Essays in Empirical Economics on the Determination of Wine Prices

Emmanuel Paroissien

UNIVERSITÉ DE BORDEAUX

THÈSE

pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

délivré par l'Ecole Doctorale Entreprise Economie et Société

Discipline: Economie

ESSAYS IN EMPIRICAL ECONOMICS ON THE DETERMINATION OF WINE PRICES

---

soutenue publiquement le 8 décembre 2017 par

Emmanuel PAROISSIEN

Directeurs:
Jean-Marie CARDEBAT, *Université de Bordeaux*
Michael VISSER, *Centre de Recherche en Économie et Statistique*

Président du Jury:
Jean-Marc FIGUET, *Université de Bordeaux*

Rapporteurs:
Julian ALSTON, *University of California at Davis*
Victor GINSBURGH, *Université Libre de Bruxelles*

UNIVERSITÉ DE BORDEAUX

THÈSE

pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

délivré par l'Ecole Doctorale Entreprise Economie et Société

Discipline: Economie

ESSAIS D'ECONOMIE EMPIRIQUE SUR LA DETERMINATION DES PRIX DU VIN

---

soutenue publiquement le 8 décembre 2017

Emmanuel PAROISSIEN

Directeurs:
Jean-Marie CARDEBAT, *Université de Bordeaux*
Michael VISSER, *Centre de Recherche en Économie et Statistique*

Président du Jury:
Jean-Marc FIGUET, *Université de Bordeaux*

Rapporteurs:
Julian ALSTON, *University of California at Davis*
Victor GINSBURGH, *Université Libre de Bruxelles*

In vino veritas

Ecole Doctorale:
Entreprise, Économie et Société
École Doctorale de Sciences Économiques, Gestion et Démographie
Université de Bordeaux
avenue Léon Duguit
33608 Pessac

Laboratoires d'Accueil:
Laboratoire d'Analyse et de Recherche en Economie et Finance Internationales
Université de Bordeaux
avenue Léon Duguit
33608 Pessac

et

Laboratoire d'Economie Industrielle
Centre de Recherche en Economie et Statistique
5 avenue Henri le Châtelier
91120 Palaiseau

Groupe de Recherche:
Bordeaux Wine Economics
Université de Bordeaux
Avenue Léon Duguit
33608 PESSAC, Franc

Remerciements

Cette thèse est le résultat de la collaboration fructueuse entre deux institutions de la recherche publique, l'Université de Bordeaux et le CREST, et une entreprise privée, le Bureau François Lillet (BFL). Quatre années ont été nécessaires à sa réalisation, quatre années durant lesquelles j'ai reçu plus d'aide et de soutiens que je n'aurais pu en espérer en me lançant dans cette aventure. Je mesure toute la chance qui m'a été donnée d'évoluer dans trois environnements si différents durant mes recherches. Entre le BFL et les deux laboratoires, j'ai eu l'occasion de fréquenter une grande diversité de profils du secteur privé ou public, bordelais ou parisiens. Chacune de ces expériences a été riche d'enseignements, tant pour la qualité de cette thèse que sur le plan personnel. En quatre ans, tant de personnes ont contribué d'une manière ou d'une autre au succès de ce projet qu'il me serait impossible d'en faire la liste exhaustive. Je voudrais simplement reconnaître ici les apports qui ont été essentiels.

La qualité de cette thèse doit tout d'abord énormément à l'implication de mes directeurs de thèse, Jean-Marie Cardebat et Michael Visser. Tous deux ont fait preuve d'une grande patience avec moi et ce succès doit beaucoup aux qualités humaines de l'un et de l'autre. Je ne saurais leur témoigner suffisamment de reconnaissance.

Jean-Marie Cardebat m'a mis le pied à l'étrier et m'a réservé un formidable accueil à l'Université de Bordeaux. Ce projet n'aurait tout simplement jamais vu le jour sans lui. Il a posé les bases de l'équilibre entre les objectifs respectifs des trois parties prenantes, parfois orthogonaux. Sa médiation active m'a permis de me focaliser sur mon travail de thèse. Jean-Marie Cardebat a également et surtout directement contribué au travail de recherche. Il est en particulier co-auteur des deux premiers chapitres et a géré leurs publications au *Journal of Wine Economics*. C'est véritablement lui qui m'a formé aux problématiques de l'économie du vin, et la qualité globale de ce mémoire lui doit énormément.

Michael Visser, avant d'accepter d'être mon directeur de thèse, a été mon premier professeur d'économétrie à l'ENSAE. Je lui dois l'essentiel de mes connaissances sur le sujet, et bien plus encore. Il a apporté une grande sérénité tout au long des quatre années qu'a duré cette collaboration, et s'est déplacé plusieurs fois à Bordeaux pour participer aux réunions de médiation. A Paris, nos conversations au CREST ont été à chaque fois des étapes clés dans l'évolution de mes recherches. Il est le co-auteur du troisième chapitre, et ses nombreuses relectures ont permis d'améliorer substantiellement la qualité de l'anglais de l'introduction et du dernier chapitre. En toute circonstance, il a su à la fois exiger un niveau de qualité ambitieux pour les recherches, et m'insuffler la confiance nécessaire pour les mener à bien.

Je voudrais remercier les autres membres du jury chargés d'évaluer cette thèse, et en particulier Julian Alston et Victor Ginsburgh pour m'avoir fait l'honneur d'être les rapporteurs de cette thèse. Je suis également reconnaissant envers Jean-Marc Figuet pour avoir accepté de présider ce jury. Jean-Marc Figuet a aussi directement contribué au premier chapitre de cette thèse et en est un co-auteur.

J'adresse mes remerciements personnels à François Lillet qui est à l'origine de ce projet de thèse, et qui m'a fait confiance pour le mener à terme. Les bonnes conditions de travail qu'il m'a offertes ont été le meilleur garant du succès de ce projet. Il m'a accueilli dans ses locaux au BFL pendant trois ans, durant lesquels j'ai joui d'une grande liberté dans l'organisation de mon travail et où j'ai pu échanger régulièrement avec les courtiers. Il s'est également personnellement impliqué dans la conduite du travail de thèse, notamment en me permettant de rencontrer plusieurs acteurs-clés de l'industrie du vin à Bordeaux. Je remercie également l'ensemble du personnel du BFL pour leur bon accueil au sein de l'équipe. En particulier, je voudrais témoigner tout mon respect et ma gratitude envers Marion Tarel et Antoine Moga qui m'ont énormément appris sur le fonctionnement du marché du vin à Bordeaux. Ils n'ont épargné ni leur temps ni leur énergie pour assurer la médiation complexe entre les intérêts des laboratoires et ceux de l'entreprise.

Je remercie Antoine Bouët qui m'a orienté vers Jean-Marie Cardebat dès 2013 alors que je cherchais un projet de thèse, et qui m'a ensuite accueilli au LAREFI dont il était le directeur en 2014. Je remercie aussi tous les professeurs du laboratoire qui se sont montrés intéressés par mon sujet de thèse et qui ont partagé leur science avec moi. Je suis particulièrement reconnaissant aux doctorants de l'Université de Bordeaux qui m'ont réservé un accueil chaleureux quand je n'étais pas familier dans l'environnement de l'université. Je remercie particulièrement Viola Lamani, Samuel Klebaner, Linda Jiao, Laurent Baratin, Thomas Humblot, Marine Coupaud et François Viaud.

Je remercie Francis Kramarz et Laurent Linnemer pour m'avoir accueilli au sein du Laboratoire d'Économie Industrielle (LEI) du CREST, d'abord par intermittence durant trois ans puis de façon permanente pour ma quatrième année. Durant cette dernière année, j'ai profité pleinement de la qualité de l'environnement du CREST. J'ai également pu collaborer plus étroitement avec Michael Visser, ce qui a permis la finalisation de cette thèse dans les meilleurs conditions. Je remercie les autres chercheurs du CREST pour leurs commentaires et conseils lors de mes présentations en séminaire, notamment Thibaud Vergé, Philippe Choné, Alessandro Iaria, Laurent Davezies et Xavier D'Haultfoeuille. Je remercie également les doctorants du LEI, Etienne Chamayou, Hugo Molina, Morgane Cure, Julien Monardo, Tomas Jagelka et Jiekai Zhang, pour leur accueil et toutes nos discussions.

Jean-Marie Cardebat m'a également permis de bénéficier des conseils des membres du groupe de recherche Bordeaux Wine Economics, qui ont affermi ma connaissance du marché du vin et de ses enjeux économiques. Je pense en particulier à Karl Storchmann, Stephen Bazen, Florine Livat-Pécheux, Benoit Faye, Eric Le Fur, Olivier Gergaud, Stéphanie Prat, Adeline Alonso Ugaglia, Philippe Masset et Jean-Philippe Weisskopf.

Cette thèse a également bénéficié des commentaires et des conseils des participants aux conférences de l'American Association of Wine Economics organisées en 2016 et 2017, et en particulier ceux d'Orley Ashenfelter et d'Eddie Oczkowski. Je remercie aussi les participants à la conférence de l'Agricultural & Applied Economics Association de 2017 pour leurs conseils et leur sympathie, notamment Daniel Sumner, Robin Goldstein et Stephen Ziliak.

Enfin, je suis reconnaissant envers Eric Giraud-Héraud, Christophe Gouel et Jean-Marc Boussard pour leur intérêt dans mes recherche et leurs conseils précieux.

La finalisation de cette thèse doit au moins autant au soutien quotidien de mes proches qu'à toutes les personnes citées précédemment. Je terminerais simplement en exprimant toute ma gratitude à celles et ceux qui m'ont accompagné durant ces quatre années.

Acknowledgements

This thesis is the result of a fruitful collaboration between two public research institutes, the University of Bordeaux and the Crest, and one firm, the Bureau François Lillet (BFL). Four years have been necessary for its completion, four years during which I have received more help and support than I could expect at the beginning of this adventure. I realize how lucky I am to have been able to work in three completely different environments. In the BFL and the two laboratories, I had the occasion to meet a wide diversity of profiles from the private or public sector, *Bordelais* or *Parisian*. I have drawn important lessons from each of those experiences, both for my research and on a personal level. In four years, so many people have contributed in one way or another to the success of this project that it would be impossible for me to make an exhaustive list. I would simply like to acknowledge the contributions that have been essential.

The quality of this thesis first of all owes a lot to the involvement of my thesis directors, Jean-Marie Cardebat and Michael Visser. Both have shown great patience with me and the general human qualities of both have been crucial in this success. I can not give them enough credit.

Jean-Marie Cardebat put me on track and gave me a wonderful welcome at the University of Bordeaux. Simply put, this project could not have been possible without him. He laid the foundation of the balance between the respective objectives of the three stakeholders, which were sometimes orthogonal. His active mediation allowed me to focus on my dissertation. Jean-Marie Cardebat has also and above all directly contributed to the research work. He is co-author of the first two chapters and has handled their publications in the textit Journal of Wine Economics. It is really him who introduced me to wine economics and the overall quality of this memory owes him a lot.

Michael Visser, before accepting to be my thesis director, was my first professor of econometrics at ENSAE. But I owe him most of my knowledge on the subject, and much more. He brought his serenity throughout the four years that this collaboration lasted, and traveled several times to Bordeaux to attend to mediation meetings. In Paris, our conversations at CREST have always been key steps in the evolution of my research. He is the co-author of the third chapter, and his many re-readings have substantially improved the quality of the English of the introduction and the last chapter. In any case, he has both demanded an ambitious level of quality for research and given me the confidence to carry it out.

I would like to thank the other members of the jury in charge of the evaluation this dissertation, and in particular Julian Alston and Victor Ginsburgh for doing me the honor of being the rapporteurs of this dissertation. I am also grateful to Jean-Marc Figuet who accepting to preside this jury. Jean-Marc Figuet has also contributed directly to the first chapter of this thesis, of which he is a co-author.

I send my personal thanks to François Lillet who is behind this thesis project and who trusted me to complete it. The good working conditions that he offered me were the best guarantee of success for this project. In particular, he welcomed me to his premises for three years during which I enjoyed great deal of freedom in my schedule, and where I was able to meet regularly with the brokers. He was also personally involved in the conduct of the thesis work, and notably introduced me to several key players in the wine industry in Bordeaux. I also thank all BFL staff for having welcomed me in the team. In particular, I would like to express my respect and gratitude to Marion Tarel and Antoine Moga who have taught me a lot about how the Bordeaux wine market works. They spared neither their time nor their energy to manage the complex mediation between the objectives of the laboratories and those of the firm.

I thank Antoine Bouët who introduced me to Jean-Marie Cardebat in 2013 while I was looking for a thesis project, and who welcomed me at the LAREFI where he was director in 2014. I also thank all the professors of the laboratory who showed interest in my thesis topic and shared their science with me. I am particularly grateful to the PhD students at the University of Bordeaux who gave me a warm welcome when I was not familiar with the environment of the university. I particularly thank Viola Lamani, Samuel Klebaner, Linda Jiao, Laurent Baratin, Thomas Humblot, Marine Coupaud and François Viaud.

I thank Francis Kramarz and Laurent Linnemer for welcoming me at the Laboratoire d'Economie Industrielle (LEI) in the CREST, first intermittently for three years and then permanently for my fourth year. During this past year, I have taken full advantage of the quality of the environment in the CREST. I was also able to collaborate more closely with Michael Visser, which permitted the finalization of this thesis under the best conditions. I thank the other CREST researchers for their comments and advice during my seminar presentations, including Thibaud Vergé, Philippe Choné, Alessandro Iaria, Laurent Davezies and Xavier D'Haultfoeuille. I also thank LEI PhD students, Etienne Chamayou, Hugo Molina, Morgane Cure, Julien Monardo, Tomas Jagelka et Jiekai Zhang, for their welcome and all our discussions.

Jean-Marie Cardebat also allowed me to benefit from the advice of members of the research group Bordeaux Wine Economics, which strengthened my knowledge of the wine market and its economic stakes. I am thinking in particular of Karl Storchmann, Stephen Bazen, Florine Livat-Pécheux, Benoit Faye, Eric Le Fur, Olivier Gergaud, Stephanie Prat, Adeline Alonso Ugaglia, Philippe Masset and Jean-Philippe Weisskopf.

This thesis also benefited from comments and advice from participants at the American Association of Wine Economics conferences held in 2016 and 2017, and in particular from those of Orley Ashenfelter and Eddie Oczkowski. I would also like to thank the participants at the 2017 Agricultural & Applied Economics Association for their guidance and sympathy, including Daniel Sumner, Robin Goldstein and Stephen Ziliak.

Finally, I am grateful to Eric Giraud-Héraud, Christophe Gouel and Jean-Marc Boussard for their interest in my research and their precious advice.

The finalization of this thesis must at least as much to the daily support of my relatives as to all the people aforementioned. I would simply end by expressing my gratitude to those who have accompanied me during these four years.

Contents

General Introduction

1 Motivation

As far back as written history goes, humanity has consumed wine. As a matter of fact, wine production has been traced back to the 4th millennium BC (Barnard et al., 2011), which makes it older than writing itself. Today, wine represents about 0.4% of global household consumption (Anderson et al., 2003), but the high geographic concentration of the wine industry makes it essential to the economy of many regions of the world. In France, the wine sector accounts for 1.2% of GDP, employs more than 550,000 employees, and provides around 15% of world production. In the emblematic French wine region of Bordeaux, the wine industry accounts for about 80% of the total value of agricultural production, and vineyards cover half of arable land. Countries specialized in wine production also enjoy substantial returns from wine exports, especially in France where wine exports bring about 8 billions euros each year. In addition to direct economic returns, the wine industry generates indirect benefits. Each year, the oenotourism industry attracts approximately 4 millions foreign tourists in France, with considerable spillovers into other tourism-related sectors such as transports and hospitality. Beyond economic considerations, wine consumption is a part of everyday life in wine countries such as France or Portugal, where wine consumption per capita approaches 43 liters per year; equivalent to one glass a day. It is a part of their cultural identity - a unique product imbued with symbolism.

From an economic research standpoint, the wine market has proven to be a fruitful application domain for various fundamental concepts, including the seminal cost-benefits analysis of Adam Smith and the theory of comparative advantages of David Ricardo (see Chaikind 2012 for a review). One feature of the wine market that makes it particularly appealing to economists is the outstanding price discrepancy among almost identical-looking products. For instance, a standard-size bottle of 0.75 liter of entry-level Bordeaux wine sells at around 5€ whereas a bottle of the famous Bordeaux-based brand Pétrus can be worth up to 5,000€[1]. This tremendous heterogeneity makes the wine market an ideal case study for a quantitative analysis of quality. Furthermore, wine quality can only be accurately assessed after consumption, which makes it an experience good in the terminology of Nelson (1970). In this respect, the efficiency of the wine market crucially depends on the accuracy of the quality signals available to consumers. Economists focusing on the wine market have thus questioned the relevance of quality signals and their impact on prices, see Storchmann et al. (2012) for a review. The accuracy of quality scoring has been a key matter of concern not just for wine economists. Even Robert Parker, arguably one of the most influential wine experts in history, has conceded that wine tasting is partly driven by "the emotion of the moment"[2]. Today, many experts award quality scores, and most premium wines are evaluated by multiple scores, which in turn has mitigated the respective influence of each expert. Nonetheless, many still question the influence of the subjective opinions of wine experts on prices. This situation has also raised the as yet unresolved issue

[1] This is the current retail price on the website www.winedecider.com for a Château Pétrus of vintage 2000, as of October 2017.

[2] The full quotation is: "I really think probably the only difference between a 96-, 97-, 98-, 99-, and 100-point wine is really the emotion of the moment.", Mobley-Martinez (2007).

concerning the comparison of the scores across different experts. Moreover, all these experts only operate on the premium segment of quality, the rest of the producers relying on coarser quality signals. For the vast majority of wine producers, the main medium of quality differentiation is to participate in wine competitions and win medals to add to their labels. However, only few studies have addressed their efficiency and influence so far. Broadly speaking, the community of economists has primarily focused on the top end quality segment, for which the large price discrepancy is better-suited to the quantitative analysis of the impact of quality. Consequently, the bulk of the wine market has been somewhat neglected in the literature so far. In particular, classical issues in agricultural economics such as the analysis of producers' price expectations and the role of storage have been overlooked. It is all the more surprising that no futures market exists for wine, and as such producers and other actors of the market have to formulate their price expectations independently. With this dissertation, I intend to bridge these gaps of the literature adopting a resolutely empirical approach.

Regarding all the aforementioned issues, the iconic wine region of Bordeaux is an especially attractive application market. Firstly, the quality heterogeneity of the region is unmatched, with about five thousands private brands called Châteaux, more than fifty appellations of origin and several official rankings. Premium Bordeaux wines also attract the largest panel of wine critics - top-end Châteaux wines usually being evaluated by more than ten different experts. The market is therefore particularly appealing to study in consideration of the influence of experts and their respective grading methods. Secondly, Bordeaux producers are particularly prevalent in wine competitions and around 20% of Bordeaux wines have won at least one medal. Lastly, Bordeaux wine producers systematically report their transactions dealt in bulk to the joint-trade organization, and have to declare their annual harvests, ending stocks and monthly deliveries to the customs. Taken together, this data provides a unique opportunity to investigate the determinants of the fluctuations of wine prices. For all these reasons, the Bordeaux wine market will be systematically taken as a case study in the four chapters of this dissertation.

The first chapter develops and applies a new method to disentangle the effects of objective quality and subjective opinions on wine prices[3]. This work has raised the issue of the comparability of scores given by different experts on different scale. In the second chapter, a solution is suggested with the introduction of the equipercentile method in the wine economics literature[4]. If these two first chapters follow the tradition of wine economists and address the narrow super-premium segment of the wine market (typically above 20€ a bottle), the rest of the dissertation investigates new issues on the wider lower-quality segment. The third chapter explores the market of wine competitions, and estimates the impact of the awards on producers' prices[5]. The original data also allows to statements to be made on the consistency of European wine competitions and on the expected profit of participation. In the fourth and last chapter, forecasting methods are applied to Bordeaux wine prices. A fruitful comparative analysis between annual and monthly forecasts is conducted in order to assess the interest of the longer data history of the annual data in comparison to the monthly data. The estimates provide several collateral contributions on the influence of macro-level determinants of wine prices such as total wine stocks, weather, harvest expectations and exchange rates.

Several areas of wine economics are not addressed in this thesis, notably the financialization of the fine wine market. My overall aim has been to improve our understanding of the

[3]This first chapter is based on a paper co-written with Jean-Marie Cardebat and Jean-Marc Figuet published in the *Journal of Wine Economics* (Cardebat et al., 2014). This chapter is an updated version that includes a few corrections.

[4]This second chapter is based on a paper co-written with Jean-Marie Cardebat and also published in the *Journal of Wine Economics* (Cardebat and Paroissien, 2015). The version presented in this chapter includes a few precisions.

[5]This third chapter is based on a paper co-written with Michael Visser.

formation of wine prices and its determinants. At the micro-level, quality in the wider sense is the key. This is the central focus of the three first chapters, which examine the means of achieving and evaluating quality differentiation. The fourth chapter widens the focus by including all the determinants of economic conditions of the wine market as a whole. Each chapter is intended to be self-sufficient so that some elements are redundant, notably parts of the different introductions.

The remainder of this general introduction is organized as follows: the first section describes the organization of the wine industry. The second section is a review of the literature on the determinants of wine prices. Finally, I summarize the four chapters of this thesis and present how they relate to the existing literature.

2 The wine industry organization

In this section, the main characteristics of the global wine market are presented. As the four chapters of this thesis examine data of the Bordeaux region only, the primary focus shall be on the case of this region. The section begins with a short summary of the steps of wine production then goes on to describe the geographic distribution of the wine industry, providing insights of its determinants. Later, the section considers the economics of wine production, industrial organization, and consumption. The last subsection lists the main long-term trends and the recent evolutions of the global wine market.

2.1 From seed to glass

The process of wine production is arguably one of the most complex processes in agriculture. It begins with the plantation of grapevines, which can take years before actually yielding exploitable grapes. The choice of the grape variety must be carefully adapted to the climate and soil of the location of production. The vines produce grapes every year for 30 to 60 years, depending on grape variety, climate, and cultivation techniques. After the harvest, the grapes are brought to the winery where winemaking takes place. The transformation of harvested grapes into wine necessitates a large number of delicate chemical reactions - so many that winemaking is often referred to as an "art". The harvested grapes are first crushed in large tanks to liberate the juice, allowing it to ferment and produce alcohol. The alcoholic degree of the final product depends on both the sugar content of the grapes and the duration of the fermentation. For white wines, the juice is extracted before the fermentation by pressing the grapes. On the contrary, red wines are obtained by letting the juice ferment with the must for a sufficient amount of time to allow the color to impregnate. The rosé color can be obtained by early pressing, by extracting only a part of the must (bled rosé), or by blending red and white wines - a process almost only used for rosé Champagne. Moreover, a second fermentation (the malolactic fermentation) is often sought for red wines and some white wines in order to reduce acidity. While these stages are the most essential, the refinements of winemaking involves many other technical operations which are beyond the scope of this presentation. Most wines are obtained from a single grape variety, but some producers prefer to blend different varieties, particularly in the Bordeaux region, which considerably increases the complexity of the whole process. The choice, the order and the duration of the production stages, together with the care brought to each of them, defines the identity of a winemaker. The quality of the final product also crucially relies on the quality of the harvested grape, but talented winemakers may be able to mitigate the defaults of a poor harvest. Finally, fine wines are often kept in barrels to be imbued with oak flavors, often for as long as eighteen months for premium Bordeaux. The final product is then usually packaged in a bottle, although certain producers opt for a cheaper bag-in-box package. The latter type of package allows a better conservation after opening, but is less suited for long-term keeping. Indeed, some wines are ready to drink right after packaging but others must be kept up to ten years before optimal consumption and must therefore be kept in a bottle. Adapted and stable conditions of temperature and humidity are necessary to ensure the wine completes its potential.

In response to the complexity of winemaking, the appreciation and judgment of wines is also often viewed as an art. To this extent, that wine consumption is more than just drinking; it is probably the consumption good for which actual consumption requires the most skill. Professional wine tasters are remunerated to comment and evaluate the appearance, aromas, flavors and aftertaste of wines. Their tasting skills are especially useful to assess the potential of wines destined to age several years before consumption, as during the *primeurs* campaign in Bordeaux. However, the vast majority of wines do not fall in that category and are quickly consumed in the few years after production. Novices in wine tasting are often insensitive to the complexity of certain aromas, so that some fine wines are only fully appreciated by highly trained palates.

2.2 Geographic distribution

The organization of the global wine industry is mostly dictated by climatic conditions. Broadly speaking, grapevine cultivation is only possible between the 30th and the 50th parallels[6] of both hemispheres, apart from some micro-climates. As these latitudes are mostly covered by oceans in the South, around 90% of the wine production is achieved in northern countries. In fact, the bulk of wine production is concentrated into a handful of countries: France, Italy and Spain account for approximately 45% of the global production in volume. Around 85% of global wine production is achieved by only ten countries: Italy (16%), Spain (16%), France (15%), USA (11%), China (6%), Argentina (5%), Chile (4%), Australia (4%), South Africa (4%) and Germany (3%)[7].

To a certain extent, the geographic distribution of wine consumption replicates that of wine production. In particular, Europe accounts for around two-thirds of both production and consumption. Some non-wine producing countries have high consumption levels. Countries such as the United Kingdom or Sweden consume over 20 liters of wine per capita each year, twice as much as the USA. Of course, the most important wine-drinking countries remain those with a large production, such as France (43 l.), Italy (34 l.) or Portugal (44 l.)[8]. Spanish domestic consumption per capita is, conversely, rather low (22 l.) relative to its production. Spain is in turn the world leader in wine exports. Wine consumption per capita is virtually null in most countries of Southern Asia and Sub-Saharan Africa. Indeed, these climates are unsuitable for viticulture, and considering wine as a luxury product, many of their inhabitants cannot afford to drink it.

Beyond climate conditions and financial constraints, cultural preferences complete the puzzle of why some populations drink wine and others do not. Religious views on alcohol have at least partly shaped regional preferences. In particular, Islam has prohibited alcohol consumption since its origin in the 6th century. In countries like Saudi Arabia, the civil law strictly observes this precept and a strict ban on all alcohols is enforced. Even when this precept is not enforced in the law, almost no wine is produced nor consumed in Muslim countries. For example, although Turkey exhibits a fine climate for viticulture and a GDP per capita comparable

[6] Notably, the 45° parallel has been coined as "the ideal latitude for fine wines" by Bernard et al. (2014), since it runs through several of the most renowned vineyards, including Bordeaux.

[7] These figures are from the Food and Agricultural Organization (FAO), and concern the production of year 2014, the last available on the FAO website.

[8] Figures are from year 2014 and published by the Californian Wine Institute.

to that of many occidental countries, its wine production is very low and its consumption per capita is virtually zero. It should also be noted that their taxes (import plus excise plus VAT) represent more than 200% of wine value, against about only 20% in France[9].

Conversely, some seminal Christian rites and beliefs actually favor wine consumption[10], as in southern Europe where wine is a part of everyday life. Even before Christianity, wine consumption was already popular there and considered spiritual[11]. Of course, the Christian religion does not intrinsically command one to drink wine due to the sinful nature of overindulgence, resulting in drunkenness[12]. Cultural preferences and religious beliefs arguably influence each other, and only partly determine consumption habits. Nonetheless, current alcohol prohibition is based on religious grounds in the Middle East whilst no religious disincentives exist in Southern Europe.

2.3 Economics of the wine market

The economics of wine production is characterized by a high level of uncertainty, for both the producers and the consumers. From the perspective of the producer, planting grapevines is a long-term investment without return in the first years, and with uncertain returns once the vines start producing grapes. Indeed, newly planted vines do not produce significant volumes for a few years and full yields are only expected after a dozen years. At maturity, the quality of the harvest is uncertain since it crucially depends on weather vagaries. Short episodes of extreme temperatures or precipitations can jeopardize a whole harvest, both in terms of volume and quality. This vintage effect is more or less intense across wine regions, depending on the year-to-year variability of the weather. For instance, the quality of Californian wines is believed to vary little across vintages, while Bordeaux vintage quality is relatively volatile. Besides, the grape harvest occurs only once a year and producers have to manage important stocks during the marketing year. The arbitrage between waiting and selling is based on their expectations about future market conditions. For most agricultural commodities, this issue is partly resolved by the existence of futures markets where producers can sell their harvest in advance and hence secure their profits. Since no such futures market exists for wine, producers cannot hedge against market fluctuations and rely solely on their expectations regarding future prices and production levels. The profits producers can expect from planting vines are therefore highly uncertain.

Like most agricultural markets, the industrial organization of the wine industry is characterized by a multitude of independent grapegrowers. In most of the New World, the grapegrowers sell their harvests to a winery which then manages the actual winemaking. By contrast, most European wine producers are grapegrowers and winemakers at the same time. Both in the Old and the New World, the majority of winemakers promote their own brand, leading to a unique diversity of products. Whether they have actually managed the grapegrowing or had grapes provided by upstream producers, a plethora of independent winemakers coexist and

[9] These figures are taken from Anderson and Nelgen (2011) and concern only non-premium wines.

[10] The seminal miracle of Christianity is the transformation of water into wine at the marriage at Cana. Furthermore, consumption of wine is an essential rite to commemorate Jesus' Last Supper.

[11] Back in Antiquity the people of Ancient Greece regularly consumed wine, and worshiped the wine god Dionisos. Later, the Romans renamed him Bacchus and continued its cult until finally adopting the Christian religion. Interestingly, several other ancient Mediterranean polytheisms also included wine deities, which they associated with fundamental attributes of humanity such as life and death (Osiris in Ancient Egypt, Sucellus in Ancient Gaul), or renewal and fertility (Dionisos, or Liber Pater in Ancient Rome).

[12] Several Protestant churches notably recommended abstinence or prohibition. They laid the cornerstone to the temperance movement of the 19th and 20th centuries in the USA and Northern Europe. Total or partial alcohol prohibition has even been enforced in the USA (1920-1933) and the Nordic countries of Iceland (1915-1922), Sweden (1914-1955), Norway (1916-1923) and Finland (1919-1932). But these periods of abstinence were too short to durably influence the drinking habits and now most of these countries exhibit a relatively high wine consumption per capita.

each label their wines under their own brand. The number of independent producers is especially large in Europe. In France for example, more than 100,000 different private wine brands coexist, against around 4,600 in California and 2,500 in Australia. The Bordeaux region alone hosts more different independent wine producers than the whole of California[13]. Among the multitude of small producers, a few massive wine companies account for a significant share of the market. In the USA, E. & J. Gallo Winery, which is often quoted as the largest wine firm (Moutot, 2017), produces about 960 millions bottles a year, more than the whole Bordeaux region. Together with two other large wine firms, namely Constellation Brands and The Wine Group, this giant accounts for about half of the wine production in the USA. By comparison, the largest French wine company, Castel Frères produces about 600 millions bottles, and the main Australian brand, Jacob's Creek, is way behind with 5-10 millions bottles[14].

For the smallest holdings, marketing their brands on their own is often inconceivable. Aside from the marketing effort, some heavy machinery such as the harvesting truck represents high fixed costs[15]. To pool the fixed costs and the marketing expenses, many winemakers take part in a cooperative winery. Working in a cooperative allows producers both to take advantage of economies of scale and to improve their bargaining power with the downstream agents[16]. In general, cooperative winemakers continue to produce their own private brands, but sales are managed collectively. Cooperative wineries are especially powerful in France, where about 600 different cooperatives ensure more than half of the production and represent around 65% of the winegrowers. Many wine cooperatives are also influential in Spain, Italy, South Africa and Argentina notably (Karlsson and Karlsson, 2014). In these cooperatives, profits are shared and the adherents are remunerated according to their respective contributions to the pool. However, the fair distribution of profits can be controversial when quality is heterogeneous among the contributors. Generally speaking, the cooperation movement faces the issue of giving the appropriate incentives for members to produce quality products.

From the consumer perspective, the key economic issue is that quality cannot be perfectly assessed before consumption, and thus before purchase[17]. The difficulty for consumers then lies in assessing the quality of the wine from the bottle's label. Even though a products' diversity is generally considered positive, as it may match the diversity of consumers' preferences, the number of different wines may be overwhelming for the non-specialized consumers. Fortunately, many labels exist to facilitate the choice of the consumers among the thousands of different producers. First, the grape variety is usually mentioned on the bottle, together with the alcoholic content. Second, the location of production is also usually mentioned, at least at the country level. These labels are mostly used in Europe, where the French concept of *terroir* prevails[18]. Several hundreds of appellations of origin exist in Europe, and European consumers

[13] Wineries are in turn much larger in California on average.

[14] Jacob's Creek is a now brand of the French company Pernod-Ricard, a world leader in wine and spirits with a production of about 300 millions bottles per year, only second to the British Diageo.

[15] Hand harvesting is more costly and only adopted by high-quality producers.

[16] See Traversac et al. (2011) for an analysis of the incentive to venture into direct sales for wine producers.

[17] Consumers buying directly from the winery can usually taste the wine before purchase.

[18] The *terroir* of a wine region is the set of characteristics influencing the tastes of the wines produced in this region: producers know-how, climatic conditions and soil composition. If the roles of climatic conditions and winemaker skill are well-documented, the influence of soil composition remains somewhat controversial. This is because the latter is often argued by top-end wine producers to justify their dominant position, and thus to discourage competitors allegedly less well-endowed. In fact, Gergaud and Ginsburgh (2008) have examined data on the Haut-Médoc region in Bordeaux and estimated that natural endowments explain little of the heterogeneity of prices and quality scores compared to technology choices. As they put it: "The French terroir legend obviously does not hold, at least in the Haut-Médoc region, which is probably one of the most famous in the world.". Nonetheless, as the vine feeds on the soil components to grow the grapes, the influence of the soil composition on the taste of the wine cannot be absolutely refuted. But the intensity of this influence has not been scientifically estimated yet.

are more or less well-aware of their preferences among the wine regions. Much like the cooperative movement, the appellations of origin can be viewed as a collective branding to pool the marketing and communication efforts. The same issue arises, as common labeling limits the individual incentive to produce above-average quality wines[19]. By contrast, especially large wine brands do not need collective branding and thus usually do not claim any geographic location of production. Indeed, appellations of origin come with many constraints on the production process, and of course limit the extension of the production. Apart from brands and appellations of origins, many other quality certifications exist for wines, including organic labels, medals won at wine competitions, quality scores given by wine critics, and official rankings.

2.4 Long-term trends and recent evolutions

This section gives a panorama of the key basic trends of the global wine market, as well as some of its current and increasingly characteristic features. Most of the figures are extracted from Anderson and Nelgen (2011), and the first points of this section are a simple summary of their work.

Fiercer competition between countries

Global wine production has been stable in the past thirty years, but the traditionally dominant countries have lost a significant market share. Figure 1, extracted from Anderson and Nelgen (2011), gives the long-term evolution of the global total production in volume across four groups of countries' aggregates. NWE8 refers to eight New World wine-exporting countries, namely Argentina, Australia, Canada, Chile, New Zealand, South Africa, USA and Uruguay. ECA refers to the wine-producing countries of Central and Eastern Europe and Central Asia, namely Bulgaria, Croatia, Georgia, Hungary, Moldova, Romania, Russia, Ukraine. The last group, EU-15 are the fifteen members of the European Union as of March 2004[20]. Figures are aggregated over four-years periods, so as to remove the year-to-year irregular components[21]. The dominant position of Southern Europe has become increasingly challenged by the producers of the New World, whose market share has jumped from 15% in the 1960s to almost 30% today. The Californian wine industry notably witnessed a sudden boom after the Judgment of Paris of 1976, a wine-tasting competition in which French wines could not defeat Californian wines[22]. By questioning the well-anchored superiority of French wines, this event strongly contributed to the popularization of Californian wines.

Smaller share in the market for alcohol

Although total wine production has remained fairly steady since 1960, humanity as a whole now consumes more than three times as much alcohol[23]. Hence, wine consumption per capita

[19]Giraud-Héraud and Soler (2003) provides a study of the implication of these appellations of origin on welfare.

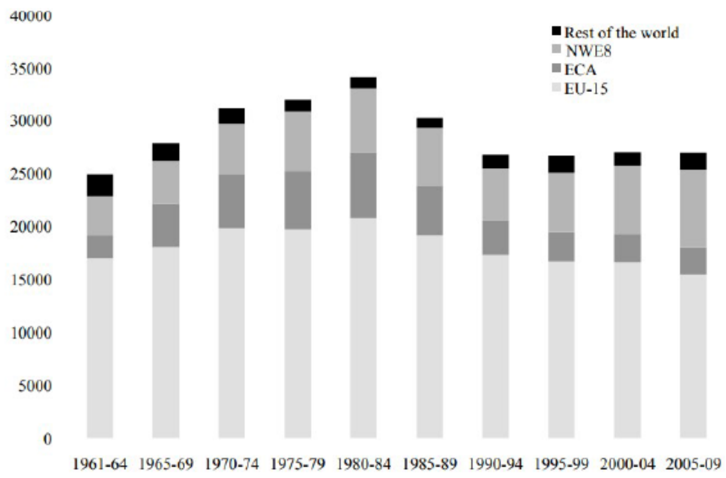
[20]Namely, those countries are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and the United Kingdom.

[21]Annual world wine production is quite irregular since about 50% is achieved in France, Italy and Spain, and thereby affected by similar weather.

[22]As an allusion to an ancient Greek myth, American and French judges were charged to evaluate the quality of a selection of top quality wines from France and California in a blind tasting. A Californian wine was unexpectedly declared the winner, although a careful analysis of the results should probably have concluded to an consensual draw between French and Californian wines (Ashenfelter and Quandt, 1999; Ashton, 2011). However the results are scrutinized, the judges did not prefer the French wines, which was a revolution in the wine world at the time. Several replications of the experiments have since led to more or less equivalent results.

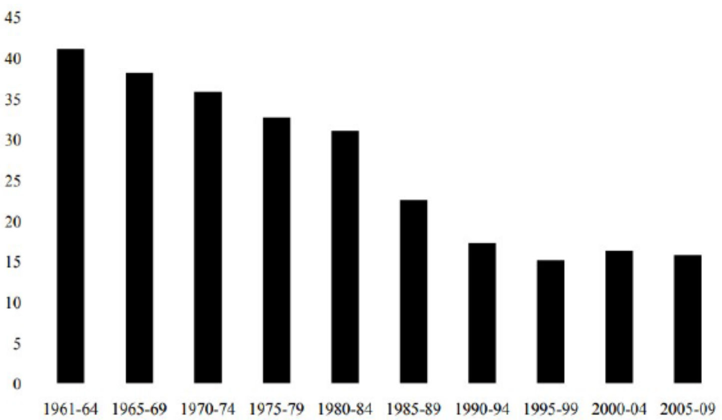
[23]Both world population and consumption per capita have increased, the former from 3 to 7.6 billions and the later from 2.1 to 2.6 liters per year (Anderson and Nelgen, 2011).

FIGURE 1 – Volume of global wine production, 1961-64 to 2005-09 (million liters)



Source: Anderson and Nelgen (2011)

FIGURE 2 – Wine's share of global recorded alcohol consumption volume, 1961-64 to 2005-09 (%)



Source: Anderson and Nelgen (2011)

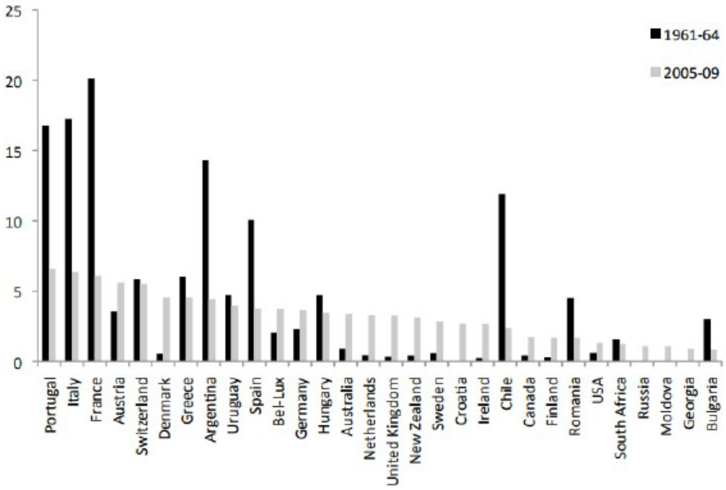
has in turn plummeted from 7.2 liters in 1961-64 to 3.4 in 2005-09. On the same period, beer consumption per capita has risen from 15 liters to 26 liters per year, and that of spirits has increased by 40% (Anderson and Magruder, 2012). As compared to other alcoholic beverages, wine has become less appealing to humanity on average. Figure 2 shows that if wine represented about 40% of world alcohol consumption in 1961-64, it only accounts for about 25% today[24]. However, this trend has now stopped, since this share has remained steady in the last thirty years.

Homogenization of consumption patterns

The apparent global decrease in the consumption of wine has mainly been driven by the process of homogenization of consumption habits. This taste convergence has been extensively documented in the case for wine and beer by Aizenman and Brooks (2008). Figure 3 shows

[24]By contrast, today wines exhibit higher alcohol content on average (Alston et al., 2015b).

FIGURE 3 – Wine consumption per adult, 1961-64 and 2005-09 (litres of alcohol per year)



Source: Anderson and Nelgen (2011)

that the main wine-producing countries used to exhibit extremely high consumption level per adult. The French used to drink about 126 liters of wine per capita in 1961 (representing about 20 liters of alcohol in figure 3) against about 43 liters of wine (about 6 liters of alcohol) today. On the other hand, Australia, New Zealand and most countries of northern Europe have considerably increased their consumption. The main long-term trend is that wine consumption per adult is less and less heterogeneous around the world. A similar homogenization is also at work on the production side. Anderson (2014) and Alston et al. (2015a) show that the national mixes of grapevine variety in Australia and in the USA is becoming closer to the global mix.

Accelerating globalization

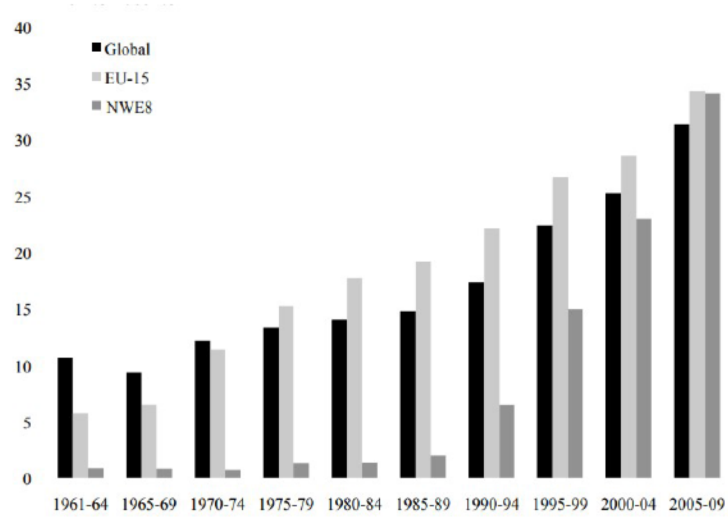
Although the desire to drink wine seems to have spread to most regions of the world, climatic constraints still geographically limits production. Consequently, the wine trade has soared in the recent years. Figure 4 gives the share of exports in wine production volumes, on average for 1961-64 and 2005-09, and among New World or Old World countries. The global share of exported volume has jumped from about 15% in the late 1980s to more than 30% today. The European countries have steadily increased their share of exports following the decrease of the domestic demand. By contrast, New World wine-producing countries have started to massively engage in international trade in the 1990s, therefore accelerating the globalization of the wine market[25]. Besides, the share of wines traded in bulk has rapidly increased. In 2010, it represented 40% of New World exports in volume against only 20% in 2001 (Rabobank, 2012).

Chinese boom

If the traditional wine markets are stagnating, others have recently shown a soaring interest in wine. The most emblematic case is the recent massive engagement of Asia in the world wine industry, especially in China (Anderson and Wittwer, 2013, 2015). Figure 5 presents the strikingly fast development of the wine industry in China, mainly driven by a skyrocketing consumption since the beginning of the 1990s. However, the rapid growth of the Chinese wine

[25]See Anderson et al. (2003) for a detailed presentation on the globalization of the wine market.

FIGURE 4 – Exports as % of wine production volume in EU-15, New World and globally, 1961-64 to 2005-09



Source: Anderson and Nelgen (2011)

market was put to an end in 2013. It is argued that the main reason is the launch of a large-scale anti-graft campaign in 2012.

Evidence of health benefits

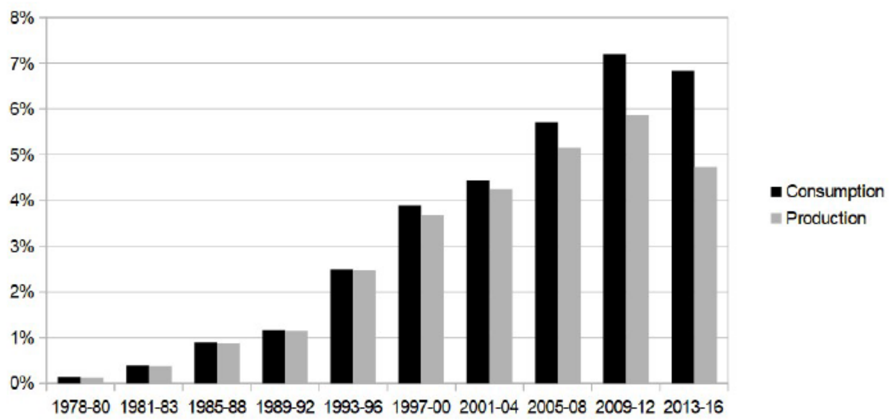
Besides the aforementioned evolutions of the wine industry, our vision of wine as a consumption good has changed in the recent years. In 1991, the medical researcher Serge Renaud presented the "French paradox" in front of a large-scale TV audience in the USA. The term referred to the observation that the French were less prone to heart diseases than what could be expected, given their high intake of fats. This emission suddenly drew considerable attention on the virtues of wine drinking habits, which appeared to be the main singular feature of the French diet. Further research has since presented evidence that moderate wine consumption can lower the incidence of various heart diseases and stimulate resistance to infection, particularly for red wines (see Bisson et al. 2002 for a short review). Although these benefits have long been popularly recognized in Europe[26], no genuinely scientific proof yet existed until the last few years. The popularization of this research has considerably altered the world vision on wine consumption. Wine is now not only sought for its recreational value, but also for its expected health benefits.

Climate change

Climate change and global warming have considerably affected the wine industry. Interestingly, Jones et al. (2005) found that this trend has so far contributed to consolidate the dominant position of some famous wine regions, such as Bordeaux and Champagne. Many European wine regions appear to be currently at or near their optimum growing season temperatures. The current hierarchy among wine regions could be disrupted if the temperature continues to increase, although to some extent grape variety can be adapted to climatic changes (see

[26]This conjecture can be traced back to Hippocrates (460 – 370 BC), often referred to as the "father of modern medicine" and attributed the following claim: "Wine is an appropriate article for mankind, both for the healthy body and for the ailing man".

FIGURE 5 – China's share in global wine production and consumption, 1978-80 to 2013-16



Source: FAO for years 1978-2011 and OIV for years 2012-2016
Note: The figures for years 2015 and 2016 are not definitive yet.

Ashenfelter and Storchmann 2016 for a review). Considering its ability to aid the production of quality wine, climate change is likely to benefit countries closer to the North and South Poles in the future. The increase of temperatures also causes wines to exhibit higher alcohol content on average. Alston et al. (2015b) estimates that alcohol contents have augmented between 0.2 and 2.0 degree-points between 1992 and 2007. Although the authors show that most of this increase is due to the producers response to evolving consumption preferences and changes in the mix of varieties, higher global temperatures lead to stronger wines *ceteris paribus*.

New highs in fine wine prices

Another recent but durable feature of the wine market is the sky-high prices fetched by the top-quality segment. Figure 6 presents the long-term rise of Bordeaux fine wine producers prices[27]. The two series represent two sets of Bordeaux Châteaux. The set LGC Sélection is composed of about 150 top-end Bordeaux wines. The Premiers Crus are the five top-ranked Châteaux of the 1855 official classification. Both series are closely linked, but the five Premiers Crus currently cost about 700 € a bottle in Bordeaux, against about 100 € for other classified Bordeaux wines[28]. The average price of Bordeaux super-premium wines has been multiplied by a factor five since 1995[29]. Interestingly, each jump of the curves of Figure 6 follows a series of consecutive good vintages (1995 and 1996, 2005 and 2006, 2009 and 2010), and is contemporaneous with a period of increasing stock prices. Indeed, those fine wines are only purchased by wealthy consumers, collectors or investors, whose wealth is largely indexed on equities values. Therefore, fine wines prices are partly driven by the level of stock markets[30].

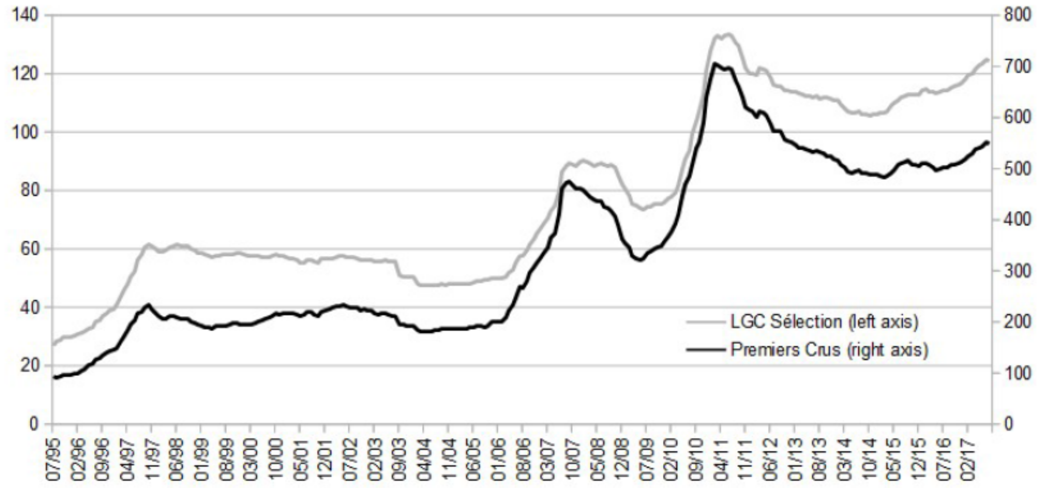
[27]The data have been provided by *Les Grands Crus*, a Bordeaux-based wine brokery office specialized in premium wines trading. The data are the monthly averages of the daily quotations estimated by the brokers for each wine. The series have been corrected for the French inflation, and prices are expressed in 2017 euro.

[28]These figures are averages among all preceding vintages, with the last vintages having a larger weight in the average.

[29]Using a long-term history of auction prices for years 1900-2012, Dimson et al. (2015) estimated a real financial return to wine investment (net of storage costs) of 4.1%, which exceeds bonds, art, and stamps.

[30]Dimson et al. (2015) also found that in the long-run, the returns to wine and stock prices are positively correlated.

FIGURE 6 – Average producer prices of premium Bordeaux Châteaux, all previous vintage mixed, July 1995 to August 2017 (€/0.75 liter bottle)



Source: Les Grands Crus

Note: Both series are monthly averages of quotations for two sets of Châteaux and all previous vintages still traded on the marketplace. The quotations are estimated by the brokers so as to reflect the current prices offered by Bordeaux-based sellers. Each quotation is weighted according to the traded volume of the Château and the vintage.

Expanding demand for experts opinion

These new highs for fine wine prices have solicited a growing demand for expert assessments of quality. Moreover, the previously mentioned Judgement of Paris of 1976 had shed serious doubts on the consensual rankings among wine producers. Experts' quality assessments were then already increasingly in demand even before the boom of the fine wine market. The 1976 tasting was followed by the timely launch of the "*The Wine Advocate*" magazine in 1978 by Robert Parker. Since the beginning of the 1980's, an increasing number of premium wines are reviewed by more and more wine-tasting experts. This thriving market has long been driven by the "wine guru" Robert Parker, but his recent retirement has left the leading role vacant. The current situation where numerous wine experts compete to provide partly subjective quality certifications raises new economic challenges for the wine industry[31].

3 The literature on the determinants of wine prices

Wine has accompanied economic research since the first writings. Chaikind (2012) has traced the mentions of the wine market in the books of most of the classical economists, back to the seminal works of Adam Smith. Today, two associations gather economists specialized in wine topics, namely the American Association of Wine Economists and the European Association of Wine Economists. The former has launched the publication of the *Journal of Wine Economics* in 2006, which has considerably accelerated the research in this field.

If wine has been merely taken as an illustration for broader economic issues, some economic mechanisms are quite specific to the wine market including the tremendous importance of

[31]This issue also applies to the market for wine medals, where the number of wine competitions has expanded in the recent years (see chapter 3).

quality and its dependence on the weather. Storchmann (2012) gives a recent and exhaustive review of the questions that have prevailed in wine economics, including those surrounding the issues of financialization, climate change and expert opinions. In what follows, I thus do not purport to provide a comprehensive panorama of the writings of economists about wine. As the focus of this dissertation is to do with wine prices, the review is focused on the past research on the determinants of wine prices. I first review the determinants of the heterogeneity of prices across producers (reputation) and time (vintage effect). Secondly, because wine prices mainly depend on quality assessments, another section is devoted to the literature on wine tasting-related issues. The last section addresses the thin literature on macro-level determinants of wine price.

3.1 Reputation, weather and experts opinions

Nelson (1970) coined the term "experience good" to describe goods for which quality can only be assessed by actual consumption. Wine arguably belongs in this category. Consumers are not usually able to get a taste before purchase, and thus rely on the informations displayed on the label to evaluate the quality inside a bottle. As a matter of fact, a first stream of research has exhibited that actual sensory attributes poorly explained the heterogeneity of wine prices, as compared to the information displayed on the labels.

Combris et al. (1997) have considered the results of an experiment in which 519 Bordeaux wines, bought at the producers' wineries, were analyzed in terms of sensory attributes in a blind tasting. The authors conducted a hedonic price regression[32], which revealed that the prices were essentially determined by the observable characteristics such as the producers' name, ranking and appellation of origin, which are constitutive of the wine's reputation[33]. Strikingly, the prevailing determinant of prices was the ranking in the 1855 classification, which has not been revised since[34]. At least two other studies have since confirmed this finding - Cardebat and Figuet (2004) also on Bordeaux wines, and Lecocq and Visser (2006a), who had included Burgundy wines to their analysis. In the latter paper, the authors also found little correlation between wine prices and the grades given by the juries. Di Vittorio and Ginsburgh (1996) also contributed to this literature by showing with a large panel of auction prices that Bordeaux wines prices still respected the hierarchy of the 1855 classification. Later, Hadj Ali and Nauges (2007) also attributed a chief role of reputation in the pricing behavior of the producers, as compared to experts grades. This series of papers drew a picture of wine prices driven by out-dated reputation, a vision supported by the results of a blind-tasting event held in San Fransisco in 1995 which reshuffled the 1855 official classification[35].

By contrast, a second seminal stream of the literature on wine prices has advanced that much of the price heterogeneity across vintages could be explained by weather, an easily observable and plausible determinant of quality. Ashenfelter et al. (1995) examined a large-scale

[32]The concept of the hedonic analysis of prices is generally attributed to Court (1939), and theoretically refined by Lancaster (1966). The theory proposes that the price of a good is a combination of the underlying prices for the attributes of the good. Given a basket of goods, a regression analysis of the prices on the goods' characteristics allows to retrieve the underlying prices of each attributes, and hence to compare their relative influences. This theory is called hedonic since the underlying prices attributes are interpreted as the respective utilities of each characteristics in the theory of revealed preferences of Samuelson (1938). Wine economists have mainly used to this framework to study the determinants of prices.

[33]See Zhao (2008), Hay (2010) and Chauvin (2013) for explanations on the sociological determinants of reputation in Bordeaux, France, and California.

[34]Since 1855, only the rank of Château Mouton Rothschild has been modified. It was upgraded from second to first rank in 1973.

[35]However, some classifications have been found to be well-grounded in objective terms. In the case of wines from the French region Mosel, Ashenfelter and Storchmann (2010) showed that the historical ranking from the 19th century could be partly explained by careful analysis of the solar radiations.

database of auction prices covering the period 1971 to 1991, and including various premium Bordeaux wines and vintages from 1950 to 1980. First, they found that Châteaux and vintage dummy variables were able to explain more than 90% of the heterogeneity of prices observed in 1990-1991. This result has highlighted the prominent vintage effect for Bordeaux wines[36]. Second, these estimated vintage effects appeared to be well-explained by simple aggregates of weather variables over the growing season such as the average temperature from April to September and the total rain in August and September. Di Vittorio and Ginsburgh (1996) give more details on the relationship between auction prices and the weather during the growing season. In particular, they describe the key role played by late frost and hailstorms. Jones and Storchmann (2001) later considerably refined the analysis of weather impact by disaggregating the weather variables across the successive phenological stages, and accounting for different grape varieties. Haeger and Storchmann (2006) and Wood and Anderson (2006) also found a strong relationship between weather and the fine wine prices, respectively for the USA and Australia.

The third and most surprising result of Ashenfelter et al. (1995) was that the prices given by the weather equation appeared to be accurate forecasts of the evolution of prices. Ashenfelter interpreted the evolution of prices as a reflection of the progressive discovery of true quality by consumers, which could have been readily assessed by examining weather information. His results suggested that the weather information was not used efficiently to set prices, which appeared to have offended part of the wine world. Notably, the increasingly influential critic Robert Parker qualified Ashenfelter's work as "a Neanderthal way of looking at wine". To some extent, the estimations of Ashenfelter questioned the relevance of having wine experts evaluating vintage quality, since weather information seemed to do the job fairly well. The relevance of expert opinion was further questioned by Ashenfelter and Jones (1998)[37], who showed how simple weather information could improve the fit of a model where prices were explained by experts ratings. Examining retail prices for wines from California and Oregon, Haeger and Storchmann (2006) also found that experts ratings add little to the overall fit after controlling for weather and other observable characteristics. Another study from Ginsburgh et al. (2013) further showed how well prices could be explained solely by producers' reputations and weather variables. Interestingly, Lecocq and Visser (2006a) found that using local weather data for subareas of the Bordeaux region instead of a single series of weather data for the whole region lead to almost identical estimates. The overall quality and global weather conditions of the vintage thus seem to predominate in the determination of prices.

