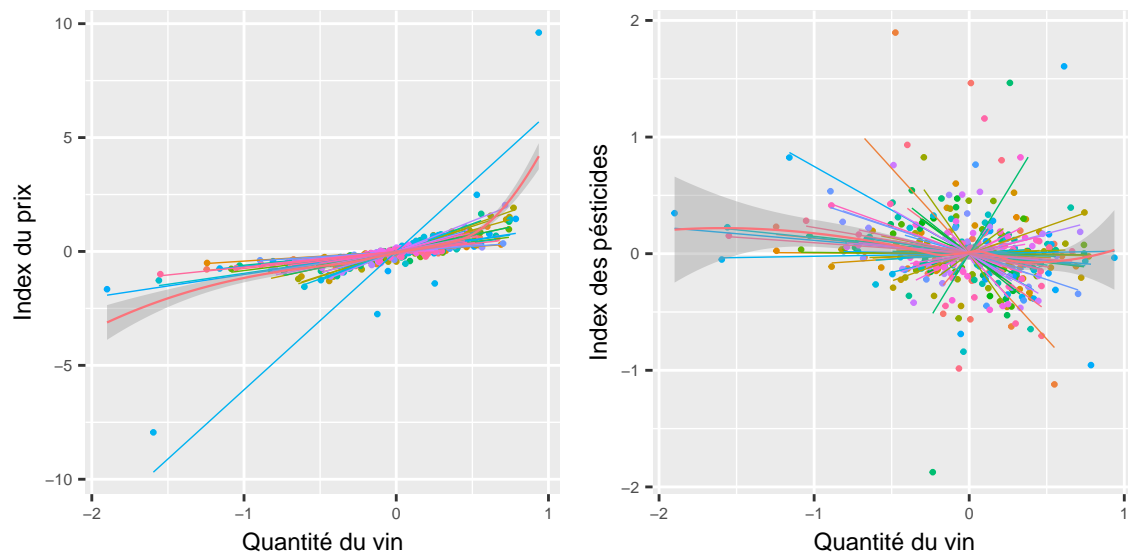


Figure 7: Relations bivariés dans le cas de transformation within, partie 1

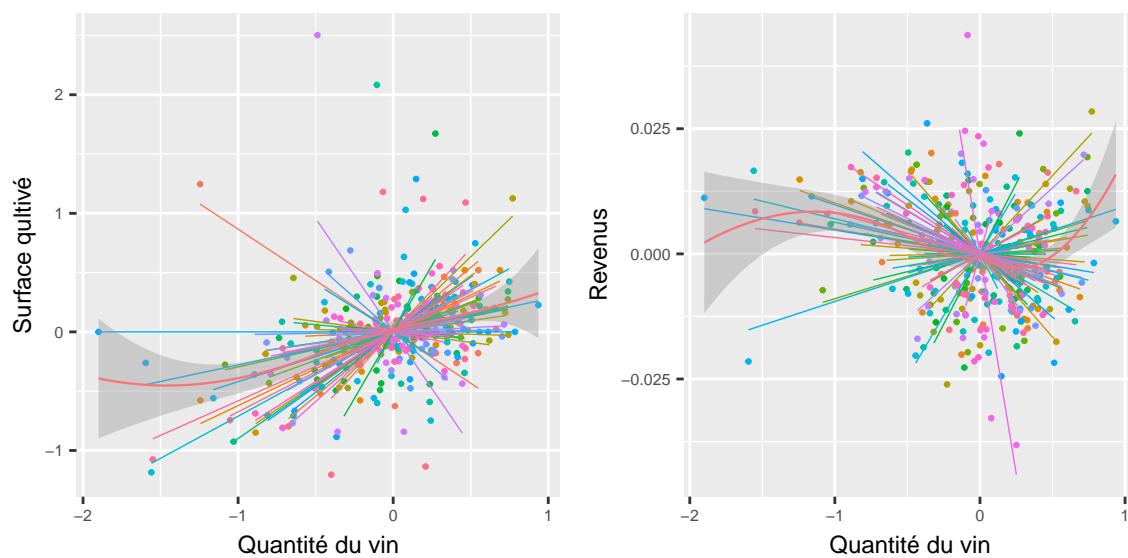


Figure 8: Relations bivariés dans le cas de transformation within, partie 2

A3 La correlation

Cas général

Le premier tableau comprend les résultats pour les données telles-queles, 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	1.0000	0.0177	0.9559	-0.0266	-0.0667	-0.0360
IP	0.0177	1.0000	-0.0513	0.0065	-0.0590	0.1082
Surface	0.9559	-0.0513	1.0000	-0.0567	-0.0486	-0.0640
Revenus	-0.0266	0.0065	-0.0567	1.0000	-0.0433	0.1188
Index pesticides	-0.0667	-0.0590	-0.0486	-0.0433	1.0000	0.2971
Temps	-0.0360	0.1082	-0.0640	0.1188	0.2971	1.0000

Transformation *Within*

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

	Quantité du vin	IP	Surface	Revenus	Index pesticides	Temps
Quantité du vin	1.0000	0.6656	0.3655	-0.1601	-0.1813	-0.1994
