

# Etude des effets des pesticides sur le marché des vins de table

---

Arnaud Blanc, Nikita Gusalov, Sasha Picon

Matière :  
Analyse empirique des marchés

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

<b>Introduction</b>	<b>1</b>
<b>1. Les pesticides</b>	<b>1</b>
<b>2. Le marché du vin français</b>	<b>2</b>
<b>3. Le cadre théorique</b>	<b>4</b>
<b>4. Les données</b>	<b>7</b>
<b>5. L'étude statistique</b>	<b>11</b>
<b>6. Modélisation</b>	<b>14</b>
<b>7. Résultats des estimations</b>	<b>16</b>
<b>9. Conclusion</b>	<b>24</b>
<b>Annexes</b>	<b>26</b>
<b>Références</b>	<b>41</b>

# Introduction

Aujourd’hui, l’utilisation des pesticides est un problème majeur de l’agriculture. Celle-ci utilise la plus grande partie des pesticides en France. Il s’agit d’un enjeu à la base 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. Ils sont supposés protéger les rendements. En effet, les aléas climatiques influencent le développement de champignons ou de maladies. 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 à 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 sensé nous donner des résultats plus précis et fiables.

## 1. Les pesticides

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 ses 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. (Butault Jean-Pierre and Guillaume 2011)

En 2008, les 30 produits les plus toxiques sont interdits. Une taxe sur les phytosanitaires a aussi été mise en place. Cette taxe est croissante avec le niveau de toxicité de ces produits. Elle devait augmenter au fil des années. De plus, l’octroi de crédits d’impôt en faveur de l’agriculture biologique devait aussi permettre de réduire l’utilisation des pesticides. (Butault Jean-Pierre and Guillaume 2011)

Malgré tous ces efforts, l’utilisation des pesticides perdurent.

En 2008, le nombre de doses unités (Nodu) a été créé pour enregistrer l’évolution de la demande de pesticides. (Butault Jean-Pierre and Guillaume 2011) 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. (Fiona and Roméo 2019) 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 (Butault Jean-Pierre and Guillaume 2011), ce qui la rend d’un intérêt particulier à étudier.

## Comment baisser l’utilisation de pesticides

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

- l'agriculture intensive, qui ne limite pas le recours aux pesticides ;
- l'agriculture raisonnée, qui limite le recours aux pesticides en fonction de seuils ;
- l'agriculture biologique, qui vise à supprimer les traitements avec des produits phytosanitaires de synthèse.

Néanmoins, ces méthodes sont difficiles à mettre en place dans la viticulture, ce qui nous amène à un problème de compréhension des raisons pour lesquels les agriculteur utilisent les pesticides. Les professionnels proposent de commencer par utiliser l'agriculture raisonnée en viticulture qui permettra 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. Ces propositions sont purement théoriques puisque l'on ne connaît pas encore les techniques qui pourraient influencer le comportement des viticulteurs. Un des mécanismes possibles est le mécanisme du marché. En manipulant l'offre et la demande du vin, le but est théoriquement atteignable.

## 2. Le marché du vin français

La France est l'un des principaux producteurs de vins. En effet, la France représente 10% de la surface des vignes mondiales. La production de vins représentait 4.6 milliards de litres. La France représentait 17% de la production totale de vins. 3% de la surface agricole française est consacrée à la production des vins. Néanmoins, le vin représente 15% de la production agricole en valeur. (Interprofessions des Vins à appellation d'origine et à indication géographique 2018) La France est aussi l'un des principaux consommateurs de vins. En effet, en France, il s'agit de la boisson alcoolisée la plus consommée. 88% des ventes de vins en France sont effectuées en grande surface. Néanmoins, la consommation française de vin baisse depuis une trentaine d'années. (Interprofessions des Vins à appellation d'origine et à indication géographique 2018)

### Utilisation des pesticides dans la viticulture

Nous avons déjà montré que la viticulture est le deuxième secteur agricole en termes d'utilisation des pesticides. En effet, elle représente plus de 14.4% des dépenses de produits phytosanitaires, en France. Néanmoins, ces pesticides ne sont pas utilisés dans la même proportion dans toutes les régions de France. (Butault Jean-Pierre and Guillaume 2011)

Les bassins viticoles Français utilisent en majorité des fongicides et des bactéricides. En effet, la vigne fait face à des aléas climatiques qui permettent le développement de champignons comme le Mildiou. (Jérôme 2017) Pour lutter contre le développement de ces champignons, les viticulteurs ne peuvent utiliser que des fongicides. En effet, ils ne peuvent pas utiliser la rotation des cultures qui pourrait réduire ou empêcher le développement de ces champignons puisque la vigne est une culture pérenne. Les pieds de vigne ne sont pas replantés chaque année. Il est donc nécessaire d'utiliser les pesticides dans la vigne pour protéger la production et éviter les pertes. En effet, les champignons s'attaquent aux feuilles de la vigne et aux fruits. Donc la pulvérisation de pesticides est un des seuls moyens pour protéger les rendements des cultures viticoles. Néanmoins, l'utilisation des pesticides a aussi un impact du côté de la demande de vin. Cet impact est plus ambigu, à cause d'un manque de transparence d'information sur les bouteilles de vin. (Robin 2018)

Un sondage de l'Ifop sur les habitudes et perceptions de consommation des Français a montré que 93% des Français considèrent que la présence de pesticides dans les aliments a un impact sur la santé. 89% des Français souhaiteraient être informés de la présence ou non de pesticides dans les produits alimentaires, à travers un étiquetage. (Ifop 2017)

Il est particulièrement intéressant et avantageux d'étudier le marché du vin afin d'identifier les mécanismes éventuels qui influencent le montant des pesticides utilisés dans la production. Dans cette étude nous visons à comprendre les mécanismes figurant dans ces relations.

## Le problème d'hétérogénéité

Mais comment étudier le marché du vin ? Le secteur du vin est constitué de produits qui sont fortement hétérogènes. En effet, il existe une forte hétérogénéité entre les différents labels (AOP, IGP, sans IG) mais aussi au sein 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 Pomeraci 2014) :

- les vins de qualité inférieure, les moins chers avec des 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 de l'agriculture et divisons le marché en deux parties :

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

La première partie est soumise à des règlements spécifiques : limitations des quantités produites, origine contrôlée, un caractère de la demande spécifique. De plus, les viticulteurs peuvent être réticents à changer leurs processus de production déjà raffiné au maximum de peur d'avoir des pertes de qualité.

La deuxième, c'est-à-dire le marché des vins moins chers, est un peu plus simple et compréhensible. Elle demeure moins hétérogène Cembalo, Caracciolo, and Pomeraci (2014). En effet, les vins qui se situent dans une fourchette de prix étroite sont quasiment homogènes. Ainsi, les vins sans indication géographique ont des attributs intrinsèques simples, une complexité de qualité faible. Il s'agit donc de vins peu différenciés. Nous avons, donc, choisi de nous concentrer sur ces vins sans indication géographique à cause de leur degré d'homogénéité qui est plus fort que pour les autres labels. Cela nous permet d'analyser le marché par département et non par marques/produits.

## Les vins de table

Le marché des vins sans indication géographique connaît de forte variation. Nous allons donc revenir sur la période qui précède notre étude. Ainsi, en 2011, les transactions de vente de vins rouges ont augmenté de 29%. Les transactions de vins rosés ont également augmenté de 13%. Les transactions de vins blancs augmentaient de 76%. Les prix de ces vins bien que faible connaissent aussi des variations importantes. Ainsi, en 2011, les trois couleurs de vins ont connus des hausses de prix. Les vins rouges ont vu leur prix moyen augmenté de 12%. Le prix moyen des vins rosés a aussi cru de 3 %. Pour finir, le prix moyen des vins blancs ont cru de 13%. Les vins de France sans indication géographique ont connu une baisse en volume des ventes de 14.6% par rapport à la moyenne des ventes sur la période 2006 à 2010. (FranceAgriMer 2011).

### 3. Le cadre théorique

#### Les hypothèses théoriques

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 moyen des marchés sont différents en raison des droits de douane à l'importation et des taxes à la consommation différentes (Anderson, Nelgen, and others 2011). De plus, la perception des produits de consommation varie d'un pays à l'autre (MÄKELÄ et al. 2006).

Le rachat du vin par les enseignes (grandes surfaces) peut avoir une impact sur l'offre du vin. Nous étudions deux cadres différents en les comparant dans ce travail. Un type d'interaction pose que les prix sont exogènes pour les fournisseurs de vin simple, comme cela fut démontré par KREMER and VIOT (2004). L'autre possibilité suppose qu'il existe quand même des interactions entre l'offre et la demande et que les prix du vin sont endogènes. Nous devrions trouver une façon pour formaliser ces deux approches différentes et pouvoir les comparer.

La plupart des bouteilles achetées le sont dans la grande distribution. Néanmoins, dans un souci de simplicité nous n'estimerons que la situation sur le marché en amont où les distributeurs achètent leurs bouteilles directement auprès du viticulteur (nous simulons l'indice des prix en amont à partir des prix sur le marché final, disponible dans la base de données de FranceAgrimer). Donc, nous supprimerons tous les intermédiaires entre le producteur et le distributeur (les grands enseignes, puisque presque la totalité du vin simple est vendu dans les grandes surfaces).

Quand aux exportations et aux importations, n'ayant pas la possibilité de les contrôler le montant des vins non IGP exportés/importés, nous laissons ces effets au terme d'erreur. Nous ignorons complètement les interactions internationales. Nous simplifions davantage notre modèle en imposant l'absence des flux du vin entre les départements (ce qui peut être justifié si les grandes enseignes rachètent le vin auprès des viticulteurs et seulement après le redistribuent au sein de leurs chaînes).

Nous supposons que les facteurs de production jouent le rôle de modificateur de l'offre (*supply shifter*) ce qui nous permet d'intégrer les variables déterminants le niveau de la production (telles que la surface ou la quantité de pesticides utilisés) directement dans l'équation d'offre. Pour information sur les facteurs de production du vin nous référençons les articles de Laporte and PICHERY (1996) et de Outreville (2010).

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 déterminée 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 racheté (consommé) au sein du même département. C'est une hypothèse très restrictive, qui nous éloigne de la réalité, mais nous devons 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 plusieurs modèles différents. Finalement, les effets que l'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 de département, si nous révélons des différences significatives entre les départements. Un autre modèle nous permettra de vérifier et de justifier cette hypothèse.

En ce qui concerne les pesticides, nous supposons d'abord, que l'utilisation des pesticides par les viticulteurs est reliée à 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 de pesticides et les agriculteurs de notre analyse. La quantité de pesticides utilisés par les agriculteurs correspond seulement à leurs besoins.

Pour résumer cette partie, 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 les différents modèles. Puis, nous pourrons choisir entre ces modèles différents le plus vraisemblable, qui nous servira à répondre à la question de recherche.

## Formalisation

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

$$Qo = \sum_{i=1}^N qo_i \quad (1)$$

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

$$qo_i = a_i + b_i Po_i + c_i X_i \quad (2)$$

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 de l'utilisation des pesticides dans la production du vin sur l'offre de ce dernier.

Cette équation permet déjà d'estimer les effets de l'utilisation des pesticides sur le marché du vin. Toutefois il existent deux possibilités sur l'organisation du marché dans ce cas. D'abord, dans le cadre le plus simple et probable du point de vue théorique nous supposons, les résultats suivants de KREMER and VIOT (2004), que les prix sont imposés par les grandes enseignes aux distributeurs de vin simple. De l'autre côté, nous pouvons supposer que les prix sont endogènes parce qu'ils sont négociés entre le distributeur et le producteur. Appelons ce modèle théorique M1 pour le référencer dans le futur, nous permettant de distinguer le cas sans interaction simultanée 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 de vin simple offerte sur le marché est déterminée seulement par le producteur et n'est pas lié à la demande. Comme nous l'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 ce cas, nous devrions prendre en compte les interactions entre l'offre et la demande. Dans ce but, nous introduisons également la demande dans notre analyse.