Although Ashenfelter and Jones (1998) showed that weather variables were better explanatory variables of fine wine prices, experts opinions have been found to significantly influence wine prices as well. In what seems to be the first hedonic analysis of wine prices, Oczkowski (1994) estimated that quality evaluations in Australian guidebooks were a strong determinant for Australian fine wine prices. Lima (2006) found significant effects of medals obtained in wine competitions on Californian wines prices. Using a natural experiment in 2003[38], Hadj Ali et al. (2008) found that the scores given by Robert Parker had a significant impact on the prices for fine Bordeaux wines at the producer level. These results were further consolidated by Dubois and Nauges (2010) who studied the same market for vintages 1994-1998. Furthermore, expert opinion has also been proven to influence the terms of international trade. Crozet et al. (2012) and Friberg and Grönqvist (2012) show that experts evaluations of quality significantly affect

[36] In a long-term analysis of prices for a famous Bordeaux Château over the period 1800-2009, Chevet et al. (2011) showed that this vintage effect has become stronger over the years.

[37] This study was republished in 2013 in the *Journal of Wine Economics* (Ashenfelter and Jones, 2013).

[38] That year, Robert Parker gave his ratings only after the producers had set their prices. In the other years, prices were set after the ratings.

the exports of Champagne and the wine imports of Sweden. Controlling for ranking and vintage fixed effects, Ashton (2016) found that the experts' ratings have a significant impact on Bordeaux producers prices for vintages 2004-2012. In an exhaustive meta-regression analysis, Oczkowski and Doucouliagos (2015) estimate that the relationship between the price of wine and its quality rating has been found to be both moderate and significant in past research.

3.2 Reliability of wine quality assessments

Experts' opinions are notoriously imperfect at indicating quality, notably in the arts (see Ginsburgh 2003 for an extensive review). However, as shown by Hadj Ali et al. (2008) and Dubois and Nauges (2010), they can cause large shifts in wine prices, especially for fine wines. Due to the impact of experts opinions on prices, this issue has been examined with considerable attention, notably by researchers in experimental economics and cognitive sciences.

Hodgson (2008) has inserted replicate samples in a Californian wine competition for three consecutive years, and found that only 18% of wine judges were able to replicate their evaluation. When attributing a medal, only 10% of the judges managed to award the same medal to both identical samples. Weil (2001, 2005) also conducted various experiments and showed that wine tasters rarely even identify differences across vintages. When they do, they are more likely to prefer the one the least appreciated by critics. In a large-scale review, Ashton (2012) has compared the accuracy of wine judging to six other fields: medicine, clinical psychology, business, auditing, personnel management, and meteorology. He found that wine exerts exhibit on average a substantially lower level of reliability, as measured by the ability for experts to replicate their evaluation, and consensus. The renowned British wine expert Jancis Robinson acknowledges the limits of rating quality by a single score on her website: "I know it would be much more convenient for everyone if there were a single objective quality scale against which every wine in the world could be measured, but I'm afraid I just don't believe such a scale exists given the myriad styles and archetypes of wine that, thank goodness, still exist.".

The task of wine tasting involves both sensory skills to collect the information on a wine's taste and cognitive skills to report that information (Ashton, 2017). In his dissertation in oenology, Brochet (2000) has highlighted the troubling influence of expectations on perceptions, a result retrieved by Ashton (2014b). See Piqueras-Fiszman and Spence (2015) for a recent review of the literature on how expectations influence taste perceptions. In addition, Morrot et al. (2001) exhibited the perceptual illusion between odor and color. In an experiment, a white wine artificially colored red was olfactorily described as a red wine by a panel of 54 tasters. The lesson of this experiment is that our senses influence each other, which limits our ability to identify the sensory attributes of a wine. Even the tremendously influential wine expert Robert Parker himself has confessed that his ratings were partly driven by "the emotion of the instant" (see the full quote in footnote 2). The individual evaluation of a wine's quality is largely context-contingent, and as such may not be an accurate representation of the tasting experience of future consumers.

Group tasting may account for idiosyncratic errors of judgments, but there are other issues involved. In wine competitions, samples are typically submitted to a number of judges who give individual evaluations of each sample. One key issue yet unresolved is the optimal way of aggregating the scores. Indeed, the overall result depends on the aggregation method, as shown by Ashenfelter and Quandt (1999) and Ashton (2011) in the case of the Judgment of Paris. Two voting methods have been proposed by Ginsburgh and Zang (2012) and Balinski and Laraki (2013), both useful in different regards. Some non-quality related aspects also influence the outcome of competitions. Ginsburgh and Van Ours (2003) highlight that the random order in which the pianists performed at the Queen Elisabeth competition affected their ranking. In wine competitions, the random order of tasting among the different flights also affects the evaluations. In particular, the last evaluations may be subject to a bias, as reported by a judge interviewed by Ashenfelter, who remarked "even though I was spitting the wines out, by this time my mouth was dry and puckered. In retrospect I don't see how anything other than a big extract, high alcohol wine could have prevailed in such a massive competition." (Ashenfelter, 2006). Hodgson (2009) also casts doubt upon the reliability of wine competitions. He has examined the number of gold medals obtained by producers who have participated in several competitions, and has concluded that the data could not reject the hypothesis that gold medals had been given at random.

These difficulties surrounding the evaluation of wine quality must not suggest the the task is useless. Ashton (2017) notes that experts opinions are always informative, since professionals in wine tasting are trained to sens and report objective information about wine taste[39]. Thus, they lessen the asymmetry of information between producers and consumers, especially in places where information on wine quality is scarce. Masset et al. (2015) demonstrated that experts opinions have greater influence on the determination of prices among Bordeaux producers not covered by the 1855 classification. Furthermore, even coarse information is better than no information at all. In fact, Harbaugh et al. (2012) have demonstrated that coarse quality certification schemes are more informative than exact quality reporting because they increase participation.

The last key issue of wine quality rating surrounds the subjectivity of tastes. Ashton (2013) notably finds little consensus between the two prevailing wine experts Robert Parker and Jancis Robinson, from the USA and the UK respectively. As the author notes: "British critics tend to agree with Robinson while American critics tend to agree with Parker, an alignment often ascribed to critics' preferences for "elegance" vs. "power."". Therefore, one somehow needs to "evaluate" each wine critic in terms of their proximity with his or her own taste so as to qualify their ratings. Yet one of the most remarkable cases of heterogeneous preferences is probably the one mentioned by Lecocq and Visser (2006b), who write that: "when non-experts blind-taste cheap and expensive wines they typically tend to prefer the cheaper ones". Goldstein et al. (2008) have examined the results of more than 6,000 blind tastings conducted by Goldstein, and concluded that "individuals on average enjoy more expensive wines slightly less, slightly more with wine training". In a similar experiment, Ashton (2014a) found no more evidence of a relationship between price and enjoyment among novices. These studies suggest that a large part of consumers are welcome to ignore the top segment of the market, since they might actually not enjoy expensive wines.

3.3 Macro-level determinants

Quality differentiation and signaling are the key determinants of the heterogeneity of wine prices in cross sections. Yet macro-level mechanisms of supply and demand are also at work on the wine market, although less documented. To my knowledge, relatively few papers deal with the macroeconomic determinant of wine prices. However, many studies are dedicated to the estimation of the demand elasticity for wine and other alcoholic beverages, usually to assist health policy makers. Fogarty (2010) provides a review of this literature and highlights the decreasing trend of the elasticities of demand and expenditures. His survey also reveals that income elasticity of the demand for wine depends on the country, making it either a luxury product or a necessity. Haeger and Storchmann (2006) relate to this literature and estimate the reaction of the prices to the quantity supplied, using retail prices given by the magazine *Wine Spectator* for California and Oregon wines over years 1998 to 2004. They regress the log of prices on the log of the volume produced and other control variables, and find an estimate

[39] The experiments on the cognitive biases of wine tasting (Weil, 2001, 2005; Brochet, 2000; Morrot et al., 2001) typically involved novices in wine tasting.

of 0.13, which indicates a high price elasticity of demand. Nerlove (1995) also estimate a large elasticity of demand, about - 1.65, for wine imports of Sweden in 1989-1990. The authors describe their estimates holding quality constant, but the mechanisms of supply and demand are especially hard to identify on the wine market. As the wine market is highly vertically differentiated, the definition of the relevant markets in terms of quality segments is especially hazardous. Nonetheless, Wittwer et al. (2003) have proposed segmentation between premium and non-premium wines in a comprehensive model of the world wine market as in 1999. Their projections from 1999 to 2005 had forecasted the fall of prices for premium wine in Europe (see chapter 4), because the growths projections of the supply exceeded that of the demand. They had also successfully forecasted the expansion of supply of premium wines in the New World countries. Due to trade accounting for an important share of production and/or consumption in the majority of countries, exchange rates are the key fluctuating macroeconomic determinants of wine prices. Anderson and Wittwer (2013) acknowledge this in the economic fluctuation of the global wine market. Finally, a few papers have mentioned the positive income effect on wine consumption, thus classifying wine in the category of luxury goods. Wittwer et al. (2003) have considered expenditure elasticities of 1.5 for premium wine and 0.6 for non-premium, based on previous estimates of the Australian Centre for International Economics. Crozet et al. (2012) have exhibited the influence of the national revenue per capita on Champagne imports using income thresholds. Dimson et al. (2015) found a weak positive correlation of wine returns with GDP growth, suggesting a close relationship between fine wine prices and stock prices. This further suggests that consumers' income has an effect on the price of wine.

4 Four essays on wine economics

The first chapter of this dissertation is derived from work undertaken in collaboration with Jean-Marc Figuet and Jean-Marie Cardebat. The main objective has been to estimate the respective impacts of objective and subjective observable information about quality on retail wine prices, namely weather and experts opinions. A large data set has been collected from 137 wine producers from Bordeaux and vintages 2000 to 2010. The retail prices have been extracted from the website winedecider.com during the last week of May 2011. All wines are graded by four experts, allowing us to discuss the influence of the level of consensus between the experts. We conduct a two-stage estimation to assess the impact of each of these experts. Firstly, all scores are projected on weather data by Ordinary Least Squares (OLS). The residuals of this estimation are assumed to reflect the respective opinions of the experts, as opposed to the fitted values that merely account for the producers and vintage effects. The two components are then included in the right-hand side of a hedonic price equation, also estimated by OLS. Our estimation exhibits the significant influence of both components on prices, and we are able to comment on the relative influences of the experts. Furthermore, we find that including the standard deviation among expert opinions in the price equation leads to a positive and strongly significant influence of the dispersion of scores on prices. We also estimate that the most favorable opinions among experts tend to be those most correlated with prices. This latter result stems from a marketing effect, causing a wider dissemination of the information about best scores.

The contributions of this first chapter are twofold. First, we introduce a two-stages methodology to disentangle the influences of experts' ratings on prices from that of other observable variables, such as weather and producer name. Secondly, it appears that less consensus between experts is associated with higher prices. This is in contradiction with the consensual view that consumers are risk-adverse. This statement is consolidated by showing that the highest scores have the strongest influence on retail prices. Our explanation based upon a strategic withholding of information suggests that we overestimate the access of consumers to existing information on quality.

The second chapter addresses the latter issue by proposing a method to aggregate the information on experts scores. This chapter has been co-written with Jean-Marie Cardebat. We have used a large-scale data set of scores from 15 wine critics regarding 4,333 wines from 447 Bordeaux Châteaux and vintages 2000 to 2014. One problem with simple averaging is the tremendous influence of experts who are used to giving extreme scores. Another difficulty is in comparing the scores on a 20-points scale given by the European critics to those on a 100-points scale given by USA critics. Both problems are solved by equating the quantiles of the expert-specific distribution functions of the scores. This method is called equipercentile equating and is generally attributed to Braun and Holland (1982) in the field of psychometry. Once the scores of all experts are correctly scaled, it is possible to directly compare the different score. In particular, we estimate that the prominent critic Robert Parker has on average given higher scores than his peers. Furthermore, we compute simple averages to aggregate all scores and shed a new light on the hierarchy among wines, producers and vintages.

Our main contribution is to have introduced the method of percentile equating to wine economics. It adds to the emerging literature on the comparison between the respective judgments of wine experts. The method is merely a way to scale scores given by different sources prior to comparison or aggregation. An increasing number of wine competitions rely on judges' scores on a 100-points scale instead of votes. Percentile equating could then be used as a prior to aggregation.

The third chapter of this dissertation provides a novel inquiry on wine competitions, and has been co-written with Michael Visser. Our primary interest has been to estimate the causal impact of the acquisition of a medal on wine price. To do so, we examine a new data set of 16,399 transactions dealt by a major Bordeaux-based brokerage office between the years 2005 and 2016. Contrary to the two previous chapters, the third chapter addresses the mass-market, illustrated by the fact that average price in our data is only 2.24€ per 0.75 liter. We have collected the exhaustive records of the eleven main wine competitions for Bordeaux wines, which has allowed precise identification of all the medals awarded to the wines present in our transaction data in terms of type (bronze, silver or gold), name of competition and date of the award. The matching between the two data sets has revealed that some wines obtained a medal only after the transaction, a fact for which we provide various rationales. We build on this unusual feature to design a consistent estimator of the price markup due to the acquisition of a medal, fixing quality. Assuming a constant markup for all medals, we estimate that producers can expect a 13% bonus from winning a medal at a wine competition. This estimate is mainly driven by the large impact of gold medals, but bronze and silver medals also exhibit significant impacts. However, the effects of quality heterogeneity are statistically indistinguishable across the three types of medals. This result suggests that the value of gold medals is somewhat overestimated. Across wine competitions, the estimated causal effect is only statistically significant for a few long-established competitions. Finally, we have collected information on the costs of participating in such competitions. This information allows us to estimate the distribution of expected profit depending on the volume of production and the expected probability of being awarded. A majority of producers is estimated to have an incentive in participating in those competitions.

This chapter provides the first impact study of European wine competitions. We conclude that there is a strong average impact on the revenues of awarded producers and find a wide incentive to participate. Secondly, we estimate a statistically significant relationship between medals and quality on average across all competitions. This result contrasts with those of previous studies of wine competitions in the USA (Hodgson, 2008, 2009). However, our estimations show that this relationship is not statistically significant for the majority of the competitions.

Even though notoriously imperfect, we thus defend the view that wine competitions are useful in a context of great uncertainty from the consumers standpoint, and are capable of identifying quality wines.

The fourth and last chapter of this dissertation somewhat departs from the three others. Combining a collection of data sets, I design, estimate and evaluate various forecasting models for the average prices of 15 main appellations of origin in Bordeaux. This chapter responds to a command of the Bordeaux wine professionals for more visibility on upcoming market conditions. The price data has been provided by the joint-trade organization of Bordeaux wine professionals. It covers the period 1982-2017 at the annual frequency, and the period 2001-2017 at the monthly frequency. The models include a large collection of exogenous determinants of prices specific to each appellation including initial stocks, harvests, weather, quality ratings, trade flows, exchange rates, GDP and interest rates. The forecasts result from a combination of various models in the time-series analysis toolbox, such as the autoregressive distributed lag models (ADL), error-correction models (ECM) and unobserved component models (UCM). One key task has been to aggregate the large data on the macro-level economic conditions into a limited selection of leading indicators. In particular, the weather data is aggregated so as to reflect the aggregate impact of the weather on expected harvest by means of an auxiliary harvest model. The fit of the harvest model satisfactorily reflects the records of official forecasts, indicating that the model correctly aggregates the weather information. The key result is that the forecasting models outperform the naive no-change expectation of prices over the last five years on average. Forecasts are especially useful for the largest regional appellation and during episodes of important supply shocks. In addition to the forecasts, the estimated models also revealed the prevailing role of stocks dynamics on the fluctuation of average prices. My estimations further establish the key role of exchange rates, a small but significant influence of weather, and a limited vintage effect among the bulk of the Bordeaux wine production.

This last chapter reconnects with the literature on agricultural price forecasting of the 1980s. This branch has been progressively deserted since agricultural economists have considered that futures prices provide satisfactory forecasts. However, these futures prices cannot purport to represent the dispersion of agricultural prices across different locations, and many agricultural markets are still not equipped with a futures market anyway, especially in the case of wine. Indeed, futures markets are of limited utility in the case of highly vertically differentiated products like wine. This chapter restores and promotes the usefulness of price forecasts on such markets. My main contribution is that well-designed forecasts can outperform the naive no-change forecasts on a regular basis, which has been found to be especially difficult to achieve for agricultural prices in the past literature. The annual price forecasting models consolidate the well-established literature on the efficiency of forecasts combination. The adopted methodology for the monthly price forecasts is innovative in several ways regarding the existing literature on price forecasting, especially in that it combines the UCM and the ECM frameworks. Complementary to Cardebat and Bazen (2016), who follow the exact same objective but focus on univariate models, my estimations highlight the importance of including exogenous predictors in price forecasting models. My estimates of the strong influence of starting inventories on wine prices add to the literature on the impact of stock dynamics on agricultural prices, which has attracted considerable interest since the food crisis of 2007-2008. Along the way, I have also estimated a model for total wine harvest per appellation of origin, which to my knowledge is novel in the literature. The truncated regression framework duly accounts for the regulated maximum yield, which sets the quality standard. The nonlinear design for the influence of temperatures from April to June on yields suggests optimal temperatures for each appellations. Estimations indicate that current average temperatures are about optimal in terms of yields, as found in the past literature with respect to quality. Hence, climate change could cause a reduction of the yields in the Bordeaux area *ceteris paribus*, although producers are believed to have options to increase their yields up to the regulated maximum.

Chapter 1

Expert Opinion and Bordeaux Wine Prices: An Attempt to Correct Biases in Subjective Judgments

This chapter has been co-written with Jean-Marie Cardebat and Jean-Marc Figuet. An earlier version has been published in 2014 in the *Journal of Wine Economics* as Cardebat et al. (2014). I here include the feedback we received after the publication, notably from Eddie Oczkowski.

1.1 Introduction

Whenever consumers have access to perfect information, the Bertrand model indicates that the equilibrium price of goods and services equals its marginal cost. In practice, however, the law of one price is the exception rather than the rule. In fact, most markets are characterized by substantial price dispersion. This is particularly true for experience goods, for which quality is known only after purchase and consumption. However, it is costly to acquire information for consumers. The pioneering research by Akerlof (1970) and Nelson (1970) established that information asymmetries that pertain to the quality of a product might influence markets in a detrimental way.

Brands (Montgomery and Wernerfelt, 1992), advertising (Ackerberg, 2003), quality labeling (Jin and Leslie, 2003), and expert endorsement (Salop, 1976) all constitute transmission channels that provide consumers with information about a product's quality. Although there are experts in a variety of domains—art, economics, weather forecasting, sport, gastronomy, cars, and electronics—it is extremely difficult to assess their influence and the quality of their opinions on products. Reinstein and Snyder (2005) concluded that movie reviews do not affect a film's box office earnings. Sorensen and Rasmussen (2004) demonstrated that book reviews, whether favorable or unfavorable, boosted sales, thereby confirming the old adage "there is no such thing as bad publicity." Indeed, empirical studies face a major methodological problem: high-quality products obtain high scores because they are, in fact, of high quality. Thus, it becomes difficult to determine the extent to which expert endorsements stimulate demand. Recent papers by Hodgson (2008, 2009) question the consistency of expert wine judges in a wine competition setting and demonstrate that wine experts make mistakes. Ashton (2011) and Lecocq and Visser (2006a) point out, however, that judgment errors can be reduced by pooling the opinions of several experts.

Bordeaux wine represents an experience good that requires a great deal of expertise in order to determine each wine's final quality and, hence, its price. This paper seeks to establish whether the experts, including the most renowned, Robert Parker, provide pertinent information for consumers and whether their scores influence wine prices.

For Ashenfelter (1989), the fallibility of Parker's judgment allows buyers to profit from his errors of judgment when wines are sold at auction. According to Ashenfelter (2008), a wine's age, the average temperature from April to September, the rainfall in August and September and then from October to March, and the vintage are the main factors behind price variations. Ashenfelter and Jones (2013) show that expert scores on Bordeaux wine do not provide useful information in poor years and correlate only with market prices, at best, in good years. Experts tend to disregard key data such as weather conditions during the growing season, which crucially determine the wine's quality. As there is detailed information about local weather conditions, in particular, available privately to each individual château (Di Vittorio and Ginsburgh, 1996), the experts merely transmit publicly available information to the consumer. Ginsburgh et al. (2013) apply the hedonic pricing model to a sample of 102 Médoc wines in order to show that expert ratings do not provide a better explanation for price than climate conditions, the 1855 classification, terroir, or production technique. Sixty-six percent of price variations could be explained by weather conditions or differences in vineyard practices. This percentage rose to 85% when the 1855 classification was taken into account. Di Vittorio and Ginsburgh (1996) come to the same conclusion. A hedonic function, calculated on the basis of the auction prices of 58 Médoc crus classés, indicate that the 1855 classification plays a greater role in explaining a wine's price than any alternative rating system drawn up by experts.

For Jones and Storchmann (2001), Parker's scores influence prices in a differentiated fashion. Their analysis suggests that a rise of one point may cause price increases between 4% and 10%, with an average increase of 7%. This result, obtained from prices for 21 prestigious Bordeaux wines, indicates that the sensitivity of a wine's price relative to Parker's scores is greater for wines made from Cabernet- Sauvignon than for those made from Merlot. Hilger et al. (2011), adopting a more experimental approach, also showed the impact of expert ratings. They analyzed wine sales in a supermarket by choosing a random sample of 150 wines from 476 rated wines and displaying each wine's score on supermarket shelves. Sales of the selected wines increased by an average of 25%, and sales of those with the best scores increased more quickly than those with lower scores. This led to the conclusion that the advertising surrounding expert endorsement produces a positive effect on global demand as it reduces information asymmetry. Storchmann et al. (2012) argued that expert opinion has a negative effect on the price dispersion of American wines evaluated by the Wine Spectator between 1984 and 2008. The authors show that expert opinion distorts the relationship between quality and price, especially in the case of poor-quality wines. Roma et al. (2013) constructed a hedonic price model to determine the variables influencing the prices in a sample of Sicilian wines. They concluded that price depends on traditional objective variables and sensorial variables as well as on the ratings published in specialized reviews. Using five years of data on expert opinions published in six Swedish periodicals, Friberg and Grönqvist (2012) showed how a positive review induced an increase in demand of 6% the week after publication. This positive effect then declined but was still significant 20 weeks later. A neutral expert opinion led to a small increase in demand, whereas a negative opinion had no effect.

The debate about the impact of expert opinion on price is even more complicated for Bordeaux wines. Bordeaux crus classés can be sold en primeur in the futures market six months after harvesting and are delivered to the purchaser only two or three years later. This creates a great deal of uncertainty concerning the wine's ultimate quality. It is the expert's role to ascertain this ultimate quality, which, consequently, influences the sale prices of primeur wines. Hadj Ali and Nauges (2007), using a sample of 108 châteaux for vintages from 1994 to 1998, showed that the price en primeur is determined chiefly by reputation. Parker's scorings have a significant but marginal effect—an increase of one point triggered a rise in price of 1.01%.

Simply put, the role of experts in influencing the price of any wine remains uncertain and differs from one study to another. Wine is not a homogeneous product but varies according to a set of characteristics. Some of these characteristics—color or grading, for example—are easy to measure inexpensively. Others, such as sensorial or taste characteristics, are difficult to measure before consumption. Expert opinions are purported to summarize quality characteristics of wine. These opinions may convey less information than a complete description of characteristics. In addition, these opinions might be imperfect because they fail to capture, for example, the quality experienced by consumers. This imperfection can have implications for prices and consumer welfare.

The present research aims to investigate the question of the impact of expert opinion on fixing the retail price of wine. It is based on exhaustive data concerning the scores attributed to different wines by a broad panel of nine[1] wine experts from three different countries over 11 years (2000 to 2010). Our main objective is to reduce the systematic econometric bias that is bound up with expert opinion and to test the impact on prices of a consensus or divergence among experts. Evidence of this bias has been revealed by Lecocq and Visser (2006a) as well as Oczkowski (2001).

As a first attempt to correct the measurement error bias, we aggregate a solid body of information from those nine experts to reduce the risk of error from any one expert. Because we use their average score, such a risk was reduced, thereby minimizing individual bias. Most other research uses data from a single expert, so this methodological approach allows us to reduce such errors of judgment (Ashton, 2011). Moreover, examining the opinions of several experts allows us to underline the specific impact of each with regard to prices. Additionally, because the key role played by Robert Parker is often highlighted, we can compare the impact of his opinion with that of other experts. According to Lecocq and Visser (2006a), use of the average score might not prevent the measurement error. Assuming that the measurement errors are independent and zero mean, the average tends to approach zero when the number of expert approaches infinity. The key issue is: how many experts are required to correctly estimate the impact of scores on price? If the number of observed scores for each wine is insufficient, the average score may still be biased. Our strategy is to decompose the scores in a first stage regression using weather data as explanatory variables. This method allows us to extract both the objective component of the scores and the measurement errors, then, to estimate their respective impacts on price in an augmented regression.

Further, the gathered data permit us to test the influence of the dispersion of the score on prices. One argument for this influence is that consumers might be wary of the true quality of a wine when its scores vary among the experts. This uncertainty is perceived negatively by risk-averse consumers, which then decreases their demand, thus decreasing the equilibrium price. We test that hypothesis by using the standard deviation of scores for each wine as a determinant of the price.

We first examine the methodology adopted and the data used before presenting the econometric results obtained and then offer our conclusion.

1.2 Model

1.2.1 The Naive Model

The hedonic model first introduced by Court (1939) and later refined by Lancaster (1966) is the traditional framework used to determine the price of agricultural produce (Costanigro and McCluskey, 2011). A hedonic function is the relation between differentiated prices for a given

[1]We first considered a large panel of 19 experts, but only 9 of them had scored a sufficient number of wines (over 300). We thus decided to focus on these 9 experts.

good and the quantity of constituent characteristics possessed by that good (Triplett, 2004)[2]. In the case of wine, prices are determined by factors including appellations, vintage, climatic conditions, expert opinions, and reputation (Benfratello et al. 2009; Cardebat and Figuet 2004, 2009; Combris et al. 1997; Landon et al. 1998; Oczkowski 1994, 2001, etc.; for a survey, see Costanigro and McCluskey 2011). Our design aims to give structure to the relation between price and its determinants. The focus here is on the relation between quality, experts' grades, and prices.

We assume here that wine prices are determined by intrinsic quality, age, and the reputation of the producers. Several variables are available to control for these factors: the names of the producers, the vintage, the experts' scores, and the weather conditions of growth. We assume that the impact of reputation is captured by a producer-specific fixed effect, instead of lagged scores, as in Oczkowski (2001). As the data are a cross section of those for 2011, the reputation of the producers is the same for all vintages: it is the reputation of the producer in 2011. The issues related to the use of a fixed-effects model have been addressed by Dubois and Nauges (2010). They explain why those fixed effects cannot be used for the purpose of controlling for quality. Therefore, we interpret these fixed effects in terms of reputation. Age is easily calculated by the vintage. The real issue is the quantitative estimation of quality. Our best indicators of quality are the scores, but they need to be corrected (see Dubois and Nauges 2010; Lecocq and Visser 2006a; Oczkowski 2001). To deal with this issue, we use the following measurement error model:

$$score_{ite} = q_{it} + o_{ite} \tag{1.1}$$

where $score_{ite}$ is the score of producer i for the vintage t with the expert e, q_{it} is the objective quality of this wine, and o_{ite} is the personal opinion of the expert e on this wine. We assume that the quantity q_{ite} is the objective component of the score, and that o_{ite} is its subjective component. Since the experts aim to evaluate the intrinsic quality of wine, their opinions o_{ite} are seen as measurement errors.

We still need to evaluate the qualities q_{ite}. A first naive method is to assume that o_{ite} are independent and identically distributed (i.i.d.), with zero mean. Under this hypothesis, we can apply the law of large numbers (LLG):

$$\underset{n\to+\infty}{plim}\,\frac{1}{n}\sum_{e=1}^{n} o_{ite} = 0$$

In this design, the average score among the nine experts for each wine is thus a consistent estimator of the objective score. We estimate the following price model:

$$ln(p_{it}) = \gamma\overline{score}_{it} + \delta t + \mu_i + \epsilon_{it} \tag{1.2}$$

where p_{it} is the price of the wine of age t of producer i, $\overline{score}_{it}$ is its average score among the nine experts, μ_i is the fixed effect of producer i[3] and ϵ_{it} are idiosyncratic shocks. These fixed effects aim to capture the effect of reputation on prices given the score and the age. The coefficient δ measures the storage cost, the quality improvements due to the keeping, and the scarcity value all at once. Remember that the data are a cross section, which means that the vintage, t of wine i, determines solely the age.

[2]This method has been used for cars (Arguea and Hsiao, 1993; Court, 1939; Griliches, 1961), real estate (Taylor, 2003), computers (Triplett, 1989), the environment (Freeman, 1993), corn (Espinosa and Goodwin, 1991), cereals (Stanley and Tschirhart, 1991), apples (Carew, 2000), and even for the French vaulting stallion semen market (Vaillant et al., 2010).

[3]Note that these producer-specific fixed effects forbid the use of ranks or appellations as additional control variables, because of perfect multicollinearity issues among the dummy variables. So using fixed effects at the level of the producer should be more efficient.

We estimate equation (2) with the ordinary least squares (OLS). We use the Newey-West variance estimator, since the residuals faced both heteroscedasticity (the variance of the errors differs among producers) and autocorrelation (there is some inertia across vintages). The γ obtained with the average score is compared to those obtained when we replace the average score with the score of a few major experts. We then use a subsample of wines that have been graded by at least the four main experts. One of these experts is Robert Parker (The Wine Advocate, WA), who enjoys a reputation as a wine guru with great influence on prices (Hadj Ali et al., 2008; Jones and Storchmann, 2001). Others include the Wine Spectator (WS), Jancis Robinson (JR), and Stephen Tanzer (International Wine Cellar, IWC). This cut leaves us with 737 prices of wines from 137 châteaux of Bordeaux, with vintages from 2000 to 2010. The use of this subsample allows us to conduct a multi-expert regression that controls for the correlations between the experts' grades. In the expert-specific regressions, the impacts of the different experts are not taken into account simultaneously, although the real impacts are indeed linked to one another. We therefore obtain more accurate estimates of experts' respective influences. This also provides an idea of the error made when only one expert is considered.

1.2.2 The Two-Stage Model

This naive model has some major limitations. The first one is the application of the LLG with at most nine experts for each wine. As Lecocq and Visser (2006a) pointed out, it seems hardly acceptable that the opinions of the nine experts correct each other perfectly. Worse, the opinions o_{ite} must be i.i.d. in order to validate the LLG. That is problematic since the grading behavior of experts depends on both their taste and their grading scale. As we shall see, some experts grade systematically using the average score and others grade systematically above the average. Given that information, this first model should be abandoned.

Another key limit of the naive model is that it assumes that the individual opinions of the experts have no influence on price. The underlying hypothesis is that the price is determined solely by the objective component of the scores, while the subjective component is irrelevant to the price equation. This point is mostly unacceptable, since the only observable score comprises both the objective and the subjective components. Moreover, some consumers might be interested in the differentiated opinions of the experts, acknowledging that each expert has his own tastes. This point has been highlighted by Lecocq and Visser (2006a). When a consumer feels well represented by one expert, he is likely to be deeply influenced by the subjective opinion of this expert and might not look at the comments of the others.

These two arguments stress the necessity of integrating the objective and subjective components in the price equation. Of course, this is achieved by using the raw scores, but this specification implies that the two components have the same coefficient. Testing this hypothesis is another goal of the present article. Hence, we need to disentangle q_{ite} from o_{ite} in equation (1.1). To this end, we use the weather data as determinants of quality and identifying variables.

It seems reasonable to assume that quality is determined solely by the soil quality, the skills of the producer (including the viticulture ability, the reaper's precision during harvest, the maturation process, and the varietal blend), and the weather conditions during growth. Making the assumption that the two first factors can be captured by producer-specific fixed effects and a trend, we design the following model of objective quality:

$$q_{it} = \beta w_{it} + \rho t + \nu_i \tag{1.3}$$

where w_{it} is the vector of the weather variables, and ν_i is the fixed effect of the producer i on quality. In this design, the fixed effects are indicators of soil quality and producer skills. It should be noted that vintage fixed effects could also be considered as an alternative to weather variables. However, following the literature[4], we contend that weather data contain additional information useful to estimate the vintage effect. Our estimations presented hereafter comfort our assumption.

The trend aims to take into account the global improvements in technology. In order to limit the number of coefficients, we assume that the impact of the weather variables on the objective scores is the same for all producers.

We obtain the reduced form of scores by combining equations (1.1) and (1.3):

$$score_{ite} = \beta w_{it} + \rho t + \mu_i + o_{ite} \tag{1.4}$$

which can be estimated using the OLS, thereby minimizing variation among the opinions. Again, we use the Newey-West variance estimator, as the opinions o_{ite} showed heteroscedasticity and autocorrelation.

This first-stage regression gives us an estimate o_{ite} of the opinions of the experts. Let o_{ite} be the vector that includes the o_{ite}. Adding this variable to the model (1.2) and replacing the average score with any expert score allows us to estimate the differentiated impacts of quality and experts' opinions on prices. This is stated formally as follows:

$$ln(p_{it}) = \gamma score_{ite_1} + \theta\hat{o}_{ite} + \lambda t + \mu_i \tag{1.5}$$

where e_1 is the chosen reference expert, and μ_i still aims to estimate the influence of reputation of producer i on price given the scores. Splitting the variable $score_{ite_1}$ into its objective component q_{it} and its subjective component o_{ite} , we get the detailed effects of scores on prices:

$$ln(p_{it}) = \gamma\hat{q}_{it} + (\gamma + \theta_1)\hat{o}_{ite_1} + \theta_2\hat{o}_{ite_2} + ... + \theta_n\hat{o}_{ite_n} + \lambda t + \mu_i$$

where n is the number of experts. That is why the choice of expert does not matter. The coefficients are estimated using the bootstrap method to correct the sampling error in the ordinary least squares (OLS) variance estimates in the second stage (due to the replacement of (q, o) by $(\hat{q}, \hat{o})$). That is, we randomly draw with replacement a same-size subsample from the initial one, we conduct the two-stage estimation with this subsample, and we obtain the second-stage estimates that are calculated using OLS. We do this procedure 1,000 times and find that the convergent bootstrap estimate is the mean value of each coefficient. The variances in the bootstrap estimates are calculated nonparametrically using the empirical variance of the 1,000 estimates for each coefficient.

At this point, we should draw the link between our approach and instrumental variable techniques. Here, we use the weather data, the age of the wine, and producers' dummy variables as instruments for the raw scores. This IV procedure was used and discussed in Haeger and Storchmann (2006). Our method is slightly different, however, since we use the residuals of this first-stage equation as a control variable for the second stage. Some tests are required to check our model:

- an endogeneity test, which is easily achieved by performing a t-test on the coefficient θ_e,
- an overidentification test, which is achieved by performing a Sargan test.

We provide these tests by bootstrapping the test statistics.

This procedure was first conducted on the entire sample by using only the mean score for each wine: this includes all wines graded by at least by one expert. Then we conduct the same analysis separately for each expert. As we did for the naive model, at the end we use

[4]See Ashenfelter et al. (1995); Ashenfelter and Jones (1998); Lecocq and Visser (2006a); Ashenfelter (2008); Ashenfelter and Jones (2013).

the subsample of the 737 wines for further inter-expert analysis and as a robustness check. We compare these results to the naive model estimates.

1.2.3 Consumer Defiance Versus Marketing Effect

Our design allows us to provide precise estimates of the individual impacts of each expert, apart from the impact of the objective component of the scores. Yet it misses one indirect effect of the scores on price. As is often argued in the literature on marketing and consumer behavior (see, e.g., Martin-Consuegra et al. 2007), the standard deviation of grades is likely to negatively affect the buyer's trust in the scores. Assuming the consumer's risk aversion, a high dispersion of the scores decreases the equilibrium price because it lowers demand.

However, another indirect effect of the standard deviation on prices might occur. As shown by Hilger et al. (2011), when a retailer displays a score for a wine, its sales (or price) increase: the higher the exhibited score, the higher the increase in sales. A high standard deviation in scores for a wine implies that at least one expert liked the wine more than the others and gave it an above-average mark. Retailers know all the scores and can choose to talk about only the best. We call this positive correlation between the standard deviation of the scores and wine prices the "marketing effect." The higher the standard deviation, the greater the likelihood that the retailer will display a good score (compared to the average) and the higher the price. Unexpectedly, the lack of consensus among experts allows retailers to improve their marketing and to increase their prices.

Our model allows us to test this hypothesis. Because the standard deviation of the scores is the standard deviation of the opinions, we add the latter to the regressors in the second stage of the multi-expert regression. To this end, we use the empirical standard deviations of estimated opinions for each wine as an estimate of the real deviation of the scores. The estimate of the coefficient and its related significance are obtained by bootstrap. We conduct this analysis on the subsample of 737 wines graded by all the experts, in order to maintain a constant number of grades per observation.

1.3 Data

Annual data were obtained for 203 wine producers, located mainly in the Bordeaux area (187 producers from 12 Appellation d'Origine Contrôlée (AOC) areas, with nine producers from Napa Valley, California, and seven from Spain) covering the period from 2000 to 2010. The prices were obtained from the website winedecider. com. This website offers prices on a wide range of wines from several countries and AOCs and is representative of the main wine sellers on the Internet, including Millesima. Here the listed price is the average retail price of a bottle packaged in a case of six or 12 bottles in 2011 prices before value-added tax (VAT) and transportation costs. Using the retail price means we can assume that these wines are priced after the experts have published their scores. This point is crucial to the relationship between wine prices and expert opinion. A retailer's pricing behavior will vary depending on whether he is aware of the expert ratings.

Table 1.1 provides descriptive statistics about the prices and the scores among the different appellations.

As in the hedonic approach, we include:

- Objective characteristics: name of the producer and vintage,
- Taste rating or subjective quality: scores from nine experts (The list of the experts is given in the Appendix A. Each wine is graded by 4.5 experts on average.),

TABLE 1.1 – 2011 Prices and Scores Across Appellations

		Prices (€/0.75 l. bottle)				Critical scores			
AOC	Number of Châteaux	Min.	Max.	Avg.	Std. Dev.	Min.	Max.	Avg.	Std. Dev.
Médoc	2	8	23	14.68	4.18	65	91	87.62	4.65
Saint-Estèphe	21	8	292	33.40	36.99	67.5	100	87.25	4.83
Pauillac	24	16	1,604	74.71	239.97	65	100	89.05	5.03
Saint-Julien	16	10	247	55.89	43.74	72.5	100	89.09	4.36
Listrac	4	8	44	14.23	7.02	70	92	83.39	4.53
Moulis	3	14	34	21.61	4.50	68.5	92	85.29	5.18
Margaux	30	11	829	58.07	100.46	65	100	87.68	4.79
Haut-Médoc	11	9	37	16.53	5.53	65	93	84.30	5.28
Pessac-Léognan	24	10	756	82.67	149.25	65	100	87.95	5.08
Sauternes	17	11	566	46.40	79.80	65	100	88.41	5.11
Pomerol	18	14	3,359	176.39	485.58	70	100	86.31	5.47
Saint-Emilion	17	11	1,501	124.62	253.65	60	100	88.04	5.51
Ribera del Duero	4	42	155	88.08	33.67	75	99	91.08	4.11
Rioja	3	15	88	37.24	19.46	80	97	90.63	4.24
Napa Valley	9	17	543	142.33	124.07	69	100	91.02	4.86
Total	203	8	3,359	83.12	203.57	60	100	87.94	5.2

Source: Winedecider.com (as of the last week of May 2011).
Note: The data counts a total of 2,087 different wines, with an average of 11 wines per producers - a few vintages are missing for certain producers. Scores are based on the average note given by the website for each château-vintage combination. For instance, the first line (Médoc) contains two wines, thus the statistics are given for 2 wines and 11 vintages (2000 to 2010), i.e., 22 observations. The prices correspond to the price of each château-vintage combination.

- Weather as a determinant of objective quality: temperature and rainfall data from several meteorological stations in the heart of the AOC, due to the great heterogeneity of local weather conditions across the vast wine-producing area of Bordeaux (discussed below).

Table 1.2 examines the evolutions of the descriptive statistics in Table 1.1 on the entire sample of wines, across the several different vintages.

As for weather data, we obtain details of daily weather conditions for the three main areas of the Bordeaux region and for the Napa Valley. We define the three main climate areas of the Bordeaux appellations: Médoc, Saint-Emilion/Pomerol, and Graves. Meteorological studies related to wine reveal significant weather variability within the Bordeaux appellation (Bois, 2007; Bois and Van Leeuwen, 2008). Table 1.3 shows the average temperatures and rainfalls across the four available stations in the areas.

Consistent with Bois and Van Leeuwen (2008), this information is crucial to our study. It is essential to correlate not only data from the main meteorological station based in Mérignac but also the meteorological data from each of these three areas in the Bordeaux region. Although Lecocq and Visser (2006a) show that the Mérignac station provided a reasonably acceptable proxy of the weather for the Bordeaux appellation as a whole between 1993 and 2002, they note the appearance of some differences: "The climate conditions prevailing in the main weather station [Mérignac] are thus clearly not representative of the Bordeaux wine region as a whole" (Lecocq and Visser, 2006a). Our model aims to use meteorological data as an instrument for the scores. Consequently, if we want to maintain some heterogeneity in our fitted scores, we cannot use only one station.

We have gathered the monthly temperatures and levels of rainfall from three stations representative of the three main wine regions of Bordeaux. For the Médoc region, we use weather data from Château Latour (which is very close to Pauillac). In the Graves region, we use

TABLE 1.2 – Prices and Scores Across Vintages

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
						Critical Score					
Min.	79	81	79	80	83	85	81	82	84	83	85
Max.	97	100	98	97	97	98	96	98	97	98	99
Avg.	89.4	88.6	87.8	88.6	88.6	90.2	88.9	88.4	89.3	90.9	90.8
S.D.	4.12	4.67	4.36	5.24	4.30	4.29	4.55	4.19	3.36	4.49	4.63
						Price (€/0.75l bottle)					
Min.	10	10	9	10	9	9	9	8	9	10	9
Max.	3,359	1,415	1,332	1,439	1,274	2,680	1,242	1,164	2,008	2,741	2,448
Avg.	1113.9	71.4	64.5	76.7	63.3	101.6	69.6	64.4	72.1	112.4	105.4
S.D.	311.5	135.7	132.2	165.2	121.6	253.7	136.3	127.3	178.8	280.7	265.5

Source: Winedecider.com
Note: Min and max are calculated on the average score given by the winedecider.com website for each château-vintage combination, rounded to the nearest integer. S.D. = standard deviation.

TABLE 1.3 – Descriptive Statistics of Weather Variables (2000-2010)

	Médoc[a]	Saint-Emilion/Pomerol[b]	Graves[c]	Napa Valley[d]
	Avg Monthly Mean Temperatures in Centigrade (resp. min., max.)[e]			
April	13.9 (11.7,16.2)	12.8 (11.7,16.2)	12.4(10.9,15.2)	13.0(11.1,15.0)
May	17.2 (11.5,19.0)	16.2 (14.7,17.1)	15.8 (14.5,16.9)	16.5 (14.3,19.0)
June	21.5 (19.7,23.9)	20.1 (18.8,22.6)	19.9 (18.5,22.4)	18.8 (17.8,19.9)
July	22.2 (20.1,26.1)	21.0 (19.3,24.3)	20.9 (19.2,23.7)	19.4 (17.9,21.5)
August	22.4 (21.2,26.5)	20.7 (19.1,24.8)	20.5 (18.5,24.6)	19.0 (17.4,20.0)
September	19.4 (17.0,23.0)	17.3 (15.8,20.0)	17.3 (15.7,19.5)	18.1 (16.3,19.4)
	Cumulative Rainfall (average) (min,max)			
December to March	311.9 (195,424)	202.0 (89,273)	233.0 (149.5,309.5)	665.4 (404.1,1,069.4)
April to June	213.5 (142,262)	231.2 (150,354.5)	228.2 (95,346.5)	100.3 (4.3,243.6)
August to September	95.4 (44.5,169.5)	113.9 (52.6,174)	101.6 (51,152)	2.1 (0,14.8)

Source: [a]: weather station of Château Latour; [b]: weather station of Château Grand Barrail; [c]: weather stations of Château Haut-Bergey; [d]: Monthly Report of California Irrigation Management Information System, Oakville weather station (CIMIS #77, Oakville); [e]: averages of monthly means over 11 years.

TABLE 1.4 – Naive Model γ Estimates of Equation (1.2)

	Average score and expert-specific regressions			Multi-expert regression	
	N. obs.	$\hat{\gamma}$	R^2	$\hat{\gamma}$	VIF
Average Score	2,172	0.048***	0.945	0.013**	9.810
IWC	1,169	0.099***	0.961	0.042***	5.361
JR	1,723	0.016***	0.944	0.004*	2.838
WA	1,644	0.063***	0.9559	0.039***	5.573
WS	1,758	0.047***	0.951	0.022***	4.466

Note: The columns $\hat{\gamma}$ contain the estimated influence of the average score and the experts on price for the two different specifications.
Significance levels: ***1%, **5%, *10%.

weather data from Château Haut-Bergey (in Léognan). For Saint-Emilion/Pomerol, we use data from Château Grand Barrail. All these weather stations are located directly in the vineyards. Our data for the Napa Valley refer to the Oakville meteorological station, which is located only in close proximity to vineyards. The exact location within certain meso-climates may explain the surprising fact that our Napa temperatures are below those in Bordeaux. We do not have weather conditions for the two Spanish appellations.

1.4 Empirical Results

1.4.1 Results from the Naive Model

Table 1.4 gathers the γ estimates for the naive model[5]. We have estimated equation (1.2) first using the average of all available scores and then using scores from each of these experts.

For these expert-specific regressions,we provide the number of observations and the coefficient of determination (R^2). Table 1.4 also lists the estimates of the multi-expert regression, including the average score. The last column shows the Variance Inflation Factors (VIFs), indicating potentially weak multicollinearity issues for the average score coefficient.

The main observation is that the estimate of γ is significantly dependent on the specification of scores. The results displayed in column 3 underline the importance of using several experts in order to model wine prices, since they all have differentiated impacts[6]. According to the naive model, a one-point-increase in the objective score of quality leads to a 4.8% increase in prices. Note that the coefficient of determination is not the highest for the average score model though it should be the best model. This suggests that the subjective opinions of the experts contribute to determining wine prices, which would invalidate the naive model.

The hierarchy of experts' influence is remarkably the same in the two models. In particular, Robert Parker (WA) is not the most influential expert: he is second to Stephen Tanzer (IWC). Both models conclude that Jancis Robinson (JR) has minimal influence and Wine Spectator (WS) has average impact. This is consistent with the results of (Ashton, 2013) regarding correlations between experts: he found that JR was the most "out of line" expert and that her taste differed the most from that of WA. However, this result may be due to a difference in the grading scale,

[5]The estimates of the trend and the fixed effects are available upon request from the authors.

[6]The differences among the $\hat{gamma}$ in column 3 are not due entirely to the differences in the samples (see the number of observations in column 2), since the distribution of the observations across appellations and vintage are actually similar for each expert. Thus we tested that assumption by conducting the six different regressions on the same sample of 700 wines. The results still indicated that the respective impacts of the experts are not equal.

TABLE 1.5 – First Stage Estimates of Equation (1.4)

	Avg. Score	IWC	JR	WE	WS	Multi-expert
Rainfalls 1	-.004***	-.004***	-.003	-.005***	-.015***	-.010***
Rainfalls 2	-.007**	-.010***	-.013***	-.017***	-.017***	-.014***
Temperature	.34***	.430***	.383 *	.537***	.375***	.446***
Trend	.315***	.379***	.441***	.351***	.453***	.444***

Note: Five separate regressions explaining each expert score and the average score by referring to weather and a trend variable; multi-expert regression explaining the expert scores altogether with the same explanatory variables.
Significance levels: ***1%, **5%, *10%.

as JR departs from WA, WS and IWC in that she uses a 20-points scale. This issue is addressed in chapter 2.

1.4.2 The Two-Stage Model Results

First Stage

Table 1.5 shows the first-stage estimates related to equation (1.4), for each single-expert regression, for the average score regression, and for the multi-expert regression.

We have tried numerous aggregated specifications of the weather data, since the estimates of the raw monthly temperatures and rainfalls coefficients were found to be rather inconsistent. We display the results using the following aggregates, which give the most robust and significant results[7]:

- The total rainfall during the first part of the growing season: April to July (referred to as Rainfall 1),
- the total rainfall during the last part of the growing season: August and September (referred to as Rainfall 2),
- the average of the monthly average temperatures during the growing season (from April to September, referred to as Temperature).

We have divided the growing season into two parts for the rainfall, because of the major expected impact of rainfall in the final months of the growth. At this point of growth, intense rainfall causes the grapes to rot, which jeopardizes the quality of the crop. To a lesser extent, rainfall still negatively affects the quality of the vintage during the rest of the growth season. The temperature showed important multicollinearity issues, which is why we focused on the average temperature during the growth season. We also show the estimated coefficients related to the trend, which is seen as the impact of overall progress in technology in the first stage estimation.

The signs of the coefficients are consistent with the literature in phenology: a good vintage is caused by a dry summer, with a high temperature. As expected, rainfall during August and September have a greater negative impact than rainfalls at the beginning of the growing season. This trend is also very significant, indicating great progress in technology through the trend in the scores.

As a clue to the accuracy of the first-stage fitted values, in Table 1.6 we provide some descriptive statistics of the residuals of the multi-expert regression. Column 2 displays the mean

[7]The several specifications we tried for the first stage all led to quasi-identical results in the second stage, supporting our conclusions.

TABLE 1.6 – Descriptive Statistics of Opinions Residuals from Multi-Expert Estimation of Equation (1.4)[a]

Expert	Mean	Mean of absolute values	Average absolute deviation from real value
IWC	1.474	1.873	2.1%
JR	-5.765	5.875	7.3%
WA	2.212	2.656	2.9%
WS	2.079	2.399	2.6%

[a]: see also last column of Table 1.5.

opinion for each expert, column 3 displays the mean of the absolute values, and column 4 gives the average deviation from the real value in percentage.

Column 1 shows that JR is, on average, far below the mean score. This is due to the fact that she originally grades on a scale of 20 and that we remapped these grades onto a 100-point scale in order to create homogeneity among the results. At the opposite end of the spectrum, WA, WS, and IWC are generally above the mean score. This regression is fairly accurate, since scores are estimated with an average error ranging from 2.1% for IWC to 7.3% for JR. Recall that our goal here is not the precise estimation of the scores but, rather, estimation of the residuals. The accuracy of the first-stage regression is not the issue, since we actually expect some heterogeneity in the residuals.

Second Stage

We now comment on the estimates of equation (1.5). Table 1.7 displays the bootstrap estimates for each expert-specific regression, for the average score regression, and for the multi-expert regression. None of the non-reported Sargan tests for overidentification allow rejection of the exogeneity of the instruments at the of 5% level. Hence, the exclusion condition for the validity of our instruments cannot be rejected.

The results definitely reject the naive model as suitable for assessing the impact of the objective component of scores. Indeed, each of the subjective components included in any of the specifications has a significant impact on wine pricing. The naive model aims to dispose of the opinions in order to focus on the objective scores, but these opinions have a significant impact on prices. The naive model is thus flawed: the $\hat{\gamma}$ displayed in Table 1.4 are all negatively biased. Two sources of bias are observed: the LLG is not valid because the opinions are not i.i.d., and the subjective components have a significant impact on prices. In addition, the systematic significance of the opinions also confirms the endogeneity of the raw scores and supports our two-stage model.

Furthermore, the opinions have different impacts. This means that one should use different experts in order to properly estimate the impact of the objective component, because they all influence wine prices in their own way. This is illustrated by the dispersion among the $\hat{\gamma}$ obtained in the average score and expert-specific regressions (upper part of Table 1.7). Note that the hierarchy of the experts is still robust: IWC and WA are the most influential experts, and JR is the least.

All the $\hat{\gamma}$ obtained with the two-stage procedure are much greater than those obtained with the naive model. This can be explained both by the downward measurement error bias (see Chesher 1991; Lecocq and Visser 2006a) and by the omitted variable bias, as the opinions are positively correlated with wine prices and mainly negatively correlated with the objective scores. It supports the use of our model to avoid those two biases. Broadly speaking, the key difference between our 2SLS estimate of γ and the naive estimate using the average score is

TABLE 1.7 – Second Stage Estimates of Equation (1.5)

	Average score and expert-specific regression				
	Avg. score	IWC	JR	WA	WS
Trend	-.044***	-.034***	-.047***	-.025***	-.021***
$\hat{\gamma}$	.170***	.183***	.128***	.109***	.088***
$\hat{\theta}$	.034***	.077***	.007***	.048***	.032***
		Multi-expert regression			
Trend	$\hat{\gamma}$	$\hat{\theta}_{IWC}$	$\hat{\theta}_{JR}$	$\hat{\theta}_{WA}$	$\hat{\theta}_{WS}$
-.022***	.137***	.045***	.004***	.041***	.015***

Note: Five separate augmented regressions explaining the prices by referring to fitted scores and first-stage residuals for each expert and the average score; multi-expert augmented regression explaining prices by referring to fitted score and all residuals for each expert from multi-expert first-stage. In the upper part, $\hat{\gamma}$ contains the estimated influence of the objective score using either only the average score or only a specific expert, leading to six different estimates of the same value. $\hat{\theta}$ contains the estimated influence of the subjective scores for each regression. In the lower part, we show the same estimates obtained with the multi-expert regression, leading to only one estimate for the influence of the objective score ($\hat{\gamma}$).
Significance levels: ***1%, **5%, *10%.

that the former better accounts for the vintage-specific effect by the use of additional weather information.

Note also that the $\hat{\gamma}$ coefficients are also greater than their related $\hat{\theta}$ coefficients. This roughly indicates that the objective component of the scores is more influential than the subjective one. In our model, a one-point increase in the objective score is estimated to lead to a 13.7% markup, whereas a one-point increase the one expert's subjective opinion have a maximum impact of 4.5% for IWC and only 0.4% for JR. It should be noted that the interpretation of the influence of each component is somewhat hazardous[8]. Indeed, the two components are purely visions and are indeed unobservable. However, our estimations do suggest that wine prices are more driven by fundamentals like weather than by the subjective opinions of experts.

Interestingly, a one-point increase of all the experts opinions together leads to a markup of 10.9%, slightly less the markup for a one-point increase of the objective component (13.7%). This consolidates the result that the use of weather variables allows to identify the vintage effect beyond that exhibited in the scores.

Another feature of these estimates is that the expert's respective influences are slightly lower in the multi-expert regression than in the naive model. This is consistent with Dubois and Nauges (2010), who also found an upward bias of the estimated influences when the unobserved quality, or the objective score as we call it here, are not controlled for.

Finally, the trend is also very significant. In the second stage multi-expert regression, we estimate that a wine becomes 2.2% more expensive each year due to storage costs, maturation, and scarcity value.

1.4.3 Marketing Effect

As an application of our model, we test the impact of the deviation among the scores on prices. Digressing slightly from our structural model, we add the empirical standard deviation of the opinions to equation (1.5) to assess the significance of its coefficient and the sign of the latter. The upper part of Table 1.8 shows the results of this estimation.

[8]Oczkowski (2016) have discussed our results presented in the published version Cardebat et al. (2014) by considering different forms for the price equation.

TABLE 1.8 – Second Stage Estimates of the Multi-expert Design

		Standard deviation regression				
Trend	$\hat{\gamma}$	$\hat{\theta}_{IWC}$	$\hat{\theta}_{JR}$	$\hat{\theta}_{WA}$	$\hat{\theta}_{WS}$	$\hat{\theta}_{\text{Standard deviation}}$
-.022***	-.138***	.035***	.046***	.030***	-.003	.093***
		Maximum score regression				
Trend	$\hat{\gamma}$	$\hat{\theta}_{IWC}$	$\hat{\theta}_{JR}$	$\hat{\theta}_{WA}$	$\hat{\theta}_{WS}$	$\hat{\theta}_{\text{Highest score}}$
-.023***	-.137***	.037***	.003***	.023***	-.006	.051***

Note: The interpretation of the estimates is the same as for the lower part of Table 1.7 except for the last column. $\hat{\theta}_{\text{Standard deviation}}$ and $\hat{\theta}_{\text{Highest score}}$ refer to the respective influence of the empirical standard deviation of the opinions and of the highest score.
Significance levels: ***1%, **5%, *10%.

The model indicates a strong positive impact of the standard deviation of experts' opinions on the prices of wines. That correlation might result from what we introduced as the "marketing effect" in section 1.2.3. In that case, we should find that the highest score has a major impact and is supposed to be the most publicized, hence the most important in the price equation.

We test this hypothesis by comparing the highest score to all the individual opinions in the two-stage model. The results of this estimation are displayed in the lower part of Table 1.8. The coefficient of the highest score is estimated to be the most important one. This is an argument in favor of the "marketing effect" interpretation. The highest score is the score most often exhibited , so it is the only hint of quality for consumers who have not searched for the other experts' grades (Hilger et al., 2011). Therefore, the highest score has the greatest influence on the price[9].

The revealed impact of the highest score sheds new light on the respective influence of the experts. The idea of a marketing effect with regard to the diffusion of scores might explain why JR is deemed to have so little influence on prices. Because she often grades below the average on her specific 20-points scale, many sellers might not be eager to exhibit her scores. The next chapter (Cardebat and Paroissien, 2015) addresses this issue and proposes a method to properly compare the scores, even on different rating scales. At the same time, the fact that WA and WS are usually above the mean score might play a role in their larger influence. This interpretation is consistent with the results of Table 1.8: the introduction of the highest score in the regression has lowered the coefficients of the above-the-mean-score experts.

1.5 Conclusion

This research assesses the role of expert opinion on Bordeaux wine prices using a methodology that, by including detailed meteorological data, fixed-effects models, and the systematic use of numerous expert scores, avoids endogeneity and bias rooted in errors of judgment. Like Dubois and Nauges (2010), Lecocq and Visser (2006a) and Oczkowski (2001), we assume that the observed scores result from an error measurement model: they can be split into an objective component shared by all experts and a subjective component specific for each expert and each wine. The latter is often seen as something that should be corrected because it obscures the signs of quality indicated by the objective component. We provide evidence, however, that in a price equation one should not try to get rid of subjective components because they significantly affect wine prices. Worse, if not handled specifically, the correction of these components leads

[9]This final result holds with all the specifications of the instruments we tried, and with the naive design, as including the maximum scores in equation (1.2) for a multi-expert regression leads to the same conclusion.

to downward-biased estimates of the impact of the scores as a quality indicator for wine prices. This result is consistent with Lecocq and Visser (2006a).

