

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

Arnaud Blanc, Nikita Gusarov, 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

Introduction	1
1. Les pesticides	1
2. Le marché du vin français	2
3. Le cadre théorique	3
4. Les données	7
5. L'étude statistique	9
6. Modélisation	11
7. Résultats des estimations	14
9. Conclusion	21
Annexes	22
References	33

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.(???)

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 le niveau de toxicité de ces produits. Cette taxe 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.(???)

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

En 2008, le nombre de doses unités 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. (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.(???)

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.

- 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.

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.

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 agricole. 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

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 pourraient 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)

Le problème d'hétérogénéité

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 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 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 contrôlé, un caractère de la demande spécifique. La deuxième, c’est-à-dire le marché des vins moins chers, est aussi complexe. Néanmoins, elle demeure moins hétérogène Cembalo, Caracciolo, and Pomarici (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, choisit 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 est non par des marques/produits.

Les vins de table

Le marché des vins sans indication géographiques 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 leurs prix moyens augmenté de 12%. Le prix moyens des vins rosés ont aussi crû de 3 %. Pour finir, les prix moyens des vins blancs ont crû de 13%. Les vins de France sans indications géographiques 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

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 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).

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 aux 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 complètement les interactions internationales.

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

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

Facteurs influençant le prix ... 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 consommée au sein du même département. C’est une hypothèse très 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 de département, 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 relié 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 des 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 pourrions 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 agrégée pour toute la France est donnée identiquement par l'équation suivante :

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

Avec la quantité offerte déterminé 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 d'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. Appelons ce modèle théorique M1 pour le référencer dans le 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 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 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 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 capter les effets au niveau national. Pour ce faire nous réécrivons les deux équation (de la demande et de l'offre respectivement) sous la forme suivante :

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

$$Q_d = \sum_{i=1}^N (\alpha_i + \beta_i P d_i + \gamma_i Z_i) = \sum_{i=1}^N \alpha_i + \sum_{i=1}^N \beta_i P d_i + \sum_{i=1}^N \gamma_i Z_i$$

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

$$Q_d = \sum_{i=1}^N \alpha_i + \sum_{i=1}^N \beta_i P d_i + \sum_{i=1}^N \gamma_i Z_i$$

$$Q_o = \sum_{i=1}^N a_i + \sum_{i=1}^N b_i P o_i + \sum_{i=1}^N c_i X$$

Néanmoins, 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 très faible). Dans le 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 :

$$q d_i = \alpha_i + \beta P d_{i,d} + \gamma Z_i$$

$$q o_i = a_i + b P o_{i,o} + c X_i$$

Où $q d_i = q o_i$ et $P d_i = P o_i$, ce qui permet de relier les équations au niveau départemental. Les coefficients b , c , β et γ 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 en notre disposition. Dans cette partie de notre travail nous allons présenter la base des données utilisée 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

Révéler 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 déflaté 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 des quantités offertes sur le marché par rapport à la quantité des pesticides utilisés.

Les propriétés de ces données sont les suivantes :

- Toutes les variables varient par département et par année.
- La 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 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 du vue géographique. Nous visualisons les valeurs moyennes par département des différentes variables (une partie des représentation se trouve dans l'annexe X).

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.

Quantité du vin produite par département

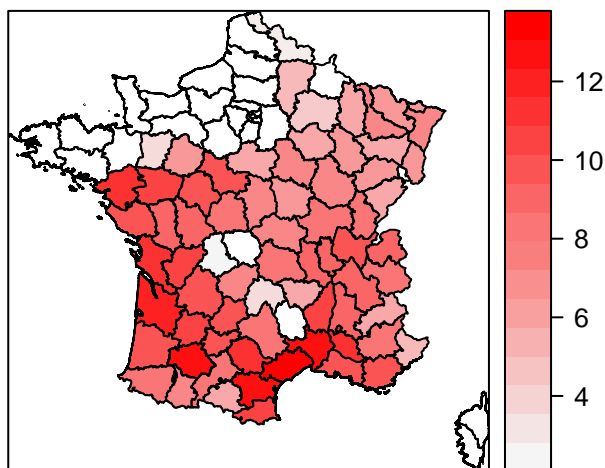


Figure 1: Les quantités du vin non-IG moyennes par département

Puis, nous observons le comportement du reste des variables (les représentations graphiques sont groupées dans l'annexe A1). L'indice des prix se comporte pratiquement comme la quantité du 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 prenons 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 significative 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: 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

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 faisons 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'effets. 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 antérieures 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 ci-dessous 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, 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 corrélation entre la quantité offerte et le prix d'équilibre. Egaleme...