La demande aggregée 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 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 un système d'équations. Il existe plusieurs façons de le faire.

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

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

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

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

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

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

Neanmoins, ce cas se révèle 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 trop restrictives, réduisant, ainsi à néant la puissance statistique de notre estimateur (ex : le nombre des observations par années est très faible). Dans les deux cas nous faisons face à une impasse.

Une des modifications 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 qui 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 de l'agrégation des effets au niveau national en estimant le coefficient moyen unique pour tous les départements nous allons réduire les biais possibles.

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

$$\begin{aligned} qd_i &= \alpha_i + \beta Pd_{i,d} + \gamma Z_i \\ qo_i &= a_i + b Po_{i,o} + c X_i \end{aligned}$$

Où  $qd_i = qo_i$  et  $Pd_i = Po_i$ , ce qui permet de relier les équations au niveau départemental. Les coefficients  $b$ ,  $c$ ,  $\beta$  et  $\gamma$  sont supposés fixes pour tous les départements. Ils nous donnent un estimateur des effets moyens au niveau de la France. L'effet des pesticides dans la production du vin sera capté 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 les différents départements à cause des spécificités des marchés locaux, géographiques ou autres ? On peut supposer, qu'il existe au moins quelques groupes majeurs ayant des caractéristiques et des 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 les  $K$  systèmes d'équations suivants :

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

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 à notre disposition. Dans cette section, nous allons présenter et commenter les principales étapes de la construction de notre base de données que nous avons utilisée pour mener notre étude. Nous commencerons par une présentation des sources principales à partir desquelles nous avons collecté nos données. Ensuite, nous procéderons à la description des méthodes et techniques utilisées pour transformer ces données et les rendre traitables. Enfin, nous nous attarderons à présenter l'ensemble des variables que nous avons jugés pertinents d'intégrer dans notre modèle économétrique. Nous préciserons notamment les effets attendus des variables exogènes sur la variable dépendante de notre modèle.

## Sources des données

Nous avons eu recours aux bases des données suivantes pour collecter les données nécessaires à notre analyse :

Les bases données de ventes de pesticides par département (Institut national de l'environnement Industriel et des Risques, (“Données de Vente de Pesticides Par Département,” n.d.)) sur la période 2008-2017, qui nous indiquent les quantités totales de chaque type de produit phytosanitaire (identifiable par leur numéro d'autorisation de mise sur le marché) vendues par département selon le type de conditionnement de ces produits (en litres ou en kilogrammes).

Des bases de données sur les prix du vin (FranceAgrimer, (“Historique Des Prix Moyens Vrac Vsig et Igp,” n.d.)), lesquelles nous transmettent notamment des informations sur les prix moyen du vin blanc puis du vin rouge-rosé entre 2009 et 2017 (selon la catégorie de vin : IG ou non IG). Plus spécifiquement, ce sont des prix moyens nationaux (ils ne varient donc pas par département) déflatés via l'indice des prix à la consommation (base 100 en 2014).

Des bases de données sur la population (INSEE, (“Revenu et Pauvreté Des Ménages” 2016)), qui contiennent notamment des données sur les diverses caractéristiques du niveau de revenu des ménages pour chaque département français sur la période 2009-2017.

Des bases de données sur la production de vin (SSM Finances Publiques, (“Statistiques Viti-Vinicoles - Relevés Annuels Des Stocks et Des Récoltes Depuis 2009,” n.d.)), nous fournissant des informations sur les quantités mensuelles produites chaque année de vin rouge-rosé puis de vin blanc (distinction entre les vins IG et non IG) de août 2009 à juillet 2019.

De plus, elles nous indiquent les surfaces totales viticoles par département qui sont dédiées à la production de différentes catégories de vins (AOC, IGP, vins de qualité supérieure, vins non IG, etc.) sur la période 2009-2018.

## Les variables utilisées pour notre modèle

Une fois l'ensemble des données collectées, un important travail de traitement des données a été nécessaire à mettre en œuvre. Tout d'abord, notre objectif principal était de concaténer les différentes bases de données entre elles. Le premier obstacle qui est apparu est que chaque base de données ne portait pas toute sur la même période. Il a fallu donc trouver une période d'étude commune entre les différentes bases. Il est apparu que la période s'écoulant entre 2009 et 2017 était celle compatible avec les différentes bases de données. Dans un premier temps, nous avons donc procédé à une élimination des observations se situant en dehors de cette période dans chacune des différentes bases de données.

Nous avons ensuite sélectionné les variables nécessaires à la mise en œuvre de notre analyse économétrique dans chaque base de données.

Au niveau des variables endogènes de notre modèle matérialisant l'équilibre sur le marché des vins sans IG, nous avons sélectionné uniquement les variables de production se rapportant aux vins non IG dans la base de données de production du vin puis nous avons pris uniquement les variables de prix se rapportant à la même catégorie de vin dans la base de données correspondante. Nous nous attendons à un effet positif du prix des vins non IG sur la quantité d'équilibre de vin dans l'équation d'offre mais à un effet négatif prix des vins non IG du côté de l'équation de demande. Nous avons également calculé les quantités totales produites de vin non IG chaque mois en additionnant les quantités produites de vin-rouge et de vin blanc pour chaque département puis nous avons ensuite

calculé les quantités moyennes de vin non IG produite chaque année par département afin de se placer au même niveau d'agrégation que dans les autres bases de données.

Au niveau de la base de données des prix, nous avons pour chaque département déterminé le prix moyen des vins non IG sur la base d'une moyenne simple entre les prix des vins rouges-rosés non IG et ceux des vins blancs non IG.

Enfin, étant donné que le prix moyen des vins non IG correspond à une moyenne nationale, nous voulions concevoir artificiellement un estimateur de ce prix moyen pour chaque département. Dans cet optique, nous avons créé en quelque sorte l'indice des prix du vin de table départementale, calculé de la manière 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 où la variable est mesurée.

En ce qui concerne les variables exogènes influant sur l'offre de vin non IG, nous avons sélectionné uniquement les surfaces viticoles dédiés à la production de vin non IG. Nous estimons que cette variable devrait avoir un impact positif sur la quantité d'équilibre de vin non IG puisque une augmentation des surfaces viticoles signifie qu'une plus grande quantité de vin non IG peut potentiellement être produite.

Au niveau de la base de données sur les ventes de pesticides, nous avons également agrégé les quantités vendues de chaque type de produit phytosanitaire afin d'obtenir à chaque fois deux valeurs pour chaque département et chaque année (2009 à 2017) : une quantité totale de pesticides vendues en litres puis une quantité totale de pesticides vendue en kilogrammes. Comme nous pouvons le constater, 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. Or, dans notre étude, nous chercher étudier l'impact de la masse totale des pesticides utilisés sur l'équilibre du marché du vin non IG. Pour pouvoir le faire, nous créons un indice qui permet de prendre en compte les évolutions des différents types des produits à la fois ce qui nous permet d'obtenir une variable unique pour représenter la quantité totale de pesticides vendue sur le marché. Cet indice simple est calculé de la manière suivante :

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

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

Nous estimons que la quantité totale de pesticides devrait exercer une influence positive sur la quantité d'équilibre des vins non IG étant donné que nous avons vu que les produits phytosanitaires constituent un intrant important pour la production de vin. En effet, ces produits sont utilisés sur les cultures viticoles pour les protéger ce qui permet d'améliorer leur rendement et par conséquent la quantité produite de vin.

En ce qui concerne les variables exogènes influant sur la demande de vin non IG, nous avons sélectionné uniquement le niveau de revenu médian des ménages, exprimé au niveau départemental (laquelle, si besoin nous pourrions facilement agrégé au niveau national) au sein de la base de données de la population que nous avons préalablement importé. Cette variable est, comme dans

notre base de données sur les prix moyens des vins, déflatée de l'indice des prix à la consommation (base 100 en 2014). En terme d'effet attendu sur l'équilibre du marché du vin non IG, nous estimons qu'elle devrait influer positivement sur la quantité d'équilibre des vins non IG. En effet, une hausse du revenu des ménages se traduirait par une hausse de leur pouvoir d'achat et donc une augmentation de la demande potentielle adressée aux vins non IG si bien que cela se répercute sur la quantité d'équilibre par une hausse de la production pour répondre à la demande supplémentaire soit sur le prix d'équilibre des vins non IG qui augmenterait.

Une fois la sélection et la construction des variables finalisée, il s'est avéré alors possible de concaténer les différentes bases pour obtenir une base de données unique en panel qui nous fait office d'appui pour la bonne conduite de notre étude économétrique et statistique. Nous avons ensuite restreint notre champ d'analyse à la période s'écoulant entre 2012 et 2016 car nous avons constaté qu'il manquait certaines données par rapport à certaines variables dans les autres périodes. Le nombre de périodes devient donc plutôt pauvre (on passe d'une échelle de 8 années à une échelle de 5 années) mais l'intérêt de cette démarche est de pouvoir disposer d'un panel « cylindré » (nombre de mesures respectifs de chaque variable de la base de données identique pour chaque département). Nous avons également éliminé les départements non producteurs de vin non IG puisqu'ils se trouvent alors en dehors du cadre d'analyse de notre étude économétrique (pas d'offre de vin non IG dans ces départements puisque nous estimons que tout le vin produit dans chaque département est consommé dans le même département : pas de mobilité des produits entre les départements). Au final, il y a 69 départements producteurs de vin non IG que nous avons inclus dans notre base de données finale. Nous avons également exclu la Corse de notre champ d'analyse même si cette région est productrice de vin car le comportement du marché du vin dans cette région est relativement atypique par rapport à l'ensemble des départements français métropolitains. L'inclure dans notre analyse risquerait alors de biaiser nos résultats d'estimation et de formuler des interprétations erronées.

En ce qui concerne la forme que doivent prendre l'ensemble des variables de notre base de données dans notre modèle économétrique, nous avons décidé de leur faire subir une transformation logarithmique, ce qui nous permet ainsi de pouvoir interpréter les effets estimés plus facilement. Nous obtenons ainsi un modèle à équations simultanées sous forme logarithmique où nous pouvons traiter les estimateurs obtenus comme l'élasticité de la demande/l'offre par rapport à des facteurs différents. Concrètement, dans notre étude économétrique, nous cherchons spécifiquement à estimer l'élasticité de la quantité d'équilibre de vin non IG par rapport à la quantité de pesticides vendue sur le marché. De plus, l'intérêt de cette transformation logarithmique réside aussi dans le fait qu'elle permet de réduire la variance des résidus de notre modèle économétrique et donc de réduire ainsi le risque d'hétéroscédasticité (variance non constante de l'erreur)

Nous pouvons résumer les propriétés de notre base de données finale de la manière suivante :

- Une base de données en panel “cylindrée”.
- Nombre d'individus large (69 départements producteurs de vin de table) combiné à un nombre de périodes pauvre (5 années : de 2012 à 2016).
- Chacune des variables de notre base de données varie par département et par année.
- La plupart des variables sont exprimées à l'échelle logarithmique (sauf les indices des prix et de quantité des pesticides utilisés)
- 2 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).
- 3 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).

## 5. L'étude statistique

Dans cette partie de l'étude nous allons mener une étude exploratoire sur les données collectées.

De l'étude de la variance pour les données en panel avec des statistiques générales, nous passerons à l'étude de l'interdépendance des variables. Puis, nous allons finir avec l'étude des données alternées par une transformation *within*.

### Visualisation au niveau de la France

Pour la première analyse il peut être intéressant de voir la situation du point de vue géographique. Nous visualisons les valeurs moyennes par département des différentes variables (une partie des représentations se trouve dans l'annexe A1). Cette représentation nous permet d'obtenir une première intuition sur la répartition spatiale des valeurs clés par département. Particulièrement la répartition de la quantité de vin simple vendu par département semble avoir une structure autocorrélée dans l'espace. Nous allons, quand même ignorer les interactions spatiales possibles, parce que notre échantillon ne comprend pas la totalité des départements français (nous avons exclus plusieurs départements qui ne produisent pas du vin simple). Le comportement des autres variables est similaire.