IP	0.6656	1.0000	0.1862	0.1119	-0.0108	0.1640
Surface	0.3655	0.1862	1.0000	-0.1657	-0.2035	-0.3103
Revenus	-0.1601	0.1119	-0.1657	1.0000	0.2103	0.6522
Index pesticides	-0.1813	-0.0108	-0.2035	0.2103	1.0000	0.4100
Temps	-0.1994	0.1640	-0.3103	0.6522	0.4100	1.0000

B Analyse des résultats OLS, WLS et SUR

B1 Indépendance des résidus

	OLS D	OLS O	WLS D	WLS O	SUR D	SUR O
Vin	0.7080	0.6932	0.7080	0.6932	0.7458	0.7454
IP	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Surface	0.2786	0.0000	0.2786	0.0000	0.3202	0.2841
Revenus	0.0000	-0.2389	0.0000	-0.2389	-0.2793	-0.3067
Pesticides	-0.1749	0.0000	-0.1749	0.0000	-0.2277	-0.2021

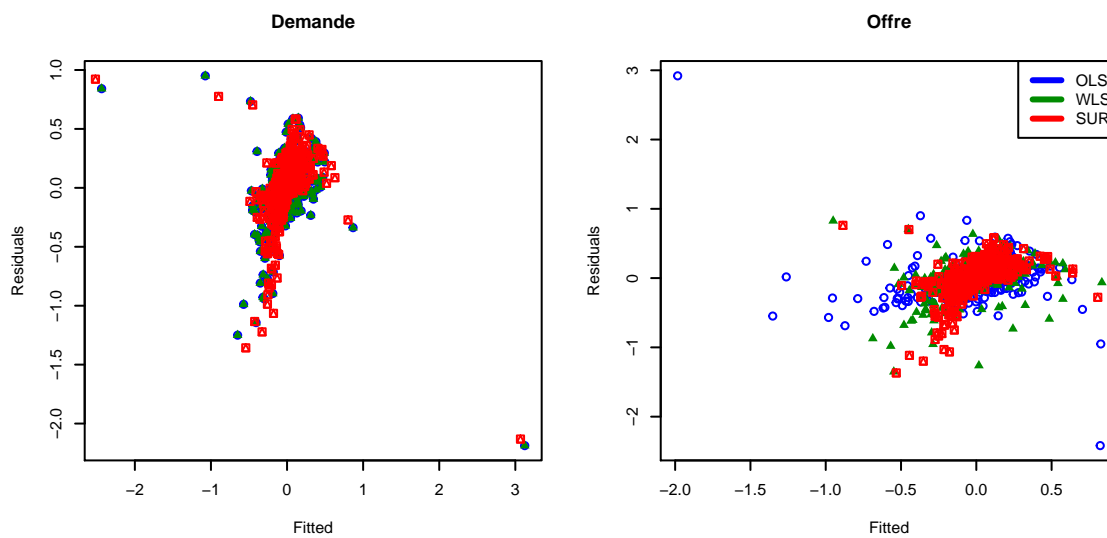


Figure 9: Les résidus contre la variable prédite

B2 L'autocorrelation

Table 10: Les statistiques test de Durbin-Watson

	OLS	WLS	SUR
Equation de demande	0.653	0.653	0.687
Equation d'offre	0.627	0.627	0.683