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 inconvénients

Ajouter des liens avec des études méthodologiques précédentes.

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 utilisées dans ce travail, suivies par la formali-

sation économétrique du modèle théorique que nous avons spécifié dans la section 5. Après, nous expliquerons la stratégie d'identification utilisée.

Presentation 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). 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. Evidement le premier cas est le plus simple, mais pour justifier l'utilisation de cette méthode nous devrions effectuer des tests d'énogénéité 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 le 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 égaux 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'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 :

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

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

Le premier prenant la forme :

$$qo_{i,t} = a_1 + b Po_{i,t} + c X_{i,t} + u_{i,t}$$

Tandis que le dernier :

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

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 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 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 production du vin simple sous l'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 variance la plus faible.

Les propriétés de cet estimateur sont :

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

Dès le début nous envisageons que cet estimateur ne reflètera pas la nature du marché. C'est pourquoi nous, dans ce travail, testons plusieurs modèles.

Parmi les inconvénients éventuels, on a également la faible représentation des effets hétérogènes entre les départements par le modèle. Nous estimons seulement les effets moyens et ainsi nous ignorons les différences des élasticités pour des départements différents. Heureusement ce problème peut être remédié 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 separerr deux cas differents. 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 grand enseignes ayant un pouvoir de négociation significatif). Cela implique que nous pouvons considérer les prix d'être exogènes dans notre modélisation d'offre et par concequant 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'equation d'offre sont endogènes. C'est-à-dire, les agriculteurs affecte les prix par son niveau de production et par les quantités emmis sur le marché final. Afin de traiter ce problème d'endogènéité nous pouvons utiliser le méthode des variables instrumentales (IV-OLS). Mais quels instruments choisir ? Sur ce point nous adresson au travail de MacKay and Miller (2018) (et plus particulièrement le travail fondateur de Hausman (1996)), où les auteurs démontrent que les instrument d'offre choisis parmi les regresseurs de l'équation de la demande sont 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'index des prix et les revenus (voir l'annexe X). Les revenus ont la moindre corrélation avec la variable explicative quand même.

Les résultats pour les deux modèles sont présenté 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 ²	0.52	-0.87
Adj. R ²	0.52	-0.89
Num. obs.	345	345
RMSE	0.29	0.58

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

Table 4: Statistical models

Avant de commenter les effets estimés, nous devons stresser que le modèle en presence d'endogénéité n'est pas correcte (les estimateurs obtenus sont biaisé et loins de la réalité). D'abord, les instruments utilisé pour l'estimer sont assez faibles et donne pas des résultats pertinents. De plus, à partir des résultats du test de Wu-Haussman nous pouvons conclure qui nous indique que les estimateurs OLS et IV-OLS sont identiquement consistant et que nous n'avons pas des raisons pour utiliser l'estimateurs de IV-OLS (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 X), bien que la fonction de répartition partielle a une forme proche à normale (annexe X).

Nous risquons également d'avoir des biais dans la variance des estimateurs, car selon le test de Bartlett sur l'hétéroscédasticité nous rejetons l'hypothèse de l'homoscedasticité des résidus (annexe X). Au moins dans ce cas nous n'avons pas des problèmes avec l'autocorrelation, ce qui peut s'expliquer par utilisation des données temporelles en série trop courte pour y pouvoir détecter l'autocorrelation. Les résidus également ne sont pas corrélés avec des variables explicatives, mais ont une corrélation forte avec la variable dépendante, ce qui nous indique sur la misspécification possible dans notre modèle (annexe X). 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 aux résultats obtenus. L'indice des prix a un faible effet positif sur l'équilibre du marché. Néanmoins il n'est pas trop sensible de commenter ces effets plus précisément, car nous n'avons toujours pas une évidence forte sur sa exogénéité. La surface dédiée aux vins simples a également un effet positif sur la quantité du vin simple vendu, ce qui est tout à fait naturelle, vu que c'est 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 et de la baisse de l'offre d'au moins de 0.0016%, toutes choses égales par ailleurs. Les résultats en termes numériques sont assez ambigus, car nous faisons face à de nombreux problèmes dans la spécification du modèle, quand même nous pouvons 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 présents

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 et 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 des preneurs des prix du consommateur final. Nous allons comparer des résultats des 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 une misspécification éventuelle. Afin de contracter la variance des estimateurs les rendant plus efficaces nous pouvons impliquer la procédure i3SLS (*iterated three step least squares*). Cette méthode nous donne des résultats similaires à ceux obtenus par FIML (*full information maximum likelihood*). Quand même l'utilisation des procédures itératives ne semble pas donner une diminution dans les variances des estimateurs significative (annexe X).