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

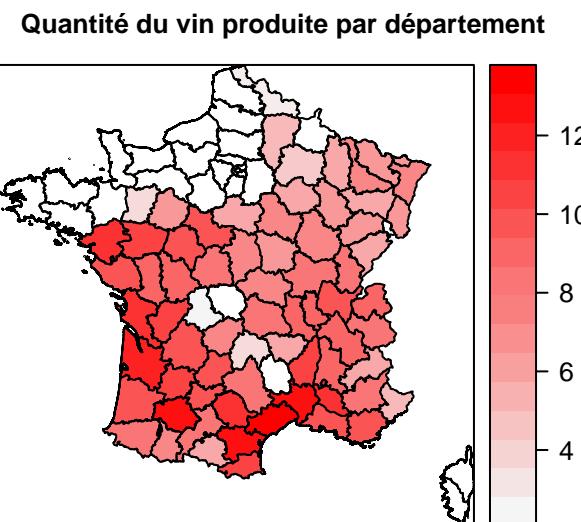


Figure 1: Les quantité du vin non-IG moyennes par département

Puis, nous observons le comportement du reste des variables (les représentations graphiques sont groupés dans l'annexe A1). L'indice des prix se comporte pratiquement comme la quantité de vin produite, car cet indice fut construit par l'intermédiaire 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 la liberté d'ignorer les effets possibles d'autocorrélation spatiale dans nos données. En effet, au moment de la construction de notre base de données, nous avons ignoré les départements ne produisant pas de vin simple. Mais, ils peuvent quand même jouer un rôle si nous prenions 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'agrégation de nos données, car il se peut que la théorie ne corresponde 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 de remarquer que la variance *between* est plus grande que la variance *within*. Cela nous amène à l'idée qu'il faut utiliser un modèle qui permettra d'estimer et de corriger ces inégalités entre les individus, car nous sommes plus intéressés par des effets individuels moyens (les effets moyens pour tous les départements). Cela est complètement conforme à l'hypothèse que l'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 aux effets fixes et aléatoires. Le tableau suivant regroupe les p-valeurs de ce test pour les différents modèles univariés.

Table 2: Pooling-test de Chow, p-valeurs

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

A part 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 faisions face à un modèle à effets fixes individuelles. Il faut quand même le justifier. Pour faire cela, nous allons effectuer le test du 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'effet. Il est évident que nous avons des effets fixes au niveau individuel ou des effets 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 évidences théoriques des études antérieurs nous décidons de ne garder que les effets fixes au niveau individuel afin de faciliter l'analyse.

### L'analyse de la corrélation

Dans le tableau A2.2 des annexes nous présentons les corrélations des variables après la correction pour les effets fixes individuels (nous effectuons la transformation *within* sur nos données en soustrayant les moyennes individuelles pour l'ensemble des variables). Dans les annexes, nous proposons également un tableau de corrélation pour les données non-transformées (A2.1), ce qui permet d'observer les inégalités et une pauvre représentativité des liens entre les variables pour les données initiales. Il est facile à remarquer la corrélation des ordres inférieurs à 0.1 pour toutes les variables sauf les interactions entre la surface cultivé et la quantité de vin offerte, ce qui explique pourquoi nous n'avons pas le choix et devons passer à la forme *within*.

En ce qui concerne les données *within*, particulièrement nous pouvons remarquer une forte corrélation entre la quantité offerte et le prix d'équilibre. Egalement nous observons une corrélation positive suffisamment significative (supérieure en grandeur de 0.2) entre la surface cultivé et la quantité du vin simple sur le marché, aussi bien qu'entre les revenus et l'indice d'utilisation des pesticides (ce

qui est probablement juste une coïncidence). Finalement, nous remarquons une corrélation négative entre la surface cultivée et l'indice d'utilisation des pesticides, ce qui est tout à fait naturel.

## 6. Modélisation

Cette partie du travail abordera la formulation économétrique de notre problème. Nous allons débuter par la présentation des notions théoriques utilisées 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 impropre pour l'estimation du marché du vin, selon Cembalo, Caracciolo, and Pomerici (2014). Dans notre étude nous allons, tout de même, utiliser une 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 de vin à utiliser les pesticides.

### Modèle économétrique

Dans cette section, nous allons présenter un par un nos modèles économétriques correspondant chacun à un des trois cadres théoriques possibles. Tous les modèles visent à estimer les effets moyens pour tous les départements sous des hypothèses différentes de fonctionnement de marché. Dans tous les cas, l'agrégation des effets au niveau national (ou au niveau des groupes) nous permet de réduire les biais éventuels, liés à la mauvaise spécification du modèle.

Pour le cadre où nous n'observons pas les interactions entre la demande et l'offre sur le marché (M1). Nous estimons 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 de prix, ce qui nous permet de traiter le prix comme une variable exogène; soit nous devrions construire un estimateur de variables instrumentales afin de traiter l'endogénéité éventuelle de l'indice des prix. Evidemment le premier cas est le plus simple, mais pour justifier l'utilisation de cette méthode nous devrions effectuer des tests d'endogénéité de prix. Le deuxième cas est beaucoup plus réaliste, puisque les viticulteurs sont rarement preneurs de 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). Particulièrement 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 l'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 égales au niveau de chaque département :  $qd_{i,t} = qo_{i,t}$ . C'est-à-dire que l'offre interne du département vise à satisfaire la demande interne du même département.

En termes d'agrégation ex-post des effets estimés, nous sommes sensés tomber sur l'équilibre au niveau du marché national. En d'autres 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 d'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 :

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

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

Pour finir cette partie, nous avons à notre disposition plusieurs chemins différents pour traiter ce modèle du point de vue économétrique. Le plus simple est d'estimer l'effet des pesticides sur l'offre de 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). De l'autre coté, nous pouvons utiliser les triples moindre carrés (nous devrions comparer les résultats obtenus avec un système d'équations non-réliées, 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 de l'interaction 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 coté du marché, à la fois. Finalement, si on trouve qu'il existe une hétérogénéité entre les départements en termes d'équilibre interne, nous pourrions réestimer les modèles en clusterisant nos *individus* (départements) par des classes différentes selon leurs attributs, pour estimer les équations par cluster.

## Hypothèses sur les résultats

Nous attendons 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 plus fiables. Cette méthode nous permet de dépasser le biais de simultanéité qui apparaît dans le cas d'estimation des systèmes d'équations liés (dans notre cas nous étudions les effets des pesticides sur l'offre et la production du vin simple sous l'hypothèse de présence des effets du marché). L'estimateur 3SLS 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 variance la plus faible.

Les propriétés de cet estimateur sont :

- La consistence ;
- L'efficience (asymptotique) ;
- La distribution pour les estimateurs suit une loi normale seulement dans des grands échantillons.

Dès le début, nous envisagions que cet estimateur ne reflète pas la nature du marché c'est pourquoi nous testons plusieurs modèles dans notre travail.

Parmi les inconvénients éventuels, on a également une faible représentation des effets hétérogènes entre les départements par le modèle. Nous estimons seulement les effets moyens et ainsi nous ignorons les différences d'élasticités pour des départements différents. Heureusement, ce problème peut être traité par l'introduction des clusters, regroupant des départements ayant des comportements similaires.

Finalement, il existe des effets que l'on ignore complètement, mais qui risquent d'intervenir. Par exemple, nous ignorons la présence d'autocorrélation spatiale et/ou temporelle dans notre modèle. Également, un nombre probablement insuffisant de facteurs est utilisé dans ce modèle, ce qui augmente le risque du biais des variables omises dans nos estimations.

## 7. Résultats des estimations

Dans cette section, nous allons présenter les résultats économétriques pour les différents modèles et les comparer.

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

- M1 : modèle simple sans interaction entre l'offre et la demande ;

- M2 : modèle complexe visant à intégrer les interactions entre l'offre et la demande en présence de variables endogènes ;
- M3 : les modèles sur les données clustérisées (M3.1 et M3.2 respectivement pour les deux cas précédents).

## M1 : Les résultats en absence d'interactions

Dans le cas des modèles sans interactions avec la demande, nous pouvons séparer deux cas différents. Le premier, le plus simple, se base sur l'hypothèse que les prix des vins simples sont imposés aux agriculteurs par les consommateurs (ou, ce qui est beaucoup plus probable, par des grandes enseignes ayant un pouvoir de négociation significatif). Cela implique que nous pouvons considérer les prix comme exogènes dans notre modélisation d'offre et par conséquent estimer notre modèle par des MCO (OLS - *ordinary least squares*) simples. En ce qui concerne le deuxième cas, nous posons que les prix dans l'équation d'offre sont endogènes. En d'autres termes, les agriculteurs affectent les prix par leur niveau de production et par les quantités écoulées de vin simple sur le marché final. Afin de traiter ce problème d'endogénéité nous pouvons utiliser la méthode des variables instrumentales (IV-OLS). Mais quels instruments choisir ? Sur ce point nous nous référerons au travail de MacKay and Miller (2018) (et plus particulièrement le travail fondateur de Hausman (1996)), où les auteurs démontrent que les instruments d'offre choisit parmi les régresseurs de l'équation de la demande constituent des instruments suffisamment pertinents. Dans notre cas nous pouvons instrumenter les prix par les revenus dans les départements étudiés, ce qui est conforme à la théorie économique, car le niveau des prix doit être corrélé avec le revenu réel.

Neanmoins, nous allons encore vérifier la validité de cet instrument et du modèle résultant. Par exemple, bien que notre supposition sur l'instrument rentre parfaitement dans le cadre théorique, nous n'observons qu'une faible corrélation entre l'indice des prix et les revenus (voir l'annexe A2.2). Les revenus ont la plus petite corrélation avec la variable explicative quand même.

Les résultats pour les deux modèles sont présentés dans le tableau ci-après :

	OLS	IV-OLS
IP	0.30*** (0.02)	-0.28 (0.25)
Surface	0.23*** (0.04)	0.47*** (0.13)
I pesticides	-0.16*** (0.05)	-0.11 (0.09)
R <sup>2</sup>	0.52	-0.87
Adj. R <sup>2</sup>	0.52	-0.89
Num. obs.	345	345
RMSE	0.29	0.58

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

Table 4: Modèles économétriques, variable dépendante - la quantité du vin vendu (en log)

Avant de commenter les effets estimés, nous devons supposer que le modèle en présence d'endogénéité n'est pas correcte (les estimateurs obtenus sont biaisés et loin de la réalité). D'abord, les instruments utilisés pour l'estimer sont assez faibles et ne donnent pas de résultats pertinents. De plus, à partir des résultats du test de Wu-Haussman, nous pouvons conclure que les estimateurs OLS et IV-OLS sont identiquement consistents et que nous n'avons pas des raisons de les utiliser (dans

cette situation, les résultats obtenus par OLS sont plus efficaces). Les résultats de ces tests sont regroupés dans l'annexe X.

Toutefois, les résultats obtenus par l'estimateur de OLS risquent d'être également biaisés. Dans l'annexe X nous pouvons voir que Shapiro-Wilk test rejette la normalité des résidus de notre modèle (annexe B4), bien que la fonction de répartition partielle a une forme proche à normale (annexe B1). Nous risquons également d'avoir des biais dans la variance des estimateurs, car selon le test de Bartlett sur l'heteroscedasticité, nous rejetons l'hypothèse de l'homoscedasticité des résidus (annexe B3). Au moins dans ce cas, nous n'avons pas de problèmes avec l'autocorrélation (les résultats du test de Durbin-Watson sont regroupés dans l'annexe B2), ce qui peut s'expliquer par utilisation des données temporelles en série trop courte pour y pouvoir détecter l'autocorrélation. Les résidus également ne sont pas correlés avec des variables explicatives, mais ont une corrélation forte avec la variable dépendante, ce qui nous indique une erreur de spécification potentielle dans notre modèle (annexe B1). L'explication la plus probable à ce problème est qu'on n'étudie pas suffisamment d'effets dans notre modèle et rencontrons par suite le biais de variable omise. Le graphique des résidus nous montre en même temps la nature quasi-aléatoire des résidus.