The most important result of our findings is the light shed on the role of the standard deviation in the price equation. We find a strong positive correlation between wine pricing and the standard deviation of the scores. Our interpretation is based on the fact that a higher standard deviation indicates that at least one score is above the others. In line with the marketing literature, this highest score might be used in an advertisement by the sellers. Hence, this particular score is likely to be the most publicized. As a result, this is certainly the only score that the average consumers have heard of. This is what we call the "marketing effect": the highest score is the most influential because it is the best known among consumers. Our interpretation is supported by the empirical analysis, since the highest score has the greatest impact on prices.

Nonetheless, we have to be cautious about this interpretation. There is another interpretation of our results: the consumers might be risk-takers. In this case, the standard deviation of the scores should also have a positive influence on demand and thus on prices. Economists generally agree that the ordinary consumer is rather risk averse, but the market we discuss here is very specific. The prices in our data range from $8 to $3,000, and the average price is $83. This is no market for the uninitiated. The consumers involved in this market are connoisseurs, professionals, or investors, and at least the last of these are likely to be risk-takers. However, we maintain our interpretation, assuming that market prices are more likely to be influenced by the marketing effect than by risk-taking behavior.

Chapter 2

Standardizing Expert Wine Scores: An Application for Bordeaux *en primeurs*

This chapter has been co-written with Jean-Marie Cardebat and published in 2015 in the *Journal of Wine Economics* as Cardebat and Paroissien (2015). This version adds the mention to the origin of equipercentile equating that we ignored at the time of the publication.

2.1 Introduction

As an experience good, the quality of a wine is only known after its consumption. In contrast to consumers, wine producers are informed about their products' quality. This information asymmetry has led to the emergence of wine experts providing information on wine quality. The contingent information market is particularly well-developed in the wine sectors where numerous experts coexist. The subjectivity of the wine quality assessment, the regional segmentations[1], or their (supposed) preferences (Storchmann, 2012) partly justify a large number of experts. Moreover, the grading systems and habits could differ from one expert to another. In particular, the European experts are used to rating wine on a 20-point scale whilst US experts use 100 points (e.g., Masset et al. 2015). The heterogeneity of the rating systems can increase the consumer's perceived uncertainty. The question of rating homogenization on the same scale of preferences is therefore at the heart of the uncertainty debate about wine quality.

The uncertainty about wine quality is particularly high during the *en primeur* campaign in the Bordeaux Region. The *primeurs* market can be seen as a forward market dedicated to fine Bordeaux wines. The *en primeur* campaign takes place during the spring, starting with a huge multi-day tasting organized by the châteaux in the first week of April. Wine merchants, wine enthusiasts, and of course, wine experts are involved in this event. They all taste the wine from the latest harvest. Therefore, the wine is not yet finished and the quality assessment is particularly difficult and uncertain. The aim of this campaign is to sell (châteaux) and buy (wine merchants)[2] before the wine is effectively released in bottles, which will happen about 18 months later. The prices and quantities exchanged are determined during the *en primeur* campaign and the wine will be delivered once it is bottled.

The economic stakes of the tasting are therefore extremely high because prices and quantities exchanged are influenced by the experts' scores. The wine economics literature has provided ample evidence of the link between *en primeur* wine prices and the experts' scores (see notably Hadj Ali and Nauges 2007; Hadj Ali et al. 2008; Masset et al. 2015). Another strand

[1] By regional segmentation we refer to the fact that not only are experts more or less specialized in wines coming from specific regions, but also that some experts target specific consumers (at least as regards the choice of the language in which they edit their comments).

[2] The wine merchants (called négociants in Bordeaux) are free to buy or not, but they receive allocations (the right to buy in a certain amount) from the châteaux and if they do not buy a specific year, the châteaux may remove their allocations for the following year.

of the literature deals with the information contained in the experts' grades (see for example Ashenfelter et al. 1995; Ashenfelter 2008; Cardebat et al. 2014), the divergence between experts (notably Ashton 2012, 2013; Hodgson 2008; Masset et al. 2015; Olkin et al. 2015) or the randomness of the tastings (e.g., Ashton 2014a; Quandt 2007; Bodington 2015).

However, no paper has tried to express the experts' scores on the same scale of preference or in the same rating system before analyzing the grades divergence or bias or impact on prices. As noted by Masset et al. (2015) "Comparisons are difficult to make, as not all experts use the same scale to establish their scores". Furthermore, as far as we know, there is no paper trying to provide a global score aggregating all the marks released by experts during the *en primeur* campaign, although a demand exists for such a global score from the professionals. However, if no academic papers exist, in the wine industry, most of the web merchants provide such aggregated scores (see, for example, wine decider or wine searcher). The website of Bertrand Leguern is also dedicated to the calculation of an aggregated score which is used by wine professionals. Nevertheless, we cannot find any information on the way these scores have been calculated. There is no transparency in their calculation, thereby reinforcing the information asymmetry instead of reducing it.

Wine professionals, particularly the négociants who buy *en primeur* wines, request aggregated and transparent information on wine quality rather than comparing numerous grades emanating from a variety of experts. This highlights the importance of reducing the information asymmetry and therefore increasing the *en primeur* market efficiency (Mahenc and Meunier, 2006). Given the pending retirement of the main expert, Robert Parker, harmonizing experts' scores appears particularly useful since Parker's disappearance will reinforce the uncertainty and the need for a reference score.

The aim of this paper is, therefore, to develop a methodology for calculating a single score aggregating the grades released by 15 experts who have traditionally been scoring Bordeaux *en primeur* wines since the beginning of the last decade. Based on a large database of Bordeaux *en primeur* expert scores, we suggest a methodology to translate the rating scale of one expert into the rating scale of another, thereby facilitating the comparability of all the experts' scores[3]. The global score is then basically calculated as a simple arithmetic average of these transformed scores. This aggregated score has the potential to be considered as a new reference score on the fine wine market.

This study may be interesting to academics who may benefit from a methodology ensuring proper expert score comparisons by taking into account the different rating systems among experts. In addition, based on this methodology, we provide wine professionals with a unique standardized wine score aggregating the information coming from all experts operating on the *en primeur* market.

The remainder of this paper is structured as follows: the next sections present our dataset, while section 2.3 displays the methodology of the standardized wine score; section 2.4 reports the standardized scores and discusses the results following different robustness checks; the last section concludes.

[3]We thank Eddie Oczkowski for having mentioned to us that this method is in fact already popular in the fields of psychology and economics of education, and referred to as equipercentile equating. The introduction of this method is generally attributed to Braun and Holland (1982), see Kolen and Brennan 2014 for an updated presentation.

TABLE 2.1 – Descriptive Statistics of Expert Score Data

Expert	Frequency	Min	Max	Mean	Median	Standard Dev.
Rene Gabriel	3,639	12	20	17.12	17	1.14
Wine Spectator	2,886	77	98.5	90.2	90	3.45
Robert Parker	2,609	71.5	99.5	90.38	90.5	3.52
Jancis Robinson	2,538	12	20	16.4	16.5	0.99
Jacques Dupont	2,156	13	20	15.82	16	1.28
Bettane & Desseauve	2,113	10	20	16.56	16.5	1.33
Neal Martin	1,711	70	99	90.03	90	3.53
Decanter20	1,615	14.5	20	16.93	17	1.04
Jean-Marc Quarin	1,497	10	20	15.74	15.75	1.07
James Suckling	1,059	84.5	100	91.25	91.5	2.72
Decanter100	1,026	81	95	88.19	88	2.91
Tim Atkin	1,011	82	100	91.37	92	3.35
La RVF	484	11.5	20	16.29	16.25	1.38
Jeannie Cho Lee	219	80	99	91.87	92	2.81
Antonio Galloni	210	79	95.5	89.24	89.5	2.77
Jeff Leve	158	83	99	90.33	90	3.01

Source: Author's calculation based on Wine services (2015) data.

2.2 Data

Our dataset contains the scores given by 15 well-known wine experts[4] during the *en primeur* campaign over the period from 2000 to 2014. All the wines rated by these experts are present in the dataset which represents 447 châteaux and 4,333 château-vintage pairs; that is, on average, each château is rated 9.7 times over the observed period.

The first column in Table 2.1 shows the number of wines effectively rated by each expert. Rene Gabriel appears to be the most productive expert with 3,639 scores over the period. Similarly, five additional experts are highly active on the wine opinion market. They all have rated more than 2000 *en primeur* wines between 2000 and 2014. In contrast, the last four experts of this list exhibit a significantly weaker activity with fewer than 500 scores each. The following columns display the traditional descriptive statistics on the experts' scores. Among the 16 (15 + 1, see footnote 1) experts, seven use a 20-point grading scale, they are all European, and nine use a 100-point scale, they are overwhelmingly American. The Chinese J. Cho Lee and the British Tim Atkin are exceptions.

The scores given by the experts seem relatively homogeneous and average between 15.74 and 17.12 for the European raters and between 89.24 and 91.87 for the U.S. experts. Interestingly, we note that the Europeans have all awarded a 20-point maximum grade at least once while only J. Suckling and Tim Atkin have handed out the maximum 100-point grade. The score range defined as the difference between the maximum and the minimum score for each expert lies between 14 and 29 for the US experts and 5.5 to 10 for the European experts.

[4]The term "expert" is used here indifferently to designate a person (James Suckling, Jancis Robinson, etc.) or an organization (i.e., magazines like Wine Spectator or La Revue du Vin de France – RVF, etc.). Decanter has a special status in the sense that we split its scores into two categories: Decanter 20 and Decanter 100 because Decanter chose to change its traditional 20-point scale for a 100-point scale during the period studied. We have therefore decided to consider its scores on a 20-point scale and on a 100-point scale as two different experts.

We note two remarkable facts. First, all experts utilize only a fraction of their scale. In comparison, the fraction utilized by U.S. experts seems to be particularly small (20 points on average). However, in absolute values this exceeds the spectrum used by European experts (7.8 points on average), giving the former a potentially higher accuracy in their rating. Second, both U.S. and European raters exhibit significant differences in the way they rate the wines: there is no homogeneity among them concerning score range they use. Therefore, the direct comparison among experts' scores is fallacious, even if they use the same rating scale. Each expert has his/her own preference space and our aim is to express all scores in the same space of preferences.

The medians also offer interesting information as they can be interpreted as a threshold between good wines and less good/bad wines. 90 points (16.5) for the U.S. (European) experts appears to be the dividing line between these two categories.

Table 2.2 presents the number of wines that have been tasted by each expert pair, i.e., by at least two experts. With 2,698 wines rated both by René Gabriel and Wine Spectator, these two experts exhibit the highest overlap. On average, Robert Parker, Neal Martin, Jancis Robinson, Wine Spectator, Bettane & Desseauve, Jacques Dupont, La Revue du Vin de France, and Rene Gabriel have rated more than 1,000 identical wines over the observed time period.

Table 3 reports a systematic positive correlation between each expert pair; however, the average correlation among experts does not exceed 0.59. Jean-Marc Quarin and Jeffe Leve exhibit the highest correlation. Jancis Robinson and Antonio Galloni exhibit the lowest correlation and therefore the lowest agreement (concordance) with the other experts. In contrast, Jeff Leve and Decanter 20 display the highest correlation and therefore the best level of concordance with the other experts. In particular, these two experts' grades are strongly correlated with those by Robert Parker. The U.S. experts seem to have higher concordance among themselves compared to the European ones. These results are in line with the work of Masset et al. (2015), even if their results suggest a high level of concordance among various wine raters. In contrast, given an average correlation of 0.59 and a high volatility of the correlation coefficients, we do not deem the level of expert concordance particularly high.

TABLE 2.2 – Wine Pairings: Number of Identical Wines Tasted by Two Experts

	RP	NM	JR	WS	AG	BD	JD	JS	JC	JL	RVF	JMQ	RG	TA	D20	D100
RP		1,361	1,833	2,168	168	1,637	1,667	706	231	568	1,317	1,041	2,443	750	81	198
NM	1,361		1,549	1,419	160	1,422	1,353	714	243	556	1,294	996	1,578	787	842	170
JR	1,833	1,549		2,049	171	1,929	1,946	730	247	561	1,440	1,268	2,361	838	898	194
WS	2,168	1,419	2,049		168	1,753	1,803	663	221	529	1,330	1,151	2,698	743	832	178
AG	168	160	171	168		161	164	177	1	173	157	167	184	158	13	183
BD	1,637	1,422	1,929	1,753	161		1,756	687	232	570	1,427	1,185	2,007	773	856	200
JD	1,667	1,353	1,946	1,803	164	1,756		650	230	515	1,316	1,139	2,039	738	796	186
JS	706	714	730	663	177	687	650		219	535	618	652	867	731	565	211
JC	231	243	247	221	1	232	230	219		158	211	186	250	227	246	1
JL	568	556	561	529	173	570	515	535	158		484	427	599	524	403	203
RVF	1,317	1,294	1,440	1,330	157	1,427	1,316	618	211	484		959	1,546	690	756	191
JMQ	1,041	996	1,268	1,151	167	1,185	1,139	653	186	427	959		1,366	657	538	206
RG	2,443	1,578	2,361	2,698	184	2,007	2,039	867	250	599	1,546	1,366		918	929	214
TA	750	787	838	743	158	773	738	731	227	524	690	657	918		670	181
D20	811	842	898	832	13	856	796	565	246	403	756	538	929	670		0
D100	198	170	194	178	183	200	186	211	1	203	191	206	214	181	0	
Average	1,262	1,054	1,340	1,329	154	1,231	1,215	602	202	473	1,008	878	1,495	656	653	180

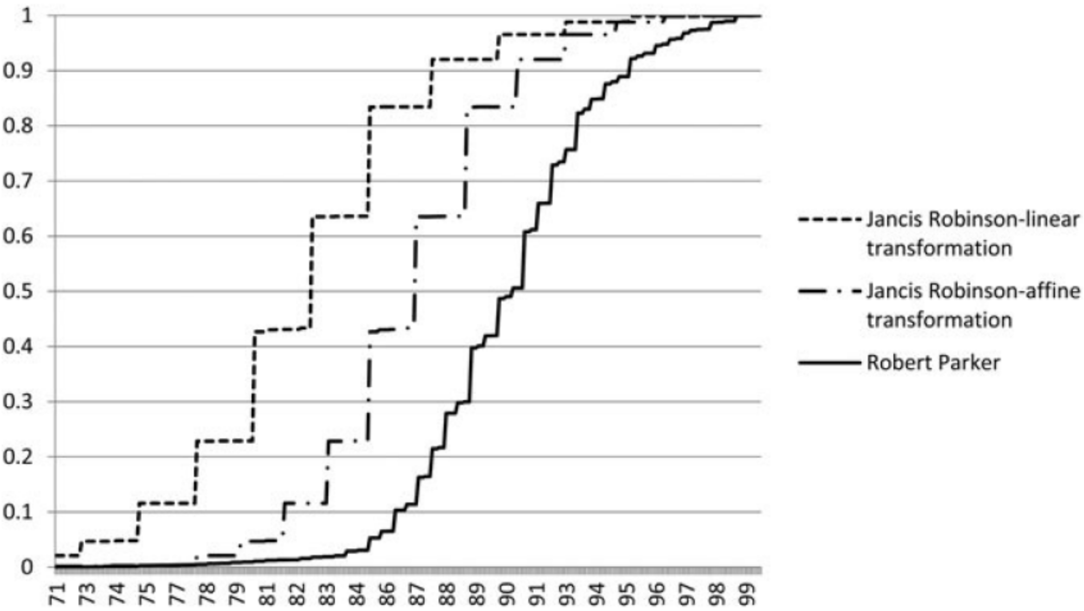
Source: Author's calculation based on Wine services (2015) data. WS: Wine Spectator; RP: Robert Parker; JR: Jancis Robinson; JD: Jacques Dupont; BD: Bettane & Desseauve; NM: Neal Martin; D20: Decanter20; JS: James Suckling; D100: Decanter100; RVF: La Revue du Vin de France; JCL: Jeannie Cho Lee; AG: Antonio Galloni; JL: Jeff Leve; JMQ: Jean-Marc Quarin; TA: Tim Atkin; RG: Rene Gabriel.

TABLE 2.3 – Expert Score Correlation Matrix

	RP	NM	JR	WS	AG	BD	JD	JS	JC	JL	RVF	JMQ	RG	TA	D20	D100
RP		.57	.43	.61	.41	.59	.50	.69	.68	.77	.60	.68	.58	.57	.71	.58
NM	.57		.49	.62	.56	.58	.50	.69	.59	.74	.59	.65	.58	.57	.67	.61
JR	.43	.49		.51	.17	.50	.39	.49	.55	.43	.46	.54	.45	.57	.63	.36
WS	.61	.62	.52		.59	.62	.47	.74	.70	.75	.60	.64	.61	.64	.69	.62
AG	.41	.56	.17	.59		.45	.35	.47		.60	.44	.52	.54	.35	.32	.56
BD	.59	.58	.50	.62	.45		.50	.65	.70	.67	.65	.69	.55	.63	.75	.74
JD	.50	.50	.39	.47	.35	.50		.59	.60	.62	.56	.62	.47	.52	.63	.67
JS	.69	.69	.48	.74	.47	.65	.59		.66	.75	.70	.70	.66	.54	.71	.60
JC	.68	.59	.55	.70		.70	.60	.66		.71	.58	.65	.64	.62	.72	
JL	.77	.74	.42	.75	.60	.67	.62	.75	.71		.73	.79	.69	.56	.68	.68
RVF	.60	.59	.46	.60	.44	.65	.56	.70	.58	.73		.72	.57	.56	.71	.78
JMQ	.68	.65	.54	.64	.52	.69	.62	.70	.65	.79	.72		.65	.67	.75	.75
RG	.58	.58	.45	.60	.54	.55	.47	.66	.64	.69	.57	.65		.57	.65	.62
TA	.57	.57	.57	.64	.35	.63	.52	.54	.62	.56	.56	.67	.57		.66	.61
D20	.71	.67	.63	.69	.32	.75	.63	.71	.72	.68	.71	.75	.65	.66		
D100	.58	.61	.36	.62	.56	.74	.67	.60		.68	.78	.75	.62	.61		
Average	.59	.60	.45	.62	.46	.59	.51	.65	.64	.69	.60	.65	.58	.57	.66	.63

Source: Author's calculation based on Wine services (2015) data. WS: Wine Spectator; RP: Robert Parker; JR: Jancis Robinson; JD: Jacques Dupont; BD: Bettane & Desseauve; NM: Neal Martin; D20: Decanter20; JS: James Suckling; D100: Decanter100; RVF: La Revue du Vin de France; JCL: Jeannie Cho Lee; AG: Antonio Galloni; JL: Jeff Leve; JMQ: Jean-Marc Quarin; TA: Tim Atkin; RG: Rene Gabriel.

FIGURE 2.1 – Distribution Functions for Each Transformation and Robert Parker's Score Distribution



Source: Author's calculation based on Wine services (2015) data.

2.3 Methodology

Robert Parker and Jancis Robinson are influential experts, in the U.S. and in England, respectively, and best embody the issue of transforming the grading scales. While Robert Parker scores out of 100 points, Jancis Robinson scores out of 20 points. Our method addresses a common quality assessment problem. Imagine a comparison between two wines where the first is graded by both experts, but the second one is only rated by Robert Parker. The key issue how to properly utilize the information given by Jancis Robinson and translate them into Parker scores.

The naive solution is the linear function by simply multiplying Jancis Robinson's scores by a factor of five. However, this solution is unsatisfactory, as it disregards the utilized score range of [12,20] for Robinson and [70,100] for Parker. In order to consider the minimums of the intervals utilized by each expert, one can employ an affine function of the Robinson's scores from the interval [12,20] into the interval [70,100][5]. The best way to judge the relevance of this transformation is to compare the respective distribution functions. Figure 2.1 displays the distribution functions of Jancis Robinson's scores after each transformation, compared to Robert Parker's score distribution function.

The distribution of Jancis Robinson's transformed scores is closer to Robert Parker's distribution with the affine function. Still, one might argue that Jancis Robinson's transformed scores are still underrated compared to the grading system of Robert Parker. More than half of Robert Parker's scores are above 90/100, against only 8% for the Robinson's scores computed with the affine function. As a result, a 90/100 for Robert Parker is a much lower evaluation of quality than a 90/100 for Jancis Robinson with the affine function. A satisfactory transformation of the scores should both put the scores on the same scale and convey the same value to each

[5]This affine conversion formula of $x \in [12; 20]$ into of $y \in [70; 100]$ is $y = \frac{30}{8}x + 25$.

score. Jancis Robinson's transformed scores should then follow the same distribution function as Robert Parker's scores. Such a function exists and is nonparametrically tractable.

The theoretical framework is the following. Posit that quality of Bordeaux wines is a random variable. The experts evaluate this quality along a scale of their choice, according to their preferences and to their utilization of their scales. Let F be the distribution function of Jancis Robinson's scores, and G be the distribution function of Robert Parker's scores. These functions express both experts' grading scales as well as their respective appreciation of Bordeaux wines. These differences in scales and in overall appreciation of Bordeaux wines tackle the comparison between grades given by two experts. The method controls for both issues at the same time.). Recall that our objective is to utilize Jancis Robinson scores and translate them into Parker scores, accounting for the fact that Jancis Robinson usually awards lower scores.

We apply the function $G^{-1} \circ F$ in order to obtain the same distribution function for the Jancis Robinson transformed scores and Robert Parker raw scores. This uses the following classical property of probability distributions. Let F_X and F_Y be the distribution of the continuous random variables X and Y, then the random variable $F_Y^{-1} \circ F_X(X)$ has the same probability distribution as Y, F_Y^{-1} being the generalized inverse of F_Y. To avoid any selection bias, the two empirical distributions are computed on a common sample, which contains all wines with a score from each of the two experts. For the chosen pair of experts, the sample includes 1,833 observations.

Let s_{ik} be the score given by expert i to wine k, and I_i be the list of the wines graded by expert i. The procedure is the following:

1. For each expert i, we compute the empirical distribution function

$$\hat{F}_i(x) = \frac{1}{card(I_i)} \sum_{k \in I_i} 1\{s_{ik} \leq x\}$$

2. For any chosen expert j (here we have chosen Parker), we compute the generalized inverse of $\hat{F}_i$:

$$\hat{F}_j^{-1} = inf\{x \in \mathbb{R} | \hat{F}_j(x) \leq y\}$$

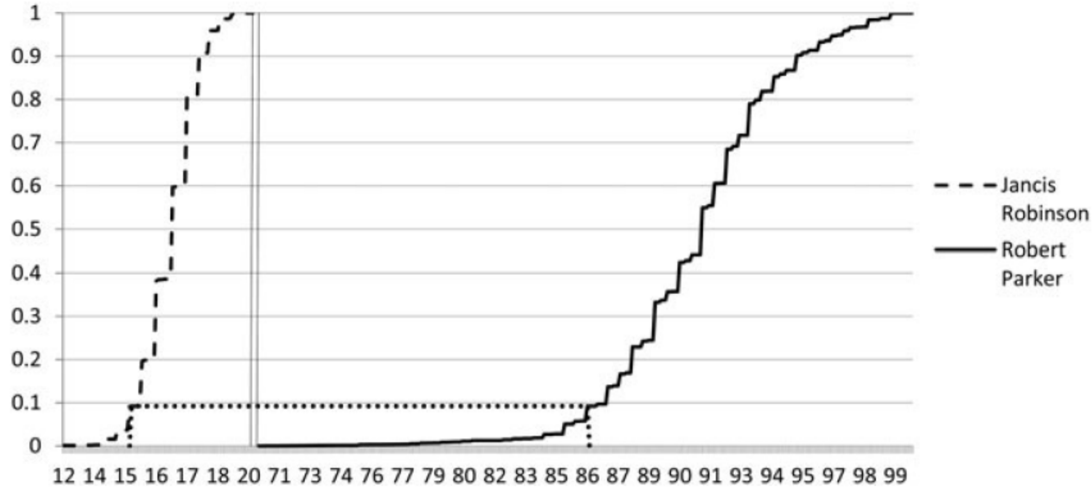
3. The conversion function of the grades of expert i into the scale of expert j is given by:

$$\phi_{ij}(x) = \hat{F}_j^{-1}(\hat{F}_i(x))$$

Figure 2.2 provides a graphical illustration of our method. As an example, we evaluate the image of a 15/20 from Jancis Robinson on the Robert Parker scale[6]. 15/20 is the quantile of

[6] The procedure is symmetrical, i.e., it is possible to turn the scores of any expert into the scale of any other expert. Also, it is self-consistent as the conversion function from expert A to expert B is the inverse of the conversion function from expert B to expert A, for all scores observed in the data. For instance, as the data contains a 90/100 from Parker, if we put this score into another expert's scale and turn it back into Parker's scale, we will always end up with a 90/100. This works for all observed scores in the data. However, to be comprehensive, it is not exactly the case with the scores that are unobserved in the data (because the empirical cumulative distribution function is not bijective). The transformation function combined with it generalized inverse does not necessarily give the exact same score. Indeed, the procedure always end up with a score that is originally observed in the data. A simple way to overcome this asymmetry would be to linearly interpolate the empirical distribution function, so as to obtain only bijective functions. As we have not meant the procedure to be applied to scores out of the sample, this is not a major issue for the scope of this paper. Furthermore, considering the large size of our database, the observed scores most likely include all potential scores, so that symmetry is guaranteed for arguably every possible score and for each expert.

FIGURE 2.2 – Method Using the Empirical Distribution Functions



Source: Author's calculation based on Wine services (2015) data.
Note: The double vertical lines stands for the gap on the x-axis between 20 and 70.

order 0.092 for Jancis Robinson's distribution function, which means that 9.2% of the Jancis Robinson scores are less than or equal to 15/20. On the Robert Parker distribution function, we read that this quantile is 86/100. We obtain that a 15/20 given by Jancis Robinson is worth a 86/100 given by Robert Parker. In the situation previously stated, this method allows the Jancis Robinson score to be turned into the Robert Parker scale. The average of the two scores is a synthetic indicator of all available information, and can be directly compared with Parker scores if Jancis Robinson scores are missing.

Applying the same method for all existing scores from Jancis Robinson, we obtain a nonparametric function which ensures that the image scores have the same distribution as the Robert Parker scores. Figure 2.3 compares the plots of three functions, i.e., linear, affine and nonparametric[7].

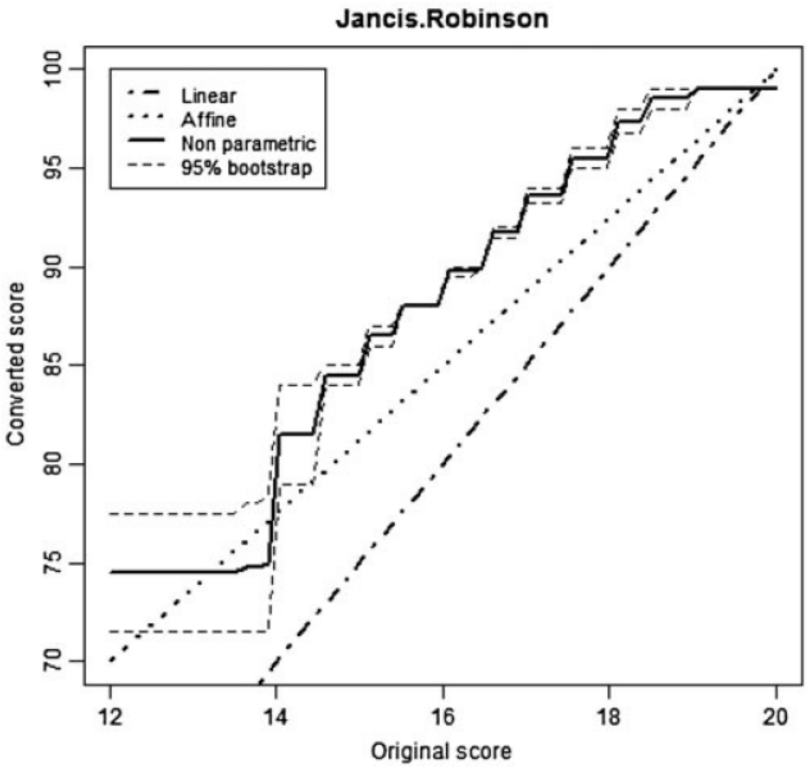
The nonparametric function is irregular on the half-open interval (12,14]. In fact, this interval only concerns five observations and 0.4% of the distribution of the Jancis Robinson scores. It corresponds to the half-open interval [70,81.5] for Robert Parker. As a result, the confidence interval is wide below 14/20, so that our conversion is not significantly different from the affine conversion for low grades. However, for high scores, the nonparametric conversion yields significantly higher grades out of 100 than the affine one.

Besides, while the correlation coefficients are neither affected by the linear nor by the affine conversion, the nonparametric method slightly alters the coefficients between the experts. The coefficients computed after conversion are given in the Appendix B. The change in the correlation coefficient provides a measure of the nonlinearity of the nonparametric conversion. This is measured by the absolute difference between the coefficients before and after conversion in the Appendix B.

This method can also be applied for two experts who both score out of 100. Figure 2.4 plots the nonparametric function which turns Neal Martin scores into the Robert Parker scale. We find the same regularity issue below 85 points, but the function suggests that Robert Parker has

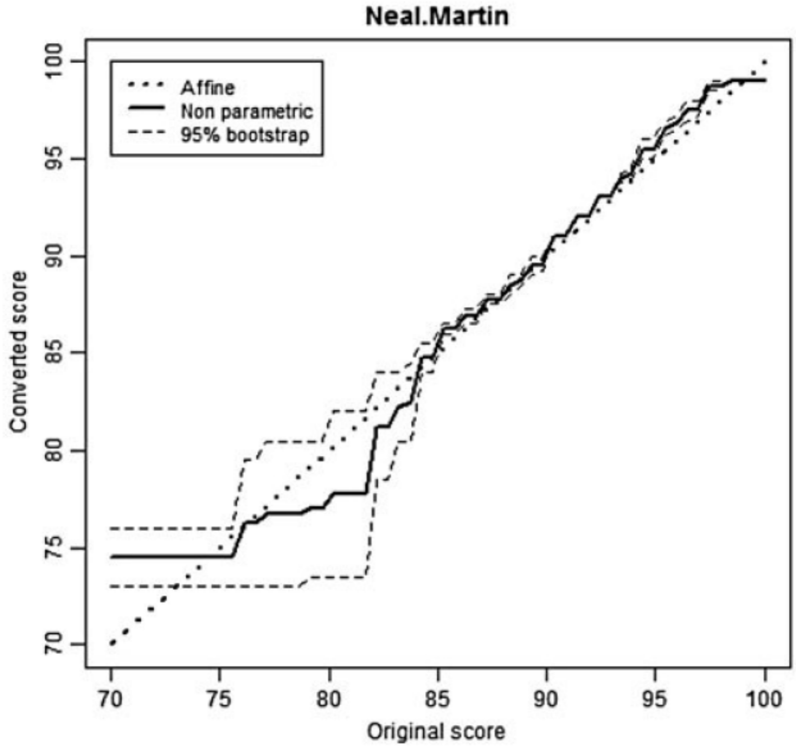
[7]The confidence bands have been obtained by bootstrapping the curve 1,000 times. That is to say, we re-sampled our data 1,000 with replacements, and conducted this procedure for each sample. For each score, we then obtained 1,000 estimates of the converted score. The bootstrap confidence interval is given by the quantiles of order 0.025 and 0.0975 of each score.

FIGURE 2.3 – Plot of Three Transformation Functions



Source: Author's calculation based on Wine services (2015) data.

FIGURE 2.4 – Conversion of Neal Martin Scores into Robert Parker Scale



Source: Author's calculation based on Wine services (2015) data.

been less reluctant than Neal Martin to grant scores above 95/100. For instance, a 97/100 by Neal Martin is as rare as a 98/100 by Robert Parker. Still, the nonparametric conversion does not represent much change compared to the identity function. Our method is more valuable for experts who do not grade on the same scale. All conversion curves are displayed in the Appendix B, along with the affine and the linear ones (which are only different for the experts who grade out of 20). For the latter in particular, the results of the nonparametric method are significantly different from the output of the affine conversion.

2.4 Example of Outcomes

Our conversion method facilitates various kinds of comparisons between scores, whether among winemakers, appellations or vintages. We hereafter provide an insight into the possible outcomes. While the general method allows the scores of any expert to be converted into any other expert's scale, we have chosen to convert all scores into the Robert Parker scale. Since he is commonly referred to as the most influential expert for Bordeaux wines (see notably Hadj Ali et al. 2008; Masset et al. 2015), we assume that his scale is the most familiar for the reader.

Table 2.4 displays all available 2013 *primeurs* scores for a subsample of twenty Bordeaux properties. Columns 2 to 4 reports the average of the available scores transformed by the linear, the affine and the nonparametric function, respectively. Our nonparametric method yields the highest scores, as it transposes the scores on the scale of Robert Parker, used to giving high scores compared to his peers. Overall, the other experts mitigate the negative opinion of Robert Parker of the 2013 vintage, as the mean score is often above Robert Parker's grade.

TABLE 2.4 – Raw *primeurs* Scores for a Subsample of Vintage 2013 and Mean Scores Computed for the Three Methods

Wine	Score Linear	Score Affine	Score Nonpar.	sd	RP	NM	JR	WS	AG	BD	JD	JS	JL	RVF	D	JMQ	RG	TA
Angelus	89.4	91.7	92.7	1.87	91.5	91	17.5		91	18.75		92.5	92	16.25	90.25	16.5	17	95
Ausone	90.3	92.6	93.7	2.27	94	92	17.5		91	19	17	91.5	94	16	92	16.75	19	94
Cheval Blanc	89.1	91.7	92.6	1.25	90	92	17	92.5	91	18	16.5	93.5	93.5	16.25	92	16.5	18	93
Clinet	86.7	89.7	90.4	1.2	92	91	16	88.5	91.5	16		90.5	92	16	89	15.5	17	91
Eglise Clinet	91.2	93,0	94,0	1.93	93	95	17.5	90.5	91.75	17.5		93.5	95	17.25	91	16.75	19	96
Evangile	88.1	90.8	91.9	1.98	88.5	92	18	91.5	93	17.75	16	90.5	91	16.25	90.25	16	17	92
Gazin	85.6	89.3	90.2	1.5	91	90	16	87.5		15.5	16.75		90.5	16.75	89	15.75	17	91
Grand Vin de Latour	89.1	91.8	92.8	1.55	89		17	91.5	92	17.25	16.75	92.5	92	17.5	94	16	18	95
Haut Brion	90.1	92.2	93.1	1.85	91	90	16.5	92.5	92.5	18	16.75	92.5	93	16.5	94		19	93
La Conseillante	86.7	89.9	90.8	1.7	90	91	15	89.5	91.5	17	15.5	90.5	92	17	90	16	17	92
La Violette	88.6	91.1	91.7	2.49	87	93			88	17.75			93.5	16.75		16.25	18	93
Lafite Rothschild	88.6	91.3	92.4	1.8	88	92	17	90.5	91.5	17.75	16.25	92.5	91	16.25	94	16	18	95
Lafleur	90.6	92.8	93.8	1.84	90	94	18		93	18		93.5	93.5	17.5	93	16.75	17	95
Le Gay	86.5	89.6	90.3	2.64	86	91	15		90.5	17		91.5	93	15	88	16	18	93
Margaux	89.9	92.1	92.9	1.53	89	92	16.5	91.5	92.5	17.25	16.5	94.5	93.5	17.5	94		18	94
Mouton Rothschild	89.6	92,0	92.8	1.79	92	93	17	92.5	89	18.25	17	92.5	93	17.5	94	16.75	17	92
Pavie	88.6	91.3	92.1	1.9	93	92	16		92.25	18		91.5	93.5	15.5	93	16.25	18	90
Petrus	90.2	92.5	93.5	1.82	91.5	91	18.5		91	18	17.5	92.5	94	17	92	16.5	18	94
Trotanoy	87.5	90.7	91.7	1.6	92.5	91	16	91.5	91	17	15	90.5	93.5	16	91	16	18	95
Vieux Château Certan	88.0	90.6	91.5	1.93	87.5	93	17.5	91.5	91.5	17.5	15	90.5	91.5		91	16.25	17	92

Source: Author's calculation based on Wine services (2015) data.
Notes: sd: standard deviation of the scores obtained from nonparametric method. RP: Robert Partker ; NM: Neal Martin; JR: Jancis Robinson; WS:Wine Spectator; AG: Antonio Galloni; BD: Bettane et Desseauve; D: Decanter; JD: Jacques Dupont; JS: James Suckling; JL: Jeffe Leve; RVF: Revue du Vin de France ; JMQ: Jean-Marc Quarin; RG: Rene Gabriel; TA: Tim Atkin.

TABLE 2.5 – Mean Vintage Score for Robert Parker and Jancis Robinson with and without Transformation

Vintage	Number of Observations	Robert Parker	Jancis Robinson Nonparametric Function	Jancis Robinson Raw Scores
2003	126	90.5	89.9	16.1
2004	69	91.3	92.0	16.6
2005	174	91.8	91.7	16.6
2006	116	91.6	92.2	16.7
2007	196	88.7	90.5	16.2
2008	198	91.0	91.9	16.6
2009	195	92.6	92.4	16.7
2010	201	92.4	92.4	16.7
2011	186	90.0	91.2	16.4
2012	194	90.5	92.3	16.7
2013	168	88.9	91.1	16.3

Source: Author's calculation based on Wine services (2015) data.
*Notes:*We lack Jancis Robinson *primeurs* scores for vintages 2000, 2001, 2002 and 2014.

The last column of Table 2.4 provides the standard deviation of the scores for each wine. As our method displays all scores on the same scale, it is now possible to compute the relevant standard deviation for each wine across experts. This provides a measure of judge concordance for each wine: the lower the deviation among the scores, the more reliable is the mean score. Château Clinet shows the highest level of agreement among the raters with a standard deviation of 1.2 while Château Le Gay shows the largest dispersion with a standard deviation of 2.64.

Another possible outcome is to facilitate the comparison between vintages for two experts. Table 2.5 displays the mean scores of vintages 2003 to 2013 for Robert Parker and Jancis Robinson with and without the transformation of Jancis Robinson's scores. Expressing the two assessments on one scale makes them comparable. Our transformation highlights that Jancis Robinson was much more lenient with the 2007 and 2013 vintages than Robert Parker, and that she apparently enjoyed vintage 2012.

2.5 Conclusion

This paper employs a simple methodology to express the scores of various wine experts on the same rating scale. It facilitates the comparability of the scores among experts and allows to calculate an average of all available wine scores.

Nevertheless, several issues still have to be addressed.Who has to be the expert of reference? Robert Parker seems to be the natural candidate but he has now retired and stopped tasting the Bordeaux *en primeur* in 2015. How to interpret the standard deviation in the cases where wines are not tasted by the same number of experts? Does a standard deviation calculated on the basis of 2 scores provide the same information as a standard deviation calculated on the basis of 15 scores in terms of consensus? Other questions will certainly have to be addressed and we hope that this paper will induce further research to improve our methodology.

Chapter 3

The Causal Impact of Medals on Wine Producers' Prices, and the Gains From Participating in Contests

This chapter has been co-written with Michael Visser.

3.1 Introduction

There are many goods whose quality is unknown until actual consumption. For instance, a book's content is uncertain until the text is read. Similarly, a film's story is only revealed when it is seen in a movie theater, and the pleasure procured by a bottle of wine can not be judged before it is uncorked, smelled, and tasted. Producers of such so-called experience goods (Nelson, 1970) face the challenge that potential purchasers must somehow be informed about the *ex ante* unknown quality. To reduce the information asymmetry between consumers and producers, the latter can spend money on advertising and marketing. The movie industry, for example, devotes substantial budget resources to promote films before they are released to the public. Consumers themselves can also contribute to spreading product information by word-of-mouth: they speak with their friends and relatives about the latest music album they have listened, or add their personal opinion on on-line music blogs. In some cases (partial) information dissemination is mandatory because laws and regulations oblige firms to disclose features of their products. Finally, hidden characteristics of goods may be revealed through awards attributed at competitions: literature lovers learn that the novel receiving the Man Booker prize is the jury's preferred one among the hundreds of new novels published each year, a signal for them that the winning book is likely of high quality. Movie fans can make analogous inferences regarding films awarded at the Oscar ceremonies or the Cannes festival.

The producers we study in this paper are Bordeaux wine makers, and for them there is basically just one way in which they can inform potential purchasers about the quality of their goods, and that is by participating in wine competitions (and win medals). One reason for this is that all forms of alcohol publicity is forbidden in France. Local regulations in Bordeaux also limit what producers are allowed to write on the bottle labels. Furthermore, the wines we are analyzing are mostly still very young and unavailable to consumers, thereby limiting customer-to-customer transmission effects. But the main reason is that the focus of our study is not the top-end segment of the market (made up of a small number of world-famous châteaux like Latour, Haut-Brion, Margaux, Mouton-Rothschild, Yquem, etc.), but the vast majority of lesser known wines. Unlike the top-notch wines, they are not actively traded in auctions throughout the world, nor are they commented and evaluated by influential critics such as Robert Parker or Jancis Robinson. In the absence of these vehicles of information transmission, the less known clarets can only hope to differentiate themselves from their numerous competitors by winning awards. Anecdotal evidence suggests that medals have strong price effects. According to *La Revue du Vin de France* (issue 600, March 2006), a leading French wine magazine, winning a medal at a wine competition allows a producer to increase its price by between 10 and 15%; in the same vein, the organizers of the *Concours de Bordeaux*, the most important competition for Bordeaux wines, state that a gold medal from this contest allows the recipient to augment its price by up to 30%.[1]

Using new data on individual transactions from a large Bordeaux-based broker (containing information on contract dates, prices and quantities, and characteristics on producers and wines) that we matched with the records of eleven important wine competitions (winners by medal color, and contest features), this paper addresses three questions. First, what is the causal impact of medals on wine producers' prices? By answering this question we formally analyze whether the above claims match the empirical findings. Identifying the causal impact of awards is challenging because there are potentially unobserved quality determinants that affect both prices and the probability to win medals. A regression of the wine price on a medal dummy (indicating whether the wine has obtained a medal prior to the transaction) would then lead to an estimate confounding the true medal effect and the effect of unobserved quality. To circumvent this omitted-variable bias, we exploit an unusual feature in our data: among the prize-winning wines in the sample, about 19% received a medal *after* the transaction. The idea is now to regress the price not just on the before-transaction medal dummy, but also on a post-transaction medal dummy. It turns out that we can consistently estimate the causal impact by taking the difference in the two dummy estimates. Two relatively weak restrictions are required to obtain this identification result. One is that the post-transaction dummy must be irrelevant for explaining the expected price of wine, once we have controlled for unobserved quality, the before-transaction dummy, and possibly other control variables. Using the terminology of Wooldridge (2002), the former dummy is thus assumed to be redundant in the structural price equation. The other restriction needed is that in the projection of quality on the medal indicators, the corresponding two projection coefficients should be equal. Loosely speaking, we assume here that the quality of a wine is the same regardless of whether it receives a prize before or after the transaction.

Second, what are the expected profits that wine producers get from participating in wine competitions? Addressing this question requires the calculation of expected costs and benefits. The former are obtained using available information on the participation fees charged by competitions, the price of medal stickers, and the costs of transporting wine samples from Bordeaux to the contest venue; the latter are obtained using observed prices, transaction volumes, our estimates of the causal impact of medals, and different values for the probabilities of winning medals (we take both small and large values, and the empirical proportions of wines awarded in each contest). The contests in our sample are quite heterogeneous. Some of them are state-owned, while others are privately run ones, and they differ in prestige, the number of participants they attract, the entry fees, the proportion of wines being awarded, and the manner in which their juries evaluate wines. It is therefore of particular interest here to show our profit calculations separately for the different competitions.

Third, are juries making efficient choices in attributing medal awards? We answer this question simply by estimating the coefficients on the post-transaction medal dummies (to account for the diversity of the competitions described above, we include in the model a dummy for each contest). Under our identification restrictions, these coefficients can be interpreted as the partial correlation between quality and the medal dummies. Checking whether the judges of a given competition make decisions that are efficient and informative amounts then to testing whether the corresponding medal indicator is statistically significant.

[1]See https://www.lenouveleconomiste.fr/lesdossiers/les-concours-14338 (downloaded May 2017).

The empirical literature on certification and quality disclosure has so far paid little attention to the price effect of awards. It has instead primarily focused on whether disclosure modifies the behavior of consumers and producers (see the survey by Dranove and Zhe Jin 2010). We are aware of only a couple of papers that look at the impact of certification effects on prices. Wimmer and Chezum (2003) compare auctions of certified and non-certified race horses and find that the former are sold at higher prices. Dewan and Hsu (2004) study stamp auctions and document that buyer prices at eBay are lower than at a specialty stamps auction (where there is lower quality uncertainty). Lima (2006) finds that wines are more expensive when they have received medals from Californian tasting events. He does not, however, account for the possible endogeneity of medal indicators.

Two closely related papers, Hadj Ali et al. (2008) and Dubois and Nauges (2010), look at the effect of grades assigned by Parker on Bordeaux wine prices. To correct for the omitted-variable bias the first paper takes advantage of a natural experiment: in one year the critic did not evaluate the wines and producers had to set prices without knowledge of his opinion. The second paper tackles the problem differently by assuming that unobserved quality is a polynomial of observed scores. Grading by wine critics differs from contest certification in the sense that the decision to evaluate a given good is taken by the experts and not the producers, and it does not entail any costs for them.

Our paper also contributes to a literature documenting that decisions taken by juries and evaluation committees are frequently influenced by factors unrelated to the quality of the objects being evaluated. Ginsburgh and Van Ours (2003) show that the random order in which pianists perform at the Queen Elisabeth competition affects their ranking. Redelmeier and Baxter (2009) find that students have a lower chance of getting admitted at the university of Toronto's medical school when interviews take place on rainy days. According to Goldin and Rouse (2000), the likelihood that female musicians get hired by symphony orchestras increases when juries use screens to conceal the gender of candidates. Our paper is also related to a series of articles showing that even highly experienced connaisseurs have difficulties in identifying and detecting the high-quality products under double-blind conditions. Fritz et al. (2012) find that professional violonists prefer new-technology violins over instruments by Stradivari and Guarneri del Gesu. Hodgson (2008) organized an experiment at a Californian wine competition in which judges had to evaluate fligths containing replicates of exactly the same wine. Only a small minority of judges were able to assign the same medal to the otherwise identical wines. Unlike these papers, we only offer suggestive evidence of the inefficiency of jury choices, through the estimation of the post-transaction medal coefficients.

In Section 3.2 we briefly describe the Bordeaux wine market and the organization of the different contests. We also explain there what are the possible reasons for observing post-transaction medals in our data. Section 3.3 contains a descriptive analysis of our data. Section 3.4 describes our estimation method and in particular our identification strategy. Section 3.5 present the results, and Section 3.6 concludes.

3.2 Institutional setting

In Section 3.2.1 we briefly present the organization of the Bordeaux wine market and the role played by brokers. In Section 3.2.2 we describe how wine contests are organized, focusing on the eleven competition from which we retrieved the medal information. Section 3.2.3 explains why it is possible that post-transaction medals are observed in the data.

3.2.1 The Bordeaux wine market and the role of brokers

Nowadays there are roughly 7,000 individual wine producers in the Bordeaux region, including two or three hundred very prestigious and internationally acclaimed châteaux (retail price of more than 50 €per bottle), and a large majority of lesser known wine-makers. Most of these producers sell their wines not directly to retailers but to local wine wholesalers called *négociants*, of which there are currently about 300 in Bordeaux. The transactions between the producers and *négociants* are typically handled by brokery offices (there are approximately 80 of them). A wine broker is a middleman who facilitates the matching between producers and *négociants*. Contrary to the latter, brokers maintain a close relationship with the producers, by regularly paying visits to the wine estates and giving advice on all aspects of wine production. While a producer treats in most cases with two or three brokers, each broker deals with hundreds of different producers and *négociants*. The everyday job of a broker is to collect the demands of the *négociants*, each demand referring to a more or less specific quality, volume and price, and to find a suitable lot within his portfolio of producers. When a broker finds a lot that possibly meets a demand, he delivers a sample of the wine to the *négociant* for tasting. If quality is satisfactory, the precise terms of the transaction are negotiated by the broker separately with the producer and the wholesaler, the main issues being the price, the quantity and the delay before the wine is available and can be delivered. Based on a historical consensus, brokers are usually remunerated at 2% of the value of each transaction they conclude. Our transaction data come from one of the largest Bordeaux-based brokers. The volume of wine traded by this broker represents about 25% of the total volume handled by all Bordeaux-based brokers, and around 15% of the annual production in Bordeaux.

Given the large number of suppliers, the Bordeaux wine market is very atomistic and competition is fierce, especially among the lesser known producers. Unlike the prestigious châteaux owners, they have few possibilities to alleviate the effects of this fierce competition and to differentiate themselves from their direct competitors.[2] One way to strengthen their market position is to join a cooperative winery.[3] This allows them not only to acquire more bargaining power vis à vis the brokers and the *négociants*, but also to share various fixed costs (e.g., the costs of harvesting machines) with other members of the cooperative. The wines are marketed under their own château names, but sales are coordinated and managed by the cooperation. The annual sale revenue is shared among the adherents depending on the quality and quantity of wine each one brought to the pool. This cooperative system offers numerous producers a form of protection while remaining somehow independent from each other. As mentioned in the introduction, the primary way for the less known wine makers to increase their market shares is to participate in wine contests and win medals.

[2]Since the early 1990s French laws severely limit all forms of publicity for alcohol products. Wine producers belonging to the top-end quality segment benefit, however, from several types of indirect publicity. Many of them are classified (e.g., according to the 1855 classification of Médoc wines, or to the 1955 classification of Saint-Emilion wines), and the rankings are mentioned on the bottle labels. Furthermore, these high-flyers are actively traded at auctions throughout the world, and get extensive media coverage from influential wine experts who taste and grade their wines. In contrast, the less known châteaux have few opportunities to advertise their products: their labels are less informative (typically only the producer name and the appellation are mentioned), and these wines are neither sold at auctions nor evaluated by the influential experts. At best some of them get mentioned and recommended in wine guides.

[3]In 2016, about half of the producers took part in one of the 36 existing cooperative wineries.

3.2.2 Wine competitions

About 130 official wine competitions are held annually in France.[4] They are organized in early spring, allowing producers to vinify the wines of their latest harvest, and participate in the competitions soon thereafter. For historical reasons, many of these contests focus exclusively on wines from a specific region of France. For instance, the *Concours de Bordeaux* is only devoted to Bordeaux wines, and the *Concours des Ligers* only to wines of the Loire region. Other competitions, such as the *Concours Général Agricole*, are nation-wide and open to wines from the whole of France. Finally there are international contests open also to non-French wines, such as the *Concours International de Lyon*.

Interestingly, the wine contests in France differ in many other respects as well. There is first of all variation in terms of the jury compositions. Most of the juries recruited in the French competitions are entirely made up of wine professionals (sommeliers, winemakers, oenologists, or *négociants*), but some contests deliberately choose to include amateur tasters as well. It is argued by the latter contests that amateurs have judgments which better reflect the tastes of everyday consumers, and that they are less prone to conflicts of interest than professionals. The contests also differ in the number of wines that each judge has to evaluate per day. This is an important issue because the accuracy of a judge is likely to decline with the number of wines that have to be tasted in a given amount of time. This is especially true if the judge is an amateur, which is perhaps why in general amateurs have less wines to taste than professional judges.

Yet another feature that distinguishes the competitions is their degree of selectivity, as measured by the share of wines that get awarded, and the proportion of medals attributed to each type of medal. French regulations prohibit contests to award more than 33% of the participating wines. Some competitions stick closely to this limit but others are more selective. The share of each type of medal also varies across competitions: some attribute for instance relatively few gold medals, while others completely ban bronze medals. Finally, the competitions vary in terms of the costs that have to be incurred by participants (participation fee, price of the medal stickers[5]), the selection of the samples,[6] and the procedures adopted by the juries to award wines. Regarding this last feature, although basically all competitions evaluate the wines in the same manner,[7] there is variation in the way judges choose winners. After evaluating the wines within a given flight, either all judges deliberate and agree orally on the laureates (attribution of awards by consensus), or they make their decisions based on the numerical grades assigned by each judge on a tasting grid (attribution by scoring).[8]

For this paper we have collected data from eleven wine competitions. Nine of them are organized in France, and two abroad. These contests are arguably the most important contests where Bordeaux wines are allowed to compete, and taken together they are responsible for about 90% of the medals that these wines win each year. The eleven competitions (abbreviations in parentheses) are: the *Concours de Bordeaux* (BOR), a regional contest devoted exclusively to Bordeaux wines; the *Concours Mondial de Bruxelles* (BRU), a Belgian international contest held

[4]Since 2000 about three new French contests have been launched each year, indicating that this is a profitable business.

[5]Medal winners who wish to disseminate this information to consumers have to pay the stickers that are put on the bottles.

[6]The samples are either chosen and sent directly by the producers, or the competition officials go to the châteaux themselves and pick the samples there. In the latter case the possibility of any manipulation of the samples is reduced.

[7]Seated at a table, the judges of a jury are served with flights of up to a dozen wines each. To the extent possible, the wines within a flight are of the same vintage and region, and the products are blind-tasted (except for the vintage and region the judges know nothing of the wines).

[8]All wines with an average score above a certain threshold get a medal, and the higher the score the better the medal. The thresholds are mostly determined at the end of the competition and enforce the 33% rule.

each year in a different country; the *Challenge International du Vin* (CHA), an international contest held in the Bordeaux region; the *Concours des Vignerons Indépendants de France* (CVI), a nation-wide contest only for individual and independent winemakers; the *Decanter World Wine Award* (DEC), a recent but large international competition organized in London by the Decanter magazine; the *Concours Mondial des Feminalise* (FEM), a recent contest that went international in 2015 and where all judges are women; the *Concours International de Lyon* (LYO), a recent international contest held in Lyon; the *Concours des Grands Vins de France à Mâcon* (MAC), an old national contest held in Mâcon; the *Concours Général Agricole* (PAR), the oldest and largest French wine contest, held in Paris; the *Vinalies Nationales* (VIN), a national contest where all judges are professional oenologists; the *Vinalies Internationales* (VII), the international counterpart of VIN.

Table 3.1 gives more details about these competitions (figures prevailing in 2016). Row 1 lists the year of creation of each contest. The most recently created ones are DEC[9], FEM, and LYO (about 10 years ago), while BOR and PAR are the two oldest ones, founded in respectively 1956 and 1870. Row 2 gives the scope of each competition. Six of them (including the two foreign competitions, BRU and DEC) are international and accept wines from all countries, four only accept French wines, and one only accepts wines from the Bordeaux region (BOR). Row 3 indicates whether the medals are granted by oral consensus or by scoring. BOR, CVI and PAR attribute the medals by consensus, and the rest of the competitions use a scoring process. Row 4 shows how the contest officials select the samples. BOR and PAR pick the samples directly in the tanks or barrels of the producers, and the other competitions have the samples sent directly by the producers.[10]

The number of wines evaluated in 2016 is given in Row 5. It varies from approximately 3,000 for VIN and VII to more than 16,000 for PAR. Row 6 gives the total share of awarded wines in 2016. All nine competitions held in France respect the 33% restriction: PAR is the most selective contest (24% of wines are awarded), and FEM the least (33%). For the two foreign competitions (recall that they are not concerned by this French regulation), BRU and DEC, the fractions are 30% and 59% respectively. The shares of each type of medal are listed in rows 7, 8, and 9. We see that BOR, FEM and LYO award relatively many gold medals (between 12% and 14% of the wines competing in these contests get gold), while DEC, VIN and VII are the ones that give few (between 3% and 7%). Three contests, BRU, LYO, and VII, give no bronze medals at all, while DEC attributes "commended" or bronze medals to almost 60% of its participating wines. Finally, BRU and VII are the most generous with silver (respectively 19% and 22% of their wines get this medal).

Row 10 indicates whether the jury is composed of wine professionals only (pro.), amateurs only (amat.), a combination of the two (mix.), or professional oenologists only (oen.). The juries of five competitions (CHA, FEM, LYO, MAC, and PAR) are mixed, and the juries of BOR, BRU and MAC are completely made up of professional judges; the juries of CVI are exclusively composed of amateurs, while those of VIN and VII only contain oenologists. Row 11 shows that the number of judges per competition ranges between 75 (VIN) and 3,227 judges (PAR), and row 12 that the contests in our sample lasted between 1 and 5 days. Row 13 gives, for each contest, the number of participating wines, divided by the number of judges times the number of days. Although this ratio does not exactly measure the number of wines tasted per judge on a given day (since each wine is typically evaluated by several judges),[11] it is a good measure

[9] The wine competition only exists since 2006, but the magazine was launched in 1975 and has since gained credit for its evaluation of fine wines. This detail will be important in the interpretation of the results in section 3.5.2.

[10] BOR and PAR are state-owned competitions, so that it is easier for them to find agents to visit the producers and collect the samples. The other competitions are organized by private firms or associations.

[11] The number of judges tasting each wine varies across contests (and even within contests) and is unknown in the data. Taking 4 judges as (a reasonable) estimate, the ratio for CHA would imply that each judge in this contest tastes 12 wines per day.