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 regresseurs endogenes de la demande sont instrumentés tout comme décrit dans Wooldridge (2005). Cette méthode donne des résultats équivalents à ILS (*indirect least squares*), une technique utilisée pour estimer les systèmes d'équations.

Dans le deuxième cas nous rejetons l'hypothèse de l'endogénéité des prix pour le marché du vin simple, en supposant qu'il est dicté par les consommateurs finaux à l'aval (le distributeur qui rachète le vin à l'agriculteur). Cela nous amène à implémenter la méthode SUR (*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 intéresse) :

	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^2	0.46	-0.41
Offre: R^2	0.46	-0.74
Demande: Adj. R^2	0.45	-0.42
Offre: Adj. R^2	0.46	-0.75
Num. obs. (total)	690	690

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

Table 5: Statistical models

Les résultats obtenus pour le cadre avec les prix endogènes (l'estimation par 3SLS) donne forcément des résultats biaisés, car comme nous avons vu dans la section précédente nous ne disposons pas d'instruments suffisamment forts pour instrumenter l'index des prix. Cela se confirme par la non-significativité des indices des prix estimés par l'approche des variables instrumentales. Ces effets sont négatifs non-différents de zéro, ce qui n'est pas normal pour un bien n'étant le bien de Giffen (bien que nous pouvons arbitrer que le vin le plus simple pourrait se comporter d'une telle façon, car c'est un des produits alimentaires les plus basiques). Des autres tests nous confirment la misspécification de ce modèle : les résidus sont parfaitement corrélés avec les valeurs de variable dépendante et les prédictions, ce qui est évident des graphiques dans l'annexe X; une forte hétéroscédasticité, qui cette fois est apparente sur les graphiques (annexes X et X), etc. Quand même les tests de spécification ne donnent pas aucune indication de misspécification du modèle 3SLS par rapport au modèle 2SLS, ce qui nous rassure dans notre supposition que nous n'utilisons 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 la proposition du vin simple sur le marché. Néanmoins, nous préférons de ne pas utiliser ce

modèle comme notre modèle de référence en futur à cause de plusieurs sources des biais.

En ce qui concerne les résultats d'estimation par SUR, nous constatons des effets ...

M3 : Les résultats pour des départements groupés

Dans cette section nous allons commencer par la présentation et la comparaison des différentes techniques pour la clustérisation de nos données. Puis nous allons procéder avec une étude et 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.

Clusterisation

Il existe plusieurs façons de séparer et clustériser les données. Nous, dans notre travail, allons implémenter la procédure *k-means* ou *k-moyennes*. Cette procédure regroupe les individus autour des *centres* (dont le nombre est k) choisis 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 exactement ce paramètre que l'on va utiliser pour comparer les différentes approches et en choisir le plus approprié.

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 proches, 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 X.

Dans ce cas le nombre des clusters optimaux est entre 3 et 8.

Within

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

Information complète

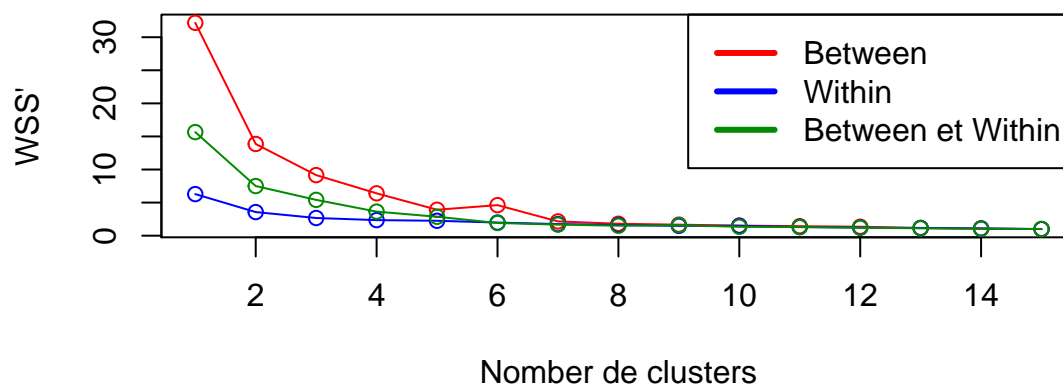
Dans le cas de l'information complète nous utilisons l'ensemble des informations disponibles sur les individus afin de construire les clusters. Nous utilisons l'approche similaire à celui utilisé dans le cas de clusterisation *within*, sauf que cette fois nous n'effectuons pas la correction sur les valeurs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

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 nombre des 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 la convergence la plus vite vers la valeur minimale de WSS. Les deux autres approches au traitement et transformation des données offrent une vitesse de convergence 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), avec la meilleure approximation absolue est atteinte pour les huit cluster et reste relativement inchangé après.