B3 Test de l'hétéroskedacité

Table 11: Les résultat du test de Bartlett sur l'heteroscedacité

	OLS	WLS	SUR
Equation de demande	0	0	0
Equation d'offre	0	0	0

B4 La normalité des résidus

Table 12: Shapiro-Wilk test de normalité des résidus

	OLS	WLS	SUR
Equation de demande	0	0	0
Equation d'offre	0	0	0

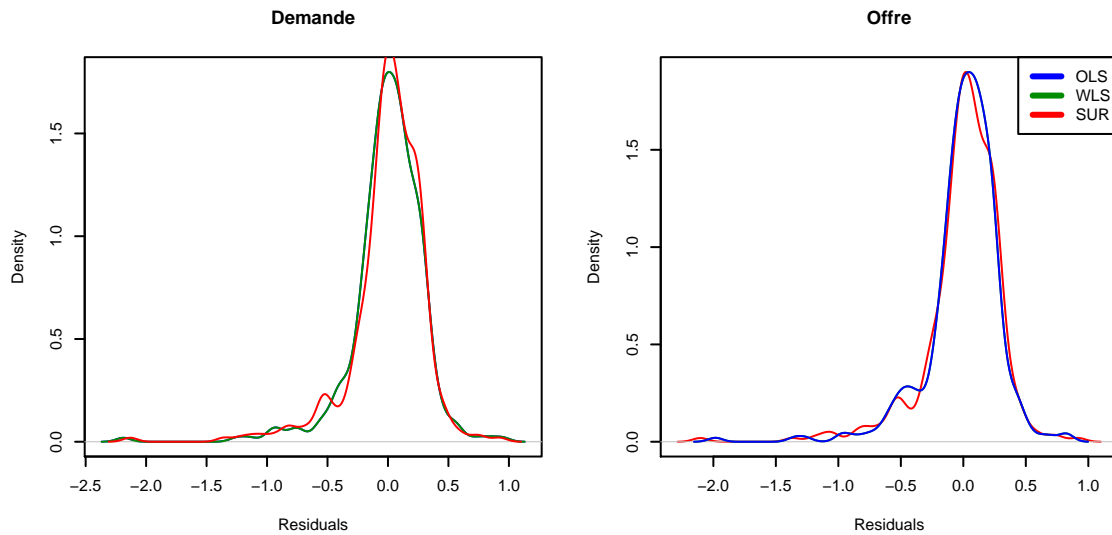


Figure 10: Les PDF des résidus

C Analyse des résultats 2SLS, W2SLS, 3SLS et i3SLS