Maintenant, passons au résultats obtenus. L'indice des prix a un faible effet positif sur l'équilibre du marché. Néanmoins, il n'est pas très pertinent pour l'instant de commenter ces effets plus précisément, car nous n'avons toujours pas une évidence forte sur leur exogénéité. La surface dédiée au vins simples a également un effet positif sur la quantité de vin simple vendue, ce qui est tout à fait naturel, vu que cela représente le facteur principal de production du vin. Finalement, nous passons à l'effet le plus intéressant dans le contexte de notre étude. Nous rappelons à notre lecteur que c'est exactement cet effet, l'effet des pesticides sur le marché du vin, qu'on vise à estimer. Conformément aux études précédentes nous découvrons que les pesticides ont un impact négatif sur l'offre du vin simple, ce qui s'explique par leur nature d'utilisation. C'est-à-dire, nous obtenons une confirmation que les pesticides sont utilisés pour minimiser les pertes par les agriculteurs. Particulièrement dans le cadre du modèle estimé, l'effet moyen pour l'ensemble des départements est qu'une augmentation de 1% d'utilisation des pesticides est une réponse des agriculteurs à des pertes d'où une réduction de l'offre d'au moins de 0.0016%, toutes choses égales par ailleurs. Les résultats en termes numériques sont assez ambigus, car nous sommes confrontés à de nombreux problèmes dans la spécification du modèle. Nous pouvons tout de même constater un effet de minimisation des pertes dans l'utilisation des pesticides.

Cette nature révèle des implications importantes en termes de traitement futur du problème d'utilisation excessif des pesticides dans la viticulture.

## **M2 : Les résultats dans le cas des effets du marché présent**

Dans cette section nous allons étudier le modèle sous l'hypothèse de présence des effets de la conjoncture sur les décisions des agriculteurs. Identiquement au cadre précédent, nous avons deux choix possibles. D'abord, sous l'hypothèse d'endogénéité des prix, nous pouvons estimer le modèle par la méthode de 3SLS, en introduisant dans notre modèle à la fois l'endogénéité des prix pour l'offre et la demande puis la corrélation entre les résidus de ces deux équations décrivant le comportement des agents du marché. Deuxièmement, nous avons la possibilité d'imposer l'exogénéité des prix dans nos équations, ce qui peut être le cas si les enseignes rachetant le vin des agriculteurs sont preneuses de prix du vin vendu au consommateur final. Nous allons comparer des résultats de plusieurs modèles afin de vérifier sa validité.

Le premier modèle estimé par les 3SLS est complexe et nous risquons d'obtenir des résultats biaisés

suite à des misspécifications eventuelles. Afin de contracter la variance des estimateurs pour les rendre plus efficace, nous pouvons recourir à la procédure i3SLS (*iterated three step least squares*). Cette méthode nous donne des résultats similaires à ceux obtenus par FIML (*full information maximum likelihood*). Pourtant, l'utilisation de procédures itératives ne semble pas provoquer une diminution significative dans les variances des estimateurs (annexe C1).

Nous pouvons également comparer les résultats obtenus par 3SLS avec le modèle en absence des interactions (sous l'hypothèse que les résidus des deux équations ont une corrélation nulle) estimé par 2SLS (*two step least squares*). Dans ce cas, les régresseurs endogènes de la demande sont instrumentés tout comme cela est décrit dans Wooldridge (2005). Ce méthode donne des résultats équivalents à ILS (*indirect least squares*), une technique utilisé pour estimer les systémés des équations.

Dans le deuxième cas, nous rejetons l'hypothèse de l'endogénéité des prix sur le marché du vin simple, en supposant qu'il est dicté par le consommateur final à des acteurs en amont (le distributeur qui rachète le vin produit par les agriculteurs). Cela nous amène à implementer la méthode SURE (*seemingly unrelated equations*) afin d'introduire la corrélation entre les résidus dans notre modèle du marché tout en ignorant l'endogénéité des prix.

Les résultats pour les deux modèles sont regroupés sous format d'un tableau ci-dessous ("D" désigne les coefficients de l'équation de la demande et "O" de l'équation de l'offre, qui nous intéressent) :

	SUR	3SLS
D : IP	0.32*** (0.02)	0.79*** (0.15)
D : Revenu	-1.10 (0.66)	-13.07*** (2.76)
O : IP	0.32*** (0.02)	-0.25 (0.25)
O : Surface	0.03 (0.02)	0.45*** (0.13)
O : I pesticides	-0.02 (0.02)	-0.17* (0.08)
Demande: R <sup>2</sup>	0.46	-0.41
Offre: R <sup>2</sup>	0.46	-0.74
Demande: Adj. R <sup>2</sup>	0.45	-0.42
Offre: Adj. R <sup>2</sup>	0.46	-0.75
Num. obs. (total)	690	690

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

Table 5: Modèles économétriques, variable dépendante - la quantité du vin vendu (en log)

Les résultats obtenus pour le cadre avec prix endogènes (l'estimation par 3SLS) donne forcément des résultats biaisés, car comme nous l'avons vu dans la section précédente, nous ne disposons pas d'instruments suffisamment fort pour instrumenter l'indice des prix. Cela se confirme par la non-significativité des indices des prix estimés par l'approche des variables instrumentales. Ces effets sont négatifs et non-différents de zéro, ce qui n'est pas normal pour un bien appartenant à la catégorie des biens de Giffen (bien que nous pourrions arbitrer que le vin le plus simple pourrait se comporter d'une telle façon, puisque c'est l'un des produits alimentaire les plus basiques). Des autres tests nous confirment la misspécification de ce modèle : les résidus sont parfaitement correlé avec les valeurs de variable dépendante et les prédictions, ce qui est évident des graphiques dans les

annexe C2; une forte hétéroscedacité, qui cette fois est apparente sur les graphiques (annexes C2 et C5), etc. Quand même les tests de spécification ne donne pas aucune indication de misspécification du modèle 3SLS par rapport au modèle 2SLS, ce qui nous rassure dans notre supposition que nous n'utilisaons pas suffisamment des variables explicatives dans notre modèle pour utiliser ces approches économétriques.

Tout de même, nous obtenons des résultats proches à ceux obtenus par OLS simples dans la section précédente ( $-0.16$  contre  $-0.17$ ). Cela confirme notre hypothèse sur le rôle joué par les pesticides dans l'offre du vin simple sur le marché. Néanmoins, nous préferions ne pas utiliser ce modèle comme notre modèle de référence à cause de plusieurs sources de biais.

En ce qui concerne les résultats d'estimation par SURE, nous constatons des effets presque identiques à ceux, obtenus par l'estimateur OLS simple. Bien que dans ce cas, les effets d'utilisation des pesticides sont un peu plus accentués, avec un comportement des résidus plus anormaux.

### **M3 : Les résultats pour des départements groupés**

Dans cette section nous allons commencer par la présentation et la comparaison des différentes techniques pour effectuer la clustérisation de nos données. Puis nous allons procéder à une étude de la modélisation des données clustérisées. Nous supposons que cette approche peut donner des résultats significativement différents, car nous supposons une forte hétérogénéité dans le comportement des différents départements.

Le regroupement des départements dans des clusters par leurs caractéristiques relativement similaires, doit nous permettre de traiter les problèmes d'hétérogénéité potentiels (nous pouvons observer des traces de ces effets sur les graphiques dans l'annexe A2). Cela devrait permettre d'améliorer les résultats des estimations en réduisant la variance de nos estimateurs.

#### **Clusterisation**

Il existent plusieurs façons de séparer et clustériser les données. Dans le cadre de notre travail, nous allons implémenter la procédure *k-means* ou *k-moyennes*. Cette procédure regroupe les individus autour des *centres* (dont le nombre est *k*) choisi par un algorithme itératif. Cet algorithme itératif minimise la distance intra-cluster. La qualité d'ajustement des clusters est évaluée par le paramètre WSS (*within sum of squares*). C'est spécifiquement ce paramètre qu'on va utiliser pour comparer les différentes approches et ainsi choisir laquelle est la plus appropriée.

En ce qui concerne les différentes méthodes de clustérisation, nous pouvons les choisir en manipulant les données à l'entrée (les *inputs*) de cet algorithme. Plus précisément nous pouvons :

- Grouper les départements en utilisant les valeurs moyennes pluriannuelles de leurs caractéristiques (ce qui est identique aux données qu'on obtient lors de la transformation *between*);
- Grouper nos individus en utilisant les variations des caractéristiques intra-annuelles (nous appliquons la transformation *within* à nos données, puis nous transformons nos données afin de conserver la dimension temporelle);
- Utiliser l'information complète pour capter les évolutions aussi bien que les différences générales entre les départements étudiés.

#### ***Between***

Nous supposons que les départements ayant des valeurs moyennes inter-annuelles proches, obtenus par la transformation *between*, ont des caractéristiques similaires, ainsi qu'un comportement presque identique. La clusterisation est effectuée sur les données *between* pour l'ensemble des départements étudiés.

Les résultats des estimations ainsi que les centres théoriques pour  $k = 3$  se trouvent dans l'annexe D1.1.

### ***Within***

Dans ce cas, nous supposons que les départements ayant des tendances et des évolutions de leurs caractéristiques identique ont un comportement similaire. Plus, précisement, nous commençons par une transformation des nos données en *within*, ce qui implique qu'on soustrait les moyennes empiriques intra-annuelles pour toutes les caractéristiques. Ensuite, nous transformons les données en créant pour chaque caractéristique et pour chaque année une variable séparée, ce qui permet de préserver la dimention temporelle de nos données.

Les centres théoriques obtenus pour  $k = 3$  se trouvent dans l'annexe D1.2.

### **Information complète**

Dans le cas de l'information complète, nous utilisons l'ensemble des informations disponibles sur les individus afin de construire les clusters. Nous utilisons l'approche similaire à celle utilisée dans le cas d'une clusterisation *within*, sauf que cette fois nous n'effectuons pas la correction sur les valeurs moyennes intra-annuelles.

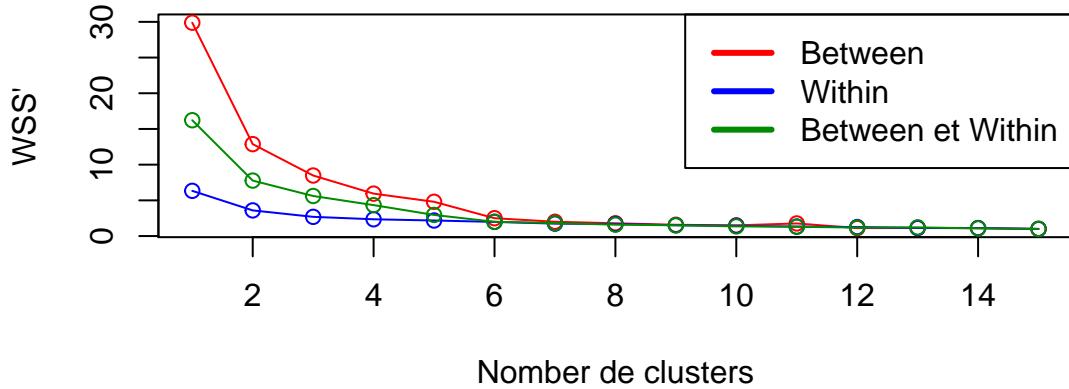
Les centres obtenus pour ce type des données sont présentés dans l'annexe D1.3.