M3.1 : Le cadre en absence des interaction 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 des interférences 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	0.69*** (0.03)	4.36 (4.60)	0.15 (0.04)	0.15 (0.04)	0.51*** (0.02)	0.22* (0.10)
Surface	0.19*** (0.04)	-0.69 (1.14)	-0.54 (0.58)	-0.59 (0.58)	0.09* (0.04)	0.19** (0.07)
I pesticides	-0.16*** (0.04)	0.14 (0.52)	7.58 (4.62)	7.73 (4.65)	-0.11* (0.04)	-0.01 (0.07)
R ²	0.76	-15.69	0.92	0.92	0.79	0.55
Adj. R ²	0.76	-15.95	0.80	0.80	0.78	0.54
Num. obs.	195	195	5	5	145	145
RMSE	0.23	1.96	0.39	0.39	0.14	0.21

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

Table 6: Statistical models

M3.2 : Le cadre d'interference avec la demande

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

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

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

Table 7: Statistical models

Nous observons que ...

9. Conclusion

Nous avons étudié et comparé plusieurs modèles différents qui visent à étudier les effets d'utilisation des pesticides sur l'offre du vin de table. Nous constatons, que parmi toutes les modèles, le meilleur estimateur des effets moyens par département est obtenu avec un simple modèle OLS en absence des effets d'interaction avec la demande ou l'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 grands enseignes achètent le vin simple sous conditions *take or leave* (ce qui n'est pas vrai pour les autres types de vin).

L'analyse des données clustérisées permet d'observer une certaine degré d'hétérogénéité entre les départements et les effets d'utilisation des pesticides sur l'offre du vin ne sont pas évidents que pour un groupe d'entités étudiées.

Les résultats obtenus dans cette étude confirment des résultats d'autres chercheurs, ainsi bien 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, champignons etc. Nous captions ce fait en estimant le modèle d'équilibre du marché de vin simple.

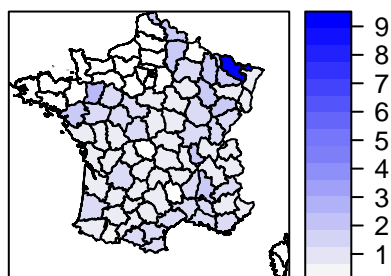
Nous devons également souligner que le modèle présenté dans ce travail est loin d'une perfection absolue. Plusieurs problèmes restent non-traités ou non-résolus. Parmi ces problèmes nous pouvons citer: un faible nombre d'observations sur la dimension temporelle, présence d'hétéroscédasticité dans les résidus, 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. Ils peuvent être intéressants d'étudier ces points et traiter ces problèmes révélés dans les études futures.

Annexes

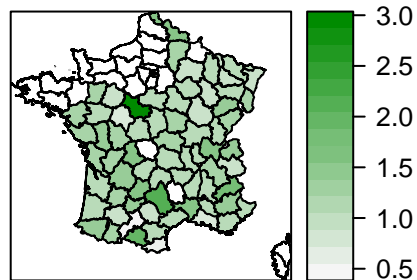
A Les statistiques descriptives

A1 Les moyennes par département

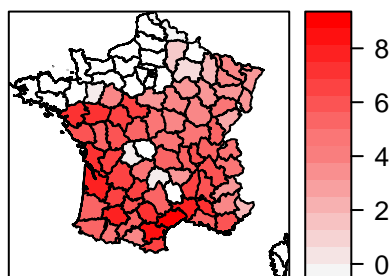
Index prix du vin par département



Index pesticides par département



Surface cultivé par département



Revenus par département

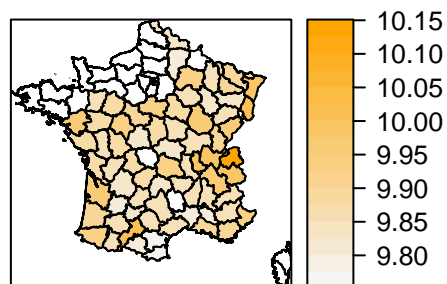


Figure 2: Les valeurs moyennes par département

A2 L'étude des interdependances

La premier de cet annexe comprend les résultats pour les données telles-quelles, le deuxieme par contre integre les résultats pour les données sous la trasformation *within*.