TABLE 3.1 – Description of the eleven wine contests (figures for 2016)

	BOR	BRU	CHA	CVI	DEC	FEM	LYO	MAC	PAR	VIN	VII
Year of creation	1956	1994	1976	1990	2006	2007	2010	1954	1870	1982	1995
Scope	Region	World	World	France	World	World	World	France	France	France	World
Consensus or Scoring	C	S	S	C	S	S	S	S	C	S	S
Samples: Picked or Sent	P	S	S	S	S	S	S	S	P	S	S
# wines	3,804	8,570	4,162	5,904	15,869	3,817	5,800^a	10,000^a	16,754	3,050	3,500^a
% medals	30%	30%	31%	25%	59%^{m}	33%	30%	30%	24%	27%	29%
% gold	12%	10%^{m}	9%	9%	3%	13%	14%^{m}	10%	10%	7%	7%
% silver	13%	19%	10%	10%	16%	11%	16%	7%	10%	12%	22%
% bronze	6%	0%	12%	7%	39%^{m}	9%	0%	13%	4%	8%	0%
Jury composition	Pro.	Pro.	Mix.	Amat.	Pro.	Mix.	Mix.	Mix.	Mix.	Oen.	Oen.
# judges	1,000^a	320	704	2,200^a	133	700^a	600^a	2,080	3,227^f	75^f	133
# days	1	4	2	5	1	1	1	1	2^f	2^f	5
# wines/(# judges $\times$ # days)	3.8	6.7	3	2.7	13.2	5.5	9.7	4.8	1.6^f	6.1^f	5.3
Participation fee (€)*	70.8$^+$	150-138	93-73	51.2	161	37.5	37	57.5	86-69	60	135-125
Sticker price (€/1,000)**	25-20	35-22	27-21	20-13	70-35	56-42	30-14	20	23-19	33.9	48

^{a}: Approximate statistic.
^{m}: The few grand gold medals awarded by BRU and LYO (in 2012 and 2013) have been merged with the gold medals. The medals with the mention "commended", awarded by DEC, have been merged with the bronze medals.
$^+$: The medal winners must also pay an additional charge: 0.6/0.4/0.25 €/100L for a gold/silver/bronze medal.
^{f}: PAR and VIN organize regional playoffs where 60% of participants were preselected for PAR and 30% for VIN. The figures are only for the national final.
*: If there are two entries the fees decrease with the number of wine samples sent by the producer, varying between high amount (first sample) and low amount (each additional sample).
**If there are two entries the marginal cost per 1,000 stickers depends on the quantity of stickers ordered, varying between the low amount and the high amount.

of the difficulty of the task faced by judges. The ratio is smallest for PAR, and largest for DEC. Finally, the last two rows give the participation fees and the prices for the medal stickers of each competition. Both figures are reported before taxes. The entry fees are not that high and range between 37 €(LYO) and 161 €(DEC). The cost of 1,000 stickers varies between 20 (CVI, MAC) and 57 €(DEC). [12]

3.2.3 Rationale for post-transaction medals

Before turning to the descriptive analysis of the data we wish to explain why, for a substantial fraction of wines in the sample, medals are attributed *after* the transactions. This feature of the data plays an important role in our identification strategy, but may seem somewhat surprising and counterintuitive at first sight. Indeed, it is not clear what are the incentives for producers to participate in contests after having sold their wine. There are four possible reasons for the phenomenon. First, wine makers typically do net sell their total production in one shot, through one broker, but mostly via multiple brokers. It can then be rational for a wine maker to sell part of the production soon after the harvest (e.g., because cash is urgently needed), say in January, participate in the competitions in spring, and sell the rest once the contest outcomes are known, say in July. Assuming that the January sale was negotiated by the broker that shared its data with us, and that in addition a a medal was obtained, this wine maker would appear in our sample as having sold its wine before obtaining an award. Second, even for wine makers who sell their total output before the contests, it may be of interest to participate in contests not to win medals but to get feedback about the quality of their latest vintage (think of producers having introduced new vinification techniques). Third, *négociants* have the right to enter wine competitions with lots they have bought from the producers (some competitions forbid this practice), and, here again, this results in the latter showing up as receiving medals after the transactions take place. Fourth, a small fraction of the transactions take the form of written contracts between producers and *négociants*, stipulating that the latter pay a price-markup to the former in case medals are awarded between the transaction date and the date of delivery/payment.[13] Such contingent contracts allow producers to sell their wines early in the season but nonetheless earn extra income in case they win prizes later on.

3.3 Descriptive Statistics

We have collected the records of our eleven wine contests for the years 2006 to 2016. For each contest and year we observe the date of the competition, the identities of all winners (i.e., the names of the châteaux and the names of the wine producers),[14] the color (bronze, silver, gold) of the medal received by each winner, and some additional competition characteristics (described in Table 3.1). The transaction data set made available to us by the broker covers the period 2005-2016. The broker excluded from the data all transactions regarding the elite châteaux mentioned in the previous section. Since these producers never participate in wine competitions, it is not problematic that they are discarded from the analysis. For each transaction we

[12]Some contests charge entry fees that decrease with the number of wine samples sent by the producer. BRU, for example, asks between 150 (first sample) and 138 €(each additional sample). Similarly, sticker prices may vary with the quantity demanded. If multiple entries are given in the table, it means that the marginal cost of 1,000 stickers varies between the lower and the higher amount.

[13]The average delay between the signature of the contract and the date of delivery is about 100 days. Payment is due 60 days after delivery.

[14]Unfortunately we have no information on the contest losers.

observe the exact transaction date, the volume of wine sold, the price of this volume, the vintage, the appellation, the type of packaging (bottled, bulk, or bottled when collected (BWC[15])), and the type of producer (individual wine maker, or wine maker belonging to a cooperative winery). From the initial sample we only kept the transactions corresponding to the 2005-2014 vintages, i.e., we excluded wines from 2015 and 2016, and those from 2004 and earlier.[16] We then matched the transaction and medal data sets on the identities of the wines, resulting in a sample of 16,399 observations.

Table 3.2 contains descriptive statistics on some of the main variables in our data set. The average price per 0.75 liters (the quantity contained in a standard bottle)[17] is 2.24 €, the minimum (resp. maximum) price is 0.05 €(40 €); 99% of prices are below 8.6 €/0.75L, and 90% below 4.9 €/0.75L, confirming that we are dealing here with the low-price segment wines. We emphasize that these prices are the ones paid by the *négociants*, final consumers pay about 30 to 40 % more at retailers. The quantities sold through the broker are substantial: on average, producers sell almost 50,000 liters. Among wines which received at least one medal prior to the transaction, the average duration between the moment the medal is awarded and the transaction is almost 14 months (if multiple medals are attributed, we pick the one such that duration between these two moments is shortest). Among those awarded at least once after the transaction, the average duration separating the transaction and award is almost 8 months (in case of multiple awards we pick again the one such that the duration is shortest).

TABLE 3.2 – Descriptive statistics

Variable	Mean	Sd. error	Min	Max
Price (€/0.75L)	2.24	1.98	0.05	40
Volume (1,000L)	48.58	66.69	0.01	1,155.56
Delay between prior medal and transaction* (months)	13.92	14.63	0	89.9
Delay between transaction and future medal* (months)	7.97	9.28	0.03	103.11
Age (months)	18.86	17.11	0	129
Vintage	2009.6	2.76	2005	2014
Delay between transaction and delivery (months)	3.11	3.35	0	37.06
Type seller: Cooperative winery	0.17	0.37	0	1
Type seller: Individual wine maker	0.83	0.37	0	1
Type packaging: Bottled	0.24	0.42	0	1
Type packaging: Bulk	0.62	0.49	0	1
Type packaging: BWC	0.14	0.35	0	1
N		16,399		

*: If the wine obtained several medals before or after the transaction, we consider the medal for which the award date is closest to the transaction date.

The remaining variables in Table 3.2 act as our control variables in the empirical analysis.[18] The wine's age (month of transaction minus September of vintage year) is around 19 months on

[15]In Bordeaux, wines are not delivered by the producers but collected by the *négociants*. When a wine is not sold in bottles, either the *négociants* come with a truck and pump up the wine from the producers' reservoirs (bulk), or bottle the wines directly at the château using bottling trucks (BWC).

[16]Each year the eleven contests attribute prizes primarily to wines of the two latest vintages (for example, in 2012, BOR awarded 87% of its prizes to the 2010 and 2011 vintages). Given that our medal data base covers the years 2006-2016, this explains why it is sufficient to drop among the recent vintages only those from 2015 and 2016. Similarly, it explains why we need to exclude all wines from 2004 and earlier.

[17]This price is calculated as the ratio of total amount paid (in euros) and volume (in liters) times 0.75.

[18]Our controls also include appellation dummies, but since there are more than 50 of them the descriptive statistics are omitted.

average, with a minimum of 0 months (corresponding to a wine sold during the harvest month) and a maximum of 129 months (almost 11 years). As explained above, the transactions in our data are chosen such that all wines are from the 2005-2014 vintages. Producers deliver their wines quickly after the transaction: on average the *négociants* receive the products slightly more than 3 months after signature of the contract. The large majority of wines (83%) are produced by individual wine makers, while the remaining 17% are made by wine makers who have joined a cooperative. The last three lines indicate the type of packaging: 24% of wines are already bottled at the transaction date, 62% are sold in bulk, and 14% are BWC.

Table C.1 in the Appendix C gives a cross-tabulation of the number of medals awarded before and after the transaction. We see that 13,298 wines in the sample have not won a medal at all in the eleven competitions. Among the 3,101 prize-winning wines (16,399-13,298), 2,711 got at least one medal before the transaction,[19] while 587 got at least one medal after the transaction. Note that there are wines that received multiple awards: for instance, 612 wines got two medals before they being sold, and 102 wines got awarded twice after the transaction date. Finally, there is a small number of wines that got prizes both before *and* after the transaction date (for example, 129 got one medal before and one after the date of sale).

Table C.2 in the Appendix C lists, for each contest, the total number of medals awarded to the wines in our sample, together with the number of awards separately for gold, silver, and bronze. We distinguish medals given before the transaction from those given thereafter. BOR is by far the competition that awarded the highest number of medals: between 2006 and 2016 it attributed a prize to 1,119 wines before they were sold, and to 178 wines after they were sold. Other competitions with many awards are MAC (735 medals before and 112 after the transaction) and PAR (727 and 69). VII is the contest which awarded the least number of medals during the observation period (30 and 11). Note that the fraction of medals attributed to the three colors is quite similar to the aggregate medal proportions reported in Table 3.1.

TABLE 3.3 – Average price by number and type of medals before/after transaction

Timing	Characteristic	Number of medals			Type of medal		
		0	1	2+	Bronze*	Silver	Gold
Before the transaction	Average Price (€/0.75L)	2.05	2.99	3.6	3.58	3.21	3.21
	Frequency	13,688	1,688	1,023	1,109	1,239	1,312
After the transaction	Average Price (€/0.75L)	2.21	3.05	3.67	3.43	3.18	3.19
	Frequency	15,812	449	138	232	260	204

*: "Commended" medals given by DEC have been merged with bronze medals.

Table 3.3 gives average prices and frequencies by number (columns 1-3) and type (columns 4-6) of medals received. The statistics are reported separately for wines sold before and after the transaction. Among the 14,212 wines which did not receive a medal before the transaction, the average price is 2.05 €/0.75L. Among the 1,688 wines with exactly one award before the transaction, the average is 2.99 €(an increase of 46%), and among the 1,023 wines with 2 awards or more 3.6 €(76%). The average price for the 15,812 wines without post-transaction prize is 2.21 €. Note that this subsample includes 2,514 wines having received a prize before

[19]The transaction data set also contains information on past medal awards. The broker did not systematically and exhaustively record this information in its archives: for 939 observations (out of 2,711) only the contest data set indicates that medals have been awarded. However, for 261 observations only the transaction data set indicates past medal awards (this concerns essentially wines awarded at MAC, a contest that only releases the producers' names of the winners (not the châteaux names), rendering matching more difficult). Our estimation results are not qualitatively different when the 261 wines are treated as if they have won no medals before the transaction date.

the transaction (see Table C.1), explaining why these wines are a bit more expensive (relatively to wines without prizes before the date of sale). Among the 449 wines with exactly 1 medal after the transaction, the average price is 3.05 €, and among the 128 wines with 2 medals or more, 3.67 €. Looking at the statistics by type of medal, we see that for the 232 post-transaction winners with at least one bronze medal the average price is 3.43 €. Surprisingly, the average price for producers winning at least one silver (resp. gold) medal is 3.18 €(resp. 3.19 €). We cannot reject, however, that mean prices differ in a statistically significant manner across the three colors. The figures are similar for producers winning prizes before the date of sale. However, the average for bronze (3.43 €) is now significantly larger than for silver and gold (both 3.21 €). In Section 3.5 we will see that this counterintuitive result disappears once we control for additional wine characteristics.

3.4 Estimation strategy

In this section we present our estimator for the causal impact of medals on prices. It is convenient to start the presentation by assuming that there are no other observed price determinants besides the medals. We thus exclude, for the moment, that variables such as the age of the wine, its appellation, or its packaging, are observed. We also assume that there exists just one type of medal and only one competition, i.e., we ignore for the moment that medals come in different colors (bronze, silver, gold), and that in practice several wine competitions coexist. Finally it is assumed that a given wine can only win a single medal before the transaction date, and/or a single medal in the future. The possibility that multiple medals of different types can be awarded will be accounted for later on.

Let the price P be generated by the following model:

$$ln(P) = \alpha_0 + \alpha_M M + Q + \epsilon = \alpha_0 + \alpha_M M + \xi \tag{3.1}$$

where M is a binary variable equal to 1 if the wine has obtained a medal before the transaction date and 0 otherwise, Q represents unobserved quality of the wine, ϵ is an error term capturing the effect of other unobserved price determinants, and $\xi = Q + \epsilon$. The parameters α_0 and α_M represent the intercept and the causal effect of the medal, respectively. Let F be a binary variable equal to 1 if the wine will get a medal somewhere after the transaction and 0 otherwise. We assume that the error term ϵ is mean-independent of M, Q, and F: $E(\epsilon|M, F, Q) = 0$. Without loss of generality it is furthermore assumed that $E(Q) = E(\epsilon) = 0$. Note that quality Q is defined in such a way that the coefficient associated with this variable is normalized to one. Note also that P is assumed to be determined only by the before-transaction medal indicator M and Q, i.e., the post-transaction medal indicator F does not affect price. To the extent that F is by definition unknown at the time of transaction, it seems natural to exclude this variable from the structural model (3.1). Note finally that our model structure is similar to the one adopted by Dubois and Nauges (2010), except that they do not observe the equivalent of the dummy F.

Let $\hat{\alpha}_M^{OLS}$ denote the OLS estimator of α_M. Since M and Q are potentially positively correlated, we expect that the probability limit of $\hat{\alpha}_M^{OLS}$ exceeds α_M. The OLS estimator is only consistent under the additional assumption that the medal indicator and unobserved quality are uncorrelated. Although this assumption is unlikely to hold in practice, we nonetheless report OLS estimates in the next section, but merely as benchmark results, which will be compared with the results produced by our estimator.

To define our estimator, we consider the linear projection of Q on F and M (see for example Wooldridge 2002 for the definition and properties of linear projections):

$$Q = \beta_0 + \beta_M M + \beta_F F + \mu \tag{3.2}$$

where β_0, β_M, and β_F are the linear prediction coefficients. The error term μ satisfies, by definition of a linear projection, $cov(M, \mu) = cov(F, \mu) = 0$. Replacing Q in equation (3.1) by (3.2) gives:[20]

$$ln(P) = (\alpha_0 + \beta_0) + (\alpha_M + \beta_M)M + \beta_F F + \epsilon + \mu. \quad (3.3)$$

Since the composite error term $\epsilon + \mu$ is uncorrelated with both M and F, the OLS estimators $(\widehat{\alpha_M + \beta_M})$ and $\hat{\beta}_F$ consistently estimate $(\alpha_M + \beta_M)$ and β_F. Under the identifying restriction $\beta_M = \beta_F$, the difference in OLS estimators is thus a consistent estimator of the causal effect α_M. This estimator is denoted $\hat{\alpha}_M^{DIF}$ (the superscript *DIF* to indicate that it is based on a difference in two estimators) and is defined as

$$\hat{\alpha}_M^{DIF} = \widehat{\alpha_M + \beta_M} - \hat{\beta}_F.$$

Note that our estimator does not require M and Q to be uncorrelated (the identifying restriction underlying the OLS estimator).[21] Instead, we need to impose the more natural and plausible restriction that the partial correlation between M and Q equals the partial correlation between F and Q.[22]

Let us now turn to the more general case where wines can win multiple medals, of different colors, and possibly from different contests. We now also account for the possibility that prices can be influenced by a vector of observable characteristics, denoted X. The analogue of the price equation (3.1) becomes

$$ln(P) = \alpha_0 + \sum_{j=1}^{J} \alpha_{M_j} M_j + \alpha_X X + Q + \epsilon \quad (3.4)$$

and the linear projection (3.2) becomes

$$Q = \beta_0 + \sum_{j=1}^{J} \beta_{M_j} M_j + \sum_{j=1}^{J} \beta_{F_j} F_j + \beta_X X + \mu \quad (3.5)$$

Here M_j equals 1 if the wine has obtained a medal of type j (i.e., of a certain color and from a specific competition) before the transaction, and 0 otherwise. The indicators F_j are similarly defined, and J represents the total number of different types of medals that can be awarded. All parameters have analogous interpretations as above. The error term μ is by definition of a projection uncorrelated with all past/future medal indicators, and with X, and has expectation equal to zero. The error terms in (3.4) are still assumed to be centered around zero: $E(Q) = E(\epsilon) = 0$. The error ϵ is now assumed to be mean-independent of all medal indicators, Q, and X: $E(\epsilon | X, Q, M_j, F_j, j = 1, ..., J) = 0$. Finally we assume that Q and X are uncorrelated: $E(Q|X) = 0$.

Estimation by OLS leads to inconsistent estimators for the same reason as previously: the indicators M_j are expected to be correlated with Q (capturing the impact of unobserved quality components after controlling for X and the J medal indicators). In particular the OLS estimators $\hat{\alpha}_{M_j}^{OLS}$ are thus likely to be inconsistent.

[20]The idea to replace Q by its projection on a set of regressors is reminiscent of Chamberlain's 1982 approach to unobserved effects models.

[21]Using (3.2) we have $cov(M, Q) = \beta_M var(M) + \beta_F cov(M, F)$. Under $\beta_M = \beta_F$, we get $cov(M, Q) = \beta_M cov(M, M + F)$, which generally differs from zero except when $\beta_M = 0$ and/or when the last covariance equals zero.

[22]The variable F is not what Wooldridge (2002) calls a proxy variable for the endogenous variable M. Although we assume that F is redundant in (3.1) (like a proxy variable), we only impose $\beta_M = \beta_F$ (while a proxy variable requires $\beta_M = 0$). F is not an instrumental variable for M either since it is correlated with Q.

To define the generalized version of our difference estimator, we substitute (3.5) in (3.4) and get

$$ln(P) = (\alpha_0 + \beta_0) + \sum_{j=1}^{J}(\alpha_{M_j} + \beta_{M_j})M_j + \sum_{j=1}^{J}\beta_{F_j}F_j + (\alpha_X + \beta_X)X + \epsilon + \mu.$$

Given our assumptions, the error term $\epsilon + \mu$ is uncorrelated with all regressors, and hence the OLS estimators of this regression model are consistent. As previously, the estimator is defined as the difference of OLS estimators: $\hat{\alpha}^{DIF}_{M_j} = \widehat{\alpha_{M_j} + \beta}_{M_j} - \hat{\beta}_{F_j}$. Under $\beta_{M_j} = \beta_{F_j}$, $j = 1, ..., J$, it is a consistent estimator of α_{M_j}. Note that the estimator of the coefficient associated with X does not allow for consistent estimation of the causal effect α_X.

An interesting byproduct of our method is that it also provides an estimate of β_{F_j} for all j. This coefficient measures the partial correlation between F_j and Q, and, given the identifying assumption, also the partial correlation between M_j and Q. Testing the hypothesis $\beta_{F_j} = 0$ then amounts to checking whether quality is uncorrelated with M_j, and testing $\beta_{F_j} > \beta_{F_{j'}} > 0$ is equivalent to checking whether the before-transaction medal indicator of type j is more strongly correlated with quality than the one of type j'.

If one is willing to make the additional assumption that μ is mean-independent of X and all medal indicators,[23] then the sum $\alpha_{M_j} + \beta_{M_j}$ has a nice interpretation. More precisely, under $H_\mu : E(\mu|X, M_j, F_j, \forall j) = 0$, we have:

$$\begin{aligned} E(\Delta ln(P)) &\equiv E(ln(P)|X, M_j = 1, M_{j'} = 0, \forall j' \neq j, F_j, \forall j) - E(ln(P)|X, M_j = 0, F_j, \forall j) \\ &= \alpha_{M_j} + E(Q|X, M_j = 1, M_{j'} = 0, \forall j' \neq j, F_j, \forall j) - E(Q|X, M_j = 0, F_j, \forall j) \\ &= \alpha_{M_j} + \beta_{M_j}. \end{aligned} \tag{3.6}$$

The expected (logarithmic) price gap between wines with a medal of type j and wines without any medal at all (conditional on X and all Fs), denoted $E(\Delta ln(P))$, can be decomposed as the sum of the causal effect of this medal, α_{M_j}, and the difference in quality between these two types of wines, β_{M_j}. Note also that our identifying restriction has a more transparent interpretation under H_μ: the expected wine quality is the same for wines receiving a medal of type j before and after the transaction.[24]

3.5 Empirical results

In Section 3.5.1 we start presenting aggregate estimation results, obtained under the assumption that medal effects are the same across the different medal colors and wine competitions. These initial results also rely implicitly on the hypothesis that winning two or more medals has the same impact as winning a single one. In Section 3.5.2 we relax these simplifying restrictions and allow for the possibility that wines can win multiple and different types of medals. This allows us to analyze how the impact of medals varies by color (bronze, silver, gold), and type of competition (prestige, participation fee, tasting method) at which they are awarded. Finally, Section 3.5.3 uses the estimated medal effects to calculate producers' expected profits from participating in a wine competition.

[23]Since (3.5) is a projection, μ is by construction centered around zero. But this error term is not necessarily mean-independent of the regressors.

[24]Given H_μ, the restriction $\beta_{M_j} = \beta_{F_j}$ is equivalent to $E(Q|X, M_j = 1, M_{j'} = 0, \forall j' \neq j, F_j = 0, \forall j) = E(Q|X, F_j = 1, F_{j'} = 0, \forall j' \neq j, M_j = 0, \forall j)$.

3.5.1 Aggregate results

All estimation results presented in this section are collected in Table 3.4. Column 1 reports the two estimates of α_M (using OLS and our alternative method), together with standard errors in parentheses, assuming that prices are generated by model (3.1). Here P is the price in €per 75 cl of wine, $M = 1$ if at least one medal has been awarded to the wine prior to the transaction date, and $M = 0$ otherwise. Note that the observed wine characteristics are not included in the model. We also report in column 1 the OLS estimates of $\alpha_M + \beta_M$, and β_F, i.e., the parameters associated with M and F in (3.3), where $F = 1$ if at least one medal will be awarded after the transaction date, and 0 otherwise. The estimate $\hat{\alpha}_M^{OLS}$ is significant at the 1% level, and suggests that a producer can get 52.4% more per bottle of wine when his product has won at least one medal before the transaction. The estimate $\hat{\alpha}_M^{DIF}$ is substantially smaller, and implies that the price-increase for medal winners is 19.3% (also significant at the 1% level). The OLS estimates $\widehat{\alpha_M + \beta_M}$ and $\hat{\beta}_F$ equal 0.512 and 0.319, respectively (both are strongly significant). Recall that the difference between the two corresponds to $\hat{\alpha}_M^{DIF}$. The R^2 in model (3.3) is 0.081.

TABLE 3.4 – Estimates of α_M

Estimate	(1)	(2)	(3)
$\hat{\alpha}_M^{OLS}$	0.524	0.192	0.173
	(0.014)	(0.007)	(0.007)
$\hat{\alpha}_M^{DIF}$	0.193	0.157	0.132
	(0.036)	(0.013)	(0.012)
$\widehat{\alpha_M + \beta_M}$	0.512	0.191	0.172
	(0.014)	(0.007)	(0.007)
$\hat{\beta}_F$	0.319	0.035	0.04
	(0.032)	(0.011)	(0.01)
Characteristics X	No	Yes	Yes
Fixed effects	No	No	Yes
N	16,399	16,399	16,399
R^2	0.081	0.904	0.924

Column 2 reports estimates when the wine/producer characteristics X are added to the model, i.e., the specification we consider now is $P = \alpha_0 + \alpha_M M + \alpha_X X + Q + \epsilon$, where M is defined as above. The variables included in X are: the age of wine at the transaction date (in months); the delay separating the transaction date and the delivery of the wine to the purchaser (in months); the producer type (a dummy indicating that the producer is an individual wine maker); the type of packaging (a dummy indicating that the wine is sold in bulk, and another one indicating that it is sold bottled); and 45 dummies indicating the appellation of each wine. Controlling for these characteristics leads to a substantial drop in the OLS estimate of α_M (it now equals 0.192); the DIF estimate remains relatively stable compared to column 1 (now 0.157). Both remain strongly significant. The estimates $\widehat{\alpha_M + \beta_M}$ and $\hat{\beta}_F$ (obtained from estimating by OLS the regression model (3.3) to which $(\alpha_X + \beta_X)$X is added) have also sharply dropped. Controlling for the characteristics of wine and producer substantially augments the R^2 (now 0.904). Column 3 lists the results when fixed effects for the transaction-year and vintage are added to the price equation. Controlling for these fixed effects reduces the magnitude of the two estimates of α_M estimates yet again, but the drop is modest compared to those reported in column 2. The estimate of $\alpha_M + \beta_M$ has slightly decreased, while the estimate of β_F has slightly increased. All estimates remain strongly significant, and the R^2 now equals 0.924.

Overall, the conclusion from Table 3.4 is that OLS produces larger estimates of α_M than our alternative method, especially when we do not control for the wine characteristics X and the fixed effects. This is consistent with our point of view that OLS overestimates the causal effect of medals while our method at least mitigates this bias. All estimates $\hat{\alpha}_M^{DIF}$ are significantly positive, implying that, ceteris paribus, a wine is more expensive when it is medaled. Even our most conservative estimate (in column 3) suggests that medal winners can augment prices by no less than 13%. Note that this estimate is between 10 and 15%, the interval of values within which the causal effect should lie according to wine magzine cited in the introduction. All OLS estimates of α_M are, however, above this interval. Since β_F is positive and significantly different from zero in all three specifications, quality and the dummy indicting a future medal award are positively correlated. Apart from column 1, $\hat{\beta}_F^{DIF}$ is much smaller than $\hat{\alpha}_M^{DIF}$. Recalling the decomposition formula (3.6), most of the expected price difference between medaled and non-medaled wines comes from the causal impact of the certification, not from the difference in quality of these wines. More precisely, taking the estimates of α_M and β_F reported in column 3, the expected price difference is 17.2%, of which 13.2 percentage points can be attributed to certification, and only 4 percentage points to quality heterogeneity.

3.5.2 Results by number of medals, color, and competition

In columns 1-3 of Table 3.5 we report estimation results for a price model which explicitly allows the medal effect to differ by the number of awards received. Specifically, we assume that prices are modeled according to (3.4), with $J = 3$ and three dummies, M_1, M_2, and M_{3+}. Here M_1 (resp. M_2) equals 1 if a wine has obtained exactly one medal (resp. two medals) prior to the transaction, and 0 otherwise; M_{3+} equals 1 if at least three medals are obtained before the transaction, and 0 otherwise. The variables F_1, F_2, and F_{3+} are defined analogously. We only report the results with X (defined as above) and fixed effects added to the specification.

TABLE 3.5 – Estimates of α_M by number and color of medals

Estimate	Number of medals			Color of the medal		
	M_1	M_2	M_{3+}	M_{gold}	M_{silver}	M_{bronze}
$\hat{\alpha}_{M_j}^{OLS}$	0.166	0.2	0.256	0.194	0.077	0.075
	(0.007)	(0.01)	(0.014)	(0.008)	(0.007)	(0.008)
$\hat{\alpha}_{M_j}^{DIF}$	0.123	0.124	0.245	0.13	0.044	0.042
	(0.013)	(0.027)	(0.038)	(0.018)	(0.017)	(0.016)
$\widehat{\alpha_{M_j} + \beta_{F_j}}$	0.163	0.201	0.257	0.194	0.077	0.076
	(0.007)	(0.01)	(0.014)	(0.008)	(0.007)	(0.008)
$\hat{\beta}_{F_j}$	0.04	0.077	0.012	0.063	0.032	0.035
	(0.011)	(0.025)	(0.036)	(0.017)	(0.015)	(0.014)
Characteristics X		Yes			Yes	
Fixed effects		Yes			Yes	
N		16,399			16,399	
R^2		0.925			0.925	

The OLS estimates of α_{M1} (coefficient associated with M_1), α_{M2} (M_2), and α_{M3+} (M_{3+}) exceed the DIF estimates of these parameters, again suggesting that the medal indicators are not exogenous, leading OLS to overestimate the causal effects. Our results show that it is relevant to let medal effects differ by the number of awards received: $\hat{\alpha}_{M1}^{DIF}$ and $\hat{\alpha}_{M2}^{DIF}$ are both around 0.12 (slightly smaller than $\hat{\alpha}_M^{DIF}$ in column 3 of Table 3.4); $\hat{\alpha}_{M3+}^{DIF}$ is 0.25 (substantially larger).

For each of the three coefficients we strongly reject the null hypothesis that they are equal to zero. Furthermore, the hypothesis $\alpha_{M2} = \alpha_{M3+}$ is rejected, but $\alpha_{M1} = \alpha_{M2}$ is not. The price markup is thus the same for wines having either one or two medals, but is significantly higher for those with at least 3 medals. The parameter β_{F2} is significantly larger than β_{F1} implying that the dummy indicating two future medals is, as expected, more strongly correlated with quality than the dummy indicating one future medal. Surprisingly, we cannot reject the null hypothesis that β_{F3+} equals zero, but this may be due to the small number of wines in the sample with three or more awards after the transaction (see Table C.1).

Columns 4-6 of Table 3.5 give estimation results for model (3.4) in which medal effects differ across color. The specification again includes $J = 3$ dummies, here defined as M_{Gold}, M_{Silver}, and M_{Bronze}, with $M_{Gold} = 1$ if a wine has won at least one gold medal in the past, and 0 otherwise, and M_{Silver} and M_{Bronze} defined analogously.The OLS estimates of α_{Mgold}, $\alpha_{Msilver}$, and $\alpha_{Mbronze}$ again exceed our alternative estimates. The latter imply that winning at least one gold medal allows the producer to augment its price by 13%; the price increases associated with silver and bronze are much smaller, at 4.4% and 4.2%, respectively. These estimates are each significantly different from zero, and we strongly reject the null hypothesis $\alpha_{Mgold} = \alpha_{Msilver} = \alpha_{Mbronze}$ (α_{Mgold} is significantly larger than $\alpha_{Msilver}$, but $\alpha_{Msilver} = \alpha_{Mbronze}$ cannot be rejected). The β_Fs are also significantly different from zero, but the hypothesis $\beta_{Fgold} = \beta_{Fsilver} = \beta_{Fbronze}$ cannot be rejected (p-value is 0.09).[25] Under H_μ the expected price gap between gold-medaled wines and non-medaled wines is 19.3%, of which 13 percentage points is due to certification and 6.3 percentage points to quality heterogeneity. The decompositions for silver and bronze are similar to each other: for both the gap is around 7.5%, with 4.5 points attributable to certification and 3.5 to quality differences. The different equality tests reported just above suggest that the larger price gap for gold is primarily due to a larger effect of certification, the effects of quality heterogeneity are statistically indistinguishable across the three types of medals and/or economically small.

Table 3.6 lists estimation results of model (3.4) allowing the impact of medals to vary across the different competitions. Hence the specification now includes $J = 11$ dummies, M_{BOR},..., M_{VII}, where, for instance, M_{BOR} equals 1 if the wine has won a medal at the contest of Bordeaux prior to the transaction, and 0 otherwise.[26] The Fs are defined analogously. A close look at the results reveals that three groups of competitions can be distinguished. A first group includes four competitions (BOR, CHA, DEC, PAR). For each of these contests the estimates of both α_M and β_F are significantly different from zero. The estimates $\hat{\alpha}_M^{DIF}$ range between 0.07 (for BOR, CHA, PAR) and 0.1 (DEC), and $\hat{\beta}_F$ between 0.04 (CHA,DEC) and 0.07 (PAR).

The second group is made up of two contests (CVI, MAC). The estimates of α_M are still significant, and of the same magnitude compared to those of the first group. Unlike the first group, however, we can no longer reject the null hypothesis $\beta_F = 0$. The third group contains the remaining five contests (BRU, FEM, LYO, VIN, VII). For BRU, LYO, VIN, and VII we can neither reject the null hypothesis $\alpha_M = 0$ nor the null $\beta_F = 0$ at conventional significance levels. For FEM we reject $\alpha_M = 0$, but the results are surprising and counterintuitive here as $\hat{\beta}_F$ is significantly negative.

Since the number of competitions in our data is limited it is not possible to formally show how contest characteristics relate to group membership. Therefore we cannot establish that say contests charging high entry fees have a statistically higher probability to belong to a particular group. What we can do, however, is to check which characteristics are shared by all (or most)

[25] Neither the hypothesis $\beta_{Fgold} = \beta_{Fsilver}$, nor $\beta_{Fsilver} = \beta_{Fbronze}$, can be rejected. However, β_{Fgold} is significantly larger than $\beta_{Fbronze}$.

[26] There are no wines in the sample with more than one medal awarded from the *same* competition. The medal dummies are therefore appropriately defined.

TABLE 3.6 – Estimates of α_M by competition

Estimate	M_{BOR}	M_{BRU}	M_{CHA}	M_{CVI}	M_{DEC}	M_{FEM}	M_{LYO}	M_{MAC}	M_{PAR}	M_{VIN}	M_{VII}
$\hat{\alpha}^{OLS}_{M_j}$	0.12	0.07	0.11	0.11	0.14	0.04	0.03	0.07	0.14	0.01	-0.06
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.05)
$\hat{\alpha}^{DIF}_{M_j}$	0.07	0.01	0.07	0.09	0.1	0.15	0	0.07	0.07	-0.1	-0.13
	(0.02)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.02)	(0.03)	(0.07)	(0.08)
$\widehat{\alpha_{M_j} + \beta_{M_j}}$	0.12	0.07	0.11	0.11	0.14	0.04	0.03	0.07	0.14	0.01	-0.07
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.05)
$\hat{\beta}_{F_j}$	0.05	0.06	0.04	0.02	0.04	-0.11	0.03	0	0.07	0.11	0.06
	(0.02)	(0.04)	(0.02)	(0.03)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.06)	(0.06)
X						Yes					
Fixed effects						Yes					
N						16,399					
R^2						0.924					

contests within each given group, and thereby determine, informally, a link between group membership and contest features.

A common feature of the contests in the first group is that they were founded a relatively long time ago. BOR, CHA and PAR are more than 40 years old, and DEC is named after a famous wine magazine launched in 1975[27]. In 40 years, these names have acquired a solid reputation. According to wine professionals, BOR and especially PAR are the most prestigious competitions. For wholesale traders in France the medals given in these two competitions are the most valuable and sought-after awards for Bordeaux wines. DEC is regarded by many as the most influential non-French wine contest in the world, while CHA is the best known international competition in France. Another common feature of the four contests is that their jury members have to evaluate relatively small number of wines on a given day (DEC is again an exception – see line 13 of Table 3.1). BOR and PAR are the only two contests where the samples are chosen and selected by the organizers themselves. Besides CVI, they are also the only ones whose judges grant medals by oral consensus.

The two contests in the second group have juries that are either fully made up of amateurs, or a mix of amateurs and professionals, and they charge the lowest entry fees and sticker prices. The five contests of the third group tend to attract the lowest number of participants (except BRU) and are, as indicated in Table 3.1, among the most recently founded competitions. The juries of VIN and VII are completely composed of oenologists, and three contests of this group (BRU, LYO, VII) do not award bronze medals, i.e., their award procedure is relatively coarse. Compared to the first group, jury members are required to evaluate more wines per day. This may diminish the accuracy of their judgments, which may in turn explain the non-significance of the quality indicator β_F for this group.

Table C.3 in the Appendix C, the last one discussed in this section, presents results of model (3.4) wherein medal effects are allowed to vary simultaneously by color and competition. It is difficult to precisely estimate the parameters now because for many medal/competition combinations the number of medal awards before and after the transactions is not sufficiently high (see Table C.2). We therefore only allow the two main French contests, BOR and PAR, to have specific medal-color effects, and for all remaining contests these effects are restricted to be the same (resulting in a specification with $J = 9$ dummies). The table only reports the results for BOR and PAR. Our alternative method produces an estimate of $\alpha_{M_{BORgold}}$ equal to 0.134, and

[27]The magazine Decanter is recognized for its review of fine wine. Consequently, its name somehow conveys the idea a high quality standard. This may have played in favor of the reputation of its wine competition, although it accepts all wines like the other wine competitions and grant much more medals than the French competitions.

the hypothesis that this parameter equals zero is strongly rejected. We thus find a causal effect of a gold medal from BOR which is much smaller than what is claimed by the organizers of this contest (30% – see the introduction), which might have overestimated their influence on producers prices. The estimate of $\alpha_{M_{PARgold}}$ implies that producers receiving a gold medal from PAR can raise their price by 13.5%. This is compatible with the contents of a contingent contract we got from the broker. This contract dates from year 2015 and concerns a sale of bulk wine of vintage 2014 from the Bordeaux appellation. It stipulates the following conditions: producer gets 1,300 €/900l if wine has received no medal or a bronze medal (regardless of the competition); 1,350 €/900l for a silver medal and 1,375 €/900l for a gold medal (both regardless of competition); 1,500 €/900l for gold from PAR.[28] The bonus for a gold medal from PAR amounts to a price increase of 15%, just above our estimate. The bonus for silver amounts to a price markup of 3.8%, again close to our estimate (4.4% – see Table 3.5). The bonuses for bronze and gold (0% and 5.7%) are, however, lower than our estimates of the respective causal effects (4% and 13%). The fact that the contract conditions and our estimates are (at least partly) in line may seem as natural, if one is willing to assume that contingent contracts have been used for wines that are representative of all wines in our sample. But it is nonetheless reassuring and gives credence to our identification strategy.

3.5.3 Producers' expected profits from participating in contests

In order to decide whether to compete in wine contests or not, producers need to compare the costs and benefits of contest participation. More precisely, this decision requires a calculation of the profit from participating in a competition. This profit is *ex ante* unknown to the producer since it depends on whether a medal will be won, and on the color of the medal. Producers can therefore only calculate the expected profit. Given that we cannot estimate sufficiently precisely the three medal-color effects separately for each contest, we shall for simplicity assume that the three types of medal have the same impact for each given contest. This amounts to saying that that there are just two states of the world: either a producer wins a medal at a competition, or wins no medal. The expected profit for producer i at a given competition is

$$E(Profit_i) = \pi V_i \left[P_i(e^{\alpha_M} - 1) - C_s\right] - C_0 \qquad (3.7)$$

where V_i is the quantity of wine i sold through the broker (measured as the number of bottles of 75 cl), P_i the price of wine i for 75 cl, C_s the cost per sticker, C_o other (fixed) costs of participating in a contest, and π the probability of winning a medal. In the results of Table C.4, we consider different values for the probabilities of winning a prize: 5%, 10%, 20% and the empirical share of awarded participants. The term $P_i(e^{\alpha_M} - 1)$ corresponds to the causal impact of the medal on the price of wine i (the expression is non-linear in α_M because the price in model (3.4) is defined in logarithms).

We have calculated $E(Profit_i)$ for all wines i in the sample[29], for each of the four contests belonging to the first group. We have taken the corresponding estimates of α_M reported in Table 3.6, and the sticker prices C_s reported in Table 3.1.[30] Other costs C_o are defined as the

[28] The broker from which we obtained the transaction data did not possess other examples of such contracts (not surprising given that they are handled and negotiated by the *négociants* and producers), but assured us that the contract conditions described in the text are representative and not atypical.

[29] For the transactions where the wine had a medal, we have divided the transaction price by $e^{\alpha_{Mj}}$. We excluded the 2,105 transactions for which V is below 1,000 liters. The calculations are thus based on 16,399-2,105=14,294 observations.

[30] For notational simplicity, the marginal cost per sticker is assumed constant in (3.7). However, in our calculation of $E(Profit)$ we allow the marginal cost to be a decreasing step function of V (see Section 3.2.2).

participation fee (also reported in Table 3.1), plus 60 €to reflect the costs of sending the samples from Bordeaux to London (DEC).

Table C.4 reports statistics on $E(Profit)$, separately for the four competitions, and different values of π: 0.05; 0.10; 0.20; and the empirical proportion of medaled wines as reported in Table 3.1. In the first three cases the results are comparable for BOR, CHA, and PAR (as expected, because all parameters determining expected profits are then similar for these contests), while those corresponding to DEC stand apart. When $\pi = 0.05$, the mean of $E(Profit)$ is positive and small in the case of BOR, CHA and PAR (around 50 €), but negative in the case of DEC (-43 €). The proportion of producers with negative expected profits is slightly higher than 50% for BOR, CHA, and PAR, and around 75% for DEC. Increasing the probability of winning a medal leads to a substantial improvement of these figures. When $\pi = 0.20$, for instance, the mean of $E(Profit)$ is around 500 €in the case of BOR, CHA, and PAR, and 570 €in the case of DEC; the fraction of producers getting negative expected profits is around 14% for the former three contests, and 36% for the latter. Replacing π by the empirical proportion of medaled participants medals in each competition (bottom panel of Table C.4), we see that the mean of $E(Profit)$ now ranges between 609 €(PAR) and 2,148 €(DEC), while the fraction of producers with negative expected profits is small, varying between 7.5% (CHA) and 16.7% (DEC). For the representative wine producer it seems thus highly attractive to participate in these wine contests.

Since a fraction of the producers do not pay the stickers themselves (they are paid instead by the *négociants*), Table C.4 presents the same statistics but corresponding to case where producers do not bear these costs (i.e., $C_s = 0$). Naturally this shifts expected profits upwards, and compared to the previous table the attractiveness of the four contests increases. For example, taking $\pi = 0.20$ implies that the mean of $E(Profit)$ ranges between 715 €(CHA) and 953 €(DEC), with the fraction of producers gaining from participation varying between 4.8% (BOR) and 14.5% (DEC). The bottom panel of the table clearly indicates that the representative producer should without any doubt send his wine to one of the four contests: DEC, for example, guarantees this wine maker an expected profit of 3,290 €, while the probability of losing money when competing at BOR is only 1.7%.

3.6 Conclusion

In this paper we first of all obtain the causal effect of medals on producers' wine prices. We adopt a novel but simple approach consisting in regressing prices on both before-transaction and post-transaction medal indicators. Under natural identifying restrictions, the difference in the estimates of the associated coefficients identifies the causal effect. Our preferred estimate indicates that a producer whose wine received a medal can augment his price by 13%. The impact for gold turns out to be much larger than for silver and bronze. When we allow the medal effect to differ across competitions, we find that only for a small group of contests there is a statistically significant effect. This group is made up of the most prestigious competitions that enjoy an old notoriety. Interestingly, their judges are required to evaluate relatively few wines per day, and they grant medals by oral consensus. Next we have calculated the profit producers may expect to get from participating in these competitions. We find that that the incentives to participate in competitions is high. Finally we contribute to a literature that sheds doubt on the reliability of juries and evaluation committees in all sorts of contexts. We find that only a minority of contests attribute medals that are significantly correlated with wine quality. Nonetheless, our estimate in the aggregate case reveals that wine competitions as a whole do identify wines of better quality, thus contrasting with the previous literature.

Chapter 4

Multivariate Forecasting of Prices of Bordeaux Wines

4.1 Introduction

Bordeaux is a flagship brand in the wine world, but especially small producers have had a hard time surviving in this competitive market. As a matter of fact, many of them have worked at loss after a price downturn in the early 2000s. So as to evaluate the severity of the situation, the local authorities in Bordeaux have evaluated the average cost of basic wines production (Bernaleau Cardinel et al., 2012). This cost has been found to be above the average price paid to the producers from 2003 to 2013 at the observed average yields. Partly as a result of this, the number of Bordeaux wine-producing firms has collapsed from 11,760 in 2000 to 6,568 in 2016[1]. Public protest following those bankruptcies has renewed the debate on the improvement of the market regulation. As noted by Cardebat and Bazen (2016), no futures market exists for wine[2]. Standardized futures contracts are in fact unpopular among wine producers who are usually vindicative about their specific attributes, especially in the Bordeaux region. Consequently, wine producers cannot hedge against the price risk by holding futures positions. As compared to the producers of staple commodities for which futures markets exist, they arguably lack information about the upcoming market conditions. This might cause a wide range of price expectations to coexist in the same period. The less well-informed agents may then adopt non-optimal behaviors and eventually obstruct the optimal smoothing of prices. Of course, prices expectations are likely to converge, all things kept equal[3]. However, the delay of convergence of these expectations depends on the speed of propagation of the news. The lack of public information about future conditions hence aggravates wine price volatility by slowing down the process of convergence of the expectations. Even if information was perfect and all wine professionals had optimal foresight, several structural rigidities prevent perfect intertemporal arbitrage. Contrary to staple agricultural products, Bordeaux wines are known for the influence of vintage quality on their prices. This causes wines produced in different years to be imperfect substitutes, which further exacerbates price volatility. Furthermore, classical market rigidities may also prevent producers to take full advantage of predictable price changes, especially in times of crisis. Storage capacities and borrowing limits became critical for many producers after the downturn caused inventories to soar and revenues to drop. Finally, the Bordeaux wine market is illiquid during the lean season and offers less possibilities for

[1]The area cultivated only fell of 8% over the same period, so that most wine producers who went bankrupt eventually found a buyer.

[2]The existing *en primeur* market for fine Bordeaux wines is sometimes regarded as a futures market. But its differs from the standard futures markets in several ways. First, it only opens during two months each year. Second, it only concerns the top 5% Bordeaux wines so that it will never be representative of the wine market as a whole. Lastly, there is no unique clearing house, but a number of independent wine brokers.

[3]In a theoretical framework, Guesnerie (1992) states the conditions of this convergence.

intra-annual arbitrage. The imperfection of the information, the vintage effect and the market rigidities all delay the impact of the shocks on the price fundamentals.

These structural rigidities limit intertemporal arbitrage and therefore cause price changes to be predictable, to some extent, and given the right information set. But price expectations differ across agents because of heterogeneous access to relevant information and to statistical facilities. In this chapter, I develop accessible forecasting methods so as to palliate to this shortcoming of the market structure, and provide transparent and objective price forecasts to Bordeaux wine professionals. I gather a collection of data sets about each of the relevant predictors of prices, and combine the latter in multivariate models. My strategy has been to rely on standard models in time series analysis such as Autoregressive Distributed Lags models (ADL), Error-Correction Models (ECM) and Unobserved Component Models (UCM), and to focus on the optimal selection of the predictors. The empirical analysis covers both annual and monthly fluctuations of average prices for the fifteen largest AOC of origins in Bordeaux. These series are freely accessible to Bordeaux wine professionals and are highly scrutinized, because they serve as benchmark for their negotiations. The main objective of the paper is to provide economically useful forecasts of these series. The secondary objective is to assess the respective influence of their determinants, such as the harvest, the stocks, the exchange rates, etc. Lastly, the monthly model allows to discuss the seasonality of the information flow, and the relevance of intra-annual price drivers, including the climatic conditions during the growth season.

Agricultural price forecasting used to be a key issue for agricultural economists. Several studies on agricultural price forecasts were published in the first half of the 20th century using econometric models (Sarle, 1925; Cox et al., 1956; L'Esperance, 1964), but price modelization and forecasting became more seriously investigated after the oil crisis of the 1970s. The following period of high volatility of commodity prices was indeed concomitant with the emergence of the modern methods in the econometrics of time series (Box et al., 1970). A stream of the literature then addressed the question of the predictability of agricultural prices (see Brandt and Bessler 1981, Brandt and Bessler 1983 and Holt and Brandt 1985, among many others). In a comprehensive review of this literature, Allen (1994) advocates combining several forecasting models and reports that the naive no-change price forecasts are difficult to outperform. Allen (1994) also argued that agricultural economists had put too much emphasis on the explanatory powers of their models, and did not sufficiently evaluate post-sample forecasts performance[4]. As a matter of fact, several studies have found that freely available price forecasts based on futures prices were the most efficient (Just and Rausser, 1981; Tomek, 1996; Kastens et al., 1998; Ahumada and Cornejo, 2016). When available, futures markets seem to provide the most accurate price forecasts, and thus publicly disclose optimal information about future price changes. According to Brorsen and Irwin (1996), research on price forecasting should therefore focus on markets unequipped with a futures market.

Even though agricultural price forecasting in the absence of futures market has been put on the agenda of agricultural economic research (Brorsen and Irwin, 1996; Tomek, 1996), only a few papers have been published on the issue of wine price forecasting. Yeo et al. (2015) uses univariate techniques to explain producer-specific prices. Ashenfelter (2008) has explained the variations of prices and quality of Bordeaux wines using meteorological data. However, neither of these articles provide an genuine analysis of out-of-sample accuracy. Lecocq and Visser (2006b) and Cardebat et al. (2014) also investigate the impacts of the determinants of Bordeaux wine prices, but still in an explanatory purpose. More broadly, agricultural economists seem to have somehow deserted the field of price forecasting. One explanation may be the "reverse backdrawer bias" proposed by Timmermann (2006), who states that researchers might have less

[4]See Shmueli (2010) for a detailed presentation of the distinction between explanatory power and predictive power.

incentives to publish successful works on price forecasting than to sell them to the private sector[5]. Another reason could be that the skills in agricultural prices forecasting are now mostly attracted in the private sector, so that private forecasters may outperform academic forecasters more often. The competition of the private forecasters may then deter researchers to invest this field of research (Brorsen and Irwin, 1996). Whatever the causes for this state of affairs, this chapter builds on the literature on agricultural price forecasting and relies on more or less standard times series methods, such as error-correction models in the tradition of Engle and Granger (1987) and unobserved component models introduced by Harvey (1989).

No future market currently exists for wine, although the idea has been discussed[6]. Contrary to staple commodities like wheat or soybean, wine is a manufactured product characterized by a high degree of vertical differentiation. Even within the localized Bordeaux region, wines are well differentiated from the generic Bordeaux wines to the top-end quality *grands crus* (Bélis-Bergouignan, 2011). Futures contracts with fixed quality standards would therefore only concern a small fraction of the market, and may not attract enough investors to provide sufficient liquidity. Besides, the 2007-2008 food crisis shed renewed doubts on the stabilizing role of futures markets (Timmer, 2010; Gutierrez, 2012). Although strong evidence has been provided to acquit the latter (Jacks, 2007; Lence, 2009; Irwin and Sanders, 2011; Wright, 2011, 2014; Sanders and Irwin, 2016), this polemic has not strengthened the trust of wine producers in the virtues of futures markets. Hence, futures contracts for wine may not be introduced in the near future, which is a motivation for the price forecasting models developed in this chapter.

Designing a forecasting model for wine prices requires to select the adequate determinants of prices, gather representative data, and estimate the price forecasting function. In that regard, explanatory analysis on past data is a prerequisite to forecasting studies. The first part of this chapter is therefore dedicated to the identification of the determinants of prices. In the economic literature on wine, the primary interest has been the relationship between cross-sectional prices, quality and reputation[7]. Because the price data I use hereafter are averages over all transactions of a given AOC, no cross-sectional analysis is provided in this chapter. However, the fluctuations of vintage quality should influence the variations of average prices, and are thus included in this study.

In the wider scope of commodity markets, the dominant theoretical model of price dynamics is the competitive storage model introduced by Gustafson (1958) and exhaustively analyzed by Williams and Wright (1991), which puts stocks at the center of price dynamics. If the first empirical studies were inconclusive (Deaton and Laroque, 1992, 1996), the empirical relevance of the competitive storage model has been recently reassessed by several papers (Cafiero et al., 2011; Guerra et al., 2014; Cafiero et al., 2015; Gouel and Legrand, 2017a,b). In the current state of research, the competitive storage model is not well-suited for operational forecasting purpose (see the discussion section 4.5). I however build on the lessons of this literature and put emphasis on the influence of initial inventories on the average prices of a given period. To my knowledge, the predictive power of stocks on prices has yet not been properly examined on real data. Symeonidis et al. (2012) and Gorton et al. (2013) found that the difference between futures prices and spot prices are well-correlated with the levels of inventories. If futures prices are arguably the best price forecasts when available, then these findings suggest that inventories are indeed a key predictor of prices. Apart from quality and stocks, many other determinants can

[5]This expression was coined after the "drawer bias" (Rosenthal, 1979), which refers to the difficulties to publish papers with insignificant results.

[6]Euronext launched a futures market dedicated to fine Bordeaux wines in September, 2001 but it quickly ended in a fiasco (Pichet, 2010).

[7]Among others, Nerlove (1995),Ashenfelter et al. (1995),Cardebat and Figuet (2004),Lecocq and Visser (2006b),Hadj Ali and Nauges (2007),Hadj Ali et al. (2008),Oczkowski (2010),Dubois and Nauges (2010),Ginsburgh et al. (2013),Cardebat et al. (2014)

be identified to play a certain role in the formation of wine prices. In a review, Balcombe (2010) has proposed a list of determinants of agricultural price volatility, which include past volatility, yields volatility, price transmission from inputs and substitutes, exchange rates volatility, and interest rates volatility. The latter are all plausible drivers of the volatility of wine prices. The influence of foreign wine productions on domestic wine prices has been documented by Wittwer et al. (2003). The interest rates and exchange rates have been found to significantly influence Bordeaux fine wine prices in Jiao (2017). These macroeconomic determinants are accounted for in this chapter using various public data sets (see section 4.4). All the considered determinants have undoubtedly a role in the formation of prices. However, they are not all relevant predictors of prices, either because they are unpredictable and only have an immediate impact[8], or because the data series are too short to accurately estimate their influence[9]. For both reasons, variable selection has been needed, and is the main focus of the empirical work of section 4.5.

The current paper finally contributes to the literature on the influence of the weather on current agricultural prices. Because they determine the next harvest, the weather conditions during the growth season are expected to influence current prices. However, the few empirical studies which have attempted to estimate this effect found it to be very low (Roll, 1984; Boudoukh et al., 2007; Chou et al., 2016; Osborne, 2004). this chapter estimates the effect of the weather conditions on Bordeaux wines prices, together with those of each of the aforementioned determinants.

The main outcome of this chapter is that the fluctuations of Bordeaux wines average prices are found to be predictable, to some extent. The forecasts are proven economically useful in that they outperform the naive no-change forecasts in a number of cases, which has been found challenging in the literature (Allen, 1994). In particular, the models would have predicted the important rise of prices concomitant with the catastrophic harvest of 2013. The forecast performance are particularly satisfactory for the AOC with the largest market share. The model thus seems to behave satisfactorily when the economic stakes are highest. I also provide a fruitful comparison between forecasts at the annual and at the monthly frequencies. The modelization choices are specific to each frequency, based on their respective data constraints in terms of number of degrees of freedom and required computation times by estimation. Monthly forecasts do better for longer time horizons. Short-run price changes are indeed more difficult to forecast since they provide more accessible arbitrage opportunities. The monthly forecasts are also more accurate during the first half of the marketing year which starts in August. Indeed, it is during this season that we have the best visibility on the upcoming volume available. However, they do not outperform annual models at forecasting the next annual average prices. These models are designed to equip the Bordeaux wine market place with price forecasts. These benchmark forecasts could play the role of futures prices, in a context where futures contracts are unpopular. The procedures developed in this chapter may also be applied to other agricultural markets where no future markets exists, and are unlikely to be introduced.

A secondary objective in this chapter has been to explain past variations of prices, and to assess the respective influences of each determinant. Choosing the statistical criterion for model selection gave rise to a discussion on the bias-variance trade-off in the context of multivariate models, with respect to both objectives of explaining past prices and forecasting future prices. My estimations provide additional evidence of the key role played by stocks in the formation of prices, which consolidates the literature on the empirical relevance of the theory of storage. By contrast, the influence of quality is found to be weak, and restricted to the high-end segment. For the basic quality segment, the market behaves much like the standard commodity

[8]Unpredictable drivers of the price can be useful for the forecasts if their influence is sufficiently lagged.

[9]This is for instance the case for long-term drivers which have remained steady over the observed period.

markets where consecutive annual productions are well-substitutable. Among the macroeconomic determinants, the exchange rates are found to be the most influential, but mostly in the long run. So as to increase the number of degrees of freedom of the estimated models, I have estimated an auxiliary harvest model which combines all weather data into an expectation of the next harvest. My estimations conclude to a significant influence of current weather conditions on spot prices during the growth season. Consistent with the literature, the explanatory power of weather is low compared to the other determinants. The estimates of this auxiliary harvest model also suggests that Bordeaux temperatures have reached their optimum for wine yields. Consistent with the conclusions drawn by Jones et al. (2005) on wine quality and climate change, a decrease in production yields can be expected if the current increasing trend in the temperatures continues.

Section 4.2 presents the structure of the wine market in Bordeaux. I first detail the institutional management of market information, and then legitimate the key role played by inventories by the structural features of the market. Section 4.3 gives summary statistics of the market data for each AOC, and compare their time series attributes to those of staple commodities with respect to the theory of storage. Section 4.4 details the auxiliary harvest model, the data collected on each of the price determinants, and how the latter is combined in leading indicators. Section 4.5 explains the methods followed for selecting the most appropriate forecasting models for annual and monthly prices. Section 4.6 provides the estimates of the representative models on the whole sample and discusses the respective influence of each driver. Finally, I evaluate the accuracy of the various price forecasts in section 4.7 and conclude in section 4.8.

4.2 The wine market information system in Bordeaux

4.2.1 Institutionalization of the market information system

When competition is perfect, market prices provide all the information about the market conditions. In real markets, the dissemination of market information is more or less institutionalized, from basic word to mouth to full economic bulletins produced by official industry representatives. Whatever their complexity, the economic role of these market information systems is to ensure that the market conditions are common knowledge, or rather to mitigate the unavoidable incompleteness of information. In the case of the Bordeaux wine industry, the market information system is well-structured and provide regular and detailed market data to all users. Hereafter, I report the content and timing of the information flow, which is a key to the evaluation of the price forecasts presented in section 4.7.

In order to maximize their inter-temporal profit, economic agents manage their inventory with respect to their expectations about future price changes (Wright and Williams, 1982). These expectations are based on what they know about the current states and the plausible future changes of the price drivers. Hence, the market performance is partly determined by the quality of the information available to the agents. Lack of market information can generate suboptimal behaviors and increase price volatility, which has been found to have exacerbated the 2007-2008 food crisis (Greenfield and Abbassian, 2011; Bobenrieth et al., 2013). The latter crisis was answered by the creation of a new United Nations institution, the Agricultural Market Information System (AMIS), designed to collect and disseminate information about the main global agricultural markets. Similarly although at a smaller scale, the high agricultural price volatility of the 1970s destabilized the French wine market, and many producers went bankrupt. In response, new regulation policies were designed. In 1976, Bordeaux producers committed to declare all transaction prices and volumes for the wines sold in bulk[10] (Smith et al., 2007). The management of the collection and broadcast of this data was assigned to the *Conseil Interprofessionel des Vins de Bordeaux* (CIVB), a joint-trade union regrouping winemakers and local wholesalers[11]. The CIVB has since then broadcast average prices and total transacted volumes at various frequencies for all *Appellation d'Origine Contrôlée* (AOC)[12] financed by a contribution collected on all the transactions in bulk. Nowadays, any economic agent involved in the bulk market of Bordeaux wines can freely access these data on the website of the CIVB. Aware of the average prices set by their competitors, the Bordeaux wine professionals are now expected to make better marketing decisions[13] and are more efficient in managing their inventories. The price and volume data are not disclosed in real time, and monthly aggregate statistics are only available with some delay. The average prices and total volume sold of month t are usually given on the CIVB's website before the 15$^{\text{th}}$ of month $t+1$[14]. There also exist weekly series of average prices, but they come with important flaws so that they are not used in this chapter, see section 4.3.