### **Comparaison des différentes méthodes**

Afin de pouvoir comparer des valeurs différents de WSS (*within summ of squares*) nous allons visualiser la valeur d'un indice :

$$WSS' = \frac{WSS}{\min(WSS)}$$

Ce qui nous permettra d'évaluer les écarts relatifs du WSS de sa valeur minimale (pour un nombre de clusters égal à 15). Le graphique suivant démontre l'évaluation des valeurs de  $WSS'$  pour les trois cas différents de la définition des clusters :



Nous pouvons observer que pour la transformation *within*, nous observons une convergence importante vers la valeur minimale de WSS. Les deux autres approches au traitement et à la transformation des données offrent une vitesse de convergence encore plus élevée, avec des valeurs relatives initiales également plus significatives.

Nous remarquons également que les valeurs optimales du nombre des clusters dans les trois cas sont à peu près identiques (autours de 3 et 5). Par conséquent, la meilleure approximation absolue est atteint pour les huit clusters et reste relativement inchangée après.

### M3.1 : Le cadre en absence des intéractions avec la demande

Nous commençons par la comparaison des résultats obtenus pour les différents clusters avec les modèles de type OLS et IV-OLS. C'est-à-dire, sous l'hypothèse de l'absence d'intéractions entre l'offre et la demande.

Nous n'évaluons pas le système en introduisant les variables de groupe (dummy variables) car cela risque de biaiser les résultats à cause d'une taille des groupes différente. Afin d'éviter ce biais, nous évaluons les modèles par cluster.

Le tableau suivant regroupe les 6 modèles estimés (3 clusters avec 2 modèles par cluster, le nombre du cluster étant affiché après *c* dans le tableau) :

	OLS c1	IV-OLS c1	OLS c2	IV-OLS c2	OLS c3	IV-OLS c3
IP	0.51*** (0.02)	0.22* (0.10)	0.15 (0.04)	0.15 (0.04)	0.69*** (0.03)	4.36 (4.60)
Surface	0.09* (0.04)	0.19** (0.07)	-0.54 (0.58)	-0.59 (0.58)	0.19*** (0.04)	-0.69 (1.14)
I pesticides	-0.11* (0.04)	-0.01 (0.07)	7.58 (4.62)	7.73 (4.65)	-0.16*** (0.04)	0.14 (0.52)
R <sup>2</sup>	0.79	0.55	0.92	0.92	0.76	-15.69
Adj. R <sup>2</sup>	0.78	0.54	0.80	0.80	0.76	-15.95
Num. obs.	145	145	5	5	195	195
RMSE	0.14	0.21	0.39	0.39	0.23	1.96

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

Table 6: Modèles économétriques, variable dépendante - la quantité du vin vendu (en log)

Nous observons que un cluster ressort d'une façon anormale car il a des valeurs d'une magnitude excessive (c'est le département numéro 57 où l'indice a une variation intra-annuelle anormale, on peut le considérer comme un outlier et ne pas le commenter). En ce qui concerne les deux autres clusters, nous observons qu'un d'entre eux comprend les départements se spécialisant dans la production des vins simples (avec des effets de prix, de surface et d'utilisation des pesticides accentué). L'autre regroupe des départements qui montre des effets moins nets et évidents.

Les diagnostics détaillés sont consultables dans l'annexe D2, où tous les tests principaux sont présentés. Les tests sur la validité des estimateurs IV sont les plus intéressants et se trouvent dans l'annexe D2.3, où nous observons que les estimations par les variables instrumentales ne sont pas toujours valides dans notre cas. On ne peut pas rejeter l'hypothèse des faibles instruments seulement pour un des clusters, et le test de Wu-Haussman donne des résultats négatifs pour tous les clusters, nous indiquant que le modèle OLS est préférable. Finalement, les résidus sont toujours non-normaux, bien que les résultats du test sont les meilleurs dans l'ensemble (annexe D2.2).

### M3.2 : Le cadre d'interférence avec la demande

Dans cette partie, nous utilisons l'approche identique à celle qu'on a déjà implementé dans la partie M2. Nous supposons, que le marché fonctionne en présence de liens entre l'offre et la demande. Les résidus des deux équations dans ce cas sont correlés entre eux.

Identiquement à la section précédente, nous regroupons les résultats d'estimation pour les 6 modèles (2 modèles par 3 clusters) sous la forme d'un tableau.

	SUR c1	3SLS c1	SUR c2	3SLS c2	SUR c3	3SLS c3
D : IP	0.52*** (0.02)	0.66*** (0.11)	0.12* (0.03)	0.12* (0.03)	0.73*** (0.03)	1.35*** (0.27)
D : Revenu	-0.86 (0.45)	-6.68*** (1.95)	18.14 (14.50)	18.43 (14.62)	-4.87*** (1.05)	-10.08** (3.51)
O : IP	0.51*** (0.02)	0.20* (0.10)	0.16 (0.04)	0.16* (0.04)	0.72*** (0.03)	4.71 (4.58)
O : Surface	0.01 (0.02)	0.18* (0.07)	-0.66 (0.57)	-0.69 (0.58)	0.05* (0.02)	-0.68 (1.14)
O : I pesticides	-0.01 (0.02)	0.03 (0.07)	4.97 (3.94)	4.99 (3.93)	-0.04 (0.03)	0.43 (0.38)
Demande: R <sup>2</sup>	0.78	0.77	0.88	0.88	0.75	0.26
Offre: R <sup>2</sup>	0.77	0.52	0.89	0.89	0.73	-19.07
Demande: Adj. R <sup>2</sup>	0.78	0.77	0.84	0.84	0.75	0.25
Offre: Adj. R <sup>2</sup>	0.77	0.52	0.78	0.78	0.73	-19.27
Num. obs. (total)	290	290	10	10	390	390

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

Table 7: Modèles économétriques, variable dépendante - la quantité du vin vendu (en log)

Nous observons que les effets des pesticides sur la quantité marchande du vin sont négligeables pour tous les clusters dans tous les modèles. La validité de ces modèles laisse donc à désirer. D'abord, pour les modèles supposant l'endogénéité des prix, nous avons retenu des instruments trop faibles ce qui rend ces modèles inefficients. De plus, l'introduction de l'hypothèse des liens entre les deux côtés du marché (introduction de la corrélation entre les résidus) n'améliore pas l'explicativité des modèles.

### Avis sur l'utilisation des clusters

En général les clusters n'améliorent pas les résultats des estimations. Nous avons réussi à identifier un département trop différent du reste dans sa nature, mais au final cette méthode n'a abouti à rien de concret.

De plus, nous retirons de ces tentatives d'amélioration de notre modèle l'évidence qu'il existe des départements étant plus inclinés à la production du vin de table que les autres, ce qui est tout à fait logique.

## 9. Conclusion

Nous avons étudié et comparé plusieurs modèles distincts qui visent à apprécier les effets d'utilisation des pesticides sur l'offre de vin de table. Nous constatons, que parmi tout ces modèles, le meilleur estimateur des effets moyens par département est obtenu à travers un simple modèle OLS en absence d'effets d'interaction avec la demande de vin de table ou d'endogénéité des prix. Ce fait est soutenu par les hypothèses théoriques sur le fonctionnement du marché des vins simples disponibles dans la littérature. Les agriculteurs proposant du vin simples sont dans la plupart des cas preneurs des prix offerts par les distributeurs, ce qui explique l'exogénéité des prix. Le même fait explique l'absence des interactions entre l'offre et la demande, car les grandes enseignes achètent le vin simple sous conditions *take or leave* (ce qui n'est pas vrai pour les autres types du vin).

L'analyse des données clusterisées permet d'observer un degré d'hétérogénéité relativement faible entre les départements. De plus, les effets d'utilisation des pesticides sur l'offre de vin ne sont clairement identifiables que pour un groupe spécifique d'entités étudiées.

Les résultats obtenus dans cette étude confirment des résultats obtenus par d'autres chercheurs, ainsi que les suppositions théoriques sur le rôle des pesticides dans le commerce du vin. Plus précisément, les pesticides sont utilisés par les viticulteurs pour minimiser les pertes causées par les maladies, les fongicides... etc. Nous captions ce phénomène en estimant le modèle d'équilibre du marché du vin simple.

Nous devons également souligner que le modèle présenté dans ce travail est loin de frôler la perfection absolue. Plusieurs problèmes restent non-traité ou non-résolus. Parmi ces problèmes nous pouvons citer: un faible nombre d'observations sur la dimension temporelle, la présence d'hétéroscedasticité dans les résidus, la non-normalité des résidus, des variables omises, des instruments faibles, etc. Toutefois, nous avons réussi à capter la tendance principale dans le comportement de l'offre face à l'utilisation des pesticides par les agriculteurs. Il peut donc s'avérer intéressant d'étudier ces aspects et de traiter ces problèmes relevés dans de futures études.

Pour de futures recherches, nous conseillerons d'étudier directement l'impact des pesticides sur la production et non sur l'équilibre du marché parce que, comme nous avons pu le constater dans nos résultats, la modélisation du marché via des équations simultanées ne nous permet pas d'améliorer significativement nos estimateurs. Ceci constitue un point important, étant donné le faible nombre de variables explicatives sélectionnées et les défauts potentiels dans la spécification de notre modèle. Dans notre étude, ces types de modèle ne donnent pas des résultats fiables et sont donc inefficaces

## Annexes

### A Les statistiques descriptives

#### A1 Les moyennes par département

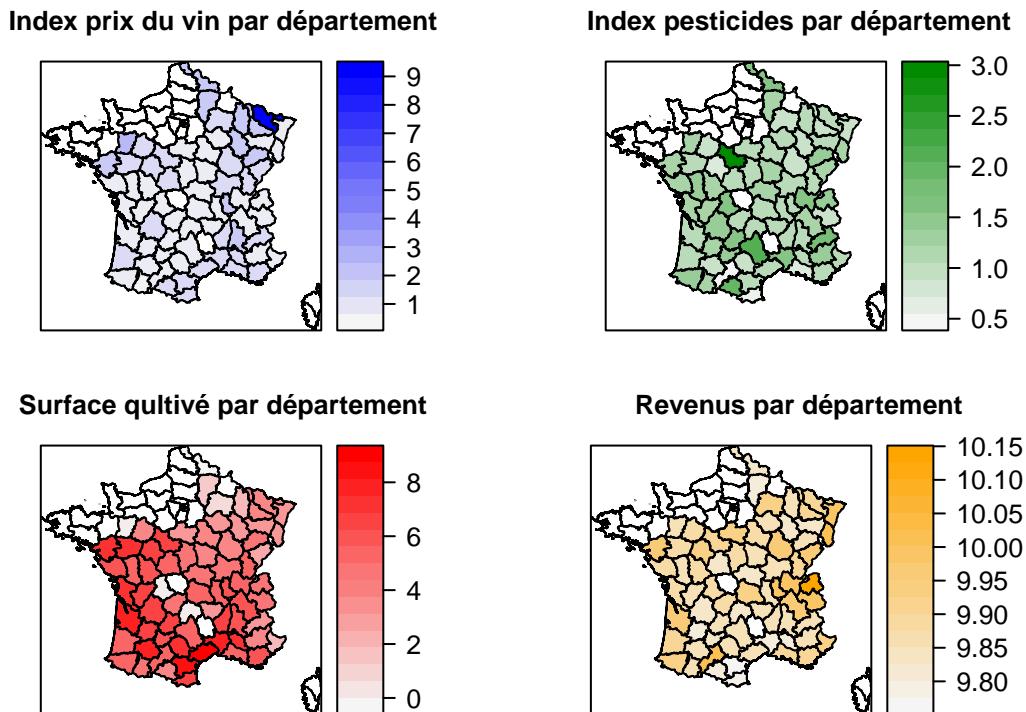


Figure 2: Les valeurs moyennes par département

## A2 L'étude des interdépendances

La première section de cette annexe comprend les résultats pour les données non transformées, le deuxième par contre intègre les résultats pour les données sous la transformation *within*.