A2.1 Information complete

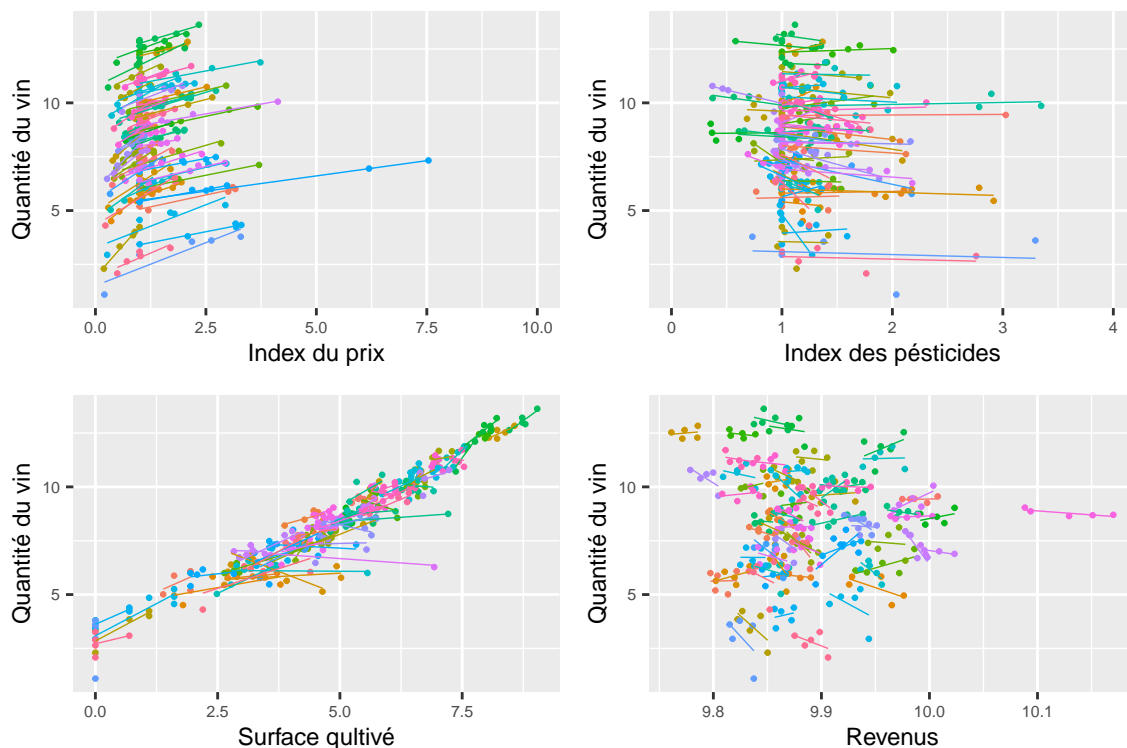


Figure 3: L'étude bivarié

	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 mieux ressortent pour les données transformées. Ce sont les données que nous intégrons dans nos modèles économétriques.