[10] About half of the wine produced in the Bordeaux region is sold in bulk. The rest is bottled by the producers and sold directly to the retailers. This arbitrage has been examined in Traversac et al. (2011).

[11] The CIVB was actually created in 1945, and was originally in charge of the fight against fraud, of the branch advertising and of the improvement of wine quality.

[12] In Bordeaux, 97% of the wines are labeled under one of the 56 different AOC, each of them being defined by a specific quality. Note that the differentiation of quality among the AOC can be vertical (AOC Bordeaux is of lower quality than AOC Bordeaux-supérieur) or horizontal (two AOC may coexist with no consensual hierarchy, like AOC Côtes-de-bourg and AOC Blaye-côte-de-bordeaux for example).

[13] In an experiment, Nakasone et al. (2013) has estimated the benefits of a similar market information system in Peru.

[14] For all market information, the delays between the collection of the disaggregated data and the disclosure of aggregated statistics are based on my personal observations since January 2014.

The CIVB also collects and discloses aggregated data from other sources, including the French customs. In France, all wine producers must declare the volume of their annual production to the customs as well as the inventory held in the end of July. There is an important delay of data treatment before the aggregated figures of the stocks and inventories are available on the CIVB's website. Usually, total inventories held at the end of July are made public before December, and the figure of the total volume harvested during year T is released before March of year $T + 1$[15]. These delays will be duly taken into account when evaluating the forecasting performance of the model in section 4.7.

In addition to the annual declaration of stocks and harvest, French wine producers must declare every month the total volume of wine that they have delivered[16]. Together with the annual declaration of production and stocks, these monthly deliveries allow to estimate monthly stocks[17]. Here again, the statistics are available only after a certain delay. The total quantity delivered for each AOC during month t is usually known by month $t + 3$. In practice, when willing to forecast the price of month $t + 1$ at the end of month t, one only knows an estimate of the ending stock of month $t - 2$. This delay will also be respected for the evaluation of the forecasts.

Finally, the customs also provide the CIVB with aggregated data on exported volumes and values by AOC, for the major importing countries. About 40% the total production in Bordeaux is exported, to more than 100 different destinations. These trade data will be used to build leading indicators of the demand is section 4.4. For each AOC, total traded volumes and values of month t are usually given on the CIVB's website around month $t + 4$.

4.2.2 Weather hazard, expected harvest and supply response

Apart from the market data provided by the CIVB, the key information flow for the agents involved in the wine business is the weather during the growth season. The weather conditions determine a large part of the upcoming wine production, and thus of the future market conditions. From March to October, extreme weather vagaries may jeopardize the quality and the volume of the harvest. In 1991 and 2017 for instance, severe frosts in late April destroyed the developing buds on a vast part of the vineyard. In 2013, extreme rain in June caused widespread flowering abortion and hail storms from July to September crushed the surviving grapes. The harvested volume was then greatly inferior to that of the preceding year, and many of the surviving buds were not fit to make top-end quality wines[18]. Interestingly, the distribution of these sudden shocks on the expected yields is asymmetric. If a harvest can be almost totally destroyed in one night of frost after budding, no weather condition can increase the expected harvest in the same proportions. Besides, the regulation of quality imposes a maximum yield, specific to each AOC. On each plot, the production exceeding the maximum yield cannot be

[15]Recall that the harvest occurs between September and November, and that winemaking is remarkably complex in Bordeaux. It is thus hard to known precisely the final volume produced for each AOC immediately after the harvest, just from weighting the harvested grapes. A good approximation of the level of the harvest is known rapidly after the harvest, but the agents must wait several months to obtain the precise total.

[16]In Bordeaux, sold wines are actually usually not delivered by the producers but collected by the buyers.

[17]Not all the wines that leave the warehouse are declared to the customs as deliveries. Little quantities are sometimes distilled in case of a large production surplus, and are not declared. Also, the personal consumption of the producer is not declared monthly. Finally, a fraction of the harvest of a given AOC is sometimes downgraded to another AOC of lesser quality, which generates mismatches between declarations. As a result, the total quantity that left the warehouses for each AOC is not exactly equal to the the monthly declared deliveries, and monthly stocks can only be computed .

[18]This depends on the color of the wine: the weather conditions of 2013 were catastrophic for the red, but the white wines were satisfying.

commercialized[19]. As the relationship between quantity and quality is decreasing[20], a maximum yield is a way to enforce a minimum quality standard (Giraud-Héraud and Soler, 2003). This maximum yield policy further skews the probability distribution of the yields to the left, therefore increasing the risks of a sudden shortage and decreasing the chances of a sudden surplus. This maximum yield is decided annually by the syndicate of each AOC[21] around July of each year, to adapt the production to the market conditions. In practice, the maximum yield does not change much, and for several AOC it has not changed in years.

The delay of the supply reaction to a demand shock is also asymmetric. As for any perennial crop, augmenting the production takes time. Newly planted vines do not produce significant quantities of grape during the first two years, and it takes around ten years before they produce at full yield. This delay makes it long for the supply to adjust to a demand increase[22]. Even when demand decreases, the adjustment of supply is slow because producers have a hard time uprooting their vine. Indeed, planting a vine-tree is a long-term investment: it produces during about sixty years, and old vines are often considered to produce better quality grapes (Zufferey and Maigre, 2008)[23]. As a result, the adjustment to a negative demand shock is also slow. This feature of the wine market generates repeated disequilibria between supply and demand, with several overproduction episodes (decades 1930, 1950, 1970 and 2000[24]) and shortages (mid-1990s and 2013). However, contrary to other products with inertial supply (like fruits or milk), Bordeaux wines are known worldwide for their ability to be kept. Even if that applies mostly to top-end Bordeaux wines, which can be kept in good conditions for up to 30 years in proper conditions, entry-level Bordeaux wines can only be stored up to three years. In the sort-run, the equilibrium between supply and demand is then primarily maintained through the management of inventories.

In order to optimally manage these stocks, the stakeholders of the wine industry evaluate the upcoming harvest by scrutinizing the weather conditions during the growth season.[25]. Based on the expected harvest, the risk of a future shortage is evaluated and buying/selling strategies are settled, therefore impacting current prices. So as to estimate the influence of weather conditions on average prices, I have estimated an auxiliary harvest model using meteorological data (see section 4.4.1). The model is designed to account for the maximum yield. The forecasts of this model are assumed to reflect the average harvest expectations at each point in time, and are plugged in the price equation (see section 4.5).

4.3 Data

4.3.1 Descriptive statistics

The CIVB has shared with me the history of several time series: the prices, the deliveries, the stocks and the harvests for all AOC and at two different time frequencies. I also have access to

[19]This policy was enforced in 1993. Before that year, producers had to respect a fixed ten-years moving average yield. More flexibility was introduced in 2010 to limit spoilage. If the wine in surplus satisfy the quality standards of the AOC, a limited amount can be kept in stock until the following year.

[20]This is true in a given state of technology and weather: the more selective the harvest process is, the better the wine will be. Of course, quantity and quality are both favored by a better technology, and both are compromised by inclement weather.

[21]Any change in the maximum yield needs to be approved by the *Institut National de l'Origine et de la Qualité* (INAO), the French authority managing the AOC.

[22]A similar asymmetry in the case of livestock dynamics has been studied by Holt and Craig (2006).

[23]When excess supply is too large, vine uprooting is supported by a public subsidy. This policy has been enforced between 2008 and 2011.

[24]See Milhau (1953) and Smith et al. (2007).

[25]Little information can be collected during winter about the quantity or the quality of the upcoming harvest, so it is neglected in the rest of this chapter.

export data, but aggregated to the color level. All exports of red and rosé AOC are aggregated in a first series, and all exports of the white AOC in a second series. The annual data (harvests, deliveries, stocks, exports and prices) cover the marketing years 1981 to 2016, which runs from August 1st to July 31st. In all the following, year T refers to the marketing year during which vintage T is harvested. For instance, the vintage 2016 has been harvest in from September to November 2016 so that "year 2016" hereafter refers to the period running from the 1st of August 2016 to the 31st of July 2017.

The monthly data (exports, deliveries and prices) only starts in August 2001, when the policy of the CIVB changed towards a closer monitoring of the market conditions, and lasts in July 2017. Of course, the monthly data are richer for the purpose of studying price volatility, but the history of the annual data covers a much longer period for which the harvest and annual initial stocks are observed. Both price series have specific interest and thus both are studied in this chapter. The annual prices are averages of all transactions prices over the marketing year. The declarations of stocks are collected by the CIVB at the end of this marketing year, that is on the 31st of July[26].

To recapitulate, the annual market data covers the years 1981-2016 for each AOC of the Bordeaux region and includes the following variables:

- average prices all vintages mixed together,
- total deliveries all vintages mixed together,
- total stocks as of July 31st of each year all vintages mixed together,
- total harvest,
- total exports all vintages mixed together, only two series for all red/rosé AOC and all white AOC.

In addition, I dispose of a collection of monthly market data for the period August 2001 to July 2017. The latter is also specific to each AOC and includes the following variables:

- average prices all vintages mixed together,
- total deliveries all vintages mixed together,
- total exports all vintages mixed together, only two series for all red/rosé AOC and all white AOC and since January 2012.

Among the 56 existing AOC of the Bordeaux region, I have selected the 15 larger ones on the bulk market for the purpose of this study. Most AOC are very small, so that few transactions are recorded by the CIVB. Because the data collected by the CIVB only concerns the latter, the monthly price series of many of the 56 AOC contain a lot of missing values, and for some prices are hardly representative. I have kept the 15 most important AOC in volume as a representative panel of Bordeaux wines. Their monthly price series contain a total of 2.7% missing values, against 59.5% for the full sample. For the estimation of the time series models presented in section 4.5, these few missing prices have been arbitrarily filled using the previous price[27]. This missing data and representativity issues is the reason why I chose not to use the available

[26] Before 2000 the end of the marketing year was the 31st of August, so that the stocks declared were somehow lower. Hence, I have artificially augmented the stocks before 2000 by a factor $\frac{12}{11}$ to roughly correct this slight structural break.

[27] Missing values are so rare that more sophisticated interpolation method should not lead to significantly different empirical results.

weekly price series mentioned in section 4.2.1. They contain 82.5% missing values on average and when they are not missing, the weekly volumes are often very low. A large part of the weekly price changes are therefore only due to changes in the basket of traded products, not to changes in the market conditions. Because only the latter is of interest for my purpose, most of the intra-month volatility of prices is irrelevant. Even at the monthly frequency, a certain share of the volatility of average prices is unrelated to the economic climate and must be filtered out. In the empirical treatment of the monthly price series (see section 4.5), this irrelevant volatility is removed using a standard time-series filtering method. Obviously, annual prices are fully representative so I do not filter the series of annual average prices.

Table 4.1 gives the characteristics of each of the 15 selected AOC[28]. Eleven of them are red wines: AOC Blaye-côtes-de-bordeaux (BLA), Bordeaux-supérieur (BSR), Bordeaux red (BR), Castillon-côtes-de-bordeaux (CAS), Côtes-de-bourg (CBO), Graves red (GRA), Haut-médoc (HME), Lussac-saint-émilion (LU), Médoc (MED), Montage-saint-émilion (MSE) and Saint-émilion (SE); three are white wines: Bordeaux white (BW), Entre-deux-mers (EDM) and Sauternes (SAU); and the last is Bordeaux rosé (BRO). Together, these AOC account for about 80% of the total annual harvest of the Bordeaux region. Column 4 gives the approximate area of production of the AOC within the region, which gives a sense of the proximities of the AOC.Closely located AOC are expected to share the same weather conditions. The four regional AOC (three colors of AOC Bordeaux and the red of AOC Bordeaux-supérieur) can be produced anywhere in the Bordeaux region, but most of the production is actually located in the eastern part of the region. For the harvest model of section 4.4.1, geographically close AOC will be attributed the same meteorological data series. Column 5 to 9 of table 4.1 give for each AOC the averages over the period 2001-2016 of the harvests, of the regulated maximum yields (see section 4.2.2), of the annual prices and of the shares sold in bulk. BR is the largest AOC and accounts for about one third of the total harvest in the region[29], and more than half of the bulk market total value. Compared to this giant, the rest of the AOC seem very small. However, even the smallest one (SAU) exhibits an average production of 3.2 Ml, representing 4.3 millions of standard bottles of 0.75 liter.

Recall from section 4.2.2 that within each AOC, the production is bounded by a maximum yield per hectare. Column 6 gives the average of the maximum yields for each AOC in hectoliter (hl) per hectare (ha). For most red AOC, the maximum yield lies around 55 hl/ha. The white wines of BW and EDM can be produced at higher yields, but the white wines of SAU are limited to an extremely low yield: 25 hl/ha[30]. Column 7 gives the average price paid to the producers between 2001 and 2016, all vintages grouped together[31]. Compared to prices commonly observed at the retail level, these production prices are low: less than €1 per bottle of 0.75 liter for six of the fifteen AOC. Indeed, these bulk prices do not include the additional costs of the bottle, the cork, the cap, the label, the transportation and the profit margins of wholesalers and

[28]As for the regional AOC Bordeaux, some AOC can exist in different colors (red, rosé and/or white). Strictly speaking, red and white wines from the AOC Bordeaux belong to the same AOC, but they represent different markets and are thus treated separately. For the sake of clarity, the different colors of a single AOC will be referred to as different AOC in the rest of the paper.

[29]As a benchmark comparison, the wine production of AOC BR slightly exceeds the production of Barefoot, which is often regarded as the largest wine private brand in volume. Interestingly, Barefoot is the leading brand of the company E.J. Gallo and the total wine production of the latter is actually comparable in size to the total wine production of the Bordeaux region as a whole. As a common brand, Bordeaux is on an equal footing with the largest private wine-producing company in terms of production.

[30]The wines of SAU are sweet white wines for which the harvest is more delicate. The grapes are left longer on the vines until the grapes are affected by a fungus, called *Botrytis cinerea*, or "noble rot", responsible for the sugary flavor of the wine.

[31]All the production is not sold in one year so that between three and five different vintages are dealt each year, depending on the AOC. The fact that all vintages are mixed together in the price data complicates the measure of the average quality of the wines transacted, as explained in section 4.4.2.

TABLE 4.1 – Characteristics of the 15 selected AOC

AOC (color for multi-color AOC)	Code	Color	Zone	Average harvest (10^6l)	Maximum yield (hl/ha)	Average price (€/0.75l)	Bulk share
Blaye-côtes-de-bordeaux (red)	BLA	red	North	26.7	52.6	0.99	0.363
Bordeaux-supérieur (red)	BSR	red	All	49.7	53.8	1	0.294
Bordeaux (red)	BR	red	All	195.6	56.6	0.86	0.662
Bordeaux (rosé)	BRO	rosé	All	17.6	57.4	0.87	0.533
Bordeaux (white)	BW	white	All	33.3	65	0.86	0.651
Castillon-côtes-de-bordeaux	CAS	red	East	10.2	52.8	0.98	0.384
Côtes-de-bourg (red)	CBO	red	North	17.4	53.7	1.04	0.395
Entre-deux-mers	EDM	white	East	7.8	63.3	0.91	0.403
Graves (red)	GRA	red	South	11.6	54.5	1.22	0.236
Haut-médoc	HME	red	West	20.8	54.4	1.74	0.102
Lussac-saint-émilion	LU	red	East	6.9	55.1	2	0.351
Médoc	MED	red	West	26.3	54.4	1.53	0.309
Montagne-saint-émilion	MSE	red	East	7.3	55.1	2.03	0.382
Saint-émilion	SE	red	East	7.1	55.1	2.71	0.451
Sauternes	SAU	white	South	3.2	25	4.02	0.52

retailers. Still, producers of the worldwide famous AOC Médoc and Saint-Emilion get on average €1.53 and €2.71 per bottle, a small fraction of the price generally paid by the consumer. Note that in compensation for the low regulated yield, producers of SAU fetch the highest prices. In the rest of the paper, price series are deflated using the French consumer price index [32] so as to obtain stationary real prices.

The last column gives the fraction of the harvest sold in bulk, varying between 10.2% for HME up to 66.2% for BR which dominates the bulk market. Although one may think that the bulk market is dedicated to low-end wines only, a few high-priced AOC like SAU of SE are in fact widely sold in bulk by producers. In total, about one third of the total harvest in Bordeaux is sold in bulk. This share is large enough to appreciate the fluctuations of the whole market. Moreover, because they are freely available with little delay, average bulk prices serve as benchmark prices in most negotiations. Forecasts of average bulk prices would therefore be helpful benchmarks for the whole market.

4.3.2 Time series characteristics of the prices and the role of storage

Columns 2 of table 4.2 gives the autocorrelation of the annual real price series[33]. In this section, I focus on the annual frequency to compare the data with that of other studies of agricultural commodities linking price and stocks. As for the majorities of commodities, the price series exhibit important autocorrelation. Most autocorrelation coefficients of the price series are between 0.7 and 0.9, close to that found in Deaton and Laroque (1992) for the main agricultural commodities (e.g. wheat, maize, soybean, sugar, etc.). Smaller autocorrelations coefficients are found for the white wines of AOC BW, EDM and SAU and for the red wines of the Saint-Emilion area (LU, MSE, SE). This autocorrelation has been explained by the role of storage (Cafiero et al., 2011), which mitigates the price drops and smooths price changes[34]. If the role

[32]source: *Institut National de la Statistique et des Etudes Economiques* (INSEE), the French administration in charge of the national accounts

[33]The autocorrelation of a times series z_t is the Pearson correlation coefficient between z_t and the lagged series z_{t-1}.

[34]The now classical competitive storage theory has been originally stated by Gustafson (1958) and then refined by Wright and Williams (1982) and Williams and Wright (1991). See Gouel (2012) for a review.

of storage is indeed at work here, other structural economic features generating price autocorrelation could be mentioned, like the autocorrelation of demand or the autocorrelation of the harvest[35]. In the case of AOC wines of the Bordeaux region, and in contrast with the commodity markets, the influence of vintage quality increases the annual volatility of the prices. Nonetheless, the series of annual prices show high autocorrelation, much in line with that of common agricultural cereals. On the contrary, the skewness coefficients[36] range between -0.61 and 0.398, whereas those reported by Deaton and Laroque (1992) range between 0.04 and 3.24. Price spikes due to shortages thus seem to be less frequent for Bordeaux wines. The skewness coefficient is however found positive for BSR, BW, EDM and MED.

The heterogeneity in the time series statistics across the AOC is partly explained by the variation in their stock levels. This is best illustrated with the use of two leading indicators, the stock-to-use and harvest-to-use ratios, hereafter referred to as the SUR and the HUR. The SUR of year T is the ratio between ending stock of the marketing year (declared on the 31$^{\text{st}}$ of July in year T) divided by the total quantity delivered during the past marketing year (between 1$^{\text{st}}$ of August of year $T-1$ and the 31$^{\text{st}}$ of July of year T) (Bobenrieth et al., 2013). This indicator provides a concise evaluation of the tension between supply and demand, and is widely used as an indicator of volatility. Notably, global SUR for major agricultural commodities are now available on the website of the AMIS[37]. The harvest-to-use ratio (HUR) is defined analogously. Column 4 of table 4.2 gives the average SUR over the sample period. They range between 0.454 for BW up to 2.519 for HME, which suggests that the inventories held at the end of the marketing year are sufficient to meet from 45% to over 250% of the annual demand depending on the AOC[38]. Apart from the good keeping potential of Bordeaux wines, another feature of the market regulation explains these high SUR. For each AOC, there exist a regulated delay between harvest and the beginning of the physical deliveries. The top-end quality AOC have to be kept longer in the producers' warehouses (until June, 15$^{\text{th}}$ for SAU[39]) than the regional AOC (November, 16$^{\text{th}}$ for BW and EDM), especially if the wine is already bottled. For ME and HME, the starting date for bulk delivery is January, 1$^{\text{st}}$, but the starting date for delivery in bottle is only on the 15$^{\text{th}}$ of June. Column 5 display the respective delays between the harvest (September) and the month of the start of the deliveries for each AOC. The longer the wines have to be kept before delivery, the larger the physical stocks declared at the end of the marketing year. This heterogeneity in the information provided by the SUR among the AOC is taken into account in section 4.6 by estimating AOC-specific models.

From the perspective of the economic literature on storage, one interesting feature of this market is that the price-smoothing role of storage is here mitigated by the influence of vintage quality, hereafter referred to as the vintage effect. All other things equal, if annual quality is more volatile and more influential on prices, then price autocorrelation is expected to be lower.

[35]Cafiero and Wright (2006) also argued that averaging over the calendar year generates an artificial price autocorrelation because the harvest occurs somewhere in the middle of the calendar year. A harvest shock thus affects the annual price averages over two calendar years in a row. The authors advocate the computation of annual price average over the marketing year, which is done in this chapter. Still, some information about the next harvest is revealed before the end of the marketing year. For instance, a late frost in April can kill an important share of the buds, and thus inform the agents of a low upcoming harvest. This pre-harvest information is also a factor of annual autocorrelation because the harvest shock also affects the prices of the preceding marketing year.

[36]The skewness coefficient, also known as the asymmetry coefficient, measures the degree of symmetry of a distribution. For a random variable X of mean μ and variance σ^2, the skewness coefficients is $\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)\right]$. A value above (resp. below) zero indicates a distribution skewed to the right (resp. to the left).

[37]Note that the inventory data used in this chapter concern a localized market, so that it is arguably more reliable than that used in the international studies. For a critic of the measure of global inventories, see Greenfield and Abbassian (2011).

[38]An unknown share of the declared stock is however already sold and waiting for delivery, which is referred to as the pipeline stock. This causes real available stocks to be somewhat overestimated.

[39]All reported starting dates for bulk deliveries are of 2015.

TABLE 4.2 – Time series attributes of the annual real prices and correlations with SUR and HUR

AOC	Price autocorrelation	Price skewness	Mean SUR	Delivery Delay	Price-SUR correlation	Price-HUR correlation
BLA	0.815	0.07	1.553	5	-0.534***	0.2
BSR	0.808	0.212	1.718	6	-0.617***	0.018
BR	0.799	0.084	0.819	4	-0.788***	0.08
BRO	0.839	-0.016	0.68	4	0.44***	0.096
BW	0.602	0.154	0.454	11	-0.26	-0.291*
CAS	0.886	-0.401	1.79	5	-0.515***	0.26
CBO	0.865	0.073	1.565	4	-0.395**	0.179
EDM	0.578	0.398	0.547	4	-0.262	-0.288*
GRA	0.816	-0.204	1.903	9	-0.626***	-0.043
HME	0.736	0.041	2.519	6	-0.325*	0
LU	0.517	-0.025	1.886	9	-0.28	-0.326*
MED	0.79	0.184	1.974	6	-0.496***	0.009
MSE	0.563	-0.357	1.914	8	-0.421**	-0.268
SE	0.496	-0.61	1.812	9	-0.614***	-0.549***
SAU	0.724	-0.064	2.507	11	-0.364***	-0.168

Significance of the correlations: ***: 1% level ; **: 5% level ; *:1% level

The figures of table 4.2 allow to roughly rank the AOC into four groups in terms of price autocorrelation, vintage effects and SUR level. The AOC of the first group, BLA, BSR, CAS, CBO, GRA, HME, MED and SAU, exhibit high price autocorrelation explained by a high average SUR level[40]. The AOC BW and EDM can be regrouped in a second category showing a low price autocorrelation which is explained by a low average SUR. The low SUR of the AOC BW and EDM may also explain the positive skewness of their price distributions, because a low SUR increases the risk of shortages and thus of price spikes. The vintage effect of the AOC of the two first groups serves as a benchmark for comparison. In a third group, we can isolate two AOC with a high price autocorrelation and a low SUR AOC (BR, BRO). The vintage effect should be particularly low for those two AOC, because their prices are highly autocorrelated even if their SUR are not particularly large. This is consistent with the fact those are the lower-end quality AOC. On the opposite, the three AOC of the fourth group (LU, MSE and SE) show low price autocorrelation and high SUR so that the vintage effect should be important. This is also consistent with the fact that those three AOC are produced in the prestigious Saint-Emilion area. These conjectures are tested in section 4.6 where I estimate the vintage effect by AOC.

Recall that the main focus of the paper is to forecast prices. For that purpose, I will use several leading indicators of the market situation, including the SUR and the HUR. To assess *a priori* the usefulness of the SUR as a predictor of the price, column 6 of table 4.2 contains the Pearson correlation coefficient between average annual prices and the SUR at the beginning of the marketing year. Except for BRO, this price-SUR correlation is negative for all AOC[41], and significant at the 5% for all but BW, EDM, HME and LU. This is consistent with the findings of Symeonidis et al. (2012) and Gorton et al. (2013) on the correlation between inventories and futures basis, and both results suggest that the SUR is a useful predictor of the price. Column 7 gives the correlation between prices and the HUR and by contrast, this correlation is actually found positive or insignificant at the 5% level for most AOC (except for SE). These are only

[40]Note that the SUR is higher for the AOC GRA, HME, MED and SAU, but that the deliveries also start later.

[41]The case of BRO is singular because the production of rosé wines have exploded over the sample period. The harvest of BRO in 2016 was 21 times larger than that of 1981. Accounting from 2000, when the BRO production had become substantial, the price-SUR correlation for BRO is significantly negative with a value of -0.28.

pairwise correlations, so these tests obviously cannot imply that the harvest has no influence whatsoever on next year's prices once we control for other price determinants. An analogous remark applies to the observed correlation between SUR and prices. However, they highlight the key role of the stocks compared to that of the harvest, and echo the results of Bobenrieth et al. (2013). In our case, the strong influence of the SUR may be explained by the important delay between the harvest and the beginning of the physical deliveries. Starting inventories must therefore meet a large part of the demand over the marketing year. For instance, SAU wines cannot be delivered before June 15th so that the stock declared at the end of the marketing year (July 31st) must be sufficient to meet the demand during 10 months and a half, before the harvest can circulate on the market. This regulation delays the influence of the harvest on prices. However, the transactions and thus the observation of the prices occur several months before the delivery. Even if the SAU harvest cannot be moved before the 15th of June, it is partly sold before that date and some transactions sometimes even occur before the harvest. This mitigates the role played by the regulation on the delay of physical deliveries.

Figure 4.1 plots together the normalized series of annual real prices, SUR and HUR for BR over the sample period. It gives a graphical illustration of the strong negative correlation between prices and SUR. We also see that the HUR is barely correlated with the price, but that the harvest drops of 1991 (late frost) and 2013 (rain and hail) caused significant price increases. To the extend that the low level of quantity harvested in 2013 was predictable before actual harvest, the price increase from 2012 to 2013 was also predictable. This question will be investigated in section 4.7. In terms of booms and busts, the recent market history can be divided into three periods. First, the market was relatively unstable and expanding between 1982 and 1997, with an increasing and exceeding demand that made inventory levels collapse until the price spikes of 1997 and 1998. Also note the higher volatility of the HUR series during this first period, caused by of lighter yield management and control (see Chevet et al. 2011 for a long-term historical analysis of the yields). After the shortage episode of 1997-1998, the market entered a long overproduction period. Between 1999 and 2005, the stock soared and the price dropped, then both remained steady between 2006 and 2010. For ten years the producers made little profit and about half of them ran bankrupt and had to sell their farms[42]. Since 2011, the economic climate has become more appealing. The huge stocks are steadily decreasing, and prices have recovered after the small harvest of 2013. However, the market is in a situation of relative vulnerability and the recent catastrophic frost of late April 2017 may increase prices back to historical highs.

4.4 Leading indicators of the fundamental determinants of prices

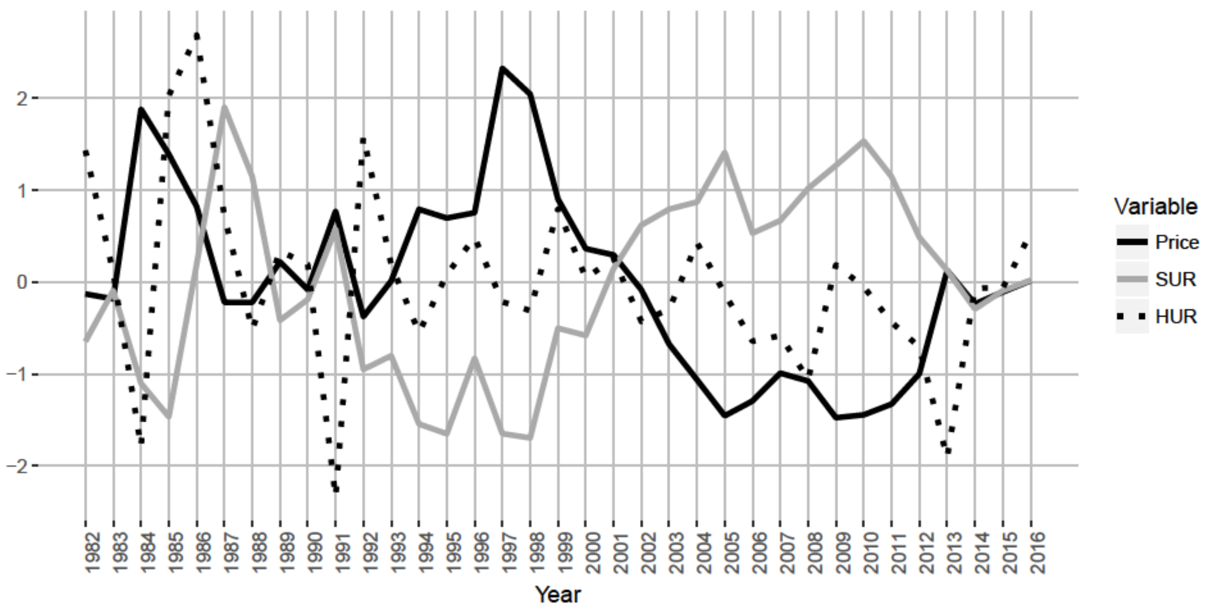
The previous section has enhanced the role played by the stocks and the harvest on the price fluctuations. Because one objective of this chapter is to explain the historical fluctuations of prices, I also consider several other potential fundamental drivers. For each of them, I have collected a specific data set and computed a leading indicator which summarizes the time variations.

Here is a list recapitulating the data collected and used for each leading indicator:

- Harvest expectations:

[42]The classical view on storage highlights the usefulness of stockpiling to mitigate price drops. However, as mentioned by Barrett (2007), Burke (2014) and Maître d'Hôtel and Le Cotty (2015), the budget constraint of the producers limits their ability to keep large inventories. During the 2000s, many little Bordeaux wine producers had to sell their production at loss so as to pay their bills, with little possibility of borrowing. Hence, there was little financial room for keeping strategic inventories and limiting the collapse of prices.

FIGURE 4.1 – Standardized series of prices, SUR and HUR for the AOC BR



- annual harvest, cultivated area and regulated maximum yield for each 15 AOC over years 1981 to 2016,
- weather data for six stations[43]:
  - daily minimum, mean and maximum temperatures,
  - daily sum of precipitations.
- Official monthly harvest forecasts produced by the French Ministry of Agriculture for years 2013 to 2016 and for the months July to November, only aggregated over the whole region. These forecasts are used to check the representativity of the constructed history of harvest expectations.

- Quality index:
  - vintage scores given by the Wine Advocate for vintages 1981-2016 for four subareas of the Bordeaux region,
  - vintage scores given by the Wine Spectator for vintages 1981-2016 for three subareas of the Bordeaux region (only one score for vintages prior to 1995),
  - harvest reports of the faculty of oenology of the University of Bordeaux, for vintages 2000 to 2016,
  - regulated delay between the harvest and the start of deliveries for each AOC.

- Exchange rates:
  - monthly averages of exchange rates over years 1981 to 2016 for 21 countries (source: European Central Bank),
  - total annual exports of Bordeaux wines for each color (red/rosé or white) for each of these 21 countries since 1981, and monthly exports since 2012.

[43] About 20% of the data are missing before 2000 for certain stations. I have replaced the missing values by those of the station of Villenave-d'Ornon for which no data is missing. The series of weather data for this station are the most correlated on average with the others, and thus the most representative.

- Consumers wealth:
  - quarterly GDP for the 16 main destinations of Bordeaux wines (source: OECD)
  - total annual exports of Bordeaux wines for each color (red/rosé or white) for each of these 16 countries since 1981, and monthly exports since 2012.
- Supply of the competitors:
  - national wine production of all countries over years 1981 to 2014 (source: FAO),
  - national wine production for the ten main countries for years 2015 and 2016 (source: OIV),
  - annual wine trade flows across all countries for years 1981 to 2014 (source: FAO)
  - total annual exports of Bordeaux wines for each color (red/rosé or white) to the 21 main countries since 1981, and monthly exports since 2012.
- Interest rates:
  - monthly average of the daily three months-ahead interest rate EURIBOR (source: Banque de France) for period from January 1999 to July 2016,
  - annual legal interest rate for years 1981 to 1999 (source: Banque de France).

4.4.1 Approximation of the harvest expectations

As previously mentioned in section 4.2.2, the wine professionals pay close attention to the weather conditions during the growth season. Based on their observations, they update their buying/selling strategies, and thus influence the current market average price. As such, the weather is potentially a relevant driver of price[44]. Because I wish to account for several other drivers and the number of observations is limited, I do not attempt to measure directly the influence of weather variable on prices and follow an indirect method. Structurally, the weather W_{it} of a given month t on the area of AOC i influences current prices P_{it} because it affects the expectation of the next harvest $E_t(H_i)$. Instead of estimating directly the impact of W_{it} on P_{it}, I first estimate the influence of W_{it} on $E_t(H_i)$, and then then the impact of $E_t(H_i)$ on P_{it}.

In this section, I present how I combine past weather data in a harvest model to obtain reasonable estimates of the monthly expectations $E_t(H_i)$. The latter series will then be used as an explaining variable of the monthly fluctuations of prices so as to control for the current weather (see section 4.5). The estimates of the harvest expectations will also be useful for the evaluation of the annual price forecasts at the beginning of the marketing year[45] in section 4.7.

Let H_{iT} be the harvest of year T for AOC i. It is given by the product of the total cultivated area A_{iT} and the average yield $\bar{Y}_{iT}$.

$$H_{iT} = \bar{Y}_{iT} A_{iT} \tag{4.1}$$

Forecasting the harvest thus requires to forecast both the average yield and the total area cultivated. Hereafter, I develop two models specific to each of these two components. Table 4.3 gives summary statistics of the two variables by AOC over the sample period. The averages of the areas and yields by AOC are in line with that of the total harvest and maximum yield given in table 4.1. Column 5 gives the average ratios between the average yield $\bar{Y}_{iT}$ and the maximum

[44]Boudoukh et al. (2007) exhibit the influence of weather in the determination of the future prices of frozen concentrated orange juice in Florida.

[45]Recall that the marketing year begins before the actual harvest, on August, 1st, so that the harvest is not known at the time of the annual price forecast.

yield Y_{iT}^{sup} by AOC. Of course, this ratio is always lower than 1 because the maximum yield applies to each producer and it is statistically impossible that all producers reach the maximum yield the same year. The average ratio attains a minimum for BRO (0.784) and a maximum for BR (0.883). Columns 3, 5 and 7 give respectively the average volatility of the area, the average yield and the average yield-to-maximum yield ratios by AOC over the sample period[46]. Except for BRO, the average yields are highly more volatile than the cultivated areas, which makes it the key variable to forecast. Column 7 indicates that the volatility of the average yield-to-maximum yield ratios is lower than that of the average yield[47]. This ratio is thus easier to forecast, and is taken as the explained variable in the model presented hereafter. Recall from section 4.2.1 that the maximum yield is decided at the beginning of June, so that it is known long before the harvest. For the real-time forecasts computed before June, the maximum yields of next harvest will be assumed to remain constant equal to those of the preceding harvest.

TABLE 4.3 – Area, Yield, and Yield-to-Maximum Yield-Ratio by AOC

AOC	Area		Yield		ratio $\bar{Y}_{iT}/Y_{iT}^{sup}$	
	Average (10^3ha)	Volatility	Average (hl/ha)	Volatility	Average	Volatility
BLA	4.5	0.07	51.1	0.28	0.873	0.22
BSR	10.3	0.13	48.4	0.3	0.838	0.27
BW	8.3	0.06	56.2	0.31	0.813	0.24
BRO	2.1	0.22	48.8	0.25	0.784	0.23
BR	34.4	0.06	55	0.25	0.883	0.22
CAS	2.6	0.06	47.9	0.42	0.821	0.39
CBO	3.5	0.04	51.8	0.24	0.882	0.19
EDM	2	0.1	55.2	0.28	0.833	0.23
GRA	2.2	0.08	47.5	0.25	0.814	0.19
HME	4.1	0.03	49.6	0.29	0.842	0.25
LU	1.3	0.03	51.4	0.43	0.879	0.36
MED	4.7	0.03	52.5	0.29	0.88	0.24
MSE	1.5	0.03	50.5	0.43	0.866	0.35
SE	1.8	0.08	50.3	0.41	0.873	0.3
SAU	1.6	0.04	19.7	0.27	0.787	0.27

The explanatory variables of the average yield are indicators of the weather conditions that are computed on data from MétéoFrance, the French national meteorological service. The raw weather data are the daily minimal, average and maximal temperatures in Celsius degree and daily total precipitations in millimeters for the period 1981-2016 from six weather stations representative of the different Bordeaux subareas. The precise geographic locations of the weather stations is given on figure D.1 in the Appendix D, together with table D.2 which details the attribution of the six weather stations among the fifteen selected AOC. All the weather data are aggregated into six different indicators. Chronologically, the first to intervene in the formation of the harvest expectations is $F4_{iT}$. I propose the indicator of the severity of the frost in late April for AOC i and year T defined in equation (4.2)

$$F4_{iT} = \sum_{d=1}^{30} d.TM_{iT4d}^2 1_{\{TM_{iT4d}<0\}} \tag{4.2}$$

[46]For a given time series Z_t, the volatility refers to the empirical standard deviation of $\frac{Z_t}{Z_{t-1}}$.

[47]For SAU, those are the same because the maximum yield has remained constant over the sample period.

where TM_{iT4d} is the minimum temperature recorded during day d of April of year T for AOC i. It is the sum of the squares of the negative daily minimum temperature of April, each multiplied by the number of the day (1 to 30) when the frost occurred. This indicator becomes larger as minimum daily temperatures fall below zero (at a quadratic speed), especially when this happens in the last days of April. As explained in section 4.2.2, the later the frost the more severe the damages to the buds, and thus the lower the expected harvest. The indicator is null for 76% of the observations, indicating that April frosts are rare. The indicator is maximum for 1991 during which a late frost had destroyed about half of the developing grapes.

Following the economic literature on the link between weather and wine harvest[48] and the advice of local researchers in oenology at the University of Bordeaux[49], I have considered three other indicators of the climatic conditions during the growth season. $T46_{iT}$ is the average of the four maximum temperatures of months April to June of year T for AOC i[50]. This indicator has a positive impact on the expected yield, until a certain threshold above which temperatures become too high for the vines (see Jones et al. 2005). To account for this non-linearity, I include a quadratic term $T46^2_{iT}$ in the yield equation. $R57_{iT}$ is the total precipitation from May to July of year T. It has a negative impact on the expected yield because it causes the development of vine diseases. Lastly, $R89_{iT}$ is the total precipitation from August to September of year T. The effect is ambiguous. On the one hand, heavy rain in this period can both make the grapes bigger and thus increase the yields, but can on the other hand develop grape diseases which decrease the yields. The overall effect depends on the terroir. In this model, I also acknowledge the recent findings of Guilpart et al. (2014) according to which the potential yield of a vintage is influenced by the weather conditions during the flowering of the preceding year. I thus add lagged variables $TX46_{iT-1}$ and $RR46_{iT-1}$ to the specification. Let $\alpha_i = (\alpha_{i0}, \alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \alpha_{i4}, \alpha_{i5}, \alpha_{i6}, \alpha_{i7}, \alpha_{i8})^t$ be a set of parameters associated with the explaining variables $X_{iT} = (1_i, T, T46_{iT-1}, R46_{iT-1}, F4_{iT}, T46_{iT}, T46^2_{iT}, R57_{iT}, R89_{iT})$. The forecasting model of the average yield $\bar{Y}_{iT}$ is given by system (4.3).

$$\begin{cases} \bar{Y}_{iT} & = Y^{sup}_{iT} Y^{ratio}_{iT} \\ ln(Y^{ratio}_{iT}) & \sim TN(\mu^Y_{iT}, (\sigma^Y)^2, 0) \\ \mu^Y_{iT} & = X_{iT}\alpha_i \end{cases} \tag{4.3}$$

The logarithm of the average yield-to-maximum yield ratio Y^{ratio}_{iT} is assumed to follow a truncated normal distribution $TN(\mu^Y_{iT}, (\sigma^Y)^2, 0)$. This is the distribution of a normal variable of mean μ^Y_{iT} and of variance $(\sigma^Y)^2$ for which only the realizations lower than zero are observed . Because the logarithm of the ratio Y^{ratio}_{iT} varies between $-\infty$ and zero, the model allows the observed average yield $\bar{Y}_{iT}$ to vary between zero and Y^{sup}_{iT}. The truncation causes most of the mass of the distribution to be located just under the maximum yield, which is consistent with the observations of table 4.3. The coefficients α_{ij} are allowed to differ among the AOC. However, because some AOC are very similar, several coefficients are expected to be very close across different AOC. In order to increase the robustness of the estimation and increase the number of degrees of freedom, I have restrained some coefficients to be equal. I propose the following algorithm to select the sets of equal coefficient:

1. estimate specific α_{ij} for each AOC by maximizing the likelihood of the truncated model regression with a common error term of variance $(\sigma^Y)^2$ across all AOC,

[48] See among others Ashenfelter et al. (1995), Chevet et al. (2011), Jones et al. (2005), Ashenfelter and Storchmann (2016).

[49] I am particularly grateful to Jean-Pascal Goutouly for his precise and detailed answers.

[50] This indicator has been found more significantly correlated with the average yields than the averages of the monthly mean temperatures, or of the monthly minimum temperatures. Detailed estimates are available upon request.

2. for all $j \in \{0, ..., 8\}$, and all couples (i_1, i_2) among the fifteen AOC, compute the statistic of the Wald test for the hypothesis: $\alpha_{i_1 j} = \alpha_{i_2 j}$,

3. for all $j \in \{0, ..., 8\}$, collect all the couples (i_{1j}, i_{2j}) for which the Wald statistics is the lowest,

4. if all the tested equalities are rejected at the 5% level, the aggregation algorithm stops.

5. Otherwise, go back to step 1 and re-estimate the model with the additional restrictions $\alpha_{i_{1j} j} = \alpha_{i_{2j} j}$ for the $j \in \{0, ..., 8\}$ such that at least one equality was rejected at the 5% level.

The outcome of this algorithm is a model in which all the estimated coefficients relative to the same variable are significantly different from each other at the 5 % level. The coefficients are estimated by maximum likelihood at each step, and the estimates of the final specification are given in table 4.4. This algorithm allows to considerably decrease the number of coefficients to estimate, from 135 to only 30, at arguably no cost because only non-significantly different coefficients are aggregated.

Table 4.4 shows that the algorithm leaves seven different values for the intercept, and five different values for the time trend. Note that the trend can be positive or negative, depending on whether the observed average yield is getting closer to or farther from the maximum yield. The rest of the coefficients are more aggregated. The aggregation algorithm leaves two different coefficients for the influence of $T46_{iT-1}$, the temperature of the preceding year, and only one for $R46_{iT-1}$ which is negative. Among the latter three, the coefficient of $T46_{iT-1}$ for the AOC BSR, BRO, EDM, GRA, MED, MSE and SE is statistically significant at the conventional 5% level, consistent with the observations of Guilpart et al. (2014). Row 5 gives the coefficients before $G4_{iT}$, the indicator of frost. Four different values are selected by the algorithm, and all are negative and strongly significant. Rows 6 and 7 respectively give the coefficient of $T46_{iT}$ and its square. The former take four different values, all significantly positive, and the latter takes three different values, all significantly negative. As expected, the influence of temperature on yields is thus shaped in an inverse U, with a maximum noted $T46^*_{iT}$ given by row 10. These maximums are the estimated ideal conditions of temperature for the yields, and are represented for each AOC on figure 4.2 together with the series of $T46_{iT}$. Note that the temperatures all exhibit an upward trend over the observation period. This trend is more pronounced for the AOC of the eastern zone than for the AOC near the ocean coast. Note that the estimated optimal temperature for potential yield is more and more often exceeded. This collateral result of the model suggests that the current temperature conditions are about optimal, but that the upward trend may cause future potential yields to decline. This is in line with the findings of Jones et al. (2005) who argue that the climatic conditions in Bordeaux are currently more or less optimal for wine quality, but may become too hot if the current trend continues. My estimations lead to the same conclusion regarding quantity. The last two coefficients for the model account for the precipitations from May to July and from August to September. The former takes two different values, both strongly significant and negative. The latter takes two values, one significantly positive and the other negative but not statistically significant. The overall effect of rain before the harvest thus seems to be positive, although difficult to measure accurately on the observed period.

TABLE 4.4 – Estimates of the yield model

Variable	BLA	BSR	BW	BRO	BR	CAS	CBO	EDM	GRA	HME	LU	MED	MSE	SE	SAU
Intercept	-19	-8.159	-1.141	-42.323	-7.402	-7.315	-7.402	-1.699	-8.159	-19	-7.315	-7.402	-1.699	-7.315	4.855
	(2.815)	(1.98)	(2.22)	(4.582)	(1.967)	(1.968)	(1.967)	(2.221)	(1.98)	(2.815)	(1.968)	(1.967)	(2.221)	(1.968)	(2.83)
Trend	.004	-.005	-.002	.012	-.002	-.005	-.002	-.002	-.002	.001	-.002	.001	-.005	-.002	-.005
	(.001)	(.001)	(.001)	(.002)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
$TX46_{iT-1}$	.001	.028	.001	.028	.001	.001	.001	.028	.028	.001	.001	.028	.028	.028	.001
	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)
$RR46_{iT-1}$	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.045	-.939
	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.089)	(.315)
$G4_{iT}$	-.014	-.003	-.006	-.003	-.003	-.026	-.006	-.006	-.014	-.006	-.026	-.006	-.026	-.026	-.003
	(.003)	(0)	(0)	(0)	(0)	(.002)	(0)	(0)	(.003)	(0)	(.002)	(0)	(.002)	(.002)	(0)
$TX46_{iT}$	1.004	1.528	.439	1.528	1.004	1.528	1.004	.439	1.004	1.528	1.004	.439	.974	.974	.439
	(.117)	(.153)	(.151)	(.153)	(.117)	(.153)	(.117)	(.151)	(.117)	(.153)	(.117)	(.151)	(.116)	(.116)	(.151)
$TX46^2_{iT}$	-.022	-.034	-.01	-.034	-.022	-.034	-.022	-.01	-.022	-.034	-.022	-.01	-.022	-.022	-.01
	(.003)	(.004)	(.004)	(.004)	(.003)	(.004)	(.003)	(.004)	(.003)	(.004)	(.003)	(.004)	(.003)	(.003)	(.004)
$RR57_{iT}$	-1.156	-.617	-.617	-.617	-.617	-1.156	-1.156	-.617	-.617	-.617	-1.156	-.617	-1.156	-1.156	-.617
	(.139)	(.111)	(.111)	(.111)	(.111)	(.139)	(.139)	(.111)	(.111)	(.111)	(.139)	(.111)	(.139)	(.139)	(.111)
$RR89_{iT}$	.369	-.067	.369	-.067	-.067	-.067	.369	.369	.369	-.067	-.067	-.067	.369	.369	-.067
	(.133)	(.122)	(.133)	(.122)	(.122)	(.122)	(.133)	(.133)	(.133)	(.122)	(.122)	(.122)	(.133)	(.133)	(.122)
$TX46^*_{iT}$	22.8	22.5	21.9	22.5	22.8	22.5	22.8	21.9	22.8	22.5	22.8	21.9	22.1	22.1	21.9
σ^Y	0.120														
R^2	0.816														

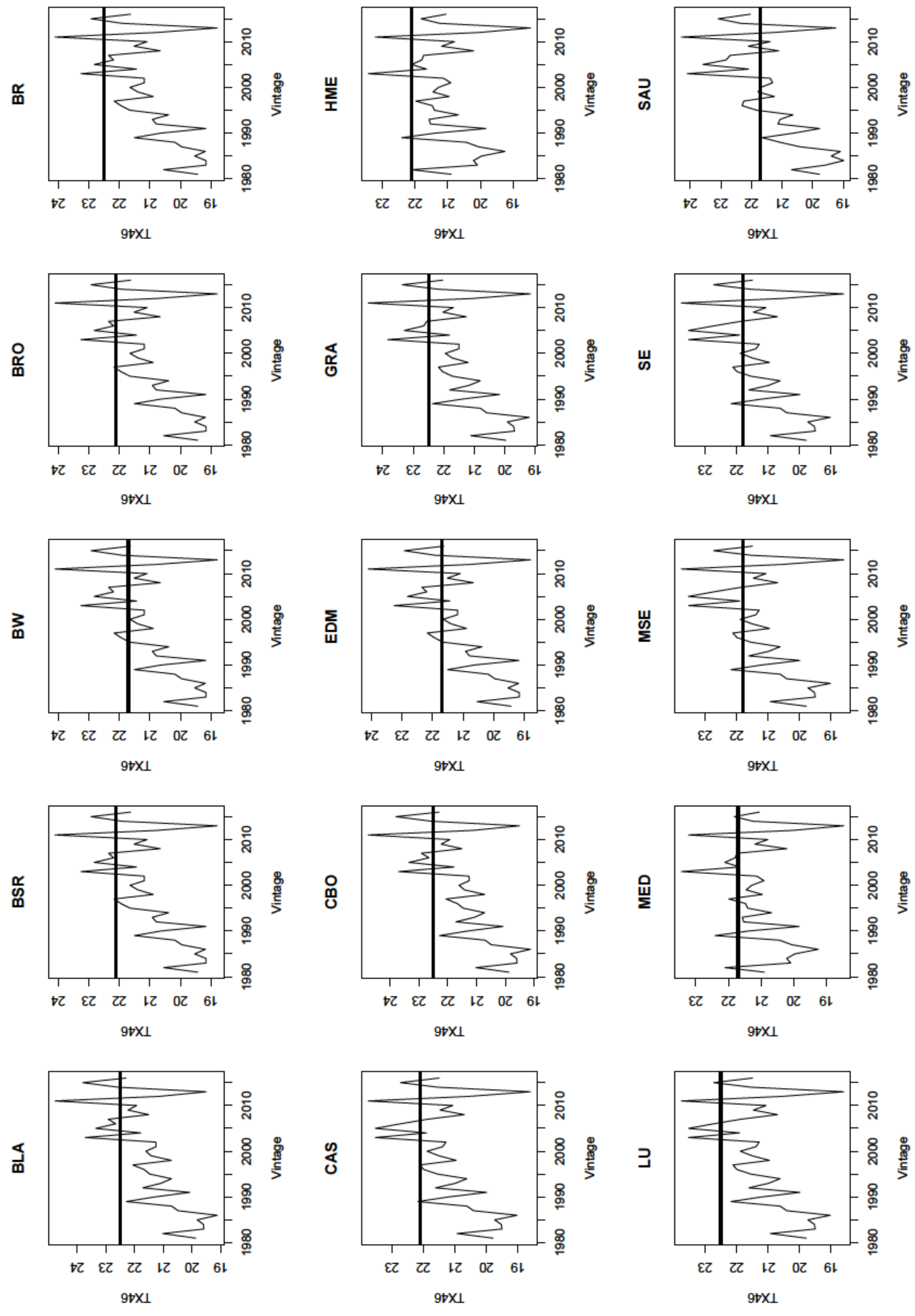


FIGURE 4.2 – History of the indicator $TX46_{iT}$ by AOC and the estimated optimal temperature levels for potential yields

With much less coefficients than the fully disaggregated model, the optimally aggregated model is used to generate monthly updates of the expected harvest for all AOC. In particular, when all the weather variables are known, the expectation of $\bar{Y}_{iT}$ is given by equation (4.4) (see the computation in the Appendix D). I assume that Y_{iT}^{sup} is known at the time of the computation of the forecast.

$$E(\bar{Y}_{iT}) = Y_{iT}^{sup} e^{\mu_{iT}^{Y} + \frac{1}{2}(\sigma_{iT}^{Y})^2} \frac{\Phi(-\frac{\mu_{iT}^{Y} + (\sigma_{iT}^{Y})^2}{\sigma_{iT}^{Y}})}{\Phi(-\frac{\mu_{iT}^{Y}}{\sigma_{iT}^{Y}})} \tag{4.4}$$

For the computation of the yield expectation at a given month m before the harvest, all the variables accounting for the weather after m are taken at their respective averages over the preceding decade[51]. Weather expectations can thus be viewed as naive, although it is unlikely to find significantly better weather forecasts at the monthly frequency.

In order to obtain the expectation of the total harvest, the yield expectations are multiplied by the expectation of the cultivated area[52]. The fluctuations of the cultivated area are very slow and regular, so that the expected area for the next period is mostly given by the recent trend. For each AOC, the estimation of this trend is achieved using a standard local-trend unobserved component model, hereafter noted UCM, and introduced by Harvey (1989)[53]. These models are estimated on the logarithm of the areas for each AOC so that the area expectations take positive values. Let a_{iT} be the logarithm of A_{iT}[54]. The forecast model for the areas is AOC-specific and given by the standard local-trend UCM detailed by system (4.5).

$$\begin{cases} A_{iT} &= e^{a_{iT}} \\ a_{iT} &= \mu_{iT}^{a} + \epsilon_{iT}^{a} \\ \mu_{iT}^{a} &= \mu_{iT-1}^{a} + \beta_{iT}^{a} + \nu_{iT}^{a} \\ \beta_{iT}^{a} &= \beta_{iT-1}^{a} + \xi_{iT}^{a} \\ \epsilon_{iT}^{a} &\sim \mathcal{N}(0, \sigma_{\epsilon_i^a}^2) \\ \nu_{iT}^{a} &\sim \mathcal{N}(0, \sigma_{\nu_i^a}^2) \\ \xi_{iT}^{a} &\sim \mathcal{N}(0, \sigma_{\xi_i^a}^2) \end{cases} \tag{4.5}$$

The local-trend UCM decomposes the times series in a level μ_{iT}^{a} and an irregular component ϵ_{iT}^{a}. The level is assumed to be a random walk of innovation ν_{iT}^{a} shifted by a stochastic trend β_{iT}^{a}, itself being a random walk of innovation ξ_{iT}^{a}. The innovations ϵ_{iT}^{a}, ν_{iT}^{a} and ξ_{iT}^{a} are assumed to follow centered normal laws of respective variances $\sigma_{\epsilon_i^a}^2$, $\sigma_{\nu_i^a}^2$ and $\sigma_{\xi_i^a}^2$. The latter are the only parameters characterizing the overall process. All variables are unobserved and estimated with the coefficients by maximizing the likelihood evaluated by the Kalman filter. The estimates are given in table 4.5. Because the equation is estimated in logarithm, the expected area in the end

[51] For instance, at the beginning of October, only half of $R89_{iT}$ is known, namely $R8_{iT}$. $R9_{iT}$ is yet unknown and thus taken at its average over the last ten years. This average is then summed with $R8_{iT}$ to obtain a plausible anticipation of the indicator $R89_{iT}$.

[52] The $\bar{Y}_{iT}$ and A_{iT} can be reasonably assumed to be independent variables, since the planting decision is made many years before the harvest.

[53] Structurally, the cultivated area is determined by planting decisions of the preceding years. Assuming profit-maximizing producers, these decisions are driven by past prices and maximum yields since those determine the expected profit of planting new vines. An econometric model incorporating these features have been estimated, but it actually provides poor fit as compared to the UCM. I therefore only present the latter in this chapter. Details and estimates of the econometric model for the area are available upon request.

[54] In what follows, variables noted in lowercase are the natural logarithms of the variables noted in uppercases.

of year $T-1$ is given by equation (4.6).

$$E_{T-1}(A_{iT}) = exp\{\mu^a_{iT-1} + \beta^a_{iT-1} + 0.5(\sigma^2_{\epsilon^a_i} + \sigma^2_{\nu^a_i} + \sigma^2_{\xi^a_i})\} \tag{4.6}$$

Note that the expectation of the next area does not evolve within a year.

TABLE 4.5 – ML estimation of local-trend UCM on log(Area) by AOC

	$\hat{\sigma}_{\xi^a_i}$	$\hat{\sigma}_{\nu^a_i}$	$\hat{\sigma}_{\epsilon^a_i}$
BLA	1.2e-04	1.9e-03	7.7e-04
BSR	1.3e-08	1.7e-03	7.4e-03
BW	5.4e-04	4.7e-04	7.1e-04
BRO	3.2e-04	3.4e-02	6.1e-03
BR	1.5e-04	1.6e-05	9.5e-04
CAS	6.1e-04	1.6e-04	1.1e-03
CBO	2.5e-05	2.5e-08	4.7e-04
EDM	4.1e-08	9.4e-03	4.2e-06
GRA	7.4e-05	7.3e-04	2e-03
HME	3.1e-05	1.5e-07	4e-04
LU	3.3e-05	9.9e-08	3e-04
MED	1.1e-04	2e-04	6.2e-05
MSE	4.8e-06	1.3e-07	3.2e-04
SE	2.2e-05	2.2e-03	2.1e-03
SAU	4e-08	1.4e-03	2.1e-07

The aim of the harvest model is not to provide post-sample forecast, but to replicate the historical expectations of harvest so as to account for weather conditions in the price model. In order to evaluate the credibility of the estimated series of expectations, I have checked its consistency with the recent history of real-time official aggregate forecasts produced by the French Ministry of Agriculture. The forecasts are based on surveys among wine producers, and are disclosed in monthly bulletins published at various months of the growth season. The available history of these forecasts on the website of the Ministry goes back to August 2013. Only the forecast of the total harvest of the Bordeaux region is disclosed all AOC combined. I thus have compared its accuracy with the outcome of the model for the total harvest across the 15 selected AOC. Table 4.6 gives comparative statistics for each observed date. The first two column give the year and the month at the beginning of which the forecasts of the next harvest are computed with the model. For instance, the first row refers to the forecasts of the harvest of vintage 2016 computed with the information available in November 2016. The in-sample percentage error of the model is given in column 3, and is on average of 3.42%. As expected, the error of the forecasts decreases from July to November for each year. Column 4 gives the publication dates of the official forecasts, and the last column contains the forecasts errors. The two errors are comparable in absolute values, with my model being less accurate for 2014 but much more precise for the small harvest of 2013, for which the official forecasts underperformed. On average, the model in-sample precision is better than that of the real-time official forecasts, with and average error of 3.42% against 5.82%. This suggest that the model overestimates the historical precision of the harvest expectation. However, the local agents that directly influence prices are likely to have better sources than official forecasts, so that the latter might underestimate the actual precision of the agents expectations. All in all, the model fitted values provide plausible estimates of the expectations of the next harvest for each month and each AOC.