C1 Indépendance des résidus

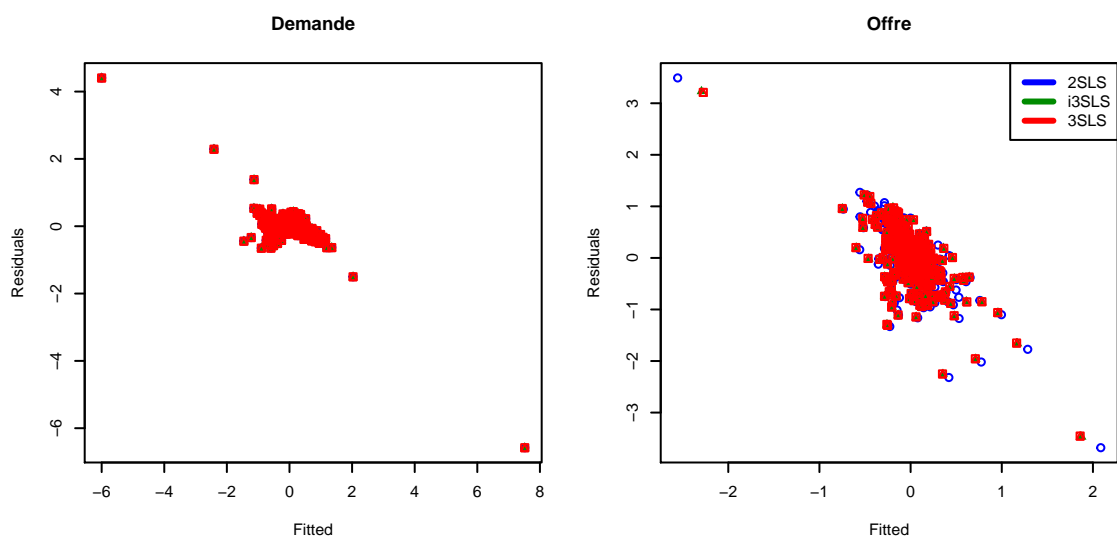


Figure 11: Les résidus contre la variable prédite

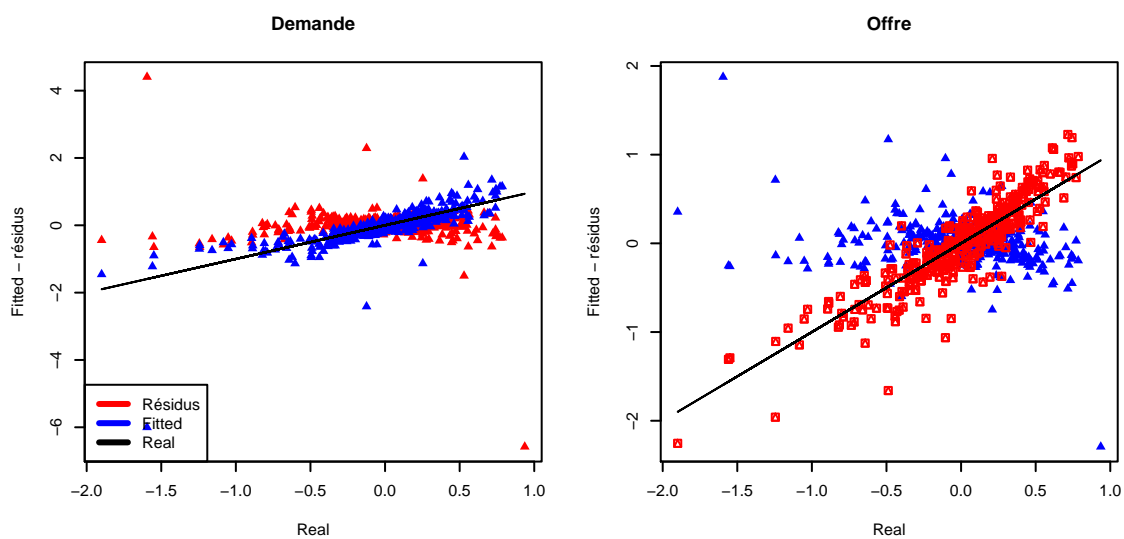


Figure 12: Les résidus et les prédictions, le cas de i3SLS

C2 L'autocorrelation

Table 13: Les resultats du test de Durbin-Watson

	2SLS	3SLS	i3SLS
Equation de demande	0.618	0.618	0.618
Equation d'offre	0.637	0.638	0.638