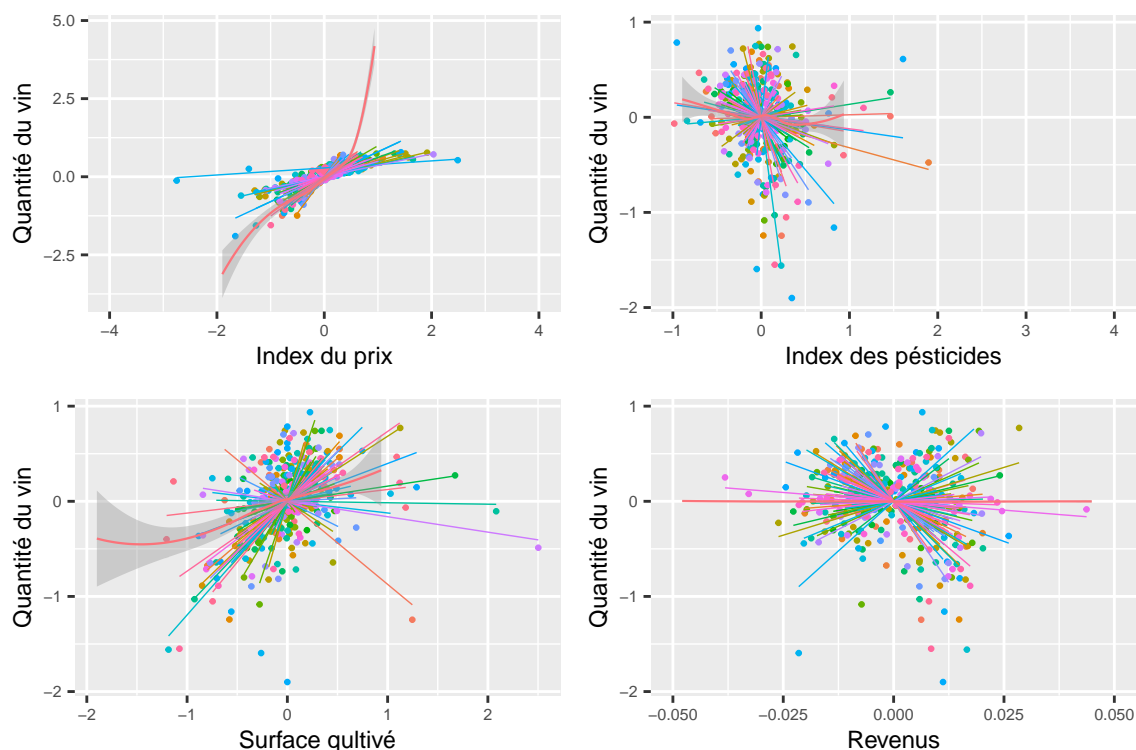


Figure 4: Relations bivariés dans le cas de transformation within

	Quantité du vin	IP	Surface	Revenus	Index pesticides	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 pesticides	-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

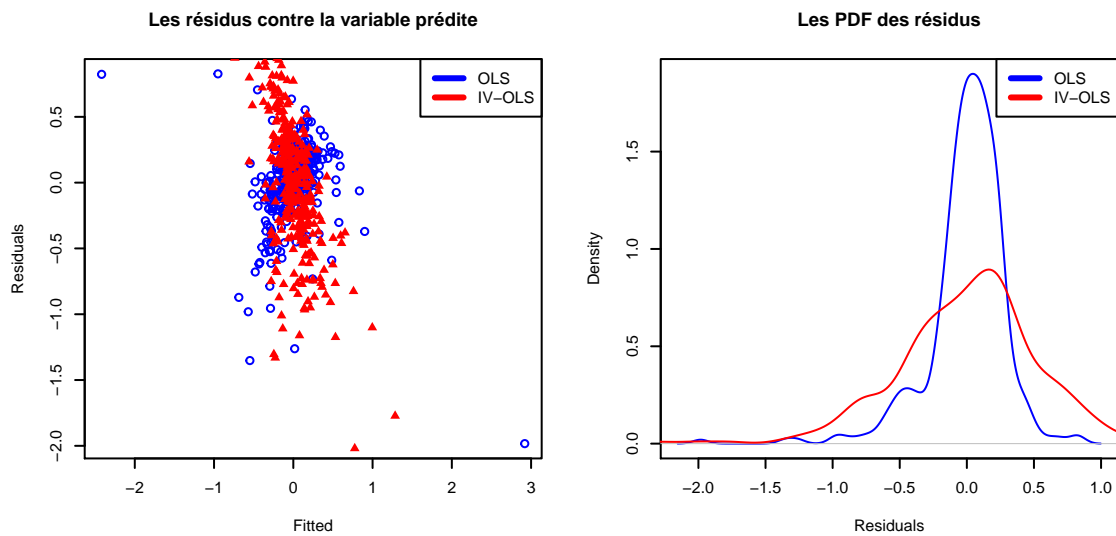


Figure 5: Le comportement des résidus

B2 L'autocorrelation

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

	OLS	IV-OLS
Equation d'offre	0.627	0.637

B3 Test de l'hétéroskedacité

Table 12: Les résultat du test de Bartlett sur l'heteroscedacité, 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 Diagnostiques IV-OLS

Table 14: Diagnostiques 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^2	0.46	-0.41	-0.41	-0.41
Offre: R^2	0.46	-0.87	-0.74	-0.75
Demande: Adj. R^2	0.45	-0.42	-0.42	-0.42
Offre: Adj. R^2	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: Statistical models