TABLE 4.6 – Aggregate in-sample accuracy of the model and of official forecasts

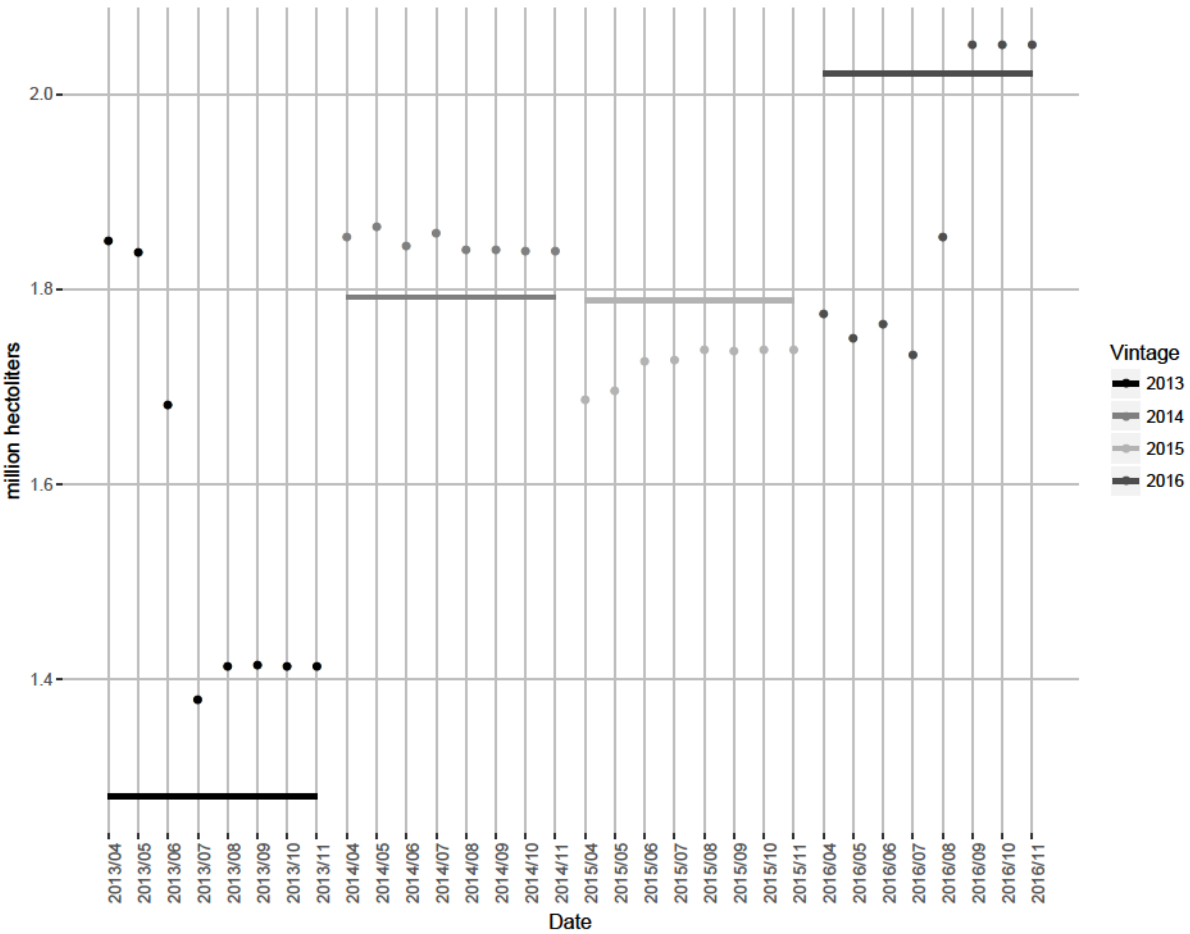
Year	Month	Model error (%)	Bulletin date	Bulletin error (%)
2016	11	1.54	01/11	-1.78
2016	10	1.54	01/10	-1.78
2016	9	2.87		
2016	8	6.59	22/08	3.78
2016	7	13.31	18/07	5.78
2015	11	-4.58	01/11	-1.39
2015	10	-4.58	01/10	-1.39
2015	9	-4.74		
2015	8	-4.53	19/08	-1.39
2015	7	-5.98	20/07	-1.39
2014	11	1.24	01/11	-4.69
2014	10	1.24	01/10	-4.95
2014	9	1.71		
2014	8	1.81	20/08	-5.93
2014	7	2.12	21/07	-4.74
2013	11	-1.11	01/11	-5.4
2013	10	-1.11	01/10	-10.29
2013	9	0.53	01/09	-10.29
2013	8	-1.17	01/08	-20.93
2013	7	6.17	01/07	-31.14
Average		3.42		5.82

Figure 4.3 plots the observed next harvest against the estimated monthly expectations for the main appellation BR. The model estimates that the small harvest of 2013 was already expected accounting from July 2013, because the weather had been very cold and rainy in June. Weather updates are thus correctly taken into account in the estimated expectations. For all months t and AOC i, let $E_t(H_i)$ be the expectation of the next harvest. This indicator is used as an explaining variable of the monthly prices in section 4.5. It is also used in the computation of the annual price forecast at the beginning of the marketing year in section 4.7.

4.4.2 Quantitative evaluation of quality

According to the literature in wine economics, quality plays an important role in the determination of prices, especially in the Bordeaux region. Although this chapter is not concerned with the analysis of cross-sectional prices, the fluctuation of wine quality across vintages can be expected to drive average prices by AOC. In this chapter, we show that prices are more driven by the quantity brought to the market than by the effect of vintage quality. To quantify this vintage effect, a numerical measure of quality is needed. Of course, no scoring method to evaluate quality is perfectly objective or unbiased. However, a few sources have acquired a solid reputation and thus influence the popular opinion, and eventually define the common knowledge about the quality of each vintage. For Bordeaux wines, the historical benchmarks are the scores attributed to the top-end wines by Robert Parker. Even though he has retired in 2013, his magazine, the Wine Advocate, continues to be influential. In the recent years, the Wine Spectator (WS) has become a serious challenger to the hegemony of the Wine Advocate, as well as a dozen of other lesser-known wine critics. In addition to individual scores for a small number of very well-known Châteaux, these two main sources publish average scores for each vintage

FIGURE 4.3 – Observed next harvest (lines) and estimated monthly expectation (dots) for appellation BR



and each subregion in Bordeaux. As a first source of quality scores of each vintage, I have used these vintage grids freely available on the websites of the Wine Advocate (WA, the magazine created by Robert Parker, who is now retired) and the WS[55]. In addition to these two sources, I have collected the harvest reports written by professors of the oenology faculty of Bordeaux since 2000, which provide first-hand and detailed information on the expected quality of each vintage just after the harvest[56]. These reports do not give numerical estimates of quality, but they do list five criteria as signals of favorable weather during the growth season (March to October)[57]. For instance, the first criterion is that the burst of the buds in March must be early and homogeneous. Based on these reports, I have attributed a value between -1 (not satisfied), 0 (average) and 1 (fully satisfied) to each criterion and each vintage for which these reports are available (2000-2016). The addition of these five scores gives a score between -5 and 5 for each vintage, which provides an objective evaluation of quality from the local professional oenologists. One advantage of the vintage grids of WE and WS is that they are more detailed by AOC, whereas the scores computed using the criteria are the same for all red wines. Table 4.7 presents how I have attributed the vintage scores of WA and WS among the fifteen selected AOC. The names of the subareas reported in the chart for WA and WS are those given on their respective websites. One major shortcoming of this attribution is that there exists no evaluation of the quality of the vintage for dry white wines of the AOC BW and EDM. I have considered the scores for the AOC Sauternes as a proxy variable for those two AOC, but it is admittedly far from perfect.

TABLE 4.7 – Matching the 15 AOC with the scores of WA, WS and harvest reports

AOC	WA	WS (>1995)	WS (<1995)	Reports
BLA	Margaux	Médoc, Pessac-Léognan	Bordeaux red	Bordeaux red
BR	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
BRO	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
BW	Barsac/Sauternes	Sauternes		
BSR	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
CAS	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
CBO	Margaux	Médoc, Pessac-Léognan	Bordeaux red	Bordeaux red
EDM	Barsac/Sauternes	Sauternes		
GRA	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
HME	Saint Julien/Pauillac/Saint Estephe	Médoc, Pessac-Léognan	Bordeaux red	Bordeaux red
LU	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
MED	Margaux	Médoc, Pessac-Léognan	Bordeaux red	Bordeaux red
MSE	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
SE	Saint-Emilion	Pomerol, Saint Emilion	Bordeaux red	Bordeaux red
SAU	Barsac/Sauternes	Sauternes		

The quality of each couple AOC-vintage is thus evaluated by up to three scores, two scores on a 100-point scale, and one on a 10-point scale (between -5 and 5). To obtain one single score for each couple AOC-vintage, I have followed the equipercentile matching method presented in Cardebat and Paroissien (2015). Hence, I obtain average scores for each each AOC-vintage. Now, the prices given in the data are averages over all transactions of a given period, which mixes all traded vintages together. The appropriate quality score for a given period is therefore the average among the scores of the vintages dealt during this period. Furthermore, the weights of each vintages should not be equal, since recent vintages have a larger share than

[55]I have also surveyed the vintage grids given by other sources, but none could provide additional information: their history is very short, less disaggregated among the AOC, and their scores are almost collinear with those the WS and the WA.

[56]See for instance Geny and Dubourdieu (2015), the harvest report of the vintage 2015.

[57]I thank Antoine Moga for having suggested me this first-hand indicator of quality.

old vintages. Unfortunately, I do not observe the precise shares of each vintage in the average prices. In order to approximate the evolution the shares of each vintage, I have assumed a weighting rule. The assumed flow function begins at zero after the harvest, increases and reaches a maximum between six and ten months after the harvest, and then slowly decreases until zero, with significant values during two to five years depending on the AOC. For each AOC, I have computed a specific monthly flow function parameterized by the delay between harvest and deliveries. Let n be the number of months since the harvest (September), and d the delay between the harvest and the start of the deliveries given in table 4.2, the weight $w(n, d)$ is given by the equation (4.7).

$$w(n, d) = \phi_{LN}(n, 1.5 + \frac{d}{10}, \frac{d}{10}) \tag{4.7}$$

where $\phi_{LN}(x, \mu, \sigma)$ is the density taken in x of log-normal density of parameter (μ, σ)[58]. The graphics of the monthly flow functions by AOC are given in the Appendix D. This weighting rule is admittedly somewhat arbitrary. But it has the merit of being consistent with the few known stylized facts about the flow of a vintage, namely: the flow first increases, reaches a maximum within the first year, and then decreases and disappears of the market in about five years[59]. Because I wish to estimate the vintage effect both on the annual data and on the monthly data[60], I have computed monthly and annual quality scores, respectively noted Q_{it} and Q_{iT}, for each AOC i. The monthly quality scores Q_{it} is computed as the weighted averages of the average scores of the last five vintages, the share of the fifth vintage being already negligible. The annual vintage score Q_{iT} is computed as the simple average of the monthly quality score over the marketing year. The latter is plotted on figure 4.4 for the two regional AOC BR and BW.

4.4.3 Leading indicators of the macroeconomic determinants

In the previous sections, I have focused on the supply-side drivers of prices. For the purpose of explaining the price history, it is obviously necessary to account for the variation in the drivers of demand. Three exogenous macroeconomic drivers of demand are considered: the revenue of consumers (as measured by the GDP of Bordeaux wine consuming countries), the exchange rates against the euro, and the quantity of wine production by the competitors of Bordeaux. For each of these three drivers, I have collected specific data and aggregated all the latter into one leading indicator. Lastly, I have also included the interest rate in the analysis, which defines the terms of the storage arbitrage and is therefore expected to shift both supply and demand.

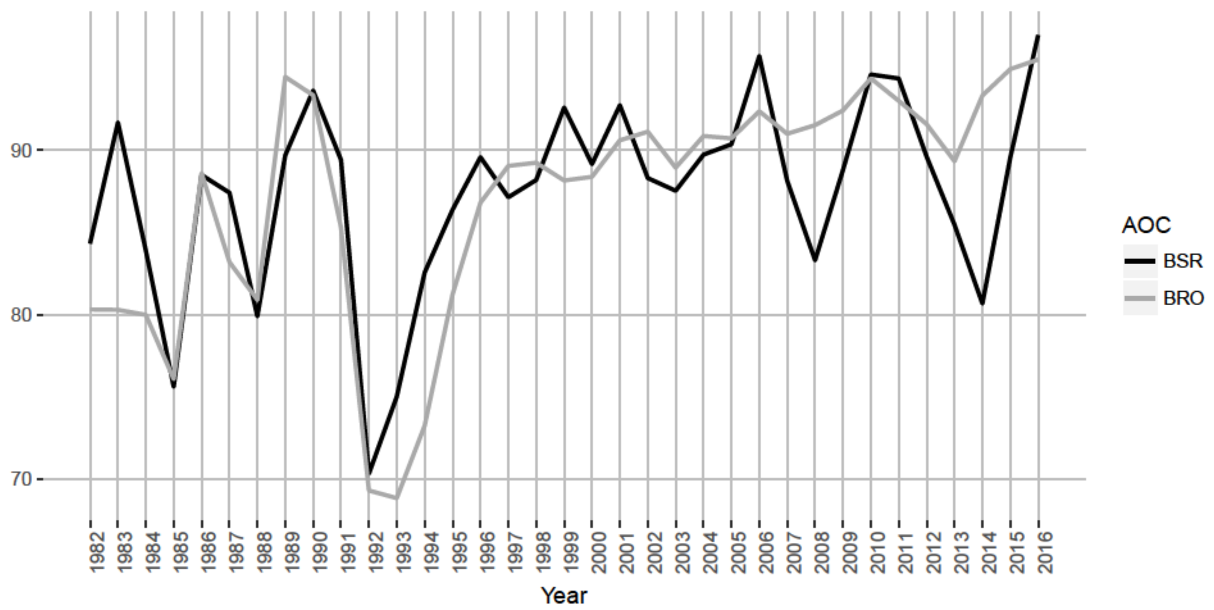
About 40% of Bordeaux wines are exported, and about one third are exported out of the euro zone. The value of the euro against the other currencies should then be an important demand shifter. So as to include this effect in the forecasting models, I have collected historical series for the main exchange rates against euro on the website fxtop.com, whose data come from the European Central Bank. I have collected the information for the period 1981-2016 at a monthly frequency for the following countries: Belgium, Cameroon, Canada, China, Germany, Denmark, Hong-Kong, Ireland, Japan, Latvia, Lithuania, Netherlands, New-Caledonia, Poland, Russia, Singapore, South Korea, Switzerland, Taiwan, the United Kingdom (UK) and

[58] In other words, $\phi_{LN}(x, \mu, \sigma)$ is the density of the exponential of a normal random variable of mean μ and standard deviation σ

[59] I take this information from the lifetime expertise of the wine brokers François Lillet and Marion Tarel.

[60] Recall that the annual data has a longer history and thus may be more relevant to estimate this kind of annual phenomenon.

FIGURE 4.4 – Annual indicator Q_{iT} of the quality for the two main AOC BR and BW



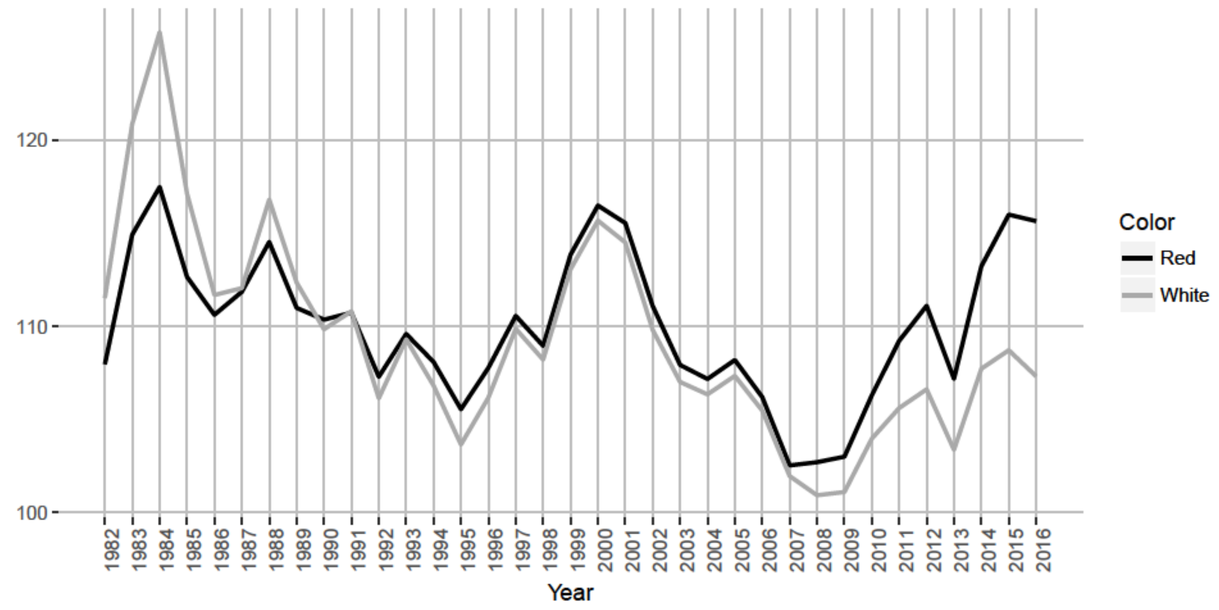
the United Sates of America (USA)[61]. These countries account for more than 90% of the exports in volume. The 21 monthly exchange rates have been aggregated into one single leading indicator for the value of the euro. This index is a Lapeyre index where the weights are the exported value to each country for the preceding month, or year for the annual index. The latter are lagged because they are arguably correlated with the price, and using the current exports in the construction of a determinant of the price would raise an endogeneity issue. Exports data are only color-specific, and that the exchange rate index is the same for all red AOC, and all white AOC. I hereafter denote E_{it} and E_{iT} the monthly and annual indicators of the exchange rates for each AOC i. The annual indicator E_{iT} is plotted on figure 4.5 for the two colors. The exchange rates are taken so as to reflect the strength of the other currencies. The higher the indicator, the weaker the French currency against the others, and the higher the Bordeaux wine prices are expected to be. The series is found well correlated with the series of annual prices for BR, see section 4.6.

In the same fashion, I account for the fluctuation of the wealth in the Bordeaux wine-consuming countries. I have collected the quarterly growth of the Gross Domestic Product (GDP) for France and 15 major Bordeaux wine-importing countries[62] on the website of the OECD. As for the exchange rates, these 16 GDP growths are averaged using the exports as weights, the weight of France being the difference between the volume exported and the total volume delivered. In order to be used as an explaining variable of the fluctuation of monthly price series, I have converted this quarterly series to a monthly frequency. For each month of a given quarter, I have simply approximated the monthly GDP growth by the cubic root of the quarterly GDP growth. As the exchange rate index, this revenue index is weighted by lagged

[61]The euro currency was launched in 2000 and several of these countries now share this currency with France. However, between 1982 and 2000, France used a national currency, the French Franc, which was allowed to fluctuate against the other european national currencies, in the limits fixed by the former European Monetary System.

[62]Belgium, Canada, China, Germany, Denmark, France, Ireland, Japan, Latvia, Lithuania, Netherlands, Poland, South Korea, Switzerland, the UK and the USA

FIGURE 4.5 – Annual indicator E_{iT} of the exchange rates for each color



Note: These annual indexes are obtained by averaging the monthly indexes, which both equal 100 in January 1982.

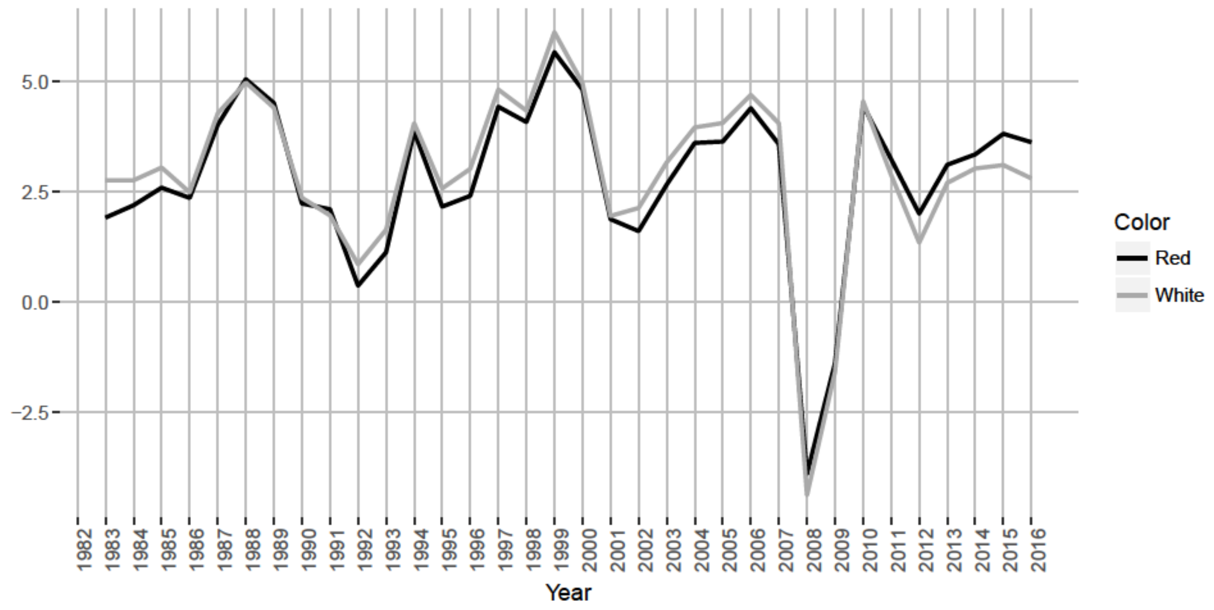
exports and is color-specific. I hereafter note respectively Y_{it} and Y_{iT} the monthly and annual indicators of the wealth of the consumers of wines from AOC i. I have represented the evolution of the growth of Y_{iT} on figure 4.6 for each of the two colors. The main stylized fact of the series of growths is the striking economic downturn of 2008.

Yet another determinant of prices at the macro level is the supply of its competitors, referring to the other wine regions in France and the other countries. For instance, when the other French wine regions enjoy a very large harvest, Bordeaux wine prices decrease because French wines are relatively close substitutes. Since Bordeaux wines are exported throughout the world, all the wine regions of the planet are potential competitors. As for the exchange rates and the GDP, I do not estimate the relative impact of the harvest of each competitor on Bordeaux wine prices, but rather compute a weighted average of the latter and add this leading indicator to the price equation. However, the weighting system for the competitors' harvests has required more attention. The goal is to estimate a weight for each competitor which reflects the intensity of the competition between the latter and the Bordeaux region. To do so, I have used the annual world wine trade matrices given on the website of the Food and Agriculture Organization of the United Nations (FAO), as well as the history of the harvests by country. The data is available on a long history, but there is delay of about two years before the national harvests are given. For the recent harvests of the ten most important wine-producing countries[63], I have used the half-yearly reports of the International Organization of Vine and Wine (OIV), freely available on its website. For the last harvest of 2016, I have consulted the world outlooks of the USDA and reports of the French Ministry of Agriculture[64].

[63] Argentina, Australia, Chile, China, France, Germany, Italy, South Africa, the USA and Spain

[64] The delay between actual harvests and the dates when the statistics are revealed on the FAO website is not taken into account in the forecasts presented in section 4.7. Obviously, the economics agents negotiating Bordeaux wine bulk prices are aware of the level of the wine production of the competitors way before the statistics are available on

FIGURE 4.6 – Annual indicator ΔY_{iT} of the GDP growth for each color



Note: This figure displays the indexes Y_{iT} taken in first difference and noted ΔY_{iT}. In the estimation procedure presented thereafter, I have considered both indexes taken in level or in first difference. These annual indexes are obtained by averaging the monthly indexes, which both equal 100 in January 1982.

I propose a weighting system for all wine-producing countries that takes into account the structure of the international wine trade. Let x_{ijT} be the volume exported from country i to country j during year T, and H^*_{iT} the wine production of country i during year T. I compute the m_{ijT} as the volume coming from country i and consumed in country j for all couples (i, j).

$$m_{ijT} = \begin{cases} x_{ijT} & \text{if } i \neq j \\ H^*_{iT} - \sum_{k \neq i} x_{ikT} & \text{if } i = j \end{cases}$$

The quantity produced and consumed in country i is the harvest minus the total exports[65]. For each country i, I then compute the share w_{ijT} representing the importance of the consumer market of each country j during year T, as follows:

$$w_{ijT} = \frac{m_{ijT}}{\sum_k m_{ikT}}$$

I finally estimate the intensity w_{iT} of the competition between country i and the Bordeaux region during year T by the average of the importance of each market j for production of country i, weighted by the importance market j for the production of Bordeaux. The latter, noted s_{jT}, is estimated by the share of Bordeaux wines deliveries consumed in country j[66]. Simply put, the weight of competitor i for year T is given by:

the FAO's website. Using the real statistics as if they were known without delay may cause a slight underestimation of the forecasting error in section 4.7.

[65]I therefore neglect the re-exportations. For platform countries like Singapore, the total wine exports actually exceeds the wine production. In this case, I attribute a value zero.

[66]Here again, I neglect the issue of re-exportations.

$$w_{iT} = \sum_j s_{jT} w_{ijT}$$

The index w_{iT} thus increases when the market power of country i increases in countries that are key destinations of Bordeaux wines. The leading indicator for global wine production is also a Lapeyre index of the fluctuation of the harvests for each country H^*_{iT}, weighted by the lagged w_{iT-1}. For the monthly index, I have first computed one value of the index by semester of the marketing year to take into account the respective timings of the grape harvest in both hemispheres. In semester 1 (from August to January), only the fluctuations of the harvests of the countries of the northern hemisphere affect the index. The grape harvests of the countries of the southern hemisphere[67] occur between January and May, so it only affect the index during the second semester. The total weights of the southern and northern hemispheres are summed to zero, so that the variations of the harvests in the northern hemisphere have more impact on the global competition index (during the first semester), than those of the southern hemisphere. This reflects the fact that northern wine productions (mostly the rest of France, Spain and Italy) are more direct competitors for the Bordeaux region. I obtain a monthly index by first attributing the same values by semester to each month of each semester, and then compute a 6-months moving average. One shortcoming of this index is that it ignores that the level of substitution between the wines produced in a given country and Bordeaux wines actually depends on how close those two are in terms of quality. To my knowledge, no quantitative indicator exists for comparing the quality of the wines produced for different country. I thus consider that the flows of international trade are sufficient indicators of the level of competition. The monthly and annual indicators are respectively noted C_{it} and C_{iT}. The annual indicator C_{iT} is plotted on figure 4.7 for the two colors. The decreasing trend is mainly due to the decrease of the French harvest, which has a weight of 0.7 on average in the index.

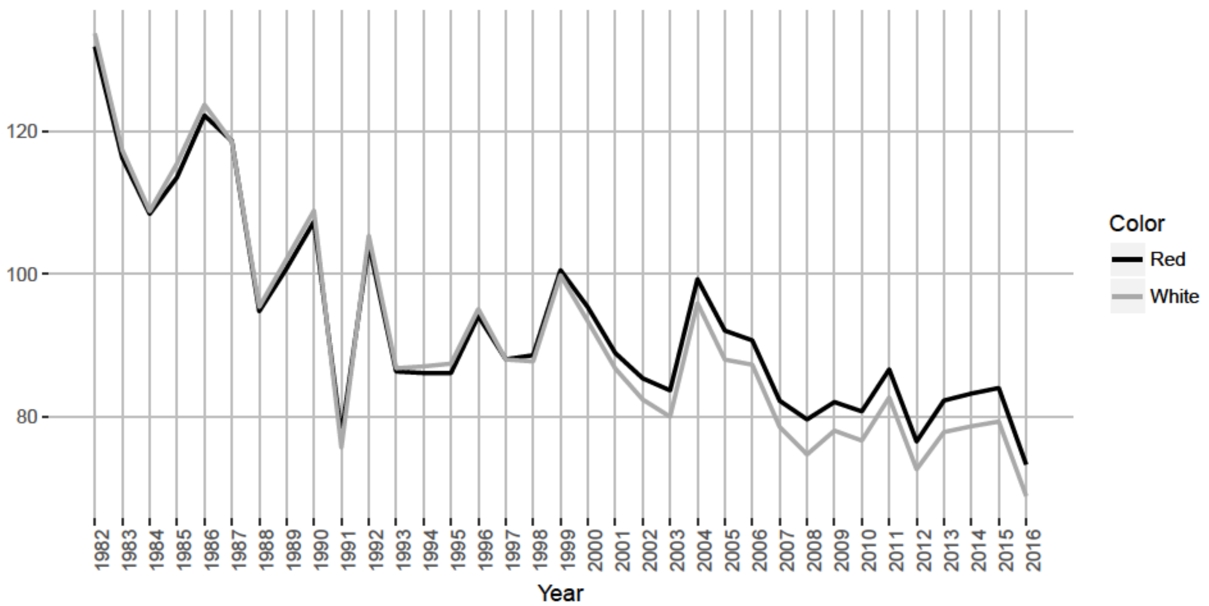
The last macroeconomic driver of Bordeaux wines prices I consider is the three months ahead interest rate EURIBOR computed by the *Banque de France* (BCF), the former French central bank. The available monthly history starts in January 1999. For the preceding years 1982-1998, I have used the legal interest rate also given by the BCF, that is precisely computed as the 12-months average of the three-month ahead interest rate. This interest rate is noted r_t for the monthly series, and r_T for the annual series. The annual indicator r_{iT} is plotted on figure 4.8.

Obviously, the exhaustive list of all the factors determining wine prices is endless. However, I trust that the above shortlist recapitulates those which are measurable, have significantly fluctuated over the sample period (1982-2016) and for which the available data have enough historical depth. One concern has been to guarantee that the wine professionals in Bordeaux can appropriate the procedure and use operational price forecasts, so that all the data used in this chapter are accessible to them[68].

[67]The only southern countries with a significant wine production are Argentina, Australia, Chile, Peru, New-Zealand, the Reunion Island, South Africa, Viet-Nam and Zimbabwe.

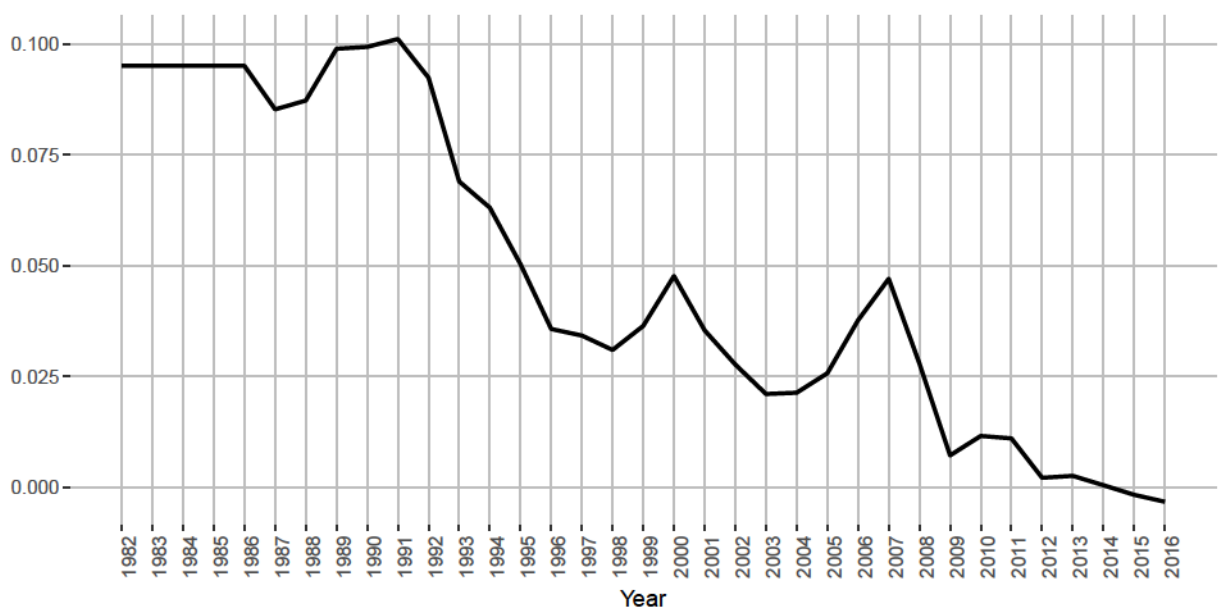
[68]I have also obtained homogeneous data on tariffs computed by the method proposed by Bouet et al. (2004). Unfortunately, these data are not publicly available and are only produced for five different years so far and are released with an important delay. However accurate, these data eventually turned out not to be helpful for the purpose of price forecasting. To my knowledge, the other sources of data on wine tariffs are even less convenient and lack historical depth. Without readily available data that can be appropriated by the professionals, I had to exclude wine tariffs from the analysis, although they arguably influence Bordeaux wine production prices to some extent.

FIGURE 4.7 – Annual indicator C_{iT} of the harvest of the competitors for each color



Note: These annual indexes are obtained by averaging the monthly indexes, which both equal 100 in January 1982.

FIGURE 4.8 – Annual indicator r_T of interest rates in France



Note: This is the reference interest rate in France for obligations due in three months, and is not corrected for inflation. A value of 0.1 stands for a rate of 10%

4.5 Model selection

4.5.1 Annual model

The two series of monthly and annual prices have demanded specific treatments to obtain effective forecasting models. I start by presenting the modeling strategy for the annual prices. Obviously, the monthly model developed in the next section allow a closer look at the price dynamics. However, because the annual data has a longer history, the latter may allow a better estimation of the respective influences of the drivers (especially for the harvest) and therefore provide better forecasts for annual prices. The respective performances of the annual model and the monthly model are presented in section 4.7.

Let P_{iT} be the average bulk prices dealt for wines of AOC i during marketing year T. I consider the following drivers for P_{iT} : the starting inventory S_{iT}, the quantity harvested H_{iT}, the representative quality Q_{iT}, the indicator E_{iT} of the value of the euro against the currencies of the main importers, the indicator Y_{iT} of the GDP of the main countries consuming wines of the AOC i, the indicator C_{iT} of the quantity harvested by the competitors, and the interest rates r_T. I divide S_{iT} and H_{iT} by the quantities delivered during the past year so as to obtain the indicators sur_{iT} and hur_{iT}, which should be better predictors according to the findings of section 4.3. Recall also from section 4.3 that the price series are highly autocorrelated, suggesting that P_{iT} may be partly determined by P_{iT-1}, even after controlling for the exogenous variables. P_{iT-1} is thus included in the set of the determinants. Finally, the four Lapeyre indexes Q_{iT}, E_{iT}, Y_{iT} and C_{iT} are taken in logarithm, and respectively noted q_{iT}, e_{iT}, y_{iT} and c_{iT}. The basic framework for the annual price model P_{iT} is given by equation (4.8)

$$P_{iT} = f_i(P_{iT-1}, sur_{iT}, hur_{iT}, q_{iT}, e_{iT}, y_{iT}, c_{iT}, r_T) + \epsilon_{iT} \tag{4.8}$$

where f_i is a function specific to AOC i, and ϵ_{iT} is the error term. Because our main purpose is forecasting, the objective is to find the function f_i for which the expected next forecast error is minimal. The search for this optimal function f_i is restricted to the space of linear and log-linear functions of the explanatory variables and their lags. The expected next error is evaluated by cross-validation by the average of the last L post-sample forecast errors. Following Armstrong and Collopy (1992), I scale the post-sample error $\tilde{\epsilon}_{iT}$ of a given year T by the relative absolute error, noted RAE_T and given by equation (4.9).

$$RAE_T = \left| \frac{\tilde{\epsilon}_{iT}}{P_{iT} - P_{iT-1}} \right| \tag{4.9}$$

The forecasting error is scaled by the error of the random walk forecast, also sometimes referred to as the naive forecast[69]. Indeed, a forecast is helpful to the extend that it is more precise than the random walk forecast. The accuracy of the forecasts must then be evaluated relatively to the accuracy of the random walk forecast, which is achieved in formula (4.9). A forecasting model with a RAE_T below 1 manages to predict the correct direction of price changes, and with a limited absolute error. In the analysis of forecasts accuracy in section 4.7, I also comment the absolute percentage errors, but the latter is not used the model selection. As advised by Armstrong and Collopy (1992), I also censor the RAE_T between 0.1 and 10 to temper the impact of the outliers. Let $\bar{T}$ be the last year of the sample period. For a given length L of the test window, the loss function used for selecting f_i is the geometric mean over the test window

[69] Another possibility would have been to consider the last observed monthly prices as the naive forecast. However, monthly prices are seasonal and very volatile, so that the last annual price is actually a more accurate annual forecast. Taking the last annual price as the naive forecast is therefore more challenging.

$\{\bar{T}-L+1,...,\bar{T}\}$ of the winsorized RAE_T on the interval $[0,10]$. This loss function is noted $MRAE_L$ and formally given by equation (4.10).

$$MRAE_L = \prod_{T=\bar{T}-L+1}^{\bar{T}} max(0.1\{min\{RAE_T, 10\}\} \tag{4.10}$$

A stream of the literature on forecasting argue in favor of combining forecasts from different models (see Timmermann (2006) for a review). I therefore consider annual price forecasts given by an average of the k models for which $MRAE_L$ is the lowest. The choice of the parameters (L,k) is based on a sensitivity analysis presented in section 4.7.

Remains the scope of the forecasting models to be considered and evaluated. Recall from section 4.3 that the price-smoothing role of storage is consistent with the time series attributes of the annual price series, so that the competitive storage model may be viewed as a natural option. Although this refined microstructure is theoretically appealing, it is actually not adapted to the purpose of this chapter in the current state of research. First, the resolution of the model is only readily available in very simple specifications[70], and requires to input several additional parameters. Second, the estimation of these additional parameters on real data has only been achieved for standard commodity markets in even simpler specifications,[71], only using price data[72] and never in the aim of operational price forecasting. By contrast, several features of the wine market are quite peculiar, such as the important pipeline stocks (see section 4.3), the vintage effect, the fact that real demand is stochastic, the lag of the supply response and the autocorrelation of the harvest. These specific dynamics compromise the application of the theory of storage in its current state of research[73]. Most importantly, it is uncertain that such a heavy structure would produce better price forecasts than standard and readily available forecasting models from time series econometrics. Furthermore, the latter come at a much reduced price in terms of computing complexity, which is a important asset for cross-validation. Loosely speaking, the empirical research on the competitive storage model is yet not mature enough to make it a better option than time series econometrics for operational forecasting. Hence, I here adopt a data-oriented approach and limit the scope of the estimations to flexible time series models.

The best practice advocated by the literature in time series econometrics is to first test for the stationarity of the series. Indeed, estimating time series models on non-stationary data may lead to spurious findings (Granger and Newbold, 1974). Series that are found non-stationary because of a stochastic trend should then be considered in first-difference in the estimation of the model[74]. Table 4.8 gives the results of the classical Augmented Dickey-Füller test (ADF), the more powerful test of Elliott et al. (1996) (ERS), and that of Kwiatkowski et al. (1992) (KPSS)[75]. Unfortunately, the results are mostly inconclusive because these three tests rarely agree. The ERS test is more powerful than the ADF test, so that it tends to reject non-stationarity more often. However, for the AOC BRO the ADF test rejects the non-stationarity hypothesis at the 5% level but the ERS test does not. For the AOC BLA, non-stationarity is rejected at the level

[70]Christophe Gouel provides a free solver for MATLAB at http://www.recs-solver.org for the canonical competitive storage model.

[71]See Deaton and Laroque (1992, 1996), Chambers and Bailey (1996), Cafiero et al. (2011), Guerra et al. (2014) and Gouel and Legrand (2017b) for staple commodity markets. Osborne (2004) somehow departs from this literature in using intra-annual price data of the Ethiopian grain market. In all these models, demand is deterministic and supply is inelastic.

[72]Gouel and Legrand (2017a) is a recent and notable exception, allowing for an elastic supply.

[73]Adapting the existing methods for the resolution an the estimation of the competitive storage model is beyond the scope of this chapter.

[74]When non-stationarity is caused by a deterministic trend, the latter can be estimated together with the other coefficient on the raw series.

[75]The data series are real prices, so no deterministic trend are considered.

10% by the ERS test, but stationarity is also rejected at the same level by the KPSS test. Because none of the two assumptions seems to be definitely better suited to the annual price series, I have estimated both models for stationary and non-stationary series for all AOC[76].

TABLE 4.8 – Unit root tests on the annual series of prices by AOC

Code	ADF	ERS	KPSS
AOC		Test statistics	
BLA	-0.135	-1.861	0.456
BSR	0.043	-1.739	0.338
BR	-0.029	-1.699	0.43
BRO	-0.251	-1.659	0.614
BW	0.277	-2.196	0.419
CAS	0.174	-2.386	0.462
CBO	-0.083	-1.865	0.507
EDM	0.264	-2.572	0.25
GRA	0.012	-1.76	0.427
HME	-0.09	-2.474	0.266
LU	0.178	-1.81	0.09
MED	0.064	-1.73	0.258
MSE	0.171	-2.007	0.073
SE	0.38	-1.371	0.231
SAU	0.204	-1.298	0.238
Level		Critical values	
1%	-2.62	-2.63	0.739
5%	-1.95	-1.95	0.463
10%	-1.61	-1.62	0.347
H_0	non-stationnarity		stationnarity

For stationary series, the standard model is the Autoregressive Distributed Lag (ADL) model. In an ADL(p,q) model, the function f_i is simply taken as linear with respectively p lags on the dependent variable and q on the exogenous drivers. Given the limitation of the data, I assume that the market conditions of year $T-2$ cannot improve the forecasts of the annual price of year T, so that I have taken $p = q = 1$. Hence, only one lag is included in every model considered for the annual price series. The same framework applies when the series is viewed as non-stationary, but for the data taken in first difference. Hereafter, the ADL(1,1) model in first difference is noted DADL(1,1). Finally, I include standard error-correction models (ECM) to the set of tested models, which are often referred to as the best option for forecasting purpose (Hendry and Clements, 2003)[77]. All three models have been estimated both separately for each of the fifteen AOC, and taking all the AOC together. The ADL(1,1) and the DADL(1,1) models are estimated by the ordinary least squares method (OLS) and the ECM(1,1) are estimated by the two-stages least squares method (2SLS) introduced by Engle and Granger (1987)[78]. A short presentation of these standard models is given in the Appendix D.

[76]This is the reason why I do not present unit root tests for the exogenous variables.

[77]Because one year-lagged substitution effect are likely to be negligible, I have not considered the Vector Autoregressive Models (VAR).

[78]I do not use the maximum likelihood estimator of Johansen (1991). Although the latter is often considered more robust, its main advantage lies in allowing for several endogenous variables, while only one variable is endogenous in my case. Besides, the computation procedure involves a numerical optimization which is considerably longer than the 2SLS estimation method of Engle and Granger (1987). This is a crucial advantage for the cross-validation analysis because computation time becomes rapidly prohibitive. Furthermore, the gains in forecasts accuracy are likely to be small since both procedures would estimate the very same price equation. I however acknowledge that without constraints on the estimation time, it is possible that using the procedure of Johansen (1991) would further improve the forecasting accuracy.

Depending on the AOC, some variables may have too little explanatory power to actually improve the forecasts. This is a classic case of the bias-variance trade-off: removing one variable increases the bias of the model, but it also reduces the variance of the other estimates and may therefore improve the overall forecast accuracy (see Shmueli (2010) for a detailed discussion). Obviously, all the considered determinants do intervene at some point in the formation of the observed prices. However, when one driver has too little explanatory power, it is actually counter-productive to include the latter in the model. The annual data size being small, computations are fast enough to consider all possible subsets of the eight explaining variables for each of the three model types (ADL(1,1), DADL(1,1) and ECM(1,1)), and for each of the fifteen AOC[79]. In each estimation, I also check that estimates of the coefficients take the expected sign. Indeed, when the number of degrees of freedom is reduced, some estimates may take a sign that is inconsistent with economic theory. To prevent the minimization of the loss function to select irrelevant forecasting models[80], I manually constrain the estimates to take the signs expected by theory. Classically, I assume that prices are high when stocks are low, when harvests are small, when quality is high, when the euro is low, when the GDP is important, when the competitors' harvests are low and when the interest rate is low[81]. For each of these variables, the corresponding estimates are constrained to take the expected signs. In addition, I restrain the lagged prices to have a positive coefficient. If some autocorrelation of the price process is left unexplained by the autocorrelation of the exogenous drivers, the autoregressive coefficient must be positive. All these conditions ensures that all the estimates account for a theoretically-based phenomenon, so that all the price forecasts are economically grounded[82]. For a given estimation, if at least one estimate takes an unexpected sign I do not compute the post-sample forecasts and this specification is eliminated of the search for the best k models.

Because the more explaining variables are included the more likely one estimate takes an unexpected sign, I limit the estimations to the subsets of five or less different variables. When more variables are included, the estimates have very few chances to take correct signs, and the resulting model has very poor forecast performance anyway. The total number of subsets of variables is thus $\sum_{k=1}^{5}\binom{8}{k} = 218$. For each AOC, I consider ADL(1,1), DADL(1,1) and ECM(1,1) models with each of the 218 subsets, estimated either only on the data of the AOC or on all AOC together. In each case, the explained variable is either the nominal price or the real price, and this variable is either taken in logarithm or not. Hence, a total of 5,232 different forecasting models are estimated for each AOC. For each of those, I compute the loss function $MRAE_L$[83]. The annual forecast is the average of the forecasts of these k models weighted by their $MRAE_L$.

4.5.2 Monthly model

The treatment of the monthly data is quite different. First, as mentioned in section 4.3, certain monthly prices are not fully representative. This is because the terms of the transactions and the qualities of the wines sold during one month can be somewhat heterogeneous, even within

[79]The same subset of variables is included both contemporaneously and with one lag. In the ECM, I consider the same variables in the long-run equation and in the short-run equation.

[80]Even economically irrelevant models may by chance generate accurate price forecasts on the sample period.

[81]As for any other product, the lower the interest rate the less profitable it is to save money over buying a wine, so demand increases when the interest rate drops. Symmetrically, when the interest rate drops, the producers have less incentive to sell and supply decreases. Both effects go in the direction of a price increase when interest rates drop.

[82]This set of conditions is also the reason why I do not use the machine learning techniques emerging in the agricultural price forecasting literature (Jha and Sinha, 2013; Yeo et al., 2015). In these generation of models, the derivatives are typically intractable and there is thus no possibility to control for their signs.

[83]The computation of the $MRAE_L$, because it involves post-sample errors, necessitate to estimate the model L times.

one AOC[84]. When the number of transactions is reasonably large, the average price can be expected to be representative of the average market conditions. However, when the number of transactions is low, the few observed prices may not be representative of the overall market conditions. This lack of representativity causes the observed monthly average prices to fluctuate around the true unobserved market price level. A part of the monthly fluctuations are thus irrelevant for forecasting. Cardebat and Bazen (2016) have considered univariate forecasts using the UCM framework described in section 4.4.1 for the forecast of the cultivated area. I build on their results and estimate the unobserved underlying levels of monthly prices with AOC-specific UCM. The relevant specification is a local-trend UCM including a seasonal component. Let P_{it} be the average of bulk prices dealt during month t for AOC i, and p_{it} its natural logarithm. The decomposition of the price is given by the system (4.11).

$$\begin{cases} p_{it} & = p^*_{it} + \gamma_{it} + \epsilon^p_{it} \\ p^*_{it} & = p^*_{it-1} + \beta^p_{it} + \nu^p_{it} \\ \beta^p_{it} & = \beta^p_{it-1} + \xi^p_{it} \\ \sum_{k=0}^{11} \gamma_{it-k} & = \omega^p_{it} \\ \epsilon^p_{it} & \sim \mathcal{N}(0, \sigma^2_{\epsilon^p_i}) \\ \nu^p_{it} & \sim \mathcal{N}(0, \sigma^2_{\nu^p_i}) \\ \xi^p_{it} & \sim \mathcal{N}(0, \sigma^2_{\xi^p_i}) \\ \omega^p_{it} & \sim \mathcal{N}(0, \sigma^2_{\omega^p_i}) \end{cases} \tag{4.11}$$

By maximum likelihood, I estimate the series of the underlying price level $\hat{p}^*_{it}$ which is more regular than the observed raw monthly average prices p_{it}. These $\hat{p}^*_{it}$ extract the core information of the observed raw price series by filtering out the seasonal and irregular components of the price series, respectively γ_{it} and ϵ^p_{it} in the system (4.11). The estimated $\hat{\gamma}_{it}$ account for seasonal variations in the transaction conditions, essentially the delivery delays, and the share of medaled wines[85]. The estimated parameters $(\hat{\sigma}^2_{\epsilon^p_i}, \hat{\sigma}^2_{\nu^p_i}, \hat{\sigma}^2_{\xi^p_i}, \hat{\sigma}^2_{\omega^p_i})$ and the estimated series $\hat{p}^*_{it}$ are given in the Appendix D for each AOC.

The estimation of the $\hat{p}^*_{it}$ is the first stage of the estimation of the monthly forecasting model. The UCM models are here only used as filters to estimate the smoother series $\hat{p}^*_{it}$. Now, I do not suppose that these $\hat{p}^*_{it}$ follow a random walk as given by the second equation of the system (4.11). In a second stage, I model the $\hat{p}^*_{it}$ and account for the influence of the exogenous determinants. This framework follows and extends the analysis of Cardebat and Bazen (2016). I suppose that their fluctuations can be explained by their lagged values and monthly versions of the annual exogenous variables. The appropriate computation of monthly versions of the exogenous variables has demanded some attention, especially for the starting stocks and the harvest expectations.

The total physical stocks S_{it} held at the beginning of month t for AOC i are obtained using the annual declaration of stocks, the harvests, and the monthly physical deliveries. The quantity harvested is added to the monthly stock at the end of the harvesting period, at the beginning of

[84]The delivery delays, possible medals won at wine competitions (see chapter 3 of this dissertation) and other various bargaining power can explain the heterogeneity of the monthly distribution of prices. Note that all transacted wines have been previously tasted by the buyer, so that information asymmetry is limited. The fact that quality is somewhat heterogeneous within an AOC does not generate information asymmetry at this stage of the chain value, but it eventually does at the retail level.

[85]Most wine competitions occur in spring, causing a seasonal increase of the average price around the end of the marketing year.

November[86]. The resulting series of raw monthly stocks is highly seasonal, and its seasonality is irregular over the sample period. This feature complicates the measure of its influence on the deseasonalized series $\hat{p}^*_{it}$ and hence the forecasts. I therefore also deseasonalize this stock series with a local-trend UCM including a seasonal component, using the form detailed in the system (4.11). As for prices, I have estimated a UCM for each AOC separately, using the logarithm of the monthly stocks (s_{it})[87]. The estimated parameters $(\hat{\sigma}^2_{\epsilon^s_i}, \hat{\sigma}^2_{\nu^s_i}, \hat{\sigma}^2_{\xi^s_i}, \hat{\sigma}^2_{\omega^s_i})$ are given in the Appendix D for each AOC. The estimated levels of the logarithm of the stocks $\hat{s}^*_{it}$ are used as an predictor of the series of the price levels $\hat{p}^*_{it}$[88].

Recall that in the annual model, one key explaining variable is the quantity harvested. Obviously, no equivalent at the monthly frequency. As explained above, the harvest is simply added to the the monthly stock in November. Besides, the finer monthly frequency allows to control for another driver of the price: the expectation of the upcoming harvest. The model described in section 4.4.1 is used to generate monthly series of expectations of the next harvest. For a given month t, the next harvest is that due between $t+1$ and $t+12$. Instead of using this raw series, I here consider the expected stocks at the beginning of month $t+12$, noted $\widehat{ES}_{it+12}$[89]. These expected stocks account for both the expected deliveries and the expected harvest, the latter being determined by the weather. The computation of the expected stocks twelve months ahead requires the expectation of the deliveries over the same period. For a given month t, let ed_{it+h} be the expectation taken at the beginning of month t of the deliveries of months $t+h$ for AOC i. For each month t of the sample period and each AOC i, the expectations ed_{it} to ed_{it+11} are approximated by the 12-months ahead forecasts of a standard local-trend UCM model estimated on the monthly series of deliveries up to month $t-1$[90]. The estimates $(\hat{\sigma}^2_{\epsilon^d_i}, \hat{\sigma}^2_{\nu^d_i}, \hat{\sigma}^2_{\xi^d_i}, \hat{\sigma}^2_{\omega^d_i})$ of the latter UCM computed for each AOC on the whole sample period are given in the Appendix D, as well as the histories of the raw and filtered monthly deliveries[91]. The resulting

[86]Note that the figures of the quantity harvested is actually not yet known by November, but only around March. In the evaluation of the forecasts, I use the expected harvest estimated by the model of section 4.4.1 to compute the expected monthly stocks.

[87]Here the main purpose of the UCM filter is to deseasonalize the series of the stocks, whereas its main purpose in the case of prices was to remove the irregular component.Removing the seasonality of the series could also have been achieved using the simple 12-lag differencing operator ($\Delta_{12}s_{it} = s_{it} - s_{it-12}$), but that would have removed 12 months from the data, and would also have removed the trend. The UCM filter manage to deseasonalize the series both keeping the whole observed sample, and keeping the trend.

[88]Recall that in the annual model, stocks are divided by past annual deliveries so as to obtain the SUR, which are more correlated with prices. At the monthly frequency, the equivalent would be to divide the monthly stocks by the total deliveries between the same month of the preceding year and the last harvest. However, because of the high pipeline stocks, these monthly SUR become disproportionate just before the harvest. The stocks held at the beginning of the month preceding the arrival of the harvest are way larger than the average monthly deliveries, so that the ratio of the former on the latter takes large values. When trying to apply the UCM filter on these monthly SUR, the estimated levels are highly irregular and are less suitable to forecast prices. Besides, using the SUR instead of the raw stocks is useful to the extent that the level of the deliveries vary over the sample period. The annual deliveries did increase considerably during the period 1982-2000, jointly with the production, so that the use of the SUR is relevant for the annual data. However, the deliveries are more stable since the downturn of the late 1990s, and the monthly data only starts in 2001. Therefore, I do not rescale the estimated levels of the stocks $\hat{s}^*_{it}$ to account for the variations of the levels of deliveries in the monthly models.

[89]Adding the raw series of expected harvest to the explaining variables leads to less accurate forecasts. This is because the series expected harvest has irregular jumps just after the month of the harvest. The series of expected stocks has jumps too, but those are more regular and thus can be easily removed by filtering.

[90]For the first months, there are obviously not enough past data to estimate the filtering parameters. I therefore assume a perfect forecasts for the 72 first months, and use the actual deliveries of month $t+h$ as the expectations at the beginning of month t. One other solution would have been to remove the first years from the data. Given that the sample is already somehow short to estimate the multivariate models, the former approximation is a better option. Furthermore, the fluctuations of deliveries are very regular, so that agents' expectations about future deliveries should in fact not be far from perfect.

[91]One constraint not enforced in the estimation is that expected deliveries cannot exceed initial stocks. This only happens in 12 occasions over the 15,480 delivery forecasts, and only during the 2013 shortage for the appellation

series of expected stocks $\widehat{ES}_{it+12}$ looks much alike the series of the monthly stocks S_{it} in terms of seasonality. In particular, the jumps of the curves after each harvest make it difficult to assess the actual underlying trend of the expectations about future market conditions. Here again, I filter the series of the logarithm of expected monthly stocks in month t, noted $\widehat{es}_{it+12}$. The estimated levels of the latter are noted $\widehat{es}^*_{it+12}$ and are included to the set of explaining variables of the series $\hat{p}^*_{it}$.

The rest of the price drivers are similar to those of the annual model: the representative quality q_{it}, the indicator e_{it} of the exchange rates, the indicator y_{it} of the consumers' wealth, the indicator c_{it} of the quantity harvested by the competitors, all taken in logarithm, and the interest rates r_t. Contrary to the annual model, I consider lagged values of the explaining variables up to last three months[92], as well as for the dependent variable. Let z_{it} be the vector of all the considered price determinants $(s^*_{it}, es^*_{it}, q_{it}, e_{it}, y_{it}, c_{it}, r_t)$. The system of equations (4.12) recapitulates the modeling strategy for the price series at the monthly frequency.

$$\begin{cases} p_{it} & = p^*_{it} + \gamma_{it} + \epsilon^p_{it} \\ p^*_{it} & = g_i(p^*_{t-1}, p^*_{t-2}, p^*_{t-3}, z_t, z_{t-1}, z_{t-2}, z_{t-3}) + \rho_{it} \\ \sum\limits_{k=0}^{11} \gamma_{it-k} & = \omega^p_{it} \\ \epsilon^p_{it} & \sim \mathcal{N}(0, \sigma^2_{\epsilon^p_i}) \\ \omega^p_{it} & \sim \mathcal{N}(0, \sigma^2_{\omega^p_i}) \end{cases} \tag{4.12}$$

The variables p^*_{it}, γ_{it}, s^*_{it} and es^*_{it} are not directly observed in the data, but only estimated by filtering and approximated by $(\hat{p}^*_{it}, \hat{\gamma}_{it}, \hat{s}^*_{it}, \widehat{es}^*_{it})$. Recall that the main objective of the price model is to produce accurate forecasts, not to estimate supposedly structural parameters. The error-in-variable problem is thus not a concern for the main objective of this chapter. However, I do acknowledge that the estimates of the coefficients given in section 4.6 may somewhat underestimate the influences of the stocks and of the expectations on current prices due to the attenuation bias.