C3 Test de l'hétéroskedacité

% 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. 23, 2019 - 20:15:19

Table 14: Test de Bartlett sur l'heteroskedacité

	2SLS	3SLS	i3SLS
Equation de demande	0	0	0
Equation d'offre	0	0	0

C4 La normalité des résidus

Table 15: Shapiro-Wilk test de normalité

	2SLS	3SLS	i3SLS
Equation de demande	0	0	0
Equation d'offre	0	0	0

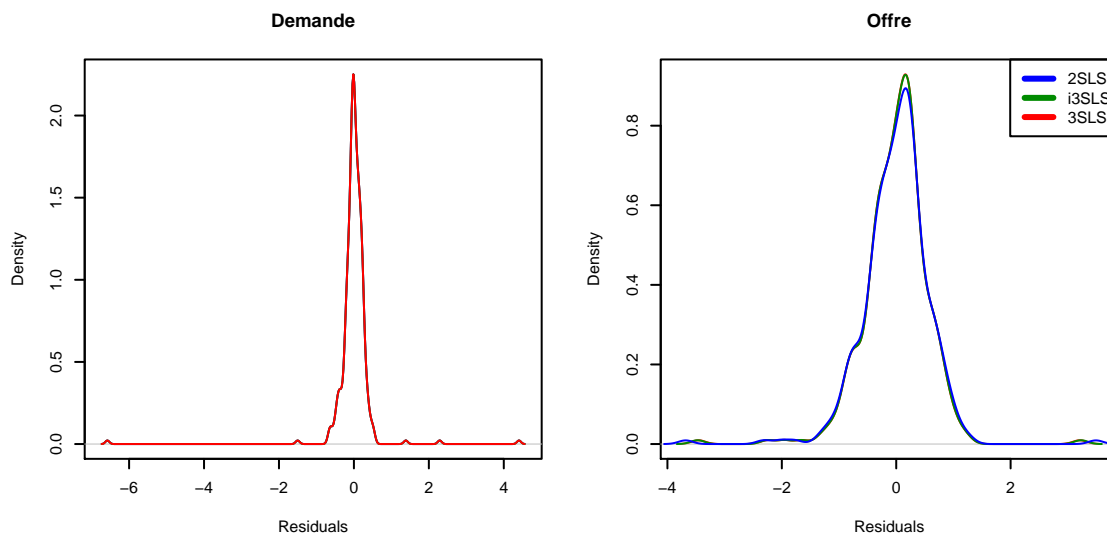


Figure 13: Les PDF des résidus

C5 Comparaison des modèles

Table 16: Hausman 3SLS consistency test

	Test	Resultats
1	2SLS contre 3SLS	0.827
2	2SLS contre i3SLS	0.910

D Clusterisation

D1 *Between* transformation

Les groupes sont définies par des caractéristiques suivantes :

Table 17: Les centres des clusters

	qi	ipi	si	ri	iki	.1
1	10.609	1.393	6.705	9.880	1.273	27
2	7.468	1.416	4.345	9.905	1.238	34
3	4.340	1.622	1.098	9.870	1.284	8

D2 *Within* transformation

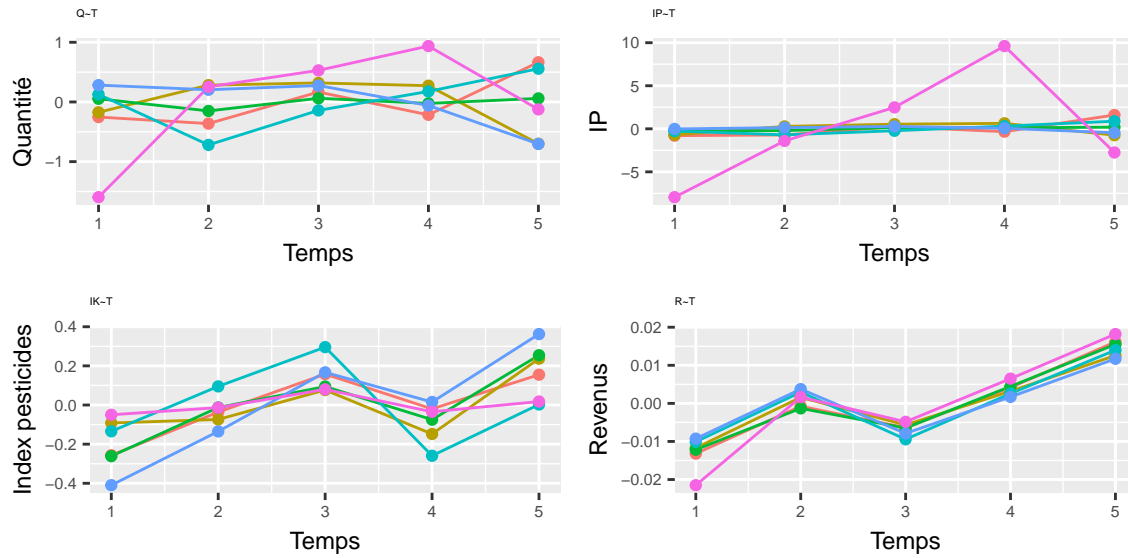
Les centres

Les groupes sont définies par des caractéristiques suivantes :