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

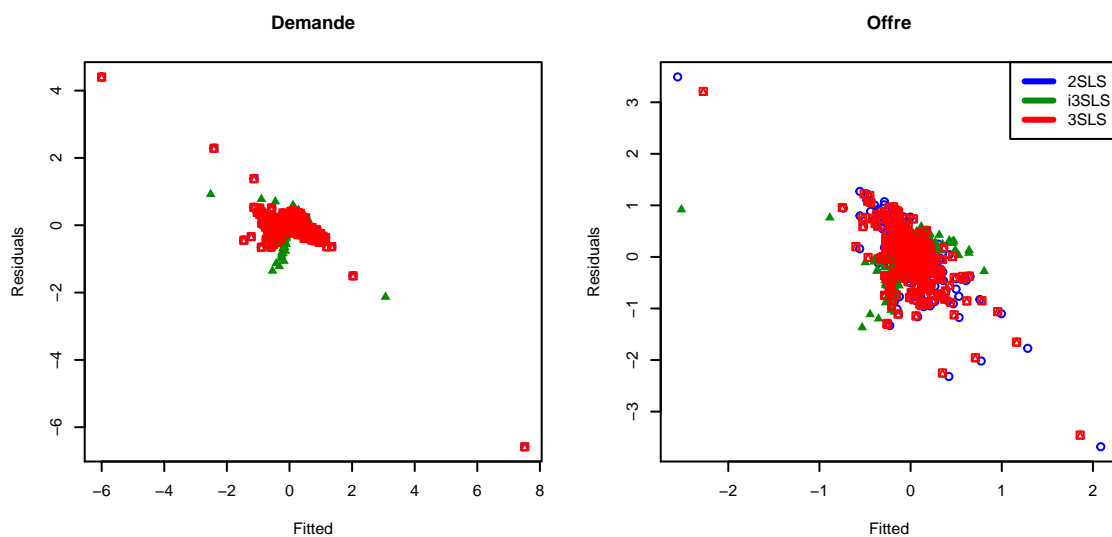


Figure 6: Les résidus contre la variable prédite

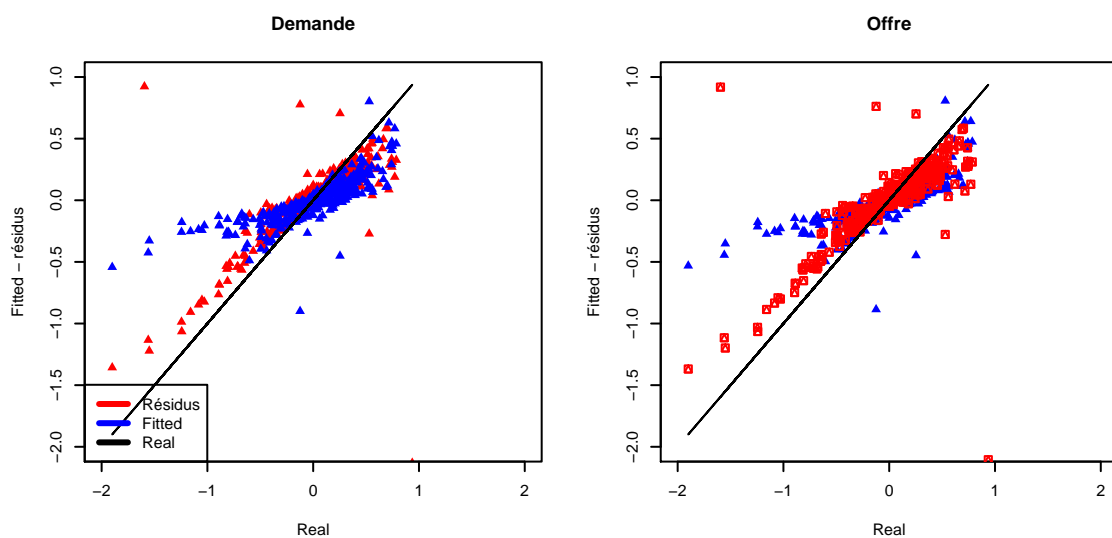


Figure 7: Les résidus et les prédictions, le cas de i3SLS

C3 L'autocorrelation

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

	SUR	2SLS	3SLS	i3SLS
Equation de demande	0.687	0.618	0.618	0.561
Equation d'offre	0.683	0.637	0.638	0.640

C4 Test de l'hétéroskedacité

Table 18: Test de Bartlett sur l'heteroskedacité, 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