As in the annual case, I now turn to the specification of the price function g_i for each AOC. I first test whether the monthly price series can be considered stationary. Table 4.9 gives results of the unit root tests for the series of price levels $\hat{p}^*_{it}$. Contrary to those of table 4.8 for the annual prices, the results here are clearly in favor of the non-stationary hypothesis. I thus restrain the search for the best model among the models suited to non-stationary series: DADL(3,3) and ECM(3,3)[93].

The selection of the best forecasting model at the monthly frequency cannot be conducted the same way as in the annual case. Indeed, minimizing a cross-validation loss function across all models and all possible subsets of predictors is no longer a reasonable option because it would be too time-consuming. First, the total number of explaining variables considered is here of 31 against only 8 in the case of the annual model. Second, the monthly data is strong of 190 months for each AOC, compared to only 35 years for the annual data, which makes each estimation longer. Third, the cross-validation loss function should be computed on much more periods to be consistent. For instance, computing the $MRAE$ for a given monthly model on the last two years necessitate to estimate the model 24 times, against only 2 at the annual frequency. This case with many variables and observation periods is actually most common in forecasting, so that simpler model selection procedures are usual.

BRO and for long-term forecasts (at least 6-months ahead). In those few cases, the expectation of delivery is taken as only half of the expected stock.

[92] This number is sufficient since fewer lags are eventually found relevant in the estimation, see section 4.6.

[93] Recall that I consider up to three lags both for prices and for the exogenous variables.

The first choice concerns the type of model: ECM(3,3) or DADL(3,3). Even without comparing the forecasts performances of each specification, there are several rationales to prefer the ECM over the DADL framework. First, the existence of a cointegration relationship is never rejected by the data at conventional levels. Second, the forecasting literature tends to conclude that ECM models are usually superior, especially for long-term forecasts (Engle and Yoo, 1987; Hoffman and Rasche, 1996). Even if the latter result is likely to be case-dependent, I also tend to prefer the ECM framework because it makes a more comprehensive use of the data structure. Indeed, the 2SLS estimation of an ECM takes both into account the structure of the raw data and that of the data taken in first difference, while the OLS estimation of a DADL model only uses the latter. For all these reasons, I only consider ECM(3,3) to model the series of monthly price levels $\hat{p}^*_{it}$.

TABLE 4.9 – Unit root tests on the unobserved levels $\hat{p}^*_{it}$ of monthly prices

Code	ADF	ERS	KPSS
AOC		Test statistics	
BLA	-0.29	-0.766	0.846
BSR	-0.163	-0.935	0.823
BR	0.765	-0.192	0.594
BRO	-0.436	-0.815	1.003
BW	-0.078	-0.938	0.822
CAS	0.489	-1.308	0.966
CBO	-0.561	-0.674	0.834
EDM	0.428	-1.303	0.341
GRA	0.029	-0.736	0.746
HME	-0.069	-0.968	1.806
LU	-0.324	-1.486	0.544
MED	-0.127	-1.092	0.842
MSE	-0.128	-1.84	0.572
SE	0.177	-1.685	0.381
SAU	-0.179	-1.492	1.745
Level		Critical values	
1%	-2.58	-2.58	0.739
5%	-1.95	-1.94	0.463
10%	-1.62	-1.62	0.347
H_0	non-stationnarity		stationnarity

The second choice is the set of explaining variables to include in the ECM. Obviously, all the 31 explaining variables of the set $(p^*_{t-1}, p^*_{t-2}, p^*_{t-3}, z_t, z_{t-1}, z_{t-2}, z_{t-3})$ cannot be included together in the models. This is because many estimates would then take inconsistent signs, and the estimation variance would be too large for the model to produce accurate forecast[94]. For the annual model, the loss function used to select the most relevant subsets of variable is the MRAE computed on the last L years. At the monthly frequency, the computation of this loss function becomes too time-consuming. The main reason is that the number of variables has tripled so that the number of possible subsets is now of 206,367 (even I limit to five the number of included variables as for the annual data). Furthermore, a cross-validation on the same window length of L years requires 12 times as many estimations. Hence, I abandon the cross-validation strategy for the selection of the explaining variables, and opt for the minimization of a more straightforward loss function. I follow the standard practice and consider the classical information criteria proposed by Akaike (1973) and Schwarz (1978), respectively noted AIC and BIC. These information criteria are a form of penalized log-likelihood. Their two formulas are given below.

[94]See Shmueli (2010) for a discussion of the bias-variance trade-off in the context of forecasting.

$$AIC = -2\lambda - 2k$$
$$BIC = -2\lambda - ln(N)k$$

where λ is the log-likelihood of the model, k is the number of degrees of freedom and N is the number of observations. The computation of AIC and BIC is straightforward under the assumption that the error terms are independent and follow the same normal distribution. Among different models, the one with the lowest information criterion is the most suitable to the data in the sense that it has a relatively high likelihood and a relatively low number of parameters, and therefore little risk of overfitting. The optimal weighting between the two objectives of maximizing the likelihood and minimizing the risk of overfitting is an illustration of the classical bias-variance trade-off in statistics. Both AIC and BIC are optimal in certain regards, and choosing which of the two criteria to minimize is sometimes referred to as the *AIC-BIC dilemna* (see Arlot and Celisse 2010 for a survey). On the one hand the BIC is consistent in the sense that its minimization would asymptotically select the real data generating process, if the latter is actually among the considered models. The AIC is not consistent, but minimizing the AIC is asymptotically equivalent minimizing the mean square error of the post-sample one-step ahead forecasts, which is another appealing property. Practically, the BIC puts a heavier weight on the number of degrees of freedom k, so that its minimization selects as many or fewer terms than that of the AIC. Each criterion is optimal in a certain sense, with no consensus about which is best. In my context, computations show that minimizing the BIC leads to the best forecast accuracy on the last five years, so that the results presented in section 4.7 are obtained with models selected by minimizing the BIC. However, minimizing the BIC also selects too few variables for a fruitful discussion of the significance of each predictor on the past history. Indeed, certain explaining variables are not selected by the BIC although they do exhibit a significant influence on the price history. Even if including the latter in the models does not allow to improve the forecasts[95], their statistical significance is informative for explanatory purpose. The specification of the monthly models presented in section 4.6 are thus obtained by minimizing the AIC.

Contrary to the annual case, I do not constraint the subsets of explaining variables to be the same in both the long-run and short-run equations of the ECM. In each equation, the variables are selected by minimizing the chosen information criterion. The minimizations are conducted by a standard stepwise algorithm for each of the two equations that compose the ECM. This algorithms is as follows.

1. First, I estimate the model with all n variables,
2. Then, I estimate the n models removing each variable separately.
3. If at least one of the n models has a lower information criterion than the one estimated at the first step, I remove the variable that allow to decrease it the most and go back to step 1 with $n-1$ variables. Otherwise, the algorithm stops.

This algorithm allows to minimize the chosen information criterion over all models without having to actually estimate all the possible specifications. At the end of this stepwise minimization algorithm, I check if all the estimates take the expected signs. If at least one does not, the variable corresponding to the estimate with a wrong sign and the larger absolute t-statistics is

[95] Including all the significant variables leads to a model that is less biased, but it may increases the estimation variance above the optimum of the bias-variance trade-off. For an optimal forecast accuracy, some significant variables must sometimes be dropped, as in the current case.

removed from the original set of 31 variables. Then, I run the stepwise minimization algorithm again starting with one less variable from the very beginning. Using this ad-hoc algorithm, I first select the subset of the seven explaining variables to include to the long-run equation[96]. Then, I select the variables to include to the short-run dynamics among the 31 that are considered. For each AOC, I thus obtain an ECM with specific sets of variables which are optimal regarding the information criterion, and for which all estimated parameters take theoretically consistent signs. The estimates for each AOC are commented in the section 4.6. The resulting model is used to generate monthly price forecasts at a 12 months horizon, which are evaluated in section 4.7.

4.6 Explaining prices

Before evaluating the forecast performances at the annual and monthly frequencies, I here present the estimates of the coefficients of the models, and the extent to which the collected data can explain the price history.

4.6.1 Estimates of the error-correction models on the annual data

As detailed in section 4.5, the annual forecasting model for year T is a combination of the k specifications that generated the best forecasts for year $(T-1,...,T-L)$. The resulting price function has a complex form and is inconvenient to present exhaustively, in particular because it mixes models with and without the logarithmic transformation. For the sake of clarity, I hereafter present the estimates of the ECM which is standard best practice in time series forecasting. The latter is effective for assessing the respective explanatory powers of each of the leading indicators, both in the short run and in the long run. All estimates are obtained by the 2SLS methods introduced by Engle and Granger (1987). In what follows, I present both the coefficients estimated on the whole sample all AOC taken altogether, and the coefficients estimated for each AOC separately. In all estimations, the explained variable is the logarithm of the real prices. The set of the explaining variables is selected by the algorithm described in the previous section, which both minimizes the AIC and enforces that all coefficients take the expected sign.

Table 4.10 presents the estimates of the long-run equilibriums of the ECM[97] for each AOC and taking all fifteen AOC in a single model. In the latter aggregate model, fifteen AOC-specific intercepts are added. The first column contains the coefficients estimated on all AOC altogether, and the other columns contain the estimates of the AOC-specific equations. The first notable feature is that the HUR variable is never selected by the algorithm, except for SE, for which the estimate is selected but not statistically significant[98]. For the other AOC, either the estimate related to the HUR takes the wrong sign, or adding the HUR to the equation does not improve the information criterion. By contrast, the SUR variable is always kept by the algorithm, except for BRO. For all other AOC, the levels of the SUR are thus very correlated with those of prices[99], even after controlling for the fluctuations of the other explanatory variables. These results are in line with the bivariate correlation statistics commented in section 4.3. Another key feature of

[96] Recall that no lagged variable is included in the long-run equation.

[97] These long-run equilibriums are standard multivariate regression models estimated by OLS on the undifferentiated data. See Appendix D for a presentation of the long-run and short-run equations that compose an ECM.

[98] Statistical significance, in the sense of the p-value of a standard Student test, is another way of selecting variables. But it is less suitable to forecasting purpose. Indeed, the statistical significance of each variable depends on the the number of degrees of freedom and on the set of the other variables, which makes it difficult to find the optimal subset in terms of statistical significance. Because information criteria evaluate the relevance of a set of variable, they are far more convenient for algorithmic model selection.

[99] All the estimates related to the SUR are statistically significant at the 5% level.

the long-run equations is that few leading indicators are actually selected by the algorithm. In each column, at most two leading indicators are kept in addition to the SUR and the HUR. This is because the sample size is so small (35 observations by AOC) that only few parameters can be added in the model without risks of overfitting.

Among the five leading indicators, the interest rate r_T and the indicator of the GDP of Bordeaux wines-consuming countries y_{iT} are never selected among the long-run equilibriums. This is because these two indicators present an important trend, which is respectively increasing for y_{iT} and decreasing for r_T (see section 4.4). r_T and y_{iT} will however play a role in the short-run dynamics, which equation is estimated on the data taken in first difference so that the trends are removed (see the presentation of the ECM in the Appendix D). By contrast, the indicator of the exchange rates e_{iT} is almost always selected, and when the related estimate is always statistically significant at the 5% level. As expected, the exchange rates are major determinants of Bordeaux wine prices, which will also be confirmed by the monthly analysis thereafter. Interestingly, the indicator c_{iT} of the harvests of the competitors is selected for white wines only (BW, EDM and SAU)[100]. This is a sign that competition with the other wine regions is more fierce for the white wines of Bordeaux than for the red[101]. Lastly, the indicator of quality q_{iT} is selected for the AOC BW and EDM but also for the AOC of the Saint-Emilion region, LU, MSE and SE, consistent with the analysis of the descriptive statistics of section 4.3. q_{iT} is also selected in the aggregate model. The last line presents the R^2 of each estimation. It is larger for the aggregate model because the inter-AOC price variance is preponderant and efficiently explained by the AOC fixed effects.

[100]The wines of AOC BW and EDM are dry whites whereas the SAU are sweet dessert white wines.

[101]Recall that c_{iT} accounts for the harvests of all wine producing countries, but also for the harvests of the other French wine regions which are more direct competitors.

TABLE 4.10 – OLS estimation of the annual long-run equilibriums on the annual data

	All	BLA	BSR	BR	BRO	BW	CAS	CBO	EDM	GRA	HME	LU	MED	MSE	SE	SAU
hur_{iT}															-.099	
															(.065)	
sur_{iT}	-.181	-.27	-.341	-.606		-.542	-.289	-.277	-.57	-.197	-.168	-.121	-.243	-.105	-.229	-.208
	(.02)	(.068)	(.08)	(.081)		(.164)	(.058)	(.088)	(.164)	(.049)	(.061)	(.046)	(.058)	(.039)	(.05)	(.077)
e_{iT}	.958	1.998	1.544	1.565	2.216		3.431	2.522		1.858	2.259	1.209	2.523	1.169		
	(.186)	(.606)	(.614)	(.48)	(.756)		(.595)	(.674)		(.596)	(.737)	(.56)	(.643)	(.577)		
y_{iT}																
c_{iT}						-.604			-.514							-.417
						(.163)			(.174)							(.28)
q_{iT}	.316					.913			.78			.679		.477	.82	
	(.106)					(.297)			(.293)			(.318)		(.318)	(.255)	
r_T																
R^2	.894	.466	.483	.716	.207	.616	.639	.413	.531	.533	.309	.294	.491	.325	.62	.189

Note: The dependent variable is the annual bulk price, all vintages mixed, with 35 observations per AOC. Column "All" contains the estimates of the system where the coefficients before each variable are constrained to be equal across all AOC. These estimates can be viewed as estimates of the average effects of each variable across all AOC. The other columns give the AOC-specific coefficients estimated on the 35 observations for each AOC. The multi-AOC estimation includes AOC fixed effects, the AOC-specific estimations include an intercept. Those are not reported in this table but are available upon request.

Table 4.11 give the estimates of the short-run dynamics. These equations are estimated by OLS on the differentiated data, where the lagged residuals $\hat{\epsilon}_{iT-1}$ of the previous long-term relationship is added to the list of the explanatory variables. This makes the whole estimation a 2SLS procedure. The latter is called the error-correction term and it constrains the short-run dynamics to be attracted toward the long-term equilibrium. One-year lagged and differentiated terms are also included for prices and for each explanatory variable. More details on the 2SLS estimation of an ECM are given in the Appendix D.

The error-correction term $\hat{\epsilon}_{iT-1}$ is not selected for the AOC CAS and EDM, meaning that the ECM representation is not relevant for these two AOC. Consequently, their short-run dynamics are represented by a DADL(1,1)[102]. The lagged variation of prices Δp_{iT-1} is also selected for most AOC, which indicates an important inter-annual inertia in the fluctuations of prices. Contrary to the results of the variable selection for the long-run equilibriums, the differentiated indicators Δhur_{iT} are selected for most AOC. As expected, the annual harvest is found to significantly influence annual prices, but only in the short-run. The exceptions are the AOC LU, MED and MSE for which the variable Δhur_{iT} is not selected. For these AOC, starting inventories are generally larger than expected annual deliveries at the beginning of the campaign, so that a large harvest is not required to meet the demand of the upcoming year[103]. On the other hand, we also observe that Δhur_{iT} is selected for HME, SE and SAU which also exhibits a high average SUR (see table 4.2 in section 4.3). But even though Δhur_{iT} is selected for these AOC, the estimates are not highly significant: not at the 5% level for HME and SE, and not even at the 10% level for SAU. In comparison, the same coefficients for appellations BR, BRO and BW, for which the average SUR is much lower, are both selected by the algorithm and significant at the 1% level.

As for the long-run equilibriums, the differentiated SUR Δsur_{iT} are selected for the majority of the AOC and for the aggregate model, so that the stocks also have an immediate influence on prices. On the opposite, the variations of the indicator of the exchange rates Δe_{iT} are not selected for most AOC, so that the influence of the exchange rates is almost solely taken into account by the long-run equilibrium. Another difference from the long-run relationship is that the indicator Δy_{iT} of the GDP growth of the Bordeaux wines-consuming countries is now selected for the aggregate model and for several AOC, including the prestigious ones from Médoc and the Saint-Emilion areas. This may stem from the fact that the wines of these AOC are mostly purchased by wealthy consumers whose revenues may be more dependent on the overall economic climate. Furthermore, the indicators Δc_{iT} of the harvests variations of the competing wine regions are still selected for the white AOC BW, EDM and SAU, but also in the aggregate model where the estimate is strongly significant. The algorithm also selects the indicator of quality Δq_{iT} for the same AOC, and for the appellation SE, as for the long-run equilibrium. Finally, the differentiated interest rate Δr_T is selected for the majority of the AOC, but often with one-year lag. One possible explanation is that the influence of the interest rates on prices is diffuse, since they only impact prices via the storing arbitrage. The last line gives the R^2 of the models which evaluate the explanatory power of the models for price changes. For white AOC and for the largest appellation BR, the model correctly fits the history of price changes. The model has less explanatory power for BSR, BRO, HME and for the AOC of the Côtes region (BLA, CAS and CBO). However, the actual forecasting models combine a number

[102]The standard ECM test does not involve AIC minimization and variable selection, but only a standard Student test for the significance of the error-correction term. When the error-correction terms are included, these tests do not reject the hypothesis that their estimates are null at conventional levels, so that the ECM is also rejected for these two AOC by the standard procedure.

[103]The average SUR over the sample period are respectively of 1.888, 1.974, and 1.914 for the AOC LU, MED and MSE. See table 4.2 in section 4.3 for the statistics on all AOC.

of different models of different specifications and variables subsets, so that the R^2 are less representative of the forecasting performance. The evaluation of the annual forecasts performance of the combination models are presented in section 4.7.

TABLE 4.11 – 2SLS estimation of the short-run dynamics on the annual data

AOC	All	BLA	BSR	BR	BRO	BW	CAS	CBO	EDM	GRA	HME	LU	MED	MSE	SE	SAU
$\hat{\epsilon}_{iT-1}$	-.263	-.376	-.453	-.501	-.259	-.36		-.323		-.486	-.338	-.683	-.245	-.717	-.823	-.257
	(.029)	(.131)	(.119)	(.113)	(.099)	(.108)		(.111)		(.131)	(.176)	(.147)	(.133)	(.146)	(.165)	(.103)
Δp_{iT-1}	.298		.369	.265	.327	.489	.392	.266	.71	.538	.487	.182	.566			.393
	(.038)		.154)	(.108)	(.16)	(.083)	(.152)	(.17)	(.071)	(.135)	(.202)	(.136)	(.137)			(.216)
Δhur_{iT}	-.126	-.09	-.179	-.203	-.123	-.248	-.086	-.158	-.412	-.232	-.182				-.181	-.087
	(.016)	(.044)	(.062)	(.043)	(.038)	(.054)	(.041)	(.071)	(.05)	(.059)	(.106)				(.091)	(.054)
Δhur_{iT-1}					-.066										-.276	
					(.038)										(.081)	
Δsur_{iT}	-.063	-.099	-.065	-.285		-.221			-.108		-.093	-.058	-.192	-.173	-.096	-.043
	(.012)	(.04)	(.05)	(.067)		(.08)			(.075)		(.074)	(.03)	(.055)	(.044)	(.04)	(.028)
Δsur_{iT-1}	-.04	-.108														
	(.012)	(.046)														
Δe_{iT}			1.144							1.321						
			(.646)							(.541)						
Δe_{iT-1}	.428						1.511		1.072							
	(.165)						(.495)		(.398)							
Δy_{iT}	.8							4.508			2.524	5.813	1.788	7.088		1.347
	(.35)							(1.659)			(1.928)	(1.721)	(1.421)	(1.896)		(1.188)
Δy_{iT-1}		2.728			3.012										3.989	
		(1.553)			(1.495)										(1.453)	
Δc_{iT}	-.201					-.597			-.51							-.423
	(.05)					(.15)			(.115)							(.233)
Δc_{iT-1}	-.222		-.373			-.648			-.935			-.377		-.522		-.355
	(.047)		(.146)			(.14)			(.127)			(.167)		(.193)		(.248)
Δq_{iT}	.261					.515			1.455						.808	.87
	(.066)					(.257)			(.198)						(.268)	(.517)
Δq_{iT-1}													.453			
													(.321)			
Δr_T								-4.198				-6.787		-7.784		
								(2.182)				(2.107)		(2.682)		
Δr_{T-1}	-1.586	-4.491	-2.981	-1.965	-4.766					-2.951	-3.935		-2.599	-1.905	-8.617	
	(.472)	(2.143)	(1.599)	(1.15)	(2.083)					(1.461)	(2.545)		(1.661)	(1.564)	(2.11)	
R^2	.515	.536	.606	.743	.574	.918	.436	.434	.945	.706	.488	.741	.744	.809	.889	.819

Note: The dependent variable is the variation of the annual bulk price, all vintages mixed, with 33 observations per AOC. Column "All" contains the estimates of the system where the coefficients before each variable are constrained to be equal across all AOC. The other columns give the AOC-specific coefficients estimated on the 33 observations for each AOC. Neither intercept nor AOC fixed effect are considered in these equations explaining prices variations, because they would imply a trend in real prices.

4.6.2 Estimates of the error-correction models on the monthly data

The monthly data is richer in terms of number of observations, but only covers the period from August 2001 to July 2016, whereas the annual data covers the period 1982-2016. Also recall that the series of monthly average real prices p_{it} have been filtered to remove the irregular and seasonal components before the multivariate analysis. For each AOC, only the estimated underlying levels $\hat{p}^*_{it}$ are assumed to be determined by the leading indicators (see section 4.5). The monthly series s_{it} of the stocks of wine for each AOC are also filtered be the same method to evaluate the underlying levels $\hat{s}^*_{it}$. The raw and filtered series are given on a figure in the Appendix D for every AOC. One additional feature of the monthly model is that the set of explanatory variables includes the expected stocks twelve months ahead, so as to account for the influence of the current weather. As for the series of prices and stocks, the series of the expected stocks are filtered and only the estimated levels noted $\widehat{es}^*_{it}$ are included to the price model. The other variables are the monthly versions of the annual indicators $(e_{iT}, y_{iT}, c_{iT}, q_{iT}, r_T)$, noted $(e_{it}, y_{it}, c_{it}, q_{it}, r_t)$. As in the annual case, the ECM are estimated for each AOC separately, and also across all AOC. Contrary to the annual case however, the model estimated on all the AOC taken together is actually not used for the monthly forecasts. Because its estimates provide a synthetic information about the relevance of each variable across all AOC, they are also commented hereafter. As in the annual case, the variables are selected by the algorithm described in section 4.5 which minimizes the information criterion AIC, and guarantees that the estimates take economically consistent signs. Models estimated on nominal prices are actually found to provide better forecasts on average (see table D.8 in the Appendix D), so I present the latter in this section[104].

Table 4.12 details the OLS estimates of the equations that represent the long-run equilibriums for each AOC. Because the monthly data is larger[105], more variables are selected by the minimization of the AIC, and the estimates are more statistically significant on average. Like in the annual case, the monthly stocks levels $\hat{s}^*_{it}$ are selected for most AOC, the exceptions being CAS, CBO, LU and SE. One key result is that the levels of the expected stocks $\widehat{es}^*_{it}$ is selected and strongly significant for most AOC and in the aggregate model. This result legitimates the harvest model described in section 4.4 and designed to summarize the weather information about the next harvest. Instead of directly considering several weather indicators in the price equation, the indicator $\widehat{es}^*_{it}$ accounts alone for the influence of the weather conditions on current prices[106]. Even if the indicator is not selected in the long-run relationship for the appellations BR, BRO, BW, EDM, and MED, it does intervene in their short-run dynamics commented thereafter. Certain observations made on the other drivers for the annual models also apply at the monthly frequency. The indicator of the exchange rates e_{it} is here again selected for most AOC, and the indicator of quality is selected for the AOC of the Saint-Emilion area LU, MSE, SE. The latter is now also selected for the appellations GRA and MED which are above the average quality. One difference is that the indicator of the harvest of the competitors is no longer selected for the white AOC BW and EDM but it is for the high quality AOC and for the appellation BRO[107]. As for the annual case, the R^2 is larger for the aggregate model due to the AOC fixed effects.

[104] The estimations with real prices lead to the same qualitative observations at the monthly frequency. This is because inflation was steady during the observation period of the monthly data (2001:2016).

[105] The monthly data contains 180 observations per AOC, against only 35 for the annual data.

[106] This aggregation leaves more degrees of freedom to estimate the influence of the other drivers. The gain is not negligible since the number of variables indicating the weather conditions actually exceeds the number of the price drivers (see section 4.4).

[107] This stems from the difference between the two estimation periods: 1982-2016 for the annual data, and 2001-2016 for the monthly data. This is checked in table D.6 in the Appendix D which gives the estimates of the long-run equilibrium estimated on the annual data limited to the period 2001-2016. This result suggests that the influence of

TABLE 4.12 – OLS estimation of the long-run equilibriums on the monthly data

	All	BLA	BSR	BR	BRO	BW	CAS	CBO	EDM	GRA	HME	LU	MED	MSE	SE	SAU
$\hat{s}^*_{it}$	-.329	-.476	-.601	-.441	-.451	-.351			-.41	-.693	-.328		-1.78	-.236		-.301
	(.017)	(.103)	(.063)	(.028)	(.026)	(.045)			(.035)	(.096)	(.207)		(.092)	(.129)		(.171)
$\widehat{es}^*_{it}$	-.057	-.207	-.447				-.133	-.168		-.239	-.322	-1.212		-1.083	-.138	-.42
	(.01)	(.099)	(.032)				(.04)	(.073)		(.053)	(.181)	(.097)		(.089)	(.063)	(.091)
e_{it}	.94	.265	1.868		.546	1.198	2.319	1.4		1.021	.772	1.579	1.081	.414	.365	
	(.049)	(.172)	(.107)		(.133)	(.179)	(.181)	(.178)		(.112)	(.173)	(.148)	(.115)	(.146)	(.251)	
y_{it}					.277						.401	1.069	.727	.58		
					(.137)						(.168)	(.23)	(.08)	(.17)		
c_{it}	-.146		-.255		-.323					-.792	-.93	-.521		-.386	-.312	-.462
	(.031)		(.073)		(.09)					(.068)	(.118)	(.122)		(.091)	(.146)	(.142)
q_{it}										.861		1.062	.489	1.902	1.476	
										(.323)		(.377)	(.349)	(.374)	(.473)	
r_t				-.016					-.014			-.01				
				(.004)					(.003)			(.006)				
R^2	.954	.552	.79	.66	.741	.498	.621	.326	.538	.818	.555	.724	.803	.8	.235	.386

Note: The dependent variable is the monthly bulk price, all vintages mixed, with 180 observations per AOC. Column "All" contains the estimates of the system where the coefficients before each variable are constrained to be equal across all AOC. The other columns give the AOC-specific coefficients estimated on the 180 observations for each AOC. The multi-AOC estimation includes AOC fixed effects, the AOC-specific estimations include an intercept. Those are not reported in this table but are available upon request.

The estimates of the equation representing the short-run dynamics at the monthly frequency are given in table 4.12. Contrary to the annual case, the ECM representation is never rejected[108]. The values of the estimates are about ten times smaller than those of the annual case, which is consistent with the difference between the two frequencies[109]. The first lagged changes in price levels $\Delta\hat{p}^*_{it-1}$ is selected and strongly significant for most AOC, the only exception being the appellation BW. As for the annual case, I thus find an important inertia in the direction of monthly price changes. Up to three lags are considered for each explanatory variables, but few are actually selected by the algorithm. The contemporary changes in the levels of stocks $\Delta\hat{s}^*_{it}$ are only selected for half of the AOC, as the changes in the excepted stocks levels $\Delta\widehat{es}^*_{it}$. The latter or its lagged values are notably selected for the AOC BR, BRO, BW, EDM, and MED for which it did not intervene in the long-run equilibrium, although the estimates are not always statistically significant at conventional levels[110]. The influence of the weather on current prices is thus found sufficiently informative to be included in every models, although in different fashions. The rest of the indicators barely affect the short-run dynamics, except for certain lags. Notably, for each AOC of the Saint-Emilion area, one lag of the changes in the quality indicator Δq_{it} is selected which consolidate the result that the vintage effect is only significant for those three. The lag distributions of the macroeconomic indicators are very unstable when the estimation period changes, especially at the monthly frequency, so that I prefer not to comment them in details. The R^2 are quite small, because the monthly price changes are more irregular

the harvests of the other wine regions was more important for white AOC before 2000, but is now stronger for the red AOC.

[108]For every AOC, the lagged error term of the long-run equation $\hat{\epsilon}_{it}$ is now selected and statistically significant.

[109]At the monthly level, an estimate of -0.05 indicates that the distance to the long-run equilibrium is reduced of 5% each month, *ceteris paribus*. Within a year, 45% of the distance to the equilibrium is covered. This corresponds to an estimate of -0.45 for the error-correction parameter at the annual frequency, so about ten times larger.

[110]Recall that, in statistical jargon, the ratio between the estimate and its standard deviation given in parenthesis must above 2.33 for the estimate to be said statistically significant at the 1% level significance. It must above 1.96 to ensure statistical significance at the 5% level, and above 1.645 for the less restrictive 10% level. The convention is to only comment the significance with respect to the 5% level. For a discussion on this convention, see notably McCloskey and Ziliak (1996).

and difficult to explain. The hierarchy of the R^2 among the AOC does not reflect the explanatory power of the model for each AOC. Indeed, the explained variable is the variation of the estimated price level $\Delta\hat{p}^*_{it-1}$, not of the observed price. These estimated levels are actually more difficult to estimate when the prices are very volatile, which causes the series of estimated price levels to be smoother when the observed prices are by contrast highly volatile[111].

TABLE 4.13 – 2SLS estimation of the short-run dynamics on the monthly data

AOC	All	BLA	BSR	BR	BRO	BW	CAS	CBO	EDM	GRA	HME	LU	MED	MSE	SE	SAU
$\hat{\epsilon}_{it-1}$	-.03	-.016	-.022	-.064	-.068	-.068	-.004	-.021	-.05	-.06	-.019	-.073	-.052	-.063	-.053	-.021
	(.003)	(.007)	(.01)	(.024)	(.018)	(.022)	(.001)	(.007)	(.016)	(.018)	(.007)	(.018)	(.018)	(.022)	(.017)	(.011)
$\Delta\hat{p}^*_{it-1}$	.397	.62	.583	.366	.494		.952	.6	.548	.554	.78	.477	.483	.418	.269	.456
	(.017)	(.053)	(.054)	(.064)	(.06)		(.013)	(.056)	(.06)	(.059)	(.044)	(.064)	(.056)	(.059)	(.069)	(.065)
$\Delta\hat{p}^*_{it-2}$																
$\Delta\hat{p}^*_{it-3}$	.072					.11						.14				
	(.017)					(.071)						(.065)				
$\Delta\hat{s}^*_{it}$	-.047		-.084	-.045	-.057		-.009		-.039					-.088	-.231	
	(.008)		(.028)	(.021)	(.016)		(.003)		(.018)					(.058)	(.086)	
$\Delta\hat{s}^*_{it-1}$	-.014					-.177										
	(.008)					(.073)										
$\Delta\hat{s}^*_{it-2}$	-.013												-.368			
	(.008)												(.1)			
$\Delta\hat{s}^*_{it-3}$																
$\Delta\widehat{es}^*_{it}$	-.005	-.043	-.045		-.004				-.025	-.056	-.087					
	(.002)	(.022)	(.018)		(.002)				(.017)	(.036)	(.035)					
$\Delta\widehat{es}^*_{it-1}$	-.004	-.032	-.028	-.03		-.1		-.045		-.063			-.169	-.17		-.077
	(.002)	(.023)	(.018)	(.018)		(.053)		(.03)		(.037)			(.08)	(.046)		(.046)
$\Delta\widehat{es}^*_{it-2}$		-.046	-.039				-.006	-.059		-.056			-.178			
		(.023)	(.018)				(.003)	(.03)		(.038)			(.082)			
$\Delta\widehat{es}^*_{it-3}$									-.026					-.081	-.125	
									(.016)					(.045)	(.046)	
Δe_{it}												.186				
												(.128)				
Δe_{it-1}											.105					
											(.054)					
Δy_{it-3}		.587														
		(.296)														
Δc_{it}				-.296												
				(.192)												
Δc_{it-1}							-.011									-.302
							(.008)									(.143)
Δc_{it-2}															-.228	
															(.159)	
Δc_{it-3}	-.077															
	(.033)															
Δq_{it-2}									.817					1.516		
									(.588)					(.488)		
Δq_{it-3}												.972			.87	
												(.434)			(.57)	
Δr_t									-.01							
									(.007)							
R^2	.248	.574	.531	.27	.349	.122	.975	.482	.437	.447	.656	.325	.495	.408	.241	.286

Note: The dependent variable is the variation of the monthly bulk price, all vintages mixed, with 176 observations per AOC. Column "All" contains the estimates of the system where the coefficients before each variable are constrained to be equal across all AOC. The other columns give the AOC-specific coefficients estimated on the 176 observations for each AOC. Neither intercept nor AOC fixed effect are considered in these equations explaining prices variations, because they would imply a trend in real prices.

[111] See the figure D.3 in the Appendix D for an illustration. The series of monthly average prices (in black) is highly volatile for the appellation CAS, whereas the series of estimated price levels (in red) is very smooth. On the opposite, the observed monthly prices of the large appellation BR are rather regular, so that the estimated levels are very close to the actual prices. As a result, the estimated price levels are more volatile for the appellation BR than for the appellation CAS.

4.7 Evaluation of the forecasts

4.7.1 Forecasts of the annual prices at the beginning of the marketing year

At the annual frequency, the forecasts are computing using a combination of the k models which gave the best post-sample forecasts on the last L years. I have conducted a sensitivity analysis to select the parameters k and L which minimize the overall forecasting error on the last five years. Table 4.14 gives the overall MRAE across all AOC and all years for different couples of parameters (L, k) for the combination scheme. To compute these statistics, I have conducted the following steps for each AOC i and marketing years T between 2012 to 2016.

- First, I have computed the post-sample forecasts of each of the 5,232 considered models (see section 4.5) for the L preceding years.
- Second, I select the best k models in terms of MRAE on the last L years.
- Third, I compute the combine the post-sample forecasts of these k models for marketing year T, and compute its MRAE noted $MRAE_{iT}$.

The figures given in table 4.14 are the geometric averages of the $MRAE_{iT}$ across all AOC i and all marketing years T, given k and L. The best forecast performances on the history are found for $L = 2$, so that the optimal test window is short. As found in the previous literature, the combination scheme allows to substantially improve the forecasts accuracy, with an optimum of $k = 250$ models, out of the 5,232 estimated specifications[112]. In comparison, only selecting the best forecasting model on the last years, i.e. selecting $k = 1$, leads to poor forecast accuracy. The combined forecast is an average of the selected specifications, where the weights are the inverses of the MRAE on the past $L = 2$ years.

The annual price models include contemporary predictors. In order to simulate the history of real-time forecasts, I have produced forecasts for each predictor. The HUR and the SUR are forecasted by the harvest model using the weather before the 1st of August, and the SUR are forecasted using the monthly stocks and the monthly deliveries forecasts generated by the UCM models. The indicator of quality is forecasted using the information about the quality of next harvest on the 1st of August[113]. The forecast of the annual indicator of the exchange rates e_{iT} is the value of the corresponding monthly indicator e_{it} in July of year $T - 1$. I also consider random walk forecasts for the indicators of the competitors' harvests and of the interest rates. The indicator of the GDP is assumed to follow the same trend as in the past years. These forecasts of the macroeconomic predictors are obviously not state-of-the-art but they are easily computable and actually difficult to outperform[114]. However, using naive forecasts for the predictors only causes to underestimate the forecasting performance of the price model. Indeed, better results can be expected if better forecasts become available for the macroeconomic predictors. In any case, my under-performing macroeconomic forecasts are sufficient to prove *a posteriori* the usefulness of the annual forecasting models on the past history.

In this section, I evaluate the accuracy of the post-sample price forecasts over the last five years of the sample, from 2012 to 2016. Table 4.15 contains the Mean Average Percentage Error (MAPE) of the price forecasts for each year and each AOC[115]. The last column and the last row give the geometric means. The global average forecast error is of 3.4% across all AOC. Among

[112] Recall that specifications for which at least one estimate takes an unexpected sign are removed from the analysis.

[113] This information is evaluated by the criteria for an ideal harvest described in Geny and Dubourdieu (2015). See section 4.4 for more details.

[114] See Kilian and Taylor (2003) in the case of exchange rates.

[115] The MAPE is the geometric average of the ratios between the absolute values of forecasts errors and the observed value of the price forecasted

TABLE 4.14 – Overall average of MRAE by parameter (L,k) of the cross-validation

	L=1	L=2	L=3
k=1	1.026	1.097	1.123
k=10	1.124	1.018	1.046
k=30	0.967	0.997	1.009
k=60	0.947	0.939	0.993
k=80	0.943	0.929	0.985
k=100	0.941	0.91	0.97
k=150	0.921	0.855	0.906
k=200	0.895	0.822	0.882
k=250	0.898	0.808	0.884
k=300	0.885	0.822	0.877
k=400	0.901	0.838	0.892
k=500	0.935	0.863	0.885

Note: The first line shows that only considering the best model ($k = 1$) in terms of out-of-sample MRAE for the previous one, two or three years ($L = 1, 2, 3$) implies forecasts that are 2.6%, 9.7% and 12.3% less precise in absolute value than the naive forecasts, on average across all AOC and for the past five years. The last line shows that considering the 500 best models ($k = 500$) for the same criterion implies forecasts that are 7.5%, 13.7%% and 11.5% more precise in absolute value than the naive forecasts, depending on the length of the test window for selection ($L = 1, 2, 3$).

TABLE 4.15 – MAPE of the annual forecasts by AOC and year

AOC	2012	2013	2014	2015	2016	Average
BLA	0.045	0.207	0.014	0.005	0.016	0.025
BSR	0.044	0.162	0.008	0.079	0.016	0.037
BR	0.043	0.177	0.024	0.003	0.012	0.024
BRO	0.066	0.112	0.026	0.02	0.002	0.024
BW	0.021	0.078	0.02	0.063	0.061	0.042
CAS	0.086	0.13	0.041	0.176	0.117	0.099
CBO	0.069	0.219	0.064	0.02	0.006	0.041
EDM	0.013	0.07	0.024	0.109	0.053	0.042
GRA	0.009	0.004	0.055	0.109	0.064	0.027
HME	0.021	0.164	0.061	0.002	0.027	0.026
LU	0.008	0.125	0.108	0.058	0.001	0.021
MED	0.041	0.14	0.103	0.031	0.023	0.053
MSE	0.014	0.11	0.093	0.027	0.026	0.04
SE	0.007	0.053	0.019	0.03	0.041	0.024
SAU	0.108	0.135	0.002	0.041	0.062	0.037
Average	0.028	0.1	0.029	0.029	0.02	0.034

Note: The top-left value of 0.045 indicates that the annual forecasting error is of 4.5% for year 2012 and the appellation BLA.

the AOC, the forecast errors are lower than 3% on average for BRO, GRA, HME and SE, and larger than 5% for CAS, CBO, LU, MSE and SAU. The errors are on the average for the main appellations BR, BW and BSR. The highest forecasts errors are found for 2013, during which the price soared rapidly due to a small harvest. However, the forecasts for 2013 were actually very useful compared to the naive forecasts. Recall from section 4.5 that a better measure of forecasting performance is the MRAE, in which the absolute error is scaled by the absolute error of a naive forecast. A value below (respectively above) 1 indicates that the forecasts are more (respectively less) accurate than the naive forecasts. The price of the past year are assumed to be the naive forecasts. This assumption is more challenging than using the last observed monthly prices, because the later are very volatile and are therefore poor forecasts for the next year.

Table 4.16 gives the values of the MRAE for each year and each AOC. The overall average is of 0.92, indicating that the annual forecasts are more accurate than naive forecasts on average. This is already a satisfactory result, given that random walk price forecasts are notoriously hard to outperform[116]. Even if the forecasts for 2013 were the least precise in terms of average MAPE, they were actually useful on average since the average MRAE for 2013 is only of 0.769. The forecasts of 2015 exhibit a poor MRAE because prices did not change much between 2014 and 2015, so that random walk forecasts are particularly hard to outperform for that year. This the main shortcoming of the MRAE: it may indicate poor forecasting performances if prices are very stable, even if the forecasts are satisfactory in absolute error. This measure is more severe with the forecasts error in the absence of price changes.

The economic stakes are the highest for the appellation BR which dominates the bulk market with more than half of the total worth. For this key AOC, the average MRAE is 0.624, so that the annual model is especially helpful here. Its MAPE is however in the average across all AOC. The model thus manages to indicate the good direction of price changes with limited absolute error. By contrast, the model underperforms for the second largest AOC, the regional whites appellation BW. As mentioned before, this is because BW annual prices did not change much between 2012 and 2016, so that the benchmark random walk forecasts was in fact accurate during this period[117]. Even for BW, the model behaved well during the shortage of 2013-2014. The annual forecasts are indeed particularly effective and useful in the case of important supply shocks.

So as to provide a more complete information to the professionals, confidence intervals can be computed by combining the previous post-sample errors for each sub-model. Only two post-sample errors by sub-model are sufficient to generate a large number of possible errors through the combination scheme, so as to consistently estimate the density of simulated forecasts. The figure D.6 in the Appendix D gives these simulated forecast density for the prices of 2016 for each AOC.

4.7.2 Forecasts of the monthly prices at various time horizons

I have computed monthly price forecasts for all the time horizons up to twelve months ahead, at each month between August 2012 to August 2016 and for each AOC. The models are AOC-specific ECM estimated by 2SLS. In both the long-run and the short-run equations, the variables are selected by the algorithm presented in section 4.5. I have tested both the AIC and the BIC criteria for the variable selection, and both nominal and deflated prices for the estimation. The results with the BIC and nominal prices where found slightly more efficient on average to

[116]This is because any predictable change of the price generates an opportunity for profit. See Allen (1994) for a review of the literature on agricultural price forecasting.

[117]The annual volatility of BW prices has been of only 0.05 between 2012 and 2016, against 0.18 over the whole period.

TABLE 4.16 – MRAE of the annual forecasts by AOC and year

AOC	2012	2013	2014	2015	2016	Average
BLA	0.578	0.814	0.396	0.396	0.591	0.535
BSR	0.467	0.642	1.366	4.63	0.495	0.988
BR	0.496	0.74	0.362	0.138	0.427	0.379
BRO	1.119	0.588	0.359	1.683	0.1	0.525
BW	1.778	0.75	0.683	4.766	2.61	1.625
CAS	0.825	0.601	0.709	2.716	0.723	0.928
CBO	4.178	0.857	1.529	0.573	0.202	0.913
EDM	0.571	0.527	3.441	4.918	2.755	1.696
GRA	0.67	0.1	1.531	1.013	0.923	0.626
HME	2.086	1.012	5.298	0.1	0.658	0.94
LU	0.296	1.029	1.057	2.989	0.1	0.626
MED	0.462	0.652	2.537	3.225	2.229	1.406
MSE	0.3	0.772	0.758	0.883	0.391	0.571
SE	0.219	0.563	0.294	10	0.539	0.721
SAU	4.147	10	0.1	0.562	0.442	1.006
Average	0.786	0.751	0.85	1.334	0.56	0.822

Note: The top-left value of 0.578 indicates that the annual forecasting error were smaller of 42.2% (1-57.8%) than the error of the naive forecast for year 2012 and the appellation BLA.

predict nominal prices in terms of MRAE [118]. The forecasts commented in this section are thus computed using this specification.

All the lags of the disclosure of the information about the market data and the predictors (see sections 4.2 and 4.4) are duly taken into account so as to replicate what would have been real-time forecasts. For instance, the stocks s_{it} held at the beginning of month t for AOC i are not known immediately because the deliveries d_{it-1} of month $t-1$ are only known after two months. Therefore, initial stocks s_{it} must be estimated using the stocks s_{it-2} that were carried two months before, and a estimation of the deliveries d_{it-2} and d_{it-1} given by UCM models. I thus make sure that the forecast of p_{it} is computed using only the information available at the beginning of month t.

Table 4.17 gives the MRAE for each AOC and each year, across all months and time horizons. The overall MRAE is of 1.039, which indicates that the monthly forecasts are not globally better than the random walk forecasts. However, the monthly forecasts outperform the random walk forecasts for the large appellations BW, BRO and especially BR, which represent the highest economic stakes in the bulk market[119]. The monthly forecasts also perform remarkably well for the quality appellations MED and SE, which significantly contribute to the prestige of the region. In fact, the average of the AOC-specific MRAE weighted by the respective market share for each AOC falls just below 1. As for the annual forecasts, the monthly forecasts would have been more informative just after the important harvest shock of 2013. Exceptions are CBO, EDM, HME and SAU for which the forecasts in 2013 showed little accuracy compared to other years. But across all AOC, and for the most important ones, the MRAE for 2013 falls well below 1. To recapitulate, the monthly models behaves well for the main AOC, and especially when an important supply shock occurs, just like the annual model. The MAPE for each year are given

[118] Table D.8 in the Appendix D gives the overall statistics in each case.

[119] The largest appellation BR represents alone more than half than the total worth of the bulk market.

in the last row, which shows again that the 2013 forecasts were not the most accurate, even if the most informative as compared to random walk forecasts. Detailed MAPE by AOC and year are given in table D.9 in the Appendix D.

TABLE 4.17 – MRAE of the monthly forecasts by AOC and year

AOC	2012	2013	2014	2015	2016	Mean
BLA	1.228	0.881	1.590	1.870	1.224	1.315
BSR	0.822	0.973	1.546	1.552	1.071	1.155
BW	0.717	0.855	1.277	0.885	1.241	0.970
BRO	0.699	0.994	1.168	0.739	0.858	0.875
BR	0.934	0.641	1.150	1.173	0.915	0.941
CAS	0.699	0.633	0.601	1.505	1.437	0.895
CBO	0.830	1.052	1.121	1.464	1.267	1.127
EDM	1.104	1.060	1.083	1.039	1.052	1.067
GRA	1.121	0.809	0.928	1.254	1.084	1.027
HME	1.008	1.129	1.086	0.962	1.215	1.076
LU	2.460	0.780	0.926	1.914	1.068	1.294
MED	0.968	0.579	0.661	1.277	1.083	0.875
MSE	1.770	0.899	1.050	1.828	0.925	1.231
SE	1.113	0.637	0.863	1.033	0.586	0.820
SAU	0.894	1.516	0.863	1.061	1.172	1.078
Mean	1.024	0.867	1.027	1.255	1.059	1.039
MAPE	0.044	0.050	0.061	0.056	0.043	0.051

Note: The top-left value of 1.228 indicates that the monthly forecasting errors were greater of 22.8% than the error of the naive forecast for year 2012 and the appellation BLA, on average across all horizons and months. The bottom-right value indicates that the monthly forecasts exhibit an average error of 5.1% across all AOC, years, months and horizons.

The previous statistics mix the MRAE at all horizons from one month ahead to twelve month ahead. Table 4.18 details the MRAE for each horizon, across all years and months. The penultimate row gives the MRAE for each time horizon across all AOC. Even if not obvious at first sight, they do exhibit a slight but significant decreasing trend when the horizon increases[120]. This means that the monthly forecasts are especially useful in distant time horizon, as compared to the random walk forecast. The last row gives the same statistics for price forecasts produced by a standard univariate local-trend UCM model. As expected, the overall MRAE is larger for the univariate forecasts. It confirms that the exogenous explaining variables are indeed useful and allow to improve the forecasts accuracy through the ECM design. Furthermore, the MRAE of the univariate UCM forecasts do not exhibit any trend[121]. It consolidates the conclusion that the multivariate monthly forecasts are all the more relevant when the forecast horizon is distant. In the same vein, Engle and Yoo (1987) and Hoffman and Rasche (1996) found that the advantage of ECM over VAR only appears in long-term forecasts. Short-term price movements are indeed the most difficult to forecast, and the usefulness of the monthly forecasting models only appears for a certain time horizon, which depends on the AOC. Of course, for certain minor AOC like LU or MSE, the monthly model does not seem

[120]A simple linear regression of the MRAE by horizon on a trend lead to an estimate of -0.0088 with a standard error of 0.0035, which indicates a statistical significance at the conventional 5% level.

[121]A linear regression on a trend leads to an estimate of -1.049e-05 with a standard error of 4.107e-03.

able to outperform the random walk forecast at any horizon. For the main appellation BR however, the MRAE falls under 1 for forecast horizon above seven months. Note that even though monthly forecasts become more useful at longer time horizon, they also become less precise in absolute value. Precisely, the MAPE increase with the forecast horizon for each AOC (see table D.10 in the Appendix D).

TABLE 4.18 – MRAE of the monthly forecasts by AOC and horizon

AOC	1	2	3	4	5	6	7	8	9	10	11	12	Mean
BLA	0.878	0.920	1.116	1.367	1.193	1.317	1.519	1.491	1.608	1.712	1.520	1.461	1.315
BSR	1.222	1.198	1.085	0.996	1.203	1.059	1.067	1.143	1.228	1.098	1.212	1.401	1.155
BW	0.911	1.117	0.782	0.926	0.929	0.865	1.074	1.004	0.870	1.004	1.070	1.170	0.970
BRO	1.017	1.135	0.902	0.878	0.865	0.718	0.732	0.724	0.742	0.845	0.917	1.170	0.875
BR	1.212	1.215	1.153	1.115	1.166	1.126	0.854	0.911	0.812	0.695	0.627	0.708	0.941
CAS	1.495	0.810	0.653	0.697	0.669	0.827	0.911	0.989	1.036	1.043	0.973	0.917	0.895
CBO	0.969	1.090	0.984	1.272	1.350	1.259	1.145	0.985	1.079	1.019	1.136	1.321	1.127
EDM	1.420	1.232	1.379	1.223	1.087	1.042	1.040	0.953	0.970	0.809	0.896	0.941	1.067
GRA	1.364	1.087	1.150	1.004	1.218	1.373	0.945	0.870	1.026	0.784	0.847	0.863	1.027
HME	1.548	1.254	0.993	1.132	1.157	1.083	0.939	1.077	1.030	0.929	0.931	0.981	1.076
LU	1.580	1.193	1.116	1.423	1.194	1.180	1.215	1.451	1.327	1.401	1.224	1.307	1.294
MED	1.044	1.027	0.944	0.819	0.812	0.801	0.772	0.767	0.893	0.894	0.979	0.804	0.875
MSE	1.211	1.229	1.370	1.269	1.288	1.212	1.117	1.259	1.292	1.170	1.168	1.205	1.231
SE	1.068	0.801	0.987	0.947	0.930	0.817	0.789	0.731	0.722	0.717	0.669	0.757	0.820
SAU	0.943	1.012	0.998	1.018	1.192	0.927	1.025	1.210	1.175	1.000	1.223	1.284	1.078
Mean	1.170	1.078	1.023	1.053	1.065	1.021	0.992	1.013	1.029	0.979	1.001	1.059	1.039
UCM Mean	1.200	1.060	1.032	1.104	1.076	1.075	1.111	1.069	1.163	1.062	1.107	1.114	1.097

Note: The top-left value of 0.878 indicates that the one-month ahead forecasting errors were smaller of 12.2% (1-87.8%) than the errors of the naive forecasts for the appellation BLA, on average across all years and months. On row "Mean", the right-end value indicates that the monthly forecasting errors produced by the model are on average 3.9% greater than the errors of the naive forecasts on average across all AOC, years, months and horizons. On row "UCM Mean", the right-end value indicates that the monthly forecasting errors produced only by the UCM are on average 9.7% greater than the errors of the naive forecasts on average across all AOC, years, months and horizons.

The last dimension of the monthly forecasts is that they are computed at different moments of the marketing year. Because the annual information flow is seasonal, the level of information about future market conditions varies during the year. Table 4.19 provides the MRAE across AOC and months of the marketing year. Last row labeled "Mean" gives the averages by month, which are below 1 for the forecasts computed at the beginning of September, October and November. This shows that the monthly forecasts have more chances to outperform random walk forecasts at the beginning of the marketing year. In other words, the direction of price changes are more easily predictable during this period. Interestingly, September to November are precisely the three months of the harvest[122]. During this period, all the remaining uncertainty about the quantity and the quality of harvest is removed, new arbitrages are made, and prices change accordingly. Although weather during the growth season convey substantial information about the next harvest, my results indicate that a large part of the overall uncertainty remains until the harvest. The forecasts produced by my models account for this information and predict correct directions of price changes. The models also outperform the univariate forecasts computed by standard local-level UCM during the harvest season. The MRAE by month of the monthly forecasts produced by the UCM are all above 1, and are given in the row labeled "UCM Mean". The multivariate analysis is therefore particularly useful at the beginning of the marketing year. On the other hand, the market conditions of the next twelve months are most

[122]Only white grapes destined to produce the sweet white wines from the AOC SAU are harvested in November.

uncertain at the beginning of March before the budding season, when very little information is known about the next harvest. For the months February to April in the middle of the marketing year, the MRAE of the model are well above 1, and barely below the MRAE of the univariate forecasts. Indeed, future market conditions are highly uncertain in the middle of the marketing year, so that naive no-change forecasts appear to be the best guess. Interestingly, during the same period the monthly forecasts are paradoxically more accurate in terms of MAPE. The latter are given in row "MAPE" for the multivariate models and in row "UCM MAPE" for the univariate UCM. As before, the explanation is that prices change less during the second half of the marketing year. The monthly volatility of the estimated price levels is steadily decreasing accounting from September and throughout the marketing year. The averages by month are given in the last row[123]. Even though the forecasts in the middle of the marketing year are more accurate, they are also more likely to indicate the wrong direction of price changes because these are of small amplitude during the lean season[124]. Overall, the seasonality of the forecast accuracy is consistent with the timing of the information flow. Hence, the information accounted for in the forecasting model appears to be representative of the knowledge of the wine professionals.

TABLE 4.19 – MRAE of the monthly forecasts by AOC and month

AOC	Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Mean
BLA	1.413	0.948	1.069	1.043	1.400	1.468	1.386	1.554	1.640	1.649	1.080	1.390	1.315
BSR	1.009	1.257	0.829	0.800	0.813	1.281	0.929	1.742	1.407	1.240	1.411	1.614	1.155
BW	1.117	0.749	0.794	0.891	1.421	0.772	0.835	1.427	1.061	1.083	0.891	0.880	0.970
BRO	0.965	0.874	1.015	0.740	1.133	0.925	0.878	0.972	0.605	0.671	0.913	0.963	0.875
BR	1.092	1.351	0.529	0.666	0.853	0.695	0.681	1.176	1.297	0.894	1.415	1.194	0.941
CAS	1.019	0.786	0.919	0.497	1.095	0.960	1.105	0.980	0.914	0.565	0.994	1.240	0.895
CBO	1.271	1.373	1.979	1.099	0.994	1.035	1.171	1.013	0.766	0.883	1.308	1.021	1.127
EDM	1.247	1.232	1.452	1.232	1.094	1.073	1.237	1.018	0.914	0.690	0.811	1.053	1.067
GRA	0.737	0.992	0.762	1.024	1.331	0.823	0.841	1.151	1.320	1.461	0.998	1.185	1.027
HME	0.891	0.818	1.329	1.687	1.130	0.756	1.169	1.125	0.992	1.247	0.921	1.159	1.076
LU	1.183	0.879	1.052	1.272	1.141	1.122	1.434	1.730	2.014	1.687	1.053	1.395	1.294
MED	1.182	0.828	0.547	0.714	0.599	0.861	0.969	1.093	0.906	0.872	0.995	1.223	0.875
MSE	1.207	1.197	1.340	1.250	1.221	1.155	1.200	1.074	1.164	1.698	1.344	1.034	1.231
SE	0.757	0.657	0.679	0.701	0.622	0.913	1.050	0.812	0.838	0.752	1.261	1.012	0.820
SAU	1.203	1.094	1.135	1.136	0.914	1.274	1.076	1.243	1.263	0.856	0.908	0.946	1.078
Mean	1.069	0.978	0.967	0.938	1.019	0.986	1.044	1.179	1.090	1.020	1.069	1.139	1.039
UCM Mean	1.191	1.061	1.097	1.019	1.053	1.006	1.127	1.203	1.073	1.070	1.100	1.183	1.097
MAPE	0.051	0.058	0.051	0.050	0.047	0.052	0.046	0.047	0.049	0.050	0.052	0.056	0.051
UCM MAPE	0.059	0.062	0.059	0.054	0.049	0.052	0.049	0.049	0.047	0.052	0.052	0.057	0.053
$\sqrt{Var(\frac{\hat{p}^*_{it}}{\hat{p}^*_{it-1}})}$	0.441	0.447	0.446	0.444	0.443	0.442	0.441	0.441	0.44	0.439	0.44	0.44	0.441

Note: The top-left value of 1.413 indicates that the monthly forecasting errors were greater of 41.3% than the error of the naive forecast when computed at the beginning of August for the appellation BLA, on average across all years and horizons.

The last comments of this section are devoted to the forecasts of the annual average prices produced by the monthly models. To compute these forecasts, I have also estimated forecasting models for the volume dealt[125] using standard local-trend UCM. The resulting price forecasts

[123] The seasonal volatility of the raw prices is quite different. Contrary to their underlying levels, raw prices are highly volatile during the second half of the marketing year. This is because fewer transactions are dealt during the lean season, so that raw prices fluctuate a lot due to the irregular component of prices.

[124] This illustrate how difficult it is to provide a synthetic evaluation of forecasts, because both MAPE and MRAE may miss a key information depending on the context.

[125] The volume dealt each month are different from the monthly deliveries. First the delivery data include both deliveries of wine in bulk or already in bottle, whereas the data on transacted volumes only concern wines sold

are evaluated by the MRAE given in table 4.20, which must be compared to those of table 4.16. The overall MRAE is worse for the annual forecasts produced by the monthly models, but the rankings of the years in the two cases are consistent. For instance, the MRAE is minimum for 2013, and maximum for 2014 and 2015 whether the annual price forecasts are produced by the annual or by the monthly model. The patterns of the forecasts accuracy of the monthly and annual models are the same, which consolidates the stylized facts described above. The monthly models are indeed richer because they cover different time horizons and are updated each month. However, the annual models developed in this chapter prove to be more accurate when it comes to forecast the average prices of the next year.

As for the annual forecasts, density forecasts can be computed for each horizon. Instead of combining historical post-sample forecast, density forecasts at the monthly frequency are more easily computed by bootstrapping the values of the exogenous predictors [126].