Table 18: Les centres des clusters

	qi	ipi	si	ri	iki	n	k	t
1	-0.252896	-0.793219	-0.07542	-0.013193	-0.256673	5	1	1
2	-0.362457	-0.716949	0.041055	-0.000747	-0.036528	5	1	2
3	0.166103	0.2392	-0.013413	-0.006388	0.157998	5	1	3
4	-0.215706	-0.332061	-0.199289	0.004172	-0.019761	5	1	4
5	0.664955	1.603029	0.247068	0.016155	0.154963	5	1	5
6	-0.173176	-0.700641	-0.063262	-0.011699	-0.091867	16	2	1
7	0.281974	0.276242	0.030031	0.001569	-0.073717	16	2	2
8	0.318986	0.534754	-0.017193	-0.005653	0.076719	16	2	3
9	0.272819	0.635118	0.096895	0.003199	-0.147998	16	2	4
10	-0.700603	-0.745473	-0.046471	0.012585	0.236862	16	2	5
11	0.053201	-0.250299	0.089817	-0.012167	-0.261237	20	3	1
12	-0.150237	-0.20443	0.168342	-0.001309	-0.012628	20	3	2
13	0.061586	0.149078	0.01519	-0.006477	0.094218	20	3	3
14	-0.023929	0.120274	-0.129925	0.004456	-0.074835	20	3	4
15	0.059379	0.185377	-0.143424	0.015498	0.254483	20	3	5
16	0.126897	-0.260636	0.047425	-0.010084	-0.134339	5	4	1
17	-0.719432	-0.703363	-0.254722	0.002955	0.095193	5	4	2
18	-0.141159	-0.225144	-0.158017	-0.009434	0.29572	5	4	3
19	0.176733	0.308934	-0.1592	0.002576	-0.259209	5	4	4
20	0.556961	0.880209	0.524514	0.013987	0.002635	5	4	5
21	0.282059	-0.014017	0.400864	-0.009269	-0.409711	22	5	1
22	0.204605	0.138187	0.285719	0.003743	-0.134925	22	5	2
23	0.274731	0.275854	0.118099	-0.007893	0.166927	22	5	3
24	-0.058265	0.063243	-0.18248	0.001702	0.015278	22	5	4
25	-0.70313	-0.463267	-0.622202	0.011716	0.362431	22	5	5
26	-1.595397	-7.941881	-0.261529	-0.02149	-0.049931	1	6	1
27	0.253285	-1.404496	-0.153318	0.001641	-0.013255	1	6	2
28	0.529538	2.487263	0.747453	-0.004844	0.080087	1	6	3
29	0.936501	9.608734	0.226164	0.006494	-0.034652	1	6	4
30	-0.123927	-2.749621	-0.55877	0.018199	0.017751	1	6	5

Representation graphique



D3 Cas d'information complete

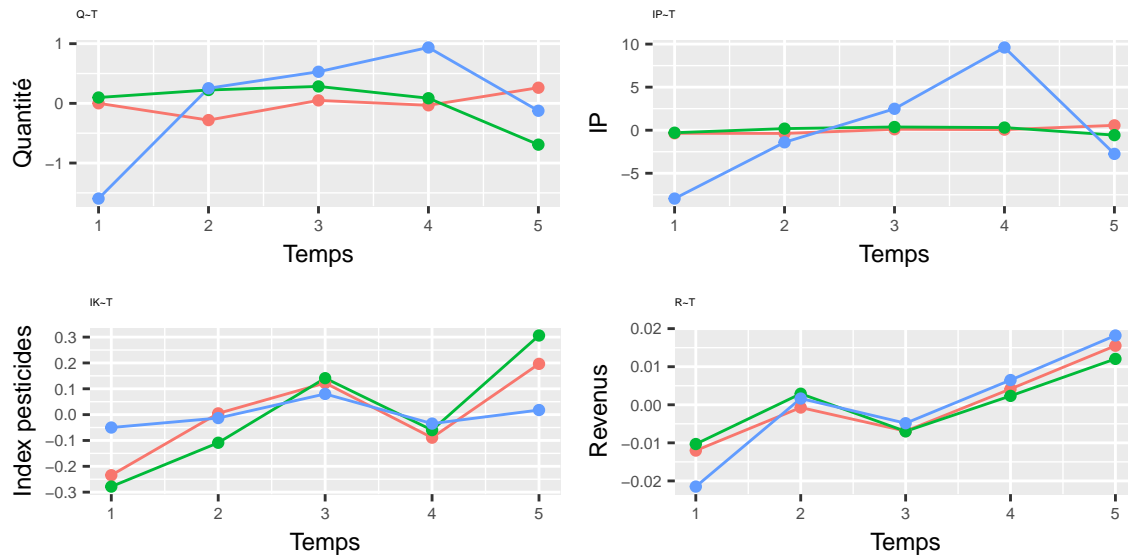
Les centres

Les groupes sont définies par des caractéristiques suivantes :