### A2.1 Information complète

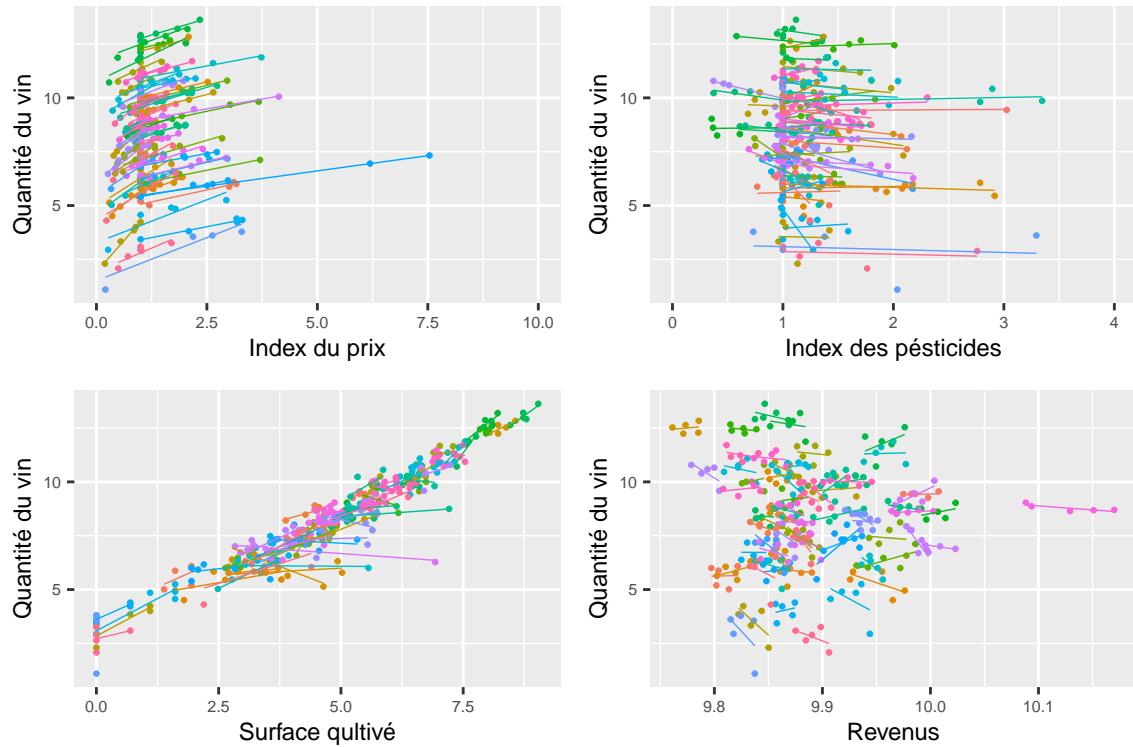


Figure 3: L'étude bivarié

	Quantité du vin	IP	Surface	Revenus	Index pesticides	Temps
Quantité du vin	1.00	0.02	0.96	-0.03	-0.07	-0.04
IP	0.02	1.00	-0.05	0.01	-0.06	0.11
Surface	0.96	-0.05	1.00	-0.06	-0.05	-0.06
Revenus	-0.03	0.01	-0.06	1.00	-0.04	0.12
I pesticides	-0.07	-0.06	-0.05	-0.04	1.00	0.30
Temps	-0.04	0.11	-0.06	0.12	0.30	1.00

Table 8: La correlation complete

## A2.2 Transformation *within*

Les relations entre les variables ressortent mieux pour les données transformées. Ce sont ces données que nous intégrons dans nos modèles économétriques.

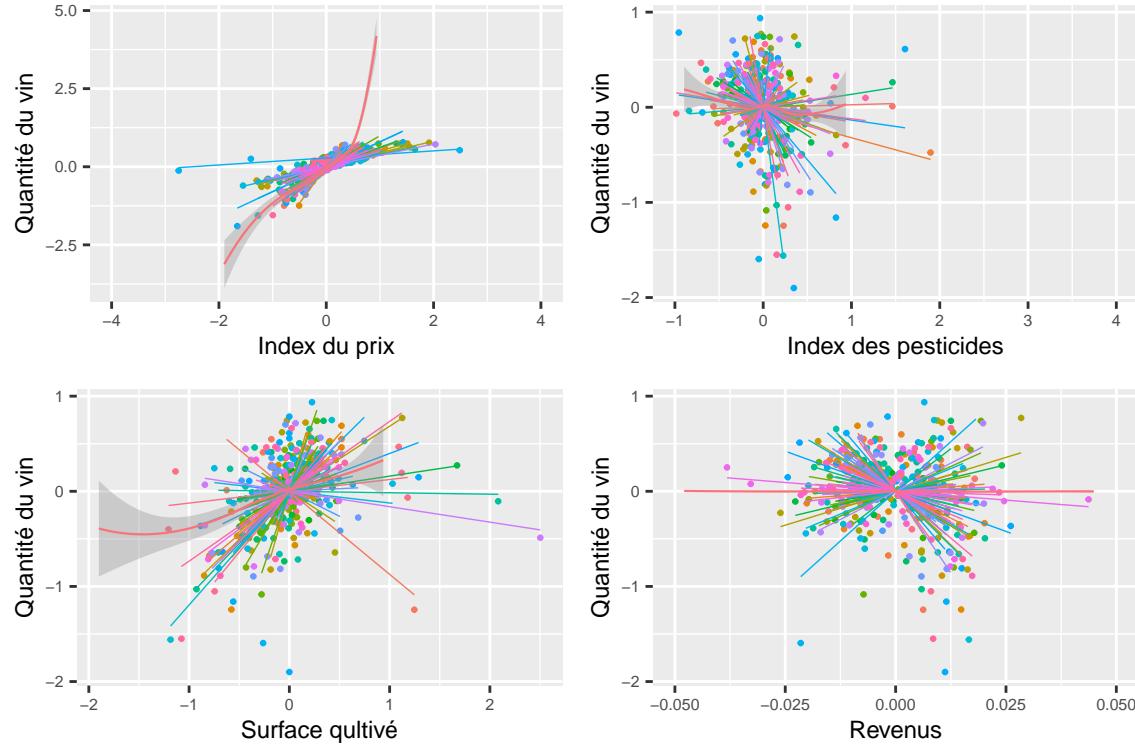


Figure 4: Relations bivariés dans le cas de transformation *within*

	Quantité du vin	IP	Surface	Revenus	Index pестициdes	Temps
Quantité du vin	1.00	0.67	0.37	-0.16	-0.18	-0.20
IP	0.67	1.00	0.19	0.11	-0.01	0.16
Surface	0.37	0.19	1.00	-0.17	-0.20	-0.31
Revenus	-0.16	0.11	-0.17	1.00	0.21	0.65
I pестициdes	-0.18	-0.01	-0.20	0.21	1.00	0.41
Temps	-0.20	0.16	-0.31	0.65	0.41	1.00

Table 9: La correlation *within*

## B Analyse des résultats du cadre M1

### B1 Le comportement des résidus

	OLS	IV-OLS
Vin	0.69	0.88
IP	-0.00	0.85
Surface	0.00	0.00
Revenus	-0.24	0.00
Pesticides	-0.00	-0.00

Table 10: Correlation des résidus

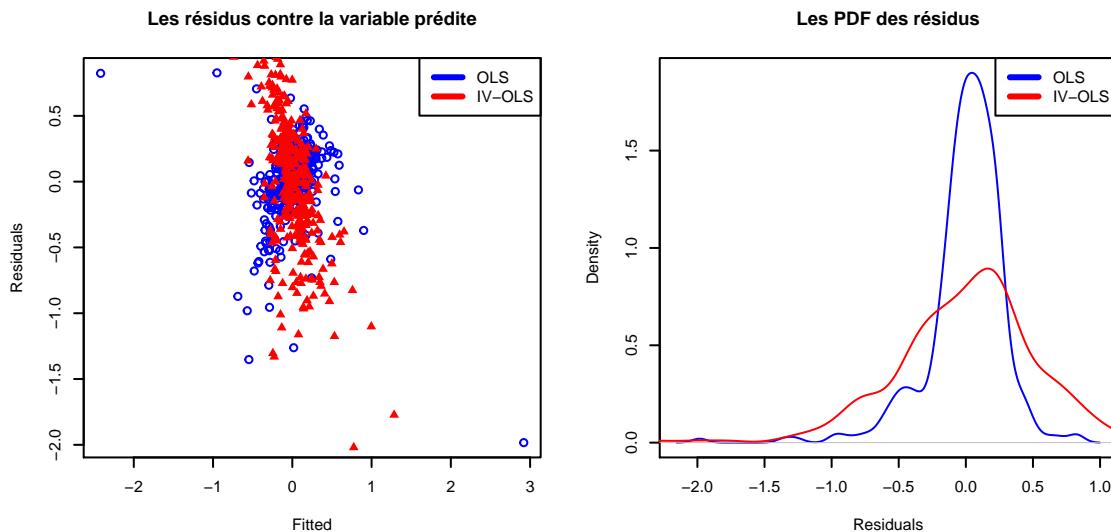


Figure 5: Le comportement des résidus

### B2 L'autocorrélation

Table 11: Les statistiques test de Durbin-Watson, t-stat

	OLS	IV-OLS
Equation d'offre	0.627	0.637

### B3 Test de l'hétéroscédasticité

Table 12: Les résultat du test de Bartlett sur l'heteroscedasticité, p-valeur

	OLS	IV-OLS
Equation d'offre	0	0

#### B4 La normalité des résidus

Table 13: Shapiro-Wilk test de normalité des résidus, p-valeur

	OLS	IV-OLS
Equation d'offre	0	0

#### B5 Diagnostics IV-OLS

Table 14: Diagnostics d'estimateur IV

	df1	df2	statistic	p-value
Weak instruments	2	341	3.645	0.027
Wu-Hausman	1	341	22.553	0.00000

## C Analyse des résultats M2

### C1 Comparaison des modèles 2SLS, 3SLS et i3SLS

	SUR	2SLS	3SLS	i3SLS
D : IP	0.32*** (0.02)	0.79*** (0.15)	0.79*** (0.15)	0.79*** (0.15)
D : Revenu	-1.10 (0.66)	-13.07*** (2.76)	-13.07*** (2.76)	-13.07*** (2.76)
O : IP	0.32*** (0.02)	-0.28 (0.25)	-0.25 (0.25)	-0.25 (0.24)
O : Surface	0.03 (0.02)	0.47*** (0.13)	0.45*** (0.13)	0.45*** (0.12)
O : I pesticides	-0.02 (0.02)	-0.11 (0.09)	-0.17* (0.08)	-0.17* (0.08)
Demande: R <sup>2</sup>	0.46	-0.41	-0.41	-0.41
Offre: R <sup>2</sup>	0.46	-0.87	-0.74	-0.75
Demande: Adj. R <sup>2</sup>	0.45	-0.42	-0.42	-0.42
Offre: Adj. R <sup>2</sup>	0.46	-0.89	-0.75	-0.76
Num. obs. (total)	690	690	690	690

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

Table 15: Modèles économétriques, variable dépendante - la quantité du vin vendu (en log)

## C2 Indépendance des résidus

	SUR D	SUR O	2SLS D	2SLS O	3SLS D	3SLS O	i3SLS D	i3SLS O
Vin	0.75	0.75	-0.13	0.88	-0.13	0.88	-0.13	0.88
IP	-0.00	-0.00	-0.80	0.85	-0.80	0.84	-0.80	0.84
Surface	0.32	0.28	0.00	0.00	0.00	-0.00	0.00	-0.00
Revenus	-0.28	-0.31	0.00	-0.00	0.00	0.00	0.00	0.00
Pesticides	-0.23	-0.20	-0.08	0.00	-0.08	0.03	-0.08	0.03