TABLE 4.20 – MRAE of annual price forecasts produced the monthly models

AOC	2012	2013	2014	2015	2016	Mean
BLA	0.878	0.56	7.758	2.746	1.74	1.787
BSR	0.65	0.78	10	3.022	1.467	1.863
BR	0.172	0.773	0.528	2.27	3.32	0.88
BRO	1.891	0.259	1.757	10	1.108	1.57
BW	5.53	0.383	4.537	5.447	3.398	2.818
CAS	0.231	0.443	2.726	0.705	1.006	0.723
CBO	4.089	0.896	10	0.428	1.62	1.91
EDM	3.314	0.441	10	6.178	0.162	1.711
GRA	1.576	0.134	1.036	0.285	0.136	0.386
HME	1.314	1.151	6.352	0.236	0.786	1.123
LU	1.65	0.547	0.959	2.861	0.507	1.047
MED	0.1	0.351	0.958	8.573	10	1.236
MSE	1.009	0.334	6.183	0.739	0.1	0.688
SE	0.1	0.323	0.966	7.085	0.192	0.532
SAU	1.717	3.087	0.1	2.058	0.18	0.722
Mean	0.895	0.522	2.329	2.005	0.759	1.106
MAPE	0.025	0.068	0.074	0.042	0.03	0.044

Note: The top-left value of 0.878 indicates that the annual forecasts produced by the monthly model were smaller of 12.2% (1-0.87.8%) than the error of the naive forecast for the appellation BLA in 2012.

in bulk. Second, there is often a delay of several months between the transaction date and the physical deliveries. The monthly deliveries d_{it} of month t for AOC i have not been dealt at the average price p_{it} observed for the same month and AOC, but at prices dealt in the preceding months.

[126]See Prescott and Stengos (1987) and Kling and Bessler (1989) who have applied the same method.

4.8 Conclusion

This chapter shows that the fluctuations of average wine prices by Bordeaux AOC can be forecasted to some extent. Forecasts are generated at the annual and the monthly frequencies, with modelization options specific to each frequency. On the one hand, the annual average prices are forecasted using a combination of a number of standard time series models. At the monthly frequency, the monthly average prices are filtered with an unobserved component model so as to extract the underlying levels from the raw series. These monthly price levels are then forecasted with AOC-specific error-correction models, hence extending the framework of Cardebat and Bazen (2016). One aim of the study has been to optimally approximate the set of information available to the agents. A key feature of the data is that the annual stocks and harvests are observed separately for each AOC. In addition, a set of exogenous predictors has been collected, including exchange rates, interest rates, GDP, country-specific wine productions, quality evaluations and meteorological data. In order to increase the number of degrees of freedom, all the meteorological data are combined in a harvest model which correctly approximates the expectations of the next harvest over the observation period. The annual forecasts generally outperform the naive random walk forecast, especially for the largest AOC Bordeaux, and for years with low yields. Monthly forecasts present the same features but are less efficient to forecast annual average prices. Interestingly, the monthly forecasts only outperform random walk forecasts at longer horizons, so that short-term forecasts are barely informative on average. The multivariate scheme also proves to be useful at longer horizons as they become more accurate than the estimated univariate models. Furthermore, the monthly forecasts only outperform the random walk forecasts during the first half of the marketing year, when future market conditions are best known. Loosely speaking, the results indicate that the annual frequency is the most relevant, since monthly forecasts only prove to be useful at the beginning of the marketing year and for long-term horizons.

These forecasts are to be used as benchmark to compensate for the absence of futures prices on the wine market. Hence, the forecasting strategy was designed for it to be easily accessible to professionals. All the data used in this study is freely available on various websites, and even if the model selection is somewhat ad hoc, all the estimated models are standard. Benchmark price forecasts could play the part of future prices in the context of wine markets where the fluctuations of quality limit the interest of future contracts. Prices are not vintage-specific so that the data are not well-suited to conclude on the significance of the vintage effect. Yet, the estimations do suggest that the vintage effect is small for the main AOC Bordeaux, which accounts for half of the bulk market. For the basic AOC, consecutive vintages seem to be well-substitutable so that quality fluctuations may not to be an obstacle to the introduction of future contracts. On the opposite, the variations in vintage quality exhibit a significant impact on prices for the fine wines of the Saint-Emilion region.

The multivariate analysis has yielded several other interesting results. They feed the recent empirical literature on storage by providing additional evidence of the central role of inventories in price dynamics. In particular, the stocks-to-use ratios advocated by Bobenrieth et al. (2013) are found to be strikingly well-correlated with prices in the long run. As Boudoukh et al. (2007) and Osborne (2004), I find that weather news significantly influence current prices, although their explanatory power is low as compared to the other drivers. Finally, the estimates of the auxiliary harvest model indicate that temperatures are currently at optimal values for the production yields. If the current increasing trend in the temperatures continues, however a decrease of the production yields should be expected. Jones et al. (2005) drew the same conclusions with respect to wine quality in Bordeaux.

Although the evaluation of the forecasts on the last five years is promising, there is room for improvement. In particular, the forecasts of the deliveries are generated by univariate models, but the exogenous drivers of the price are likely to influence the deliveries as well. Adding exogenous drivers to the model of the deliveries may improve the forecasts of deliveries and stocks, and in turn improve the forecasts of the prices. Furthermore, the harvest forecasts may also be improved. Indeed, even if it behaves correctly at the aggregate level, the Bordeaux wine professionals dispose of better information about the incoming harvest. They should be able to correct the limitations of the model, and thus to obtain more accurate price forecast than those evaluated in this chapter. On the other hand, if the models presented in this chapter highlight regular and profitable arbitrages, prices should adjust more rapidly and the arbitrage opportunities might disappear. To quote the expression coined by Timmermann and Granger (2004), the "self destruction of predictability" could render these forecasting models obsolete in the long run. However, that would mean that expectations would then converge more rapidly, and therefore that welfare has improved due to this work.

General Conclusion

This dissertation is composed of four empirical papers examining the micro-level and macro-level determinants of wine prices.

The first chapter assesses the role of expert opinion on wine prices using a methodology that, by including detailed meteorological data, fixed-effects models, and the systematic use of numerous expert scores, avoids endogeneity and bias rooted in errors of judgment. The observed scores are assumed to be generated with a measurement error; they can be split into an objective component shared by all experts and a subjective component specific to each expert and wine. The latter is often seen as something that should be corrected as it obscures the signs of quality indicated by the objective component. We provide evidence, however, that in a price equation one should endeavor not to ignore subjective components, as they significantly affect wine prices. The most important result of our findings is the light shed on the role of the standard deviation in the price equation. We find a strong positive correlation between wine pricing and the standard deviation of the scores. Our interpretation is based on the fact that a higher standard deviation indicates that a high maximum score is likely. In line with the marketing literature, this highest score may be used as an advertisement for the sellers. Hence, this particular score is likely to be the most publicized. As a result, this is certainly the only score that the average consumers may have heard of. This is what we refer to as the "marketing effect"; the highest score is the most influential as it is the best known among consumers. Our interpretation is supported by the empirical analysis, since the highest score has the greatest impact on prices.

This first chapter has opened several research paths. Firstly, the method proposed in the paper has already been adopted and discussed in subsequent research. Oczkowski (2016) has notably applied this 2SLS method to the Australian wine market to assess the respective influences of objective quality and subjective opinions. Secondly, the prevailing influence of the highest score is susceptible to generate ratings shopping behaviors, in which Châteaux would accumulate scores until satisfied. The degree to which fine wine purchasers have access to the entire set of scores should be assessed. Finally, as fine wines are considerably expensive, wine experts largely rely on the invitations of the wine producers to taste their wines, as it is the case in Bordeaux during the *primeurs* campaign. Potential conflicts of interest should consequently be seriously examined.

In the second chapter of this dissertation, the equipercentile methodology is employed to express the scores given by various wine experts to Bordeaux wines on the same rating scale. It facilitates a comparison between scores among experts, and allows the calculation of a transparent and synthetic average of all available wine scores. This nonparametric method highlights that the famous wine expert Robert Parker has, on average, given higher scores than his peers who rate on the 100-points scale. Echoing the results of the first chapter, this finding underlines the necessity to scale the scores, as experts may be tempted to inflate their scores to gain publicity. This scaling method allow for the computation of relevant standard deviations for each wine across experts. When computed after scaling, the latter constitutes a better measure of judge concordance for each wine.

This chapter opens the path to future research on how to best aggregate the information contained in wine scores. In particular, the uncertainty of the true quality as measured by the standard deviation of the scores should be included in the information given to potential purchasers. The method could also be used to track the experts' preferences among the wine regions or grape varieties.

The third chapter addresses the as yet relatively unexplored market of wine competitions. An original data set has been obtained by matching new data on individual transactions from a large Bordeaux-based broker (containing information on contract dates, prices and quantities, and characteristics on producers and wines) with the records of eleven important wine competitions (winners by medal color, and contest features). The first outcome is the estimation of the causal effect of medals on producers' wine prices. The econometric approach consists in regressing prices on both before-transaction and post-transaction medal indicators. Under simple identifying restrictions, the difference in the estimates of the associated coefficients identifies the causal effect. Our preferred estimate indicates that the acquisition of a medal causes a price increase of about 13%. As expected, the impact for gold is found to be much larger than for silver and bronze. Only a small group of contests causes a statistically significant price increase. This group is constituted of the oldest and most highly renowned competitions. Interestingly, their judges are required to evaluate relatively few wines per day, and they grant medals by oral consensus. Secondly, we have computed the profit producers can expect from participating in the most important wine competition. Only small producers in terms of volumes have no incentive to participate, and the expected profit is very high for large producers. The last contribution of this chapter is the finding of a statistically significant overall link between medals and quality. However, only a minority of contests are found to attribute medals that are significantly correlated with wine quality.

Due to the boom in this market for wine awards in recent years, several research issues have arisen and should be further examined. Firstly, the reliability of wine tasting and the consistency of the attribution of medals should be more systematically measured as an evaluation of the differing wine competitions. Notably, the number of wine samples tasted by each judge per day is expected to be influential on the consistency of the competitions. Furthermore, the comments made in the first chapter concerning the top-end wines and the wine critics have greater relevance here. As wine producers are not required to disclose the number of wine competitions they have entered, and because expected profits from participation are estimated to be high, a generalized medal-shopping behavior should be expected. In the same vein, some wine competitions may be tempted to grant better medals so as to attract more participants. Since 2013, French competitions cannot award medals to more than one third of the participants, but the Decanter competition held in London does not respect that limit, although its awards are very influential on prices. The consistency of this particular competition, often qualified as the world largest, should be examined. From the perspective of industrial organization, the wine competitions are peculiar in the sense that they both charge the participation and the award. The consequences of this certification process could be theoretically investigated. Finally, the current situation where several certifiers compete to grant (sell) awards of unknown true value gives rise to the recursive issue of how to certify the certifiers. Interestingly, two institutions already compete to certify the wine competitions: The International Organisation of Vine and Wine and the International Union of Oenologists have both proposed criteria to grant their labels. Yet almost no wine competitions have committed to these standards so far. A lively stream of research continues to investigate how best to optimally assess and label wine quality.

The last chapter examines the determinants of the fluctuations of Bordeaux wines average prices, and the extent to which the latter can be forecasted. This work responds to a direct demand of Bordeaux wine professionals who lack visibility in a context where no futures market exists for wine, nor is it foreseeable in the near future. Different models are estimated and used to simulate forecasts at the annual and the monthly frequencies, with modelization options specific to each frequency. One aim of this paper has been to best approximate the set of information available to the agents. Annual stocks and harvests are observed individually for each appellation or origin of Bordeaux, to allow for a deeper examination of each appellation. In addition to the market data, an extended data set of exogenous predictors has been collected, including exchange rates, interest rates, GDP, country-specific wine productions, quality evaluations and meteorological data. The main outcome of this paper is that the annual forecasts generally outperform the naive random walk forecast, especially for the largest appellation and during shortages. Monthly forecasts broadly present the same features, but they are less effective at forecasting annual average prices. The multivariate analysis also yielded several collateral outcomes. The estimates provide additional evidence of the central role of inventories in agricultural price dynamics. I also find that weather news significantly influences current prices, although their explanatory power is low as compared to the other drivers. Finally, the estimates of an auxiliary harvest model indicate that temperatures are currently close to their optimal values for the production yields. If an increase in temperatures continues, a decrease in the production yields should be expected.

This last chapter contributes to several areas of wine economics, and thus leads the way to various directions of research. Forecasts can always be improved, and many options remains to be tested. In particular, price forecasts depends on delivery forecasts, which could be improved using the exogenous drivers of the price. Besides, I have chosen to guarantee that the estimated models comply to the theoretically expected influences of all determinants. As such, I have deliberately excluded machine-learning methods for which these influences are troublesome to characterize. Although these methods are poorly suited to the purpose of multivariate time series forecasting with exogenous predictors, it cannot be excluded that they could somehow improve the forecasts. Furthermore, the effect of the harvest, of the stocks and of weather news on prices could be estimated with more detail, using the structural framework of the competitive storage model as in Osborne (2004). Finally, the estimates of the aggregate harvest model call for further studies of the expected influence of climate change on wine production yield, depending on grape variety and geographic locations.

Bibliography

Ackerberg, D. A. (2003). Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review*, 44(3):1007–1040.

Ahumada, H. and Cornejo, M. (2016). Out-of-sample testing price discovery in commodity markets: the case of soybeans. *Agricultural Economics*, 47(6):709–718.

Aizenman, J. and Brooks, E. (2008). Globalization and Taste Convergence: the Cases of Wine and Beer. *Review of International Economics*, 16(2):217–233.

Akaike, H. (1973). Information theory and an extension of the Maximum likelihood Principal. In B. N. Petrox and F. Caski, editors, *2nd International Symposium on Information Theory*, page 267, Budapest. Akademiai Kiado.

Akerlof, G. A. (1970). The Market for Lemons: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Allen, P. G. (1994). Economic forecasting in agriculture. *International Journal of Forecasting*, 10(1):81–135.

Alston, J. M., Anderson, K., and Sambucci, O. (2015a). Drifting Towards Bordeaux? The Evolving Varietal Emphasis of U.S. Wine Regions. *Journal of Wine Economics*, 10(03):349–378.

Alston, J. M., Fuller, K. B., Lapsley, J. T., Soleas, G., and Tumber, K. P. (2015b). Splendide Mendax: False Label Claims About High and Rising Alcohol Content of Wine. *Journal of Wine Economics*, 10(03):275–313.

Anderson, K. (2014). Changing Varietal Distinctiveness of the World's Wine Regions: Evidence from a New Global Database. *Journal of Wine Economics*, 9(4):249–272.

Anderson, K. and Nelgen, S. (2011). *Global wine markets, 1961 to 2009: a statistical compendium*. University of Adelaide Press, Adelaide.

Anderson, K., Norman, D., and Wittwer, G. (2003). Globalisation of the World's Wine Markets. *The World Economy*, 26(5):659–687.

Anderson, K. and Wittwer, G. (2013). Modeling Global Wine Markets to 2018: Exchange Rates, Taste Changes, and China's Import Growth. *Journal of Wine Economics*, 8(02):131–158.

Anderson, K. and Wittwer, G. (2015). Asia's evolving role in global wine markets. *China Economic Review*, 35:1–14.

Anderson, M. and Magruder, J. (2012). Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database. *Economic Journal*, 122(563):957–989.

Arguea, N. and Hsiao, C. (1993). Econometric issues of estimating hedonic price functions. *Journal of Econometrics*, 56(1-2):243–267.

Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79.

Armstrong, J. S. and Collopy, F. (1992). Error Measures for Generalizing About Forecasting Methods: Empirical Comparisons. *International Journal of Forecasting*, 8(8):69–80.

Ashenfelter, O. (1989). How Auctions Work for Wine and Art. *The Journal of Economic Perspectives*, 3(3):23–36.

Ashenfelter, O. (2006). Tales from the Crypt: Auctioneer Bruce Kaiser Tells Us about the Trials and Tribulations of a Wine Judge. *Journal of Wine Economics*, 1(2):173–175.

Ashenfelter, O. (2008). Predicting the quality and prices of bordeaux wine. *The Economic Journal*, 118(529):174–184.

Ashenfelter, O., Ashmore, D., and Lalonde, R. (1995). Bordeaux Wine Vintage Quality and the Weather. *Chance*, 8(4):7–14.

Ashenfelter, O. and Jones, G. V. (1998). The demand for expert opinion: Bordeaux wine. In *VDQS Annual Meeting*, Ajaccio.

Ashenfelter, O. and Jones, G. V. (2013). The Demand for Expert Opinion: Bordeaux Wine. *Journal of Wine Economics*, 8(3):285–293.

Ashenfelter, O. and Quandt, R. (1999). Analyzing a Wine Tasting Statistically. *Chance*, 12(3):16–20.

Ashenfelter, O. and Storchmann, K. (2010). Using Hedonic Models of Solar Radiation and Weather to Assess the Economic Effect of Climate Change: The Case of Mosel Valley Vineyards. *Review of Economics and Statistics*, 92(2):333–349.

Ashenfelter, O. and Storchmann, K. (2016). Climate Change and Wine : A Review of the Economic Implications. *Journal of Wine Economics*, 11(1):105–138.

Ashton, R. H. (2011). Improving Experts' Wine Quality Judgments: Two Heads Are Better than One. *Journal of Wine Economics*, 6(02):160–178.

Ashton, R. H. (2012). Reliability and Consensus of Experienced Wine Judges: Expertise Within and Between? *Journal of Wine Economics*, 7(01):70–87.

Ashton, R. H. (2013). Is There Consensus Among Wine Quality Ratings of Prominent Critics? An Empirical Analysis of Red Bordeaux, 2004–2010. *Journal of Wine Economics*, 8(2):225–234.

Ashton, R. H. (2014a). Wine as an Experience Good: Price Versus Enjoyment in Blind Tastings of Expensive and Inexpensive Wines. *Journal of Wine Economics*, 9(2):171–182.

Ashton, R. H. (2014b). "Nothing Good Ever Came from New Jersey": Expectations and the Sensory Perception of Wines. *Journal of Wine Economics*, 9(03):304–319.

Ashton, R. H. (2016). The Value of Expert Opinion in the Pricing of Bordeaux Wine Futures. *Journal of Wine Economics*, 11:1–28.

Ashton, R. H. (2017). Dimensions of Expertise in Wine Evaluation. *Journal of Wine Economics*, 12(01):59–83.

Balcombe, K. (2010). The nature and determinants of volatility in agricultural prices: an empirical study from 1962-2008. *FAO Commodity Market Review*, (2009/2010):1–24.

Balinski, M. and Laraki, R. (2013). *How Best to Rank Wine: Majority Judgement*. Springer.

Barnard, H., Dooley, A. N., Areshian, G., Gasparyan, B., and Faull, K. F. (2011). Chemical evidence for wine production around 4000 BCE in the Late Chalcolithic Near Eastern highlands. *Journal of Archaeological Science*, 38(5):977–984.

Barrett, C. B. (2007). Displaced distortions: Financial market failures and seemingly inefficient resource allocation in low-income rural communities. In Bulte, E. and Ruben, R., editors, *Development economics between markets and institutions: Incentives for growth, food security and sustainable use of the environment*. Wageningen: Wageningen Academic Publishers.

Bélis-Bergouignan, M.-C. (2011). Bordeaux Wines: An Archetypal Terroir Cluster? *The Open Geography Journal*, 4:73–90.

Benfratello, L., Piacenza, M., and Sacchetto, S. (2009). Taste or reputation: what drives market prices in the wine industry? Estimation of a hedonic model for Italian premium wines. *Applied Economics*, 41(17):2197–2209.

Bernaleau Cardinel, N., Lamoureux, F., Delpech, E., Cathelineau, S., Laveau, E., Michaud, M.-C., Montagon, R., and Samie, B. (2012). Référentiel Technico-Economique du Vigneron Bordelais. Technical report, Chambre d'Agriculture de la Gironde, Bordeaux.

Bernard, O., Dussard, T., Kauffmann, J.-P., and Cazin, C. (2014). *La magie du 45e parallèle: Latitude idéale des grands vins du monde*. Editions Féret.

Bisson, L. F., Waterhouse, A. L., Ebeler, S. E., Walker, M. A., and Lapsley, J. T. (2002). The present and future of the international wine industry. *Nature*, 418(6898):696–699.

Bobenrieth, E., Wright, B., and Zeng, D. (2013). Stocks-to-use ratios and prices as indicators of vulnerability to spikes in global cereal markets. *Agricultural Economics*, 44(s1):43–52.

Bodington, J. C. (2015). Evaluating Wine-Tasting Results and Randomness with a Mixture of Rank Preference Models. *Journal of Wine Economics*, 10(1):31– 46.

Bois, B. (2007). *Cartographie agroclimatique à méso-échelle: méthodologie et application à la variabilité spatiale du climat en Gironde*. PhD thesis, University of Bordeaux 1 and INRA.

Bois, B. and Van Leeuwen, C. (2008). Variabilité climatique dans la zone de production des vins de Bordeaux. *ENITA Bordeaux*, pages 45–50.

Boudoukh, J., Richardson, M., Shen, Y. Q. J., and Whitelaw, R. F. (2007). Do asset prices reflect fundamentals? Freshly squeezed evidence from the OJ market. *Journal of Financial Economics*, 83(2):397–412.

Bouet, A., Decreux, Y., Fontagné, L., Jean, S., and Laborde, D. (2004). A consistent, ad-valorem equivalent measure of applied protection across the world: The MacMap-HS6 database. *CEPII Working Paper*, 2004-22.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (1970). *Time series analysis: forecasting and control*. John Wiley & Sons, San Fransiso.

Brandt, J. A. and Bessler, D. A. (1981). Composite forecasting: An application with US hog prices. *American Journal of Agricultural Economics*, 63(1):135–140.

Brandt, J. A. and Bessler, D. A. (1983). Price Forecasting and Evaluation: An Application in Agriculture. *Journal of Forecasting*, 2(3):1986–237.

Braun, H. I. and Holland, P. W. (1982). Observed-Score Test Equating : A Mathematical Anaysis of Some ETS Equating Procedures. In Holland, P. W. and Rubin, D. B., editors, *Test equating*, pages 9–49. Academic Press, New York.

Brochet, F. (2000). *La dégustation: étude des représentations des objets*. PhD thesis, Université de Bordeaux.

Brorsen, B. W. and Irwin, S. H. (1996). Improving the relevance of research on price forecasting and marketing strategies. *Agricultural and Resource Economics Review*, 25(1):68–75.

Burke, M. (2014). Selling low and buying high: An arbitrage puzzle in Kenyan villages.

Cafiero, C., Bobenrieth, E. S., Bobenrieth, J. R., and Wright, B. D. (2015). Maximum likelihood estimation of the standard commodity storage model:evidence from sugar prices. *American Journal of Agricultural Economics*, 97(1):122–16.

Cafiero, C., Bobenrieth, H., Eugenio, S., Bobenrieth, H., Juan, R., and Wright, B. D. (2011). The empirical relevance of the competitive storage model. *Journal of Econometrics*, 162(1):44–54.

Cafiero, C. and Wright, B. D. (2006). Is the storage model a 'closed' empirical issue? The empirical ability of the storage model to explain price dynamics. In A. Sarris and D. Hallam, editor, *Agricultural commodity markets and trade. New approaches to analyzing market structure and instability*, chapter 162, pages 89–114. Northampton: FAO/Edward Elgar.

Cardebat, J.-M. and Bazen, S. (2016). Forecasting bordeaux aoc wine prices using state space methods. In *10th Annual Conference of the American Association of Wine Economists*, Bordeaux.

Cardebat, J.-M. and Figuet, J.-M. (2004). What explains Bordeaux wine prices? *Applied Economics Letters*, 11(5):293–296.

Cardebat, J.-M. and Figuet, J.-M. (2009). Estimation of a hedonic price equation for Alsace, Beaujolais and Provence wines. *Applied Economics Letters*, 16(9):921–927.

Cardebat, J.-M., Figuet, J.-M., and Paroissien, E. (2014). Expert opinion and Bordeaux wine prices: an attempt to correct biases in subjective judgments. *Journal of Wine Economics*, 9(03):282–303.

Cardebat, J.-M. and Paroissien, E. (2015). Standardizing expert wine scores: An application for Bordeaux en primeur. *Journal of Wine Economics*, 10(3):329–348.

Carew, R. (2000). A Hedonic Analysis of Apple Prices and Product Quality Characteristics in British Columbia. *Canadian Journal of Agricultural Economics*, 48(3):241–257.

Chaikind, S. (2012). The Role of Viticulture and Enology in the Development of Economic Thought: How Wine Contributed to Modern Economic Theory. *Journal of Wine Economics*, 7(02):213–225.

Chamberlain, G. (1982). Multivariate Regression Models for Panel Data. *Journal of Econometrics*, 18:5–46.

Chambers, M. J. and Bailey, R. E. (1996). A theory of commodity price fluctuations. *Journal of Political Economy*, 104(5):924–957.

Chauvin, P.-M. (2013). The social fabric of prices: Institutional factors and reputation work in the bordeaux wine futures campaign. *economic sociology_the european electronic newsletter*, 15(1):12–21.

Chesher, A. (1991). The effect of measurement error. *Biometrika*, 78(3):451–462.

Chevet, J.-M., Lecocq, S., and Visser, M. (2011). Climate, grapevine phenology, wine production, and prices: Pauillac (1800–2009). *The American Economic Review*, 101(3):142–146.

Chou, P. H., Hsieh, C. H., and Shen, C. H. H. (2016). What explains the orange juice puzzle: Sentiment, smart money, or fundamentals? *Journal of Financial Markets*, 29:47–65.

Combris, P., Lecocq, S., and Visser, M. (1997). Estimation of a hedonic price equation for Bordeaux wine: does quality matter? *The Economic Journal*, 107(441):390–402.

Costanigro, M. and McCluskey, J. J. (2011). Hedonic price analysis in food markets. In Lusk, J., Roosen, J., and Shogren, J. F., editors, *The Oxford handbook of the economics of food consumption and policy*, page 152–180. Oxford University Press, Oxford.

Court, A. (1939). Hedonic price indexes with automotive examples. In *The Dynamics of Automobile Demand.*, page 99–117. General Motors Corporation, New York.

Cox, C. B., Luby, P. J., Pearson, F. A., Vial, E. E., and Rudd, R. W. (1956). Predicting hog prices. *Journal of Farm Economics*, 38(4):931–939.

Crozet, M., Head, K., and Mayer, T. (2012). Quality sorting and trade: Firm-level evidence for French wine. *Review of Economic Studies*, 79(2):609–644.

Deaton, A. and Laroque, G. (1992). On the Behaviour of Commodity Prices. *The Review of Economic Studies*, 59(1):1–23.

Deaton, A. and Laroque, G. (1996). Competitive storage and commodity price dynamics. *Journal of Political Economy*, 104(5):896–923.

Dewan, S. and Hsu, V. (2004). Adverse Selection in Electronic Markets: Evidence From Online Stamp Auctions. *Journal of Industrial Economics*, 52(4):497–516.

Di Vittorio, A. and Ginsburgh, V. (1996). Des enchères comme révélateurs du classement des vins. Les grands crus du Haut-Médoc. *Journal de la société française de statistique*, 137:19–49.

Dimson, E., Rousseau, P. L., and Spaenjers, C. (2015). The price of wine. *Journal of Financial Economics*, 118(2):431–449.

Dranove, D. and Zhe Jin, G. (2010). Quality Disclosure and Certification: Theory and Practice. *Journal of Economic Literature*, 484:935–963.

Dubois, P. and Nauges, C. (2010). Identifying the effect of unobserved quality and expert reviews in the pricing of experience goods: empirical application on Bordeaux wine. *International Journal of Industrial Organization*, 28(3):205–212.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64(4):813–836.

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55(2):251–276.

Engle, R. F. and Yoo, B. S. (1987). Forecasting and testing in co-integrated systems. *Journal of econometrics*, 35(1):143–159.

Espinosa, J. and Goodwin, B. (1991). Hedonic price estimation for Kansas wheat characteristics. *Western Journal of Agricultural Economics*, 16:72–85.

Ezekiel, M. (1938). The Cobweb Theorem. *The Quarterly Journal of Economics*, 52(2):255.

Fogarty, J. (2010). The Demand for Beer, Wine and Spirits: A Survey of the Literature. *Journal of Economic Surveys*, 24(3):428–478.

Freeman, A. (1993). *The measurement of environmental and resource values: theory and methods*. Resources for the Future, Washington, DC.

Friberg, R. and Grönqvist, E. (2012). Do expert reviews affect the demand for wine? *American Economic Journal: Applied Economics*, 4(1):193–211.

Fritz, C., Curtin, J., Poitevineau, J., Morrel-Samuels, P., and Tao, F.-C. (2012). Player Preferences Among New and Old Violins. *Proceedings of the National Academy of Sciences*, 109:760–763.

Geny, L. and Dubourdieu, D. (2015). Le millésime 2015 à Bordeaux. Technical report, Institut des Sciences de la Vigne et du Vin de l'Université de Bordeaux, Bordeaux.

Gergaud, O. and Ginsburgh, V. (2008). Natural Endowments, Production Technologies and the Quality of Wines in Bordeaux. Does Terroir Matter? *The Economic Journal*, 118(529):F142–F157.

Ginsburgh, V. (2003). Awards, Success and Aesthetic Quality in the Arts.

Ginsburgh, V., Monzak, M., and Monzak, A. (2013). Red Wines of Médoc: What is Wine Tasting Worth? *Journal of Wine Economics*, 8(2):159–188.

Ginsburgh, V. and Zang, I. (2012). Shapley Ranking of Wines. *Journal of Wine Economics*, 7(02):169–180.

Ginsburgh, V. A. and Van Ours, J. C. (2003). Expert Opinion and Compensation: Evidence from a Musical Competition. *The American Economic Review*, 93:289–296.

Giraud-Héraud, M. E. and Soler, L.-G. (2003). Quelle légitimité à des mécanismes de régulation de l'offre dans les Appellations d'Origine Protégée? *Economie rurale*, 277(1):123–134.

Goldin, C. and Rouse, C. (2000). Orchestrating Impartiality: The Impact of Blind Auditions on Female Musicians. *American Economic Review*, 90:715–741.

Goldstein, R., Almenberg, J., Dreber, A., Emerson, J. W., Herschkowitsch, A., and Katz, J. (2008). Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings. *Journal of Wine Economics*, 3(01):1–9.

Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G. (2013). The fundamentals of commodity futures returns. *Review of Finance*, pages 35–105.

Gouel, C. (2012). Agricultural price instability: A survey of competing explanations and remedies. *Journal of Economic Surveys*, 26(1):129–156.

Gouel, C. and Legrand, N. (2017a). Bayesian Estimation of the Storage Model using Information on Quantities. *TSE working paper*, 17-776.

Gouel, C. and Legrand, N. (2017b). Estimating the competitive storage model with trending commodity prices. *Journal of Applied Econometrics*, 32(4):744–763.

Granger, C. W. and Newbold, P. (1974). Spurious Regressions in Econometrics. *Journal of econometrics*, 2(2):111–120.

Greenfield, J. and Abbassian, A. (2011). Strengthening global food market monitoring. In Prakash, A., editor, *Safeguarding Food Security in Volatile Global Markets*, pages 441–459, Rome. FAO.

Griliches, Z. (1961). Hedonic price indexes for automobiles: An econometric of quality change. In *The Price Statistics of the Federal Goverment*, page 173–196. NBER, New York.

Guerra, V. E. A., Bobenrieth H, E. S. A., Bobenrieth H, J. R. A., and Cafiero, C. (2014). Empirical commodity storage model: the challenge of matching data and theory. *European Review of Agricultural Economics*, 42(4):607–623.

Guesnerie, R. (1992). An exploration of the eductive justifications of the rational-expectations hypothesis. *The American Economic Review*, 82(5):1254–1278.

Guilpart, N., Metay, A., and Gary, C. (2014). Grapevine bud fertility and number of berries per bunch are determined by water and nitrogen stress around flowering in the previous year. *European Journal of Agronomy*, 54:9–20.

Gustafson, R. L. (1958). Carryover levels for grains: a method for determining amounts that are optimal under specified conditions. Technical report, US Department of Agriculture.

Gutierrez, L. (2012). Speculative bubbles in agricultural commodity markets. *European Review of Agricultural Economics*, 40(2):217–238.

Hadj Ali, H., Lecocq, S., and Visser, M. (2008). The impact of Gurus: Parker grades and en primeur wine prices. *Economic Journal*, 118:158–173.

Hadj Ali, H. and Nauges, C. (2007). The pricing of experience goods: the example of en primeur wine. *american Journal of agricultural Economics*, 89(1):91–103.

Haeger, J. W. and Storchmann, K. (2006). Prices of American Pinot Noir wines: climate, craftsmanship, critics. *Agricultural Economics*, 35(1):67–78.

Harbaugh, R., Rasmusen, E., Dalton Professor, C. M., Baye, M., Connell, C., Lubensky, D., and Klerman, D. (2012). Coarse Grades: Informing the Public by Withholding Information.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge.

Hay, C. (2010). The political economy of price and status formation in the Bordeaux en primeur market: The role of wine critics as rating agencies. *Socio-Economic Review*.

Hendry, D. F. and Clements, M. P. (2003). Economic forecasting: Some lessons from recent research. *Economic Modelling*, 20(2):301–329.

Hilger, J., Rafert, G., and Villas-Boas, S. (2011). Expert Opinion and the Demand for Experience Goods: An Experimental Approach in the Retail Wine Market. *Review of Economics and Statistics*, 93(4):1289–1296.

Hodgson, R. T. (2008). An Examination of Judge Reliability at a major U.S. Wine Competition. *Journal of Wine Economics*, 3(02):105–113.

Hodgson, R. T. (2009). An Analysis of the Concordance Among 13 U.S. Wine Competitions. *Journal of Wine Economics*, 4(01):1–9.

Hoffman, D. L. and Rasche, R. H. (1996). Assessing forecast performance in a cointegrated system. *Journal of Applied Econometrics*, 11(5):495–517.

Holt, M. T. and Brandt, J. A. (1985). Combining price forecasting with hedging of hogs: an evaluation using alternative measures of risk. *Journal of Futures Markets*, 5(3):297–309.

Holt, M. T. and Craig, L. A. (2006). Nonlinear dynamics and structural change in the US hog—corn cycle: A time-varying STAR approach. *American Journal of Agricultural Economics*, 88(1):215–233.

Irwin, S. H. and Sanders, D. R. (2011). Index funds, financialization, and commodity futures markets. *Applied Economic Perspectives and Policy*, 33(1):1–31.

Jacks, D. S. (2007). Populists versus theorists: Futures markets and the volatility of prices. *Explorations in Economic History*, 44(2):342–362.

Jha, G. K. and Sinha, K. (2013). Agricultural Price Forecasting Using Neural Network Model: An Innovative Information Delivery System. *Agricultural Economics Research Review*, 26(2):229–239.

Jiao, L. (2017). Macroeconomic determinants of wine prices. *International Journal of Wine Business Research*, 29(3):234–250.

Jin, G. Z. and Leslie, P. (2003). The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards. *The Quarterly Journal of Economics*, 118(2):409–451.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6):1551–1580.

Jones, G. V. and Storchmann, K.-H. (2001). Wine market prices and investment under uncertainty: an econometric model for Bordeaux Crus Classes. *Agricultural Economics*, 26(2):115–133.

Jones, G. V., White, M. A., Cooper, O. R., and Storchmann, K. (2005). Climate change and global wine quality. *Climatic Change*, 73(3):319–343.

Just, R. E. and Rausser, G. C. (1981). Commodity price forecasting with large-scale econometric models and the futures market. *American Journal of Agricultural Economics*, 63(2):197–208.

Karlsson, P. and Karlsson, B. (2014). Cooperative Wine, Catching On and Catching Up.

Kastens, T. L., Schroeder, T. C., and Plain, R. (1998). Evaluation of extension and USDA price and production forecasts. *Journal of Agricultural and Resource Economics*, 21(1):244–261.

Kilian, L. and Taylor, M. P. (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics*, 60(1):85–107.

Kling, J. L. and Bessler, D. A. (1989). Calibration-based predictive distributions: An application of prequential analysis to interest rates, money, prices, and output. *Journal of Business*, pages 477–499.

Kolen, M. J. and Brennan, R. L. (2014). Nonequivalent Groups: Equipercentile Methods. In *Test Equating, Scaling, and Linking*, pages 143–169. Springer New York, New York, NY.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3):159–178.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2):132–157.

Landon, S., Smith, C. E., and Monzak, A. (1998). Quality Expectations, Reputation, and Price. *Southern Economic Journal*, 64(3):628.

Lecocq, S. and Visser, M. (2006a). Spatial variations in climate and Bordeaux wine prices. *Journal of Wine Economics*, 1(1):42–56.

Lecocq, S. and Visser, M. (2006b). What determines wine prices: Objective vs. sensory characteristics. *Journal of Wine Economics*, 1(01):42–56.

Lence, S. H. (2009). Do futures benefit farmers? *American Journal of Agricultural Economics*, 91(1):154–167.

L'Esperance, W. L. (1964). A Case Study in Prediction: The Market for Watermelons. *Econometrica*, 32(12):163–173.

Lima, T. (2006). Price and Quality in the California Wine Industry: An Empirical Investigation. *Journal of Wine Economics*, 1(2):176–190.

Mahenc, P. and Meunier, V. (2006). Early Sales of Bordeaux grands crus. *Journal of Wine Economics*, 1(01):57–74.

Maître d'Hôtel, E. and Le Cotty, T. (2015). Farm storage and asymmetric price volatility.

Martin-Consuegra, D., Molina, A., and Esteban, A. (2007). An integrated model of price, satisfaction and loyalty: an empirical analysis in the service sector. *Journal of Product & Brand Management*, 16(7):459–468.

Masset, P., Weisskopf, J.-P., and Cossutta, M. (2015). Wine tasters, ratings, and en primeur prices. *Journal of Wine Economics*, 10(1):75–107.

McCloskey, D. and Ziliak, S. (1996). The Standard Error of Regressions. *Journal of Economic Literature*, 34(1):97–114.

Milhau, J. (1953). L'avenir de la viticulture française. *Revue Economique*, 4(5):700–738.

Mobley-Martinez, T. (2007). The lone wolf: A conversation with wine critic Robert Parker.

Montgomery, C. A. and Wernerfelt, B. (1992). Risk Reduction and Umbrella Branding. *The Journal of Business*, 65(1):31–50.

Morrot, G., Brochet, F., and Dubourdieu, D. (2001). The Color of Odors. *Brain and Language*, 79(2):309–320.

Moutot, A. (2017). Gallo, ce géant californien devenu premier producteur de vin du monde.

Nakasone, E., Torero, M., Leonard, K., Hoffman, V., Martins-Filho, C., Almanzar, M., and Frisancho, V. (2013). The Role of Price Information in Agricultural Markets: Experimental Evidence from Rural Peru. In Agricultural and Applied Economics Association, editor, *2013 Annual Meeting, August 4-6, 2013, Washington, DC.*

Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy*, 78(2):311–329.

Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers. *European Economic Review*, 39(9):1697–1716.

Oczkowski, E. (1994). A hedonic price function for Australian premium table wine. *Australian Journal of Agricultural Economics*, 38(1):93–110.

Oczkowski, E. (2001). Hedonic Wine Price Functions and Measurement Error. *Economic Record*, 77(239):374–382.

Oczkowski, E. (2010). Hedonic wine price predictions and nonnormal errors. *Agribusiness*.

Oczkowski, E. (2016). The preferences and prejudices of Australian wine critics. *Journal of Wine Research*, pages 1–12.

Oczkowski, E. and Doucouliagos, H. (2015). Wine Prices and Quality Ratings: A Meta-regression Analysis. *American Journal of Agricultural Economics*, 97(1):103–121.

Olkin, I., Lou, Y., Stokes, L., and Cao, J. (2015). Analyses of Wine-Tasting Data: A Tutorial. *Journal of Wine Economics*, 10(01):4–30.

Osborne, T. (2004). Market News in Commodity Price Theory: Application to the Ethiopian Grain Market. *Review of Economic Studies*, 71(1):133–164.

Pichet, E. (2010). Who Killed the WINEFEX?: Autopsy of the Bordeaux Wine Futures Contract.

Piqueras-Fiszman, B. and Spence, C. (2015). Sensory expectations based on product-extrinsic food cues: An interdisciplinary review of the empirical evidence and theoretical accounts.

Prescott, D. M. and Stengos, T. (1987). Bootstrapping confidence intervals: An application to forecasting the supply of pork. *American Journal of Agricultural Economics*, 69(2):266–273.

Quandt, R. E. (2007). On Wine Bullshit: Some New Software? *Journal of Wine Economics*, 2(02):129–135.

Rabobank (2012). The Incredible Bulk The Rise in Global Bulk Wine Trade. Technical report, Rabobank Industry.

Redelmeier, D. A. and Baxter, S. D. (2009). Rainy Weather and Medical School Admission Interviews. *CMAJ*, 181:933.

Reinstein, D. and Snyder, C. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The journal of industrial economics*, 53:27–51.

Roll, R. (1984). Orange juice and weather. *The American Economic Review*, 74(5):861–880.

Roma, P., Di Martino, G., and Perrone, G. (2013). What to show on the wine labels: a hedonic analysis of price drivers of Sicilian wines. *Applied Economics*, 45(19):2765–2778.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3):638–641.

Salop, S. (1976). Information and Monopolistic Competition. *The American Economic Review, Papers and Proceedings*, 66(2):240–245.

Samuelson, P. A. (1938). A Note on the Pure Theory of Consumer's Behaviour. *Economica*, 5(17):61.

Sanders, D. R. and Irwin, S. H. (2016). The “Necessity” of new position limits in agricultural futures markets: The verdict from daily firm-level position data. *Applied Economic Perspectives and Policy*, 38(2):292–317.

Sarle, C. F. (1925). The forecasting of the price of hogs. *The American Economic Review*, 15(3):1–22.

Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3):289–310.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1):1.

Smith, A., De Maillard, J., and Costa, O. (2007). *Vin et politique: Bordeaux, la France, la mondialisation*. Les Presses de Sciences Po.

Sorensen, A. T. and Rasmussen, S. J. (2004). Is Any Publicity Good Publicity? A Note on the Impact of Book Reviews.

Stanley, L. R. and Tschirhart, J. (1991). Hedonic Prices for a Nondurable Good: The Case of Breakfast Cereals. *The Review of Economics and Statistics*, 73(3):537.

Storchmann, K. (2012). Wine Economics. *Journal of Wine Economics*, 7(01):1–33.

Storchmann, K., Mitterling, A., and Lee, A. (2012). The detrimental effect of expert opinion on price-quality dispersion: Evidence from the wine market.

Symeonidis, L., Prokopczuk, M., Brooks, C., and Lazar, E. (2012). Futures basis, inventory and commodity price volatility: An empirical analysis. *Economic Modelling*, 29(6):2651–2663.

Taylor, L. O. (2003). The Hedonic Method. pages 331–393. Springer, Dordrecht.

Timmer, C. P. (2010). Did speculation affect world rice prices? In Dawe D., editor, *The Rice Crisis: Markets, Policies and Food Security*, pages 29–60. Routledge, Rome and London, fao and earthscan edition.

Timmermann, A. (2006). Forecast combinations. In G. Elliot, C. W. J. G. and Timmermann, A., editors, *Handbook of economic forecasting*, pages 135–196. Elsevier, Amsterdam.

Timmermann, A. and Granger, C. W. J. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting*, 20(1):15–27.

Tomek, W. G. (1996). Commodity futures prices as forecasts. *Review of Agricultural Economics*, 19(1):23–44.

Traversac, J.-B., Rousset, S., and Perrier-Cornet, P. (2011). Farm resources, transaction costs and forward integration in agriculture: Evidence from French wine producers. *Food Policy*, 36(6):839–847.

Triplett, J. (1989). Price and technological change in a capital good: A survey of research on computers. In D. Jorgenson and R. Landau, editor, *Technology and capital formation*, page 127–213. MIT Press, Cambridge.

Triplett, J. (2004). *Handbook on hedonic indexes and quality adjustments in price indexes: special application to information technology products*. OECD, Paris.

Vaillant, N. G., Lesot, P., Bonnard, Q., and Harrant, V. (2010). The use of expert opinion, quality and reputation indicators by consumers: evidence from the French vaulting stallion semen market. *Applied Economics*, 42(6):739–745.

Weil, R. L. (2001). Parker v. Prial: The Death of the Vintage Chart. *Chance*, 14(4):27–31.

Weil, R. L. (2005). Analysis of Reserve and Regular Bottlings: Why Pay for a Difference Only the Critics Claim to Notice? *CHANCE*, 18(3):9–15.

Williams, J. C. and Wright, B. D. (1991). *Storage and Commodity Markets*. Cambridge University Press, New York.

Wimmer, B. S. and Chezum, B. (2003). An Empirical Examination of Quality Certification in a Lemons Market. *Economic Inquiry*, 41(2):279–291.

Wine services (2015). http://www.wine-services.com.

Wittwer, G., Berger, N., and Andersona, K. (2003). A model of the world's wine markets. *Economic Modelling*, 20(3):487–506.

Wood, D. and Anderson, K. (2006). What Determines the Future Value of an Icon Wine? New Evidence from Australia. *Journal of Wine Economics*, 1(02):141–161.

Wooldridge, J. (2002). *Econometrics Analysis of Cross Section and Panel Data*. MIT Press.

Wright, B. (2014). Global biofuels: key to the puzzle of grain market behavior. *The Journal of Economic Perspectives*, 28(1):73–97.

Wright, B. D. (2011). The economics of grain price volatility. *Applied Economic Perspectives and Policy*, 33(1):32–58.

Wright, B. D. and Williams, J. C. (1982). The economic role of commodity storage. *The Economic Journal*, 92(367):596–614.

Yeo, M., Fletcher, T., and Shawe-Taylor, J. (2015). Machine Learning in Fine Wine Price Prediction. *Journal of Wine Economics*, 10(2):151–172.

Zhao, W. (2008). Social Categories, Classification Systems, and Determinants of Wine Price in the California and French Wine Industries. *Sociological Perspectives*, 51(1):163–199.

Zufferey, V. and Maigre, D. (2008). Age de la vigne. II. Influence sur la qualité des raisins et des vins. *Revue suisse de viticulture arboriculture horticulture*, 40(4):241–246.

Appendix A

Appendix of Chapter 1

TABLE A.1 – Acronyms and abbreviations

Short name	Full name
LLG	Law of Large Numbers
OLS	Ordinary Least Squares
AOC	*Appellation d'Origine Contrôlée*
VAT	Value-Added Tax
VIF	Variance Inflation Factor

TABLE A.2 – Experts' Scores

Code	Expert	Number of observed scores
AP	André Proensa	480
IWC	International Wine Cellar (Stephen Tanzer)	1,169
GM	Gault Millau	352
JMQ	Jean-Marc Quarin	1,857
JR	Jancis Robinson	1,723
RVF	Revue des Vins de France	217
WA	The Wine Advocate (Robert Parker)	1,644
WS	Wine Spectator	1,758
920R	920-Revue	313

Appendix B

Appendix of Chapter 2

FIGURE B.1 – Conversion Functions into Parker's Scale

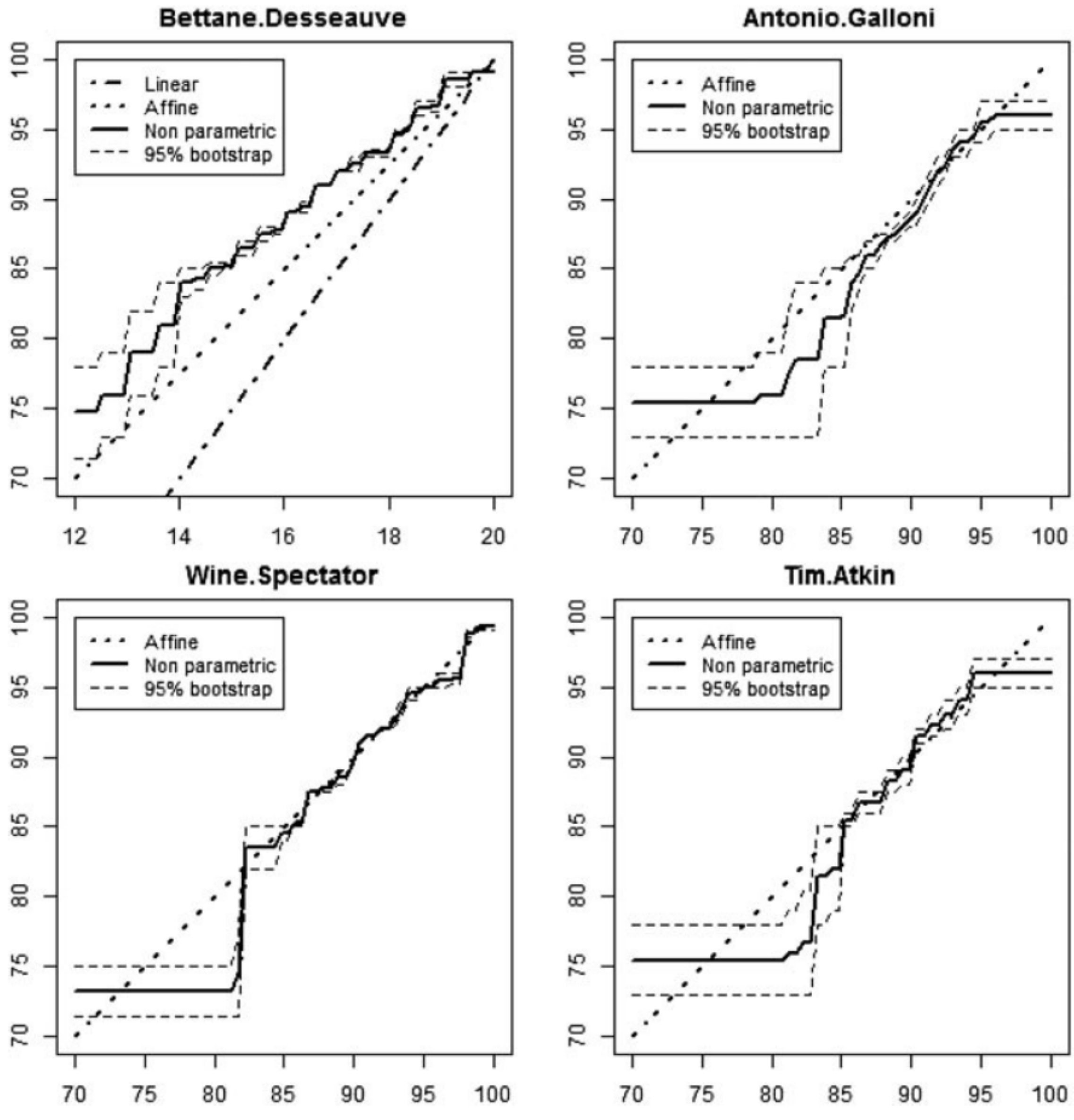


FIGURE B.2 – Conversion Functions into Parker's Scale

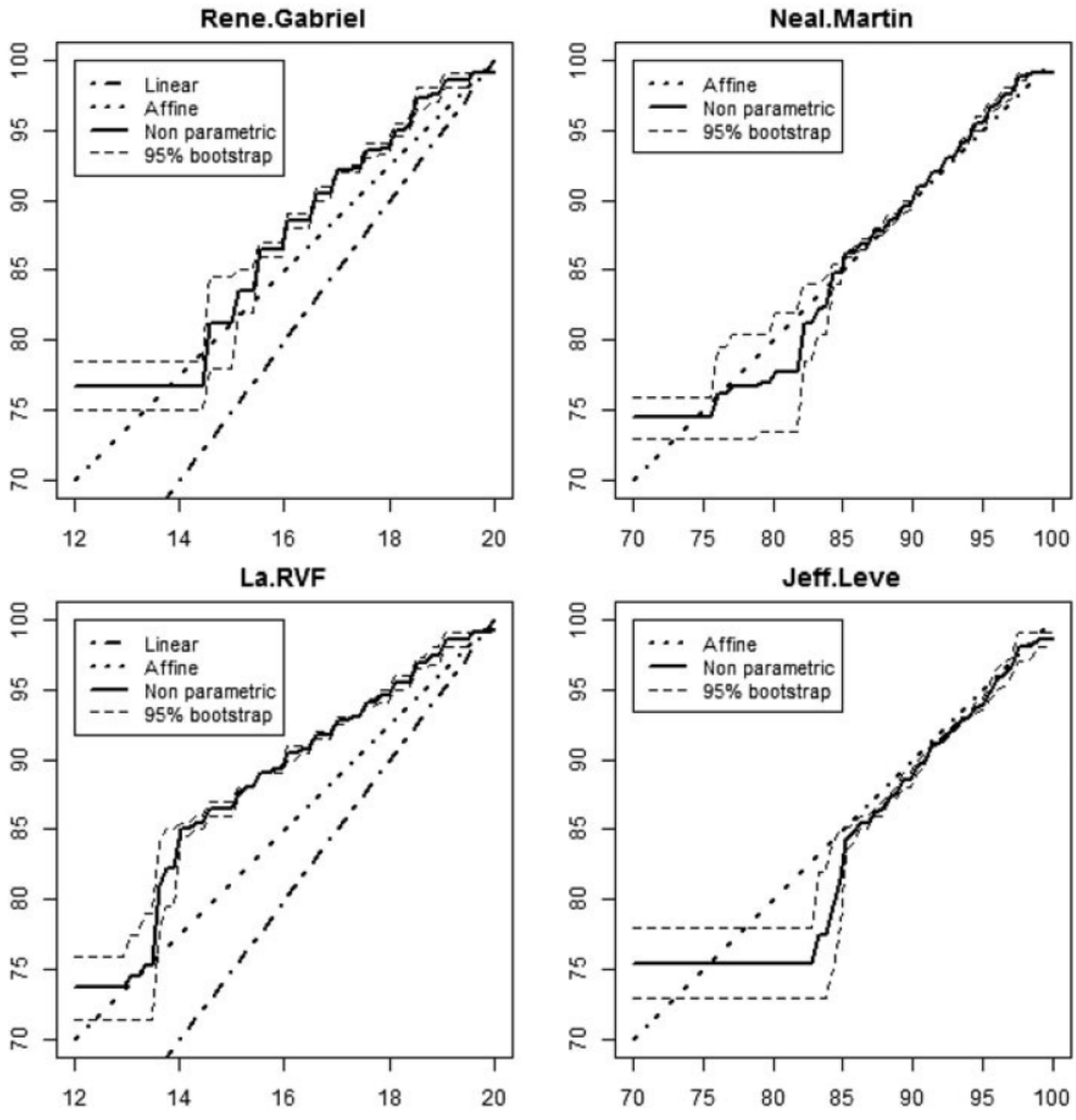


FIGURE B.3 – Conversion Functions into Parker's Scale

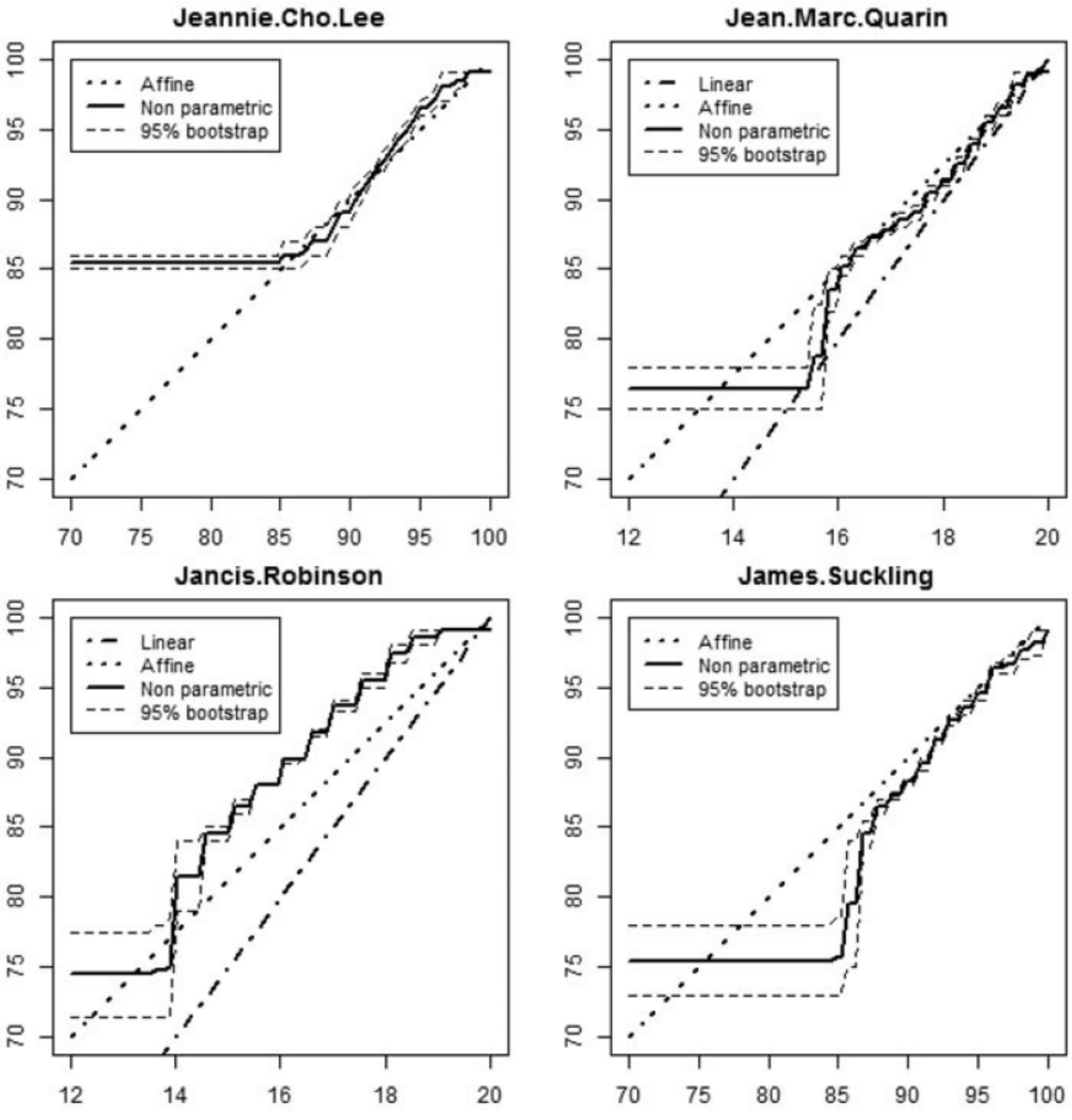


FIGURE B.4 – Conversion Functions into Parker's Scale