Table 19: Les centres des clusters

	qi	ipi	si	ri	iki	n	k	t
1	0.001186	-0.359222	0.04895	-0.012025	-0.234518	29	1	1
2	-0.280464	-0.376425	0.078291	-0.000723	0.004804	29	1	2
3	0.049743	0.107731	-0.01841	-0.006913	0.122468	29	1	3
4	-0.032737	0.065232	-0.147634	0.004155	-0.089222	29	1	4
5	0.262272	0.562683	0.038803	0.015506	0.196468	29	1	5
6	0.098311	-0.291704	0.206247	-0.01031	-0.27852	39	2	1
7	0.223902	0.18426	0.174215	0.002905	-0.108882	39	2	2
8	0.283634	0.373141	0.059067	-0.006982	0.141058	39	2	3
9	0.086134	0.306445	-0.065995	0.002333	-0.059896	39	2	4
10	-0.691981	-0.572142	-0.373534	0.012054	0.30624	39	2	5
11	-1.595397	-7.941881	-0.261529	-0.02149	-0.049931	1	3	1
12	0.253285	-1.404496	-0.153318	0.001641	-0.013255	1	3	2
13	0.529538	2.487263	0.747453	-0.004844	0.080087	1	3	3
14	0.936501	9.608734	0.226164	0.006494	-0.034652	1	3	4
15	-0.123927	-2.749621	-0.55877	0.018199	0.017751	1	3	5

Representation graphique



E Analyse des résultats pour information clusterisée OLS, 2SLS et 3SLS

Etudions la validité du modèle 3SLS :

Table 20: Hausman 3SLS consistency test

Test		Resultats
1	2SLS contre 3SLS	0

La normalité des résidus :

Table 21: Shapiro-Wilk normality test

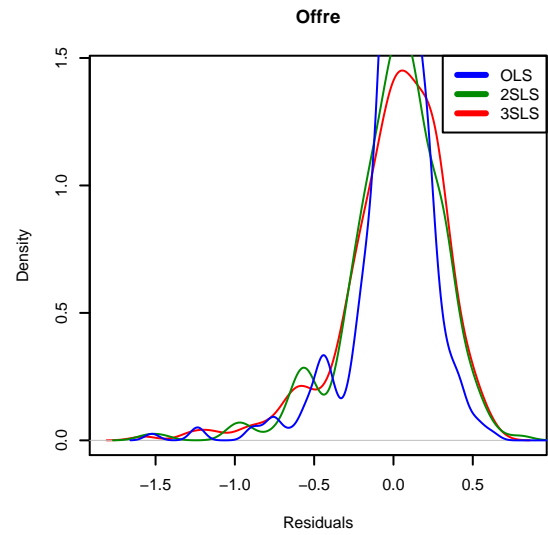
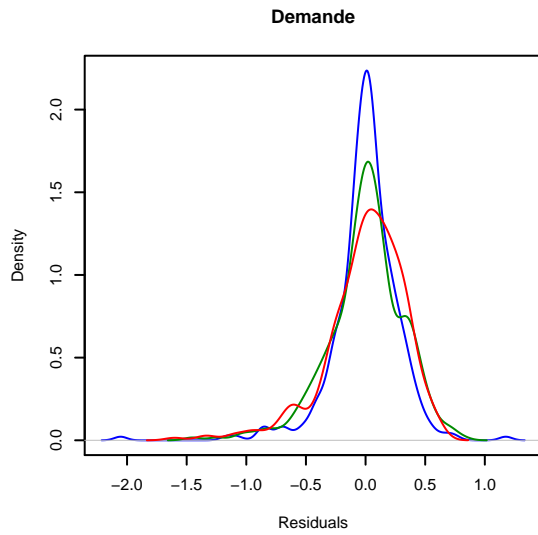
	OLS	2SLS	3SLS
Equation de demande	0	0.00000	0
Equation d'offre	0	0	0