Table 16: Correlation des résidus

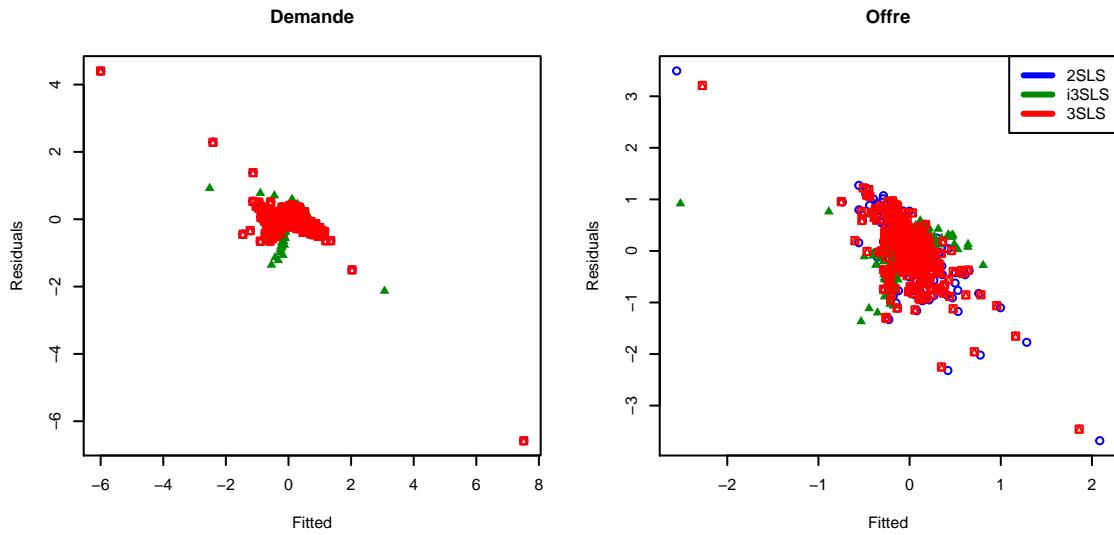


Figure 6: Les résidus contre la variable prédictive

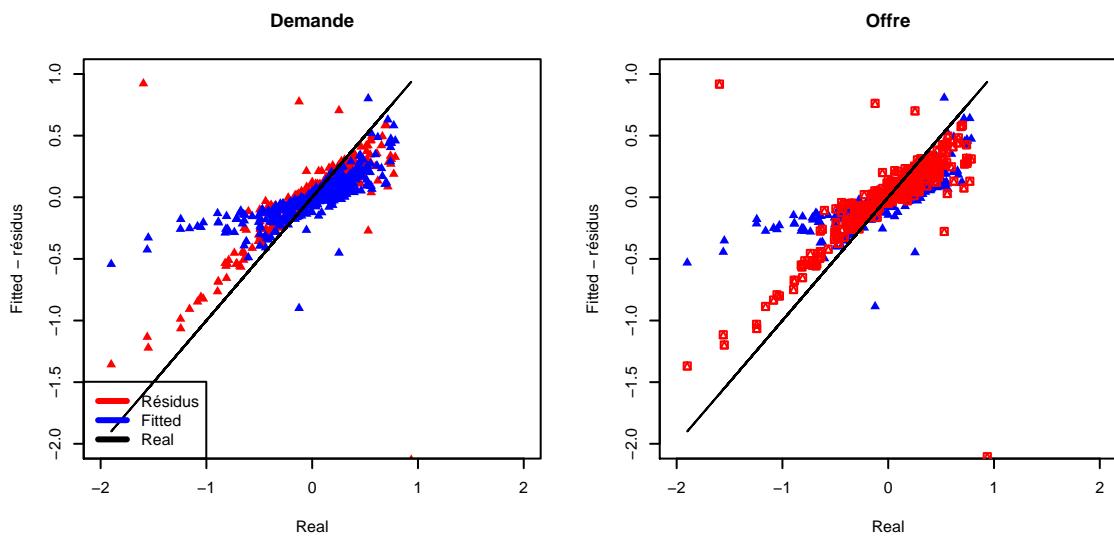


Figure 7: Les résidus et les prédictions, le cas de SUR

### C3 L'autocorrélation

Table 17: Les résultats du test de Durbin-Watson, t-stat

	SUR	2SLS	3SLS	i3SLS
Equation de demande	0.687	0.618	0.618	0.618
Equation d'offre	0.683	0.637	0.638	0.638

### C4 Test de l'hétéroscedasticité

Table 18: Test de Bartlett sur l'hétéroscedasticité, p-valeur

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

### C5 La normalité des résidus

Table 19: Shapiro-Wilk test de normalité, p-valeur

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

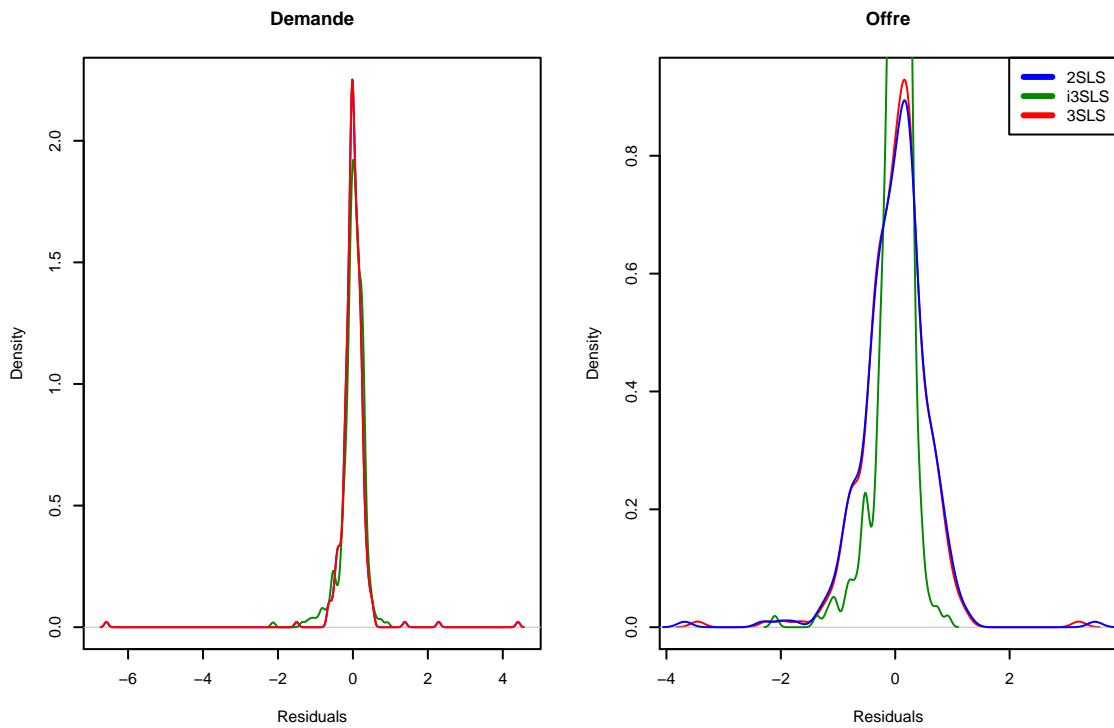


Figure 8: Les PDF des résidus

## C6 Comparaison des modèles

Table 20: Hausman 3SLS consistency test, p-valeur

	Test	Résultats
1	3SLS contre 2SLS	1
2	3SLS contre SUR	0

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	8	816.10			
2	6	-514.91	-2	2662.02	0.0000
3	8	-504.97	2	19.89	0.0000
4	8	-505.66	0	1.38	0.0000

Table 21: ML test de spécification

## D Clusterisation

### D1 Méthodes de clusterisation

#### D1.1 Les centres pour les données *between*

Les groupes sont définis par des caractéristiques suivantes :

Table 22: Les centres des clusters en between

Quantité du vin	IP	Surface	Revenu	I pesticides	N	Cluster
7.468	1.416	4.345	9.905	1.238	34	1
4.340	1.622	1.098	9.870	1.284	8	2
10.609	1.393	6.705	9.880	1.273	27	3

#### D1.2 Les centres pour les données *within*

Les groupes sont définis par les caractéristiques suivantes :

Table 23: Les centres des clusters en within

Quantité du vin	IP	Surface	Revenu	I pesticides	N	Cluster	Année
0.001186	-0.359222	0.04895	-0.012025	-0.234518	29	1	2012
-0.280464	-0.376425	0.078291	-0.000723	0.004804	29	1	2013
0.049743	0.107731	-0.01841	-0.006913	0.122468	29	1	2014
-0.032737	0.065232	-0.147634	0.004155	-0.089222	29	1	2015
0.262272	0.562683	0.038803	0.015506	0.196468	29	1	2016
0.098311	-0.291704	0.206247	-0.01031	-0.27852	39	2	2012
0.223902	0.18426	0.174215	0.002905	-0.108882	39	2	2013
0.283634	0.373141	0.059067	-0.006982	0.141058	39	2	2014
0.086134	0.306445	-0.065995	0.002333	-0.059896	39	2	2015
-0.691981	-0.572142	-0.373534	0.012054	0.30624	39	2	2016
-1.595397	-7.941881	-0.261529	-0.02149	-0.049931	1	3	2012
0.253285	-1.404496	-0.153318	0.001641	-0.013255	1	3	2013
0.529538	2.487263	0.747453	-0.004844	0.080087	1	3	2014
0.936501	9.608734	0.226164	0.006494	-0.034652	1	3	2015
-0.123927	-2.749621	-0.55877	0.018199	0.017751	1	3	2016

### D1.3 Les centres pour l'information complète

Les groupes sont définis par les caractéristiques suivantes :

Table 24: Les centres des clusters pour l'information complète

Quantité du vin	IP	Surface	Revenu	I pesticides	N	Cluster	Année
0.001186	-0.359222	0.04895	-0.012025	-0.234518	29	1	1
-0.280464	-0.376425	0.078291	-0.000723	0.004804	29	1	2
0.049743	0.107731	-0.01841	-0.006913	0.122468	29	1	3
-0.032737	0.065232	-0.147634	0.004155	-0.089222	29	1	4
0.262272	0.562683	0.038803	0.015506	0.196468	29	1	5
-1.595397	-7.941881	-0.261529	-0.02149	-0.049931	1	2	1
0.253285	-1.404496	-0.153318	0.001641	-0.013255	1	2	2
0.529538	2.487263	0.747453	-0.004844	0.080087	1	2	3
0.936501	9.608734	0.226164	0.006494	-0.034652	1	2	4
-0.123927	-2.749621	-0.55877	0.018199	0.017751	1	2	5
0.098311	-0.291704	0.206247	-0.01031	-0.27852	39	3	1
0.223902	0.18426	0.174215	0.002905	-0.108882	39	3	2
0.283634	0.373141	0.059067	-0.006982	0.141058	39	3	3
0.086134	0.306445	-0.065995	0.002333	-0.059896	39	3	4
-0.691981	-0.572142	-0.373534	0.012054	0.30624	39	3	5