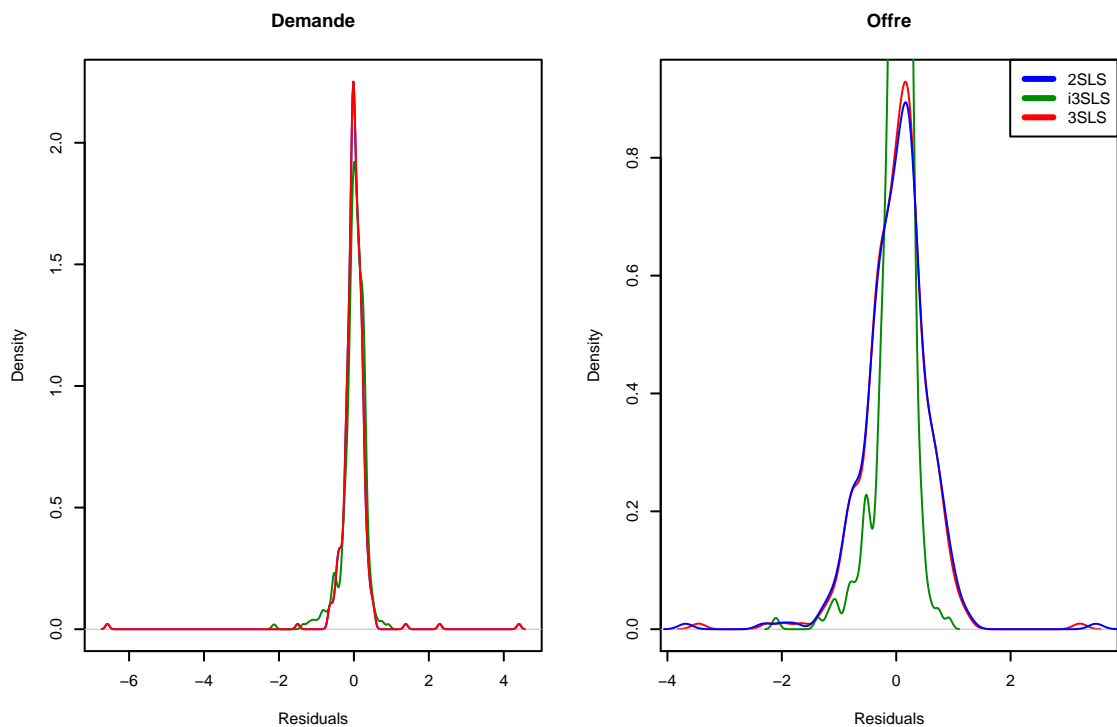


Figure 8: Les PDF des résidus

C6 Comparaison des modèles

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

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

#Df	LogLik	Df	Chisq	Pr(>Chisq)
8	816.0971	NA	NA	NA
6	-514.9105	-2	2662.0151	0
8	-504.9661	2	19.8887	0
8	-505.6559	0	1.3797	0

D Clusterisation

D1 *Between* transformation

Les groupes sont définies par des caractéristiques suivantes :

Table 22: Les centres des clusters

	qi	ipi	si	ri	iki	.1
1	5.445	1.872	2.395	9.882	1.290	18
2	10.940	1.332	7.000	9.867	1.290	21
3	8.245	1.235	4.914	9.914	1.213	30

D2 *Within* transformation

Les groupes sont définies par des caractéristiques suivantes :

Table 23: Les centres des clusters

	qi	ipi	si	ri	iki	n	k	t
1	-1.595397	-7.941881	-0.261529	-0.02149	-0.049931	1	1	1
2	0.253285	-1.404496	-0.153318	0.001641	-0.013255	1	1	2
3	0.529538	2.487263	0.747453	-0.004844	0.080087	1	1	3
4	0.936501	9.608734	0.226164	0.006494	-0.034652	1	1	4
5	-0.123927	-2.749621	-0.55877	0.018199	0.017751	1	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	0.001186	-0.359222	0.04895	-0.012025	-0.234518	29	3	1
12	-0.280464	-0.376425	0.078291	-0.000723	0.004804	29	3	2
13	0.049743	0.107731	-0.01841	-0.006913	0.122468	29	3	3
14	-0.032737	0.065232	-0.147634	0.004155	-0.089222	29	3	4
15	0.262272	0.562683	0.038803	0.015506	0.196468	29	3	5

D3 Cas d'information complete

Les groupes sont définies par des caractéristiques suivantes :

Table 24: Les centres des clusters

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

References

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.

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/filieres-Vin-et-cidre/Vin/En-un-clic/Dossiers-des-Conseils-et-comites?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.

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’.

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 Culturelles.” *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 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.

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>’.

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.