L'hétéroscédacité :

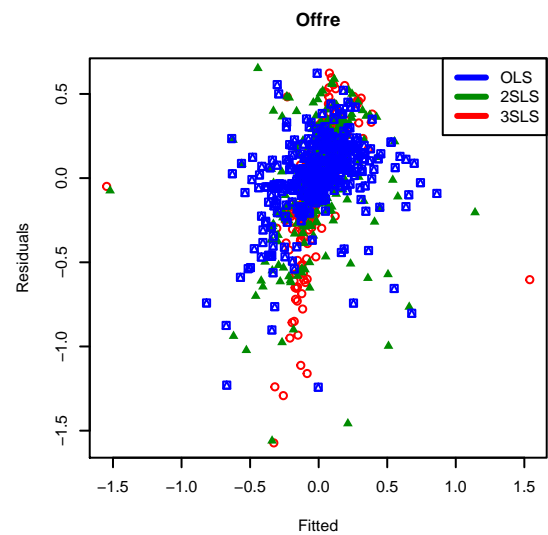
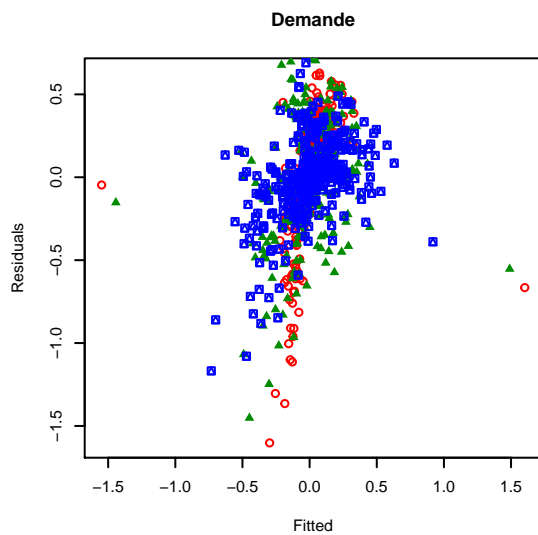
Table 22: Bartlett heteroscedasticity test

	OLS	2SLS	3SLS
Equation de demande	0	0.00002	0.00000
Equation d'offre	0	0.00000	0.00000

Les PDF des résidus :



Les résidus contre les variables prédites :



F Dictionnaire des variables

Finalement, nous offrons au lecteur un tableau de reference pour notre base des données finale.

Table 23: Ditionnaire des varibales

		Variable
		année
	ndep	
	si	
	qi	
	ipi	
	ri	
	iki	
	t	

References

- Anderson, Kym, Signe Nelgen, and others. 2011. *Global Wine Markets, 1961 to 2009: A Statistical Compendium*. University of Adelaide Press.
- 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.
- KREMER, Florence, and Catherine VIOT. 2004. “Conflit et Coopération Au Sein Du Canal: L’interaction Stratégique Entre La Grande Distribution et Les Producteurs de La Filière Viti-Vinicole.”
- Laporte, Catherine, and Marie-Claude PICHERY. 1996. “Production costs of AOC Burgundy wines.” Research Report. Laboratoire d’analyse et de techniques économiques(LATEC). <https://hal.archives-ouvertes.fr/hal-01526958>.
- MacKay, Alexander, and Nathan H Miller. 2018. “Estimating Models of Supply and Demand: Instruments and Covariance Restrictions.”
- MÄKELÄ, PIA, GERHARD GMEL, ULRIKE GRITTNER, HERVÉ KUENDIG, SANDRA KUNTSCHKE, KIM BLOOMFIELD, and ROBIN ROOM. 2006. “DRINKING PATTERNS AND THEIR GENDER DIFFERENCES IN EUROPE.” *Alcohol and Alcoholism* 41 (October): i8–i18. <https://doi.org/10.1093/alcalc/agl071>.
- Outreville, J François. 2010. “Les Facteurs Déterminant Le Prix Du Vin.” *Enometrica* 3 (1): 25–33.
- Steiner, Bodo. 2004. “French Wines on the Decline? Econometric Evidence from Britain.” *Journal of Agricultural Economics* 55 (2): 267–88.