## D2 Les modèles en absence d' interactions du marché

### D2.1 Le comportement des résidus

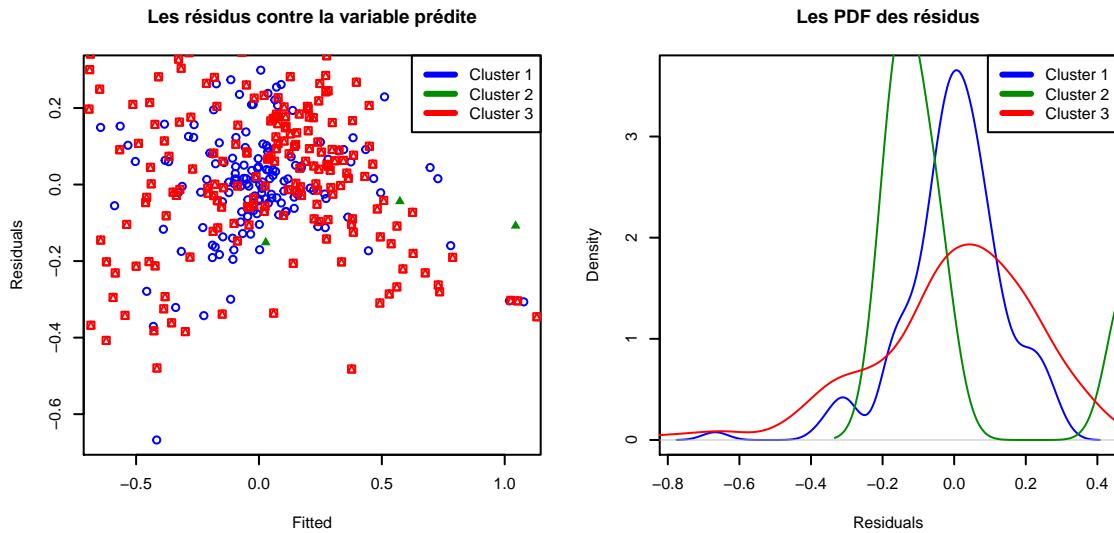


Figure 9: Le comportement des résidus pour le modèle OLS

### D2.2 La normalité des résidus

Table 25: Shapiro-Wilk test de normalité des résidus, p-valeur

	OLS c1	IV-OLS c1	OLS c2	IV-OLS c2	OLS c3	IV-OLS c3
Equation d'offre	0.00004	0.0003	0.014	0.009	0.00003	0.044

### B2.3 Diagnostics IV-OLS

Table 26: Diagnostics d'estimateur IV, cluster 1

	df1	df2	statistic	p-value
Weak instruments	2	141	8.980	0.0002
Wu-Hausman	1	141	23.590	0.00000

Table 27: Diagnostics d'estimateur IV, cluster 2

	df1	df2	statistic	p-value
Weak instruments	2	1	51.783	0.098
Wu-Hausman	1	1	1,511,030,356,715,090,395,088,268,284,282	0

Table 28: Diagnostics d'estimateur IV, cluster 3

	df1	df2	statistic	p-value
Weak instruments	2	191	0.322	0.725
Wu-Hausman	1	191	58.351	0

## D3 Analyse des résultats M2

### D3.1 Indépendance des résidus

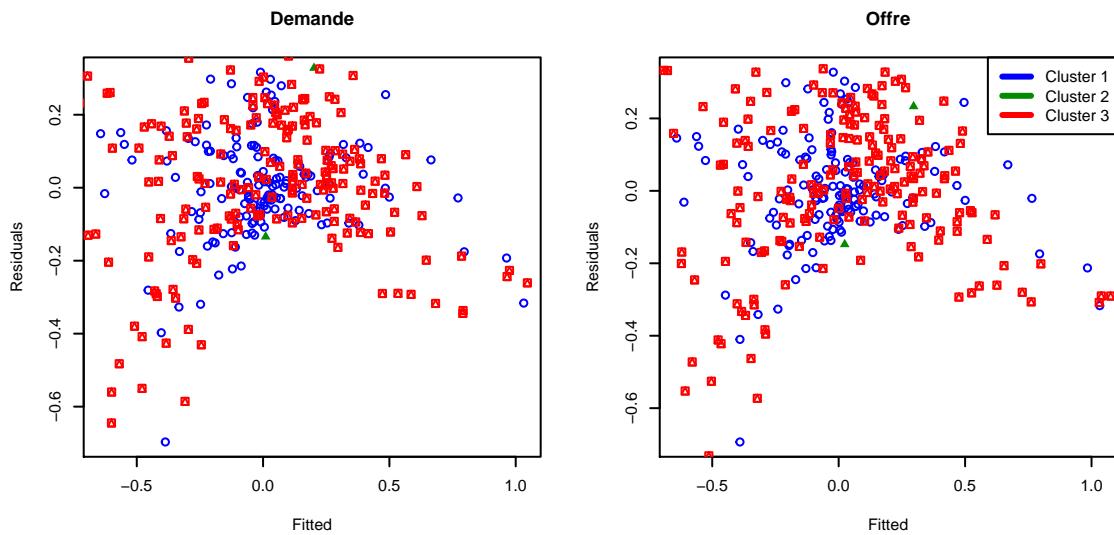


Figure 10: Les résidus contre la variable prédite, le cadre SUR

### D3.2 La normalité des résidus, le cadre de SURE

Table 29: Shapiro-Wilk test de normalité, p-valeur

	Cluster 1	Cluster 2	Cluster 3
Equation de demande	0.00004	0.115	0.00003
Equation d'offre	0.0001	0.190	0.0001

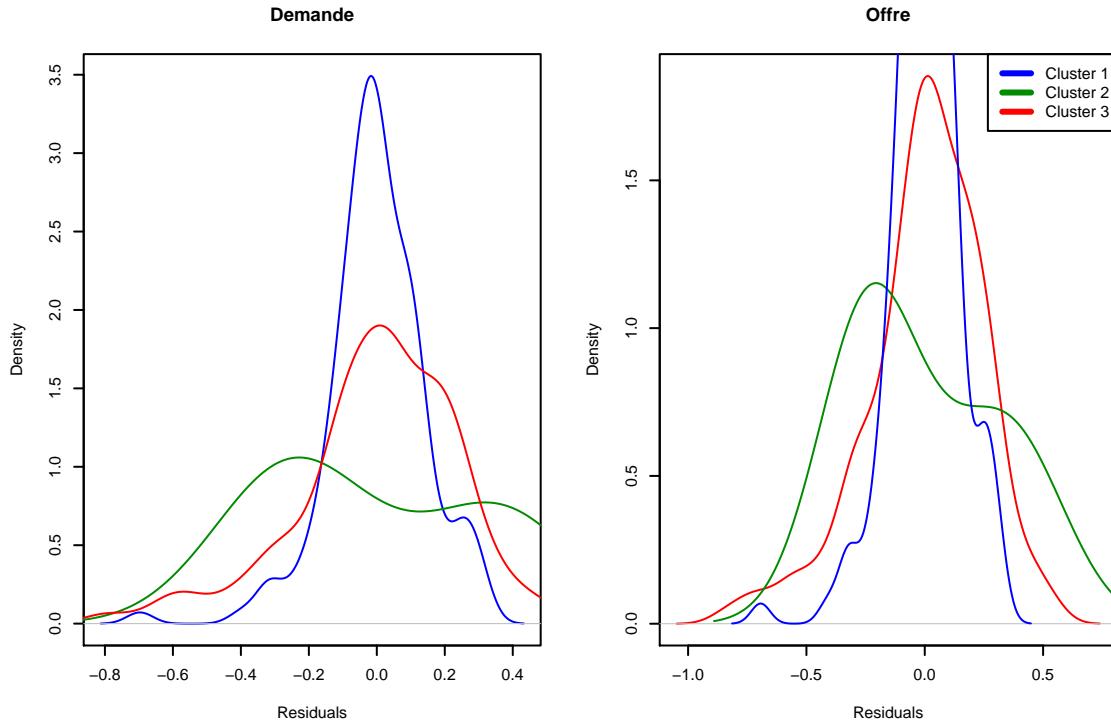


Figure 11: Les PDF des résidus

## Références

- Anderson, Kym, Signe Nelgen, and others. 2011. *Global Wine Markets, 1961 to 2009: A Statistical Compendium*. University of Adelaide Press.
- Butault Jean-Pierre, Jacquet Florence, Delame Nathalie, and Zardet Guillaume. 2011. “L’utilisation Des Pesticides En France : État Des Lieux et Perspectives de Réduction.” *Note et études Socio-économiques*, no. 35: 7–26.
- 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.
- “Données de Vente de Pesticides Par Département.” n.d. *INERIS*. [https://www.data.gouv.fr/fr/datasets/donnees-de-vente-de-pesticides-par-departement/#\\_](https://www.data.gouv.fr/fr/datasets/donnees-de-vente-de-pesticides-par-departement/#_).
- Fiona, Moghaddam, and Van Mastrigt Roméo. 2019. “Comment L’utilisation Des Pesticides N’a Cessé d’évoluer Ces 10 Dernières Années.” <https://www.franceculture.fr/ecologie-et-environnement/comment-lutilisation-de-pesticides-na-cesse-devoluer-ces-dix-dernieres-annees>.
- FranceAgriMer. 2011. “Note de Conjoncture.” <https://www.franceagrimer.fr/filières-Vin-et-cidre/Vin/En-un-clic/Dossiers-des-Conseils-et-comités?moteur%5BfiltreFiliere%5D=1506&moteur%5BfiltreTypeContenu%5D=analyse&page=6>.
- Hausman, Jerry A. 1996. “Valuation of New Goods Under Perfect and Imperfect Competition.” In *The Economics of New Goods*, 207–48. University of Chicago Press.
- “Historique Des Prix Moyens Vrac Vsig et Igp.” n.d. *FranceAgriMer*. [http://visionet.franceagrimer.fr/Pages/OpenDocument.aspx?fileurl=SeriesChronologiques/productions%20végétales/vin%20et%20cidriculture/prix%20moyens/prix%20moyens%20vrac%20VSIG%20et%20IGP/COT-VIN-HISTORIQUE\\_ACTVITE\\_PRIX\\_DEPUIS2000-C15-16.xls&telechargersanscomptageoui](http://visionet.franceagrimer.fr/Pages/OpenDocument.aspx?fileurl=SeriesChronologiques/productions%20végétales/vin%20et%20cidriculture/prix%20moyens/prix%20moyens%20vrac%20VSIG%20et%20IGP/COT-VIN-HISTORIQUE_ACTVITE_PRIX_DEPUIS2000-C15-16.xls&telechargersanscomptageoui).
- Ifop. 2017. “Les Français, La Consommation écoresponsable et La Transition écologique.” [https://www.wwf.fr/sites/default/files/doc-2017-10/171010\\_sondage\\_wwf\\_ifop\\_agriculture%202.pdf](https://www.wwf.fr/sites/default/files/doc-2017-10/171010_sondage_wwf_ifop_agriculture%202.pdf).
- Interprofessions des Vins à appellation d’origine et à indication géographique, Comité National des. 2018. “Chiffres Clés.” <https://www.intervin.fr/etudes-et-economie-de-la-filiere/chiffres-cles>.
- Jérôme, Pujol. 2017. “Apports de Produits Phytosanitaires En Viticulture et Climat : Une Analyse à Partir Des Enquêtes Pratiques Culturales.” *Agreste Les Dossiers*, no. 39: 3–25.
- 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 GRITTLER, HERVÉ KUENDIG, SANDRA KUNTSCHE, 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/alc/alc071>.
- Outreville, J François. 2010. “Les Facteurs Déterminant Le Prix Du Vin.” *Enometrica* 3 (1): 25–33.
- “Revenu et Pauvreté Des Ménages.” 2016. *INSEE*. <https://www.insee.fr/fr/statistiques/4190004># consulter.
- Robin, Prudent. 2018. “Enquête Franceinfo. Additifs, Pesticides... Le Vin Que Vous Buvez Ne Contient Pas Que Du Raisin : Découvrez Le Résultat de Nos Analyse.” <https://www.francetvinfo.fr/economie/emploi/metiers/agriculture/enquete-franceinfo-additifs-pesticides-le-vin-que-vous-buvez-ne-contient-pas-2957897.html>.
- “Statistiques Viti-Vinicoles - Relevés Annuels Des Stocks et Des Récoltes Depuis 2009.” n.d. *SSM Finances Publiques*. <https://www.data.gouv.fr/fr/datasets/statistiques-viti-vinicoles-releves-annuels-des-stocks-et-des-recoltes-depuis-2009/>
- Steiner, Bodo. 2004. “French Wines on the Decline? Econometric Evidence from Britain.” *Journal of Agricultural Economics* 55 (2): 267–88.
- Wooldridge, Jeffrey M. 2005. “Instrumental Variables Estimation with Panel Data.” *Econometric Theory* 21 (4): 865–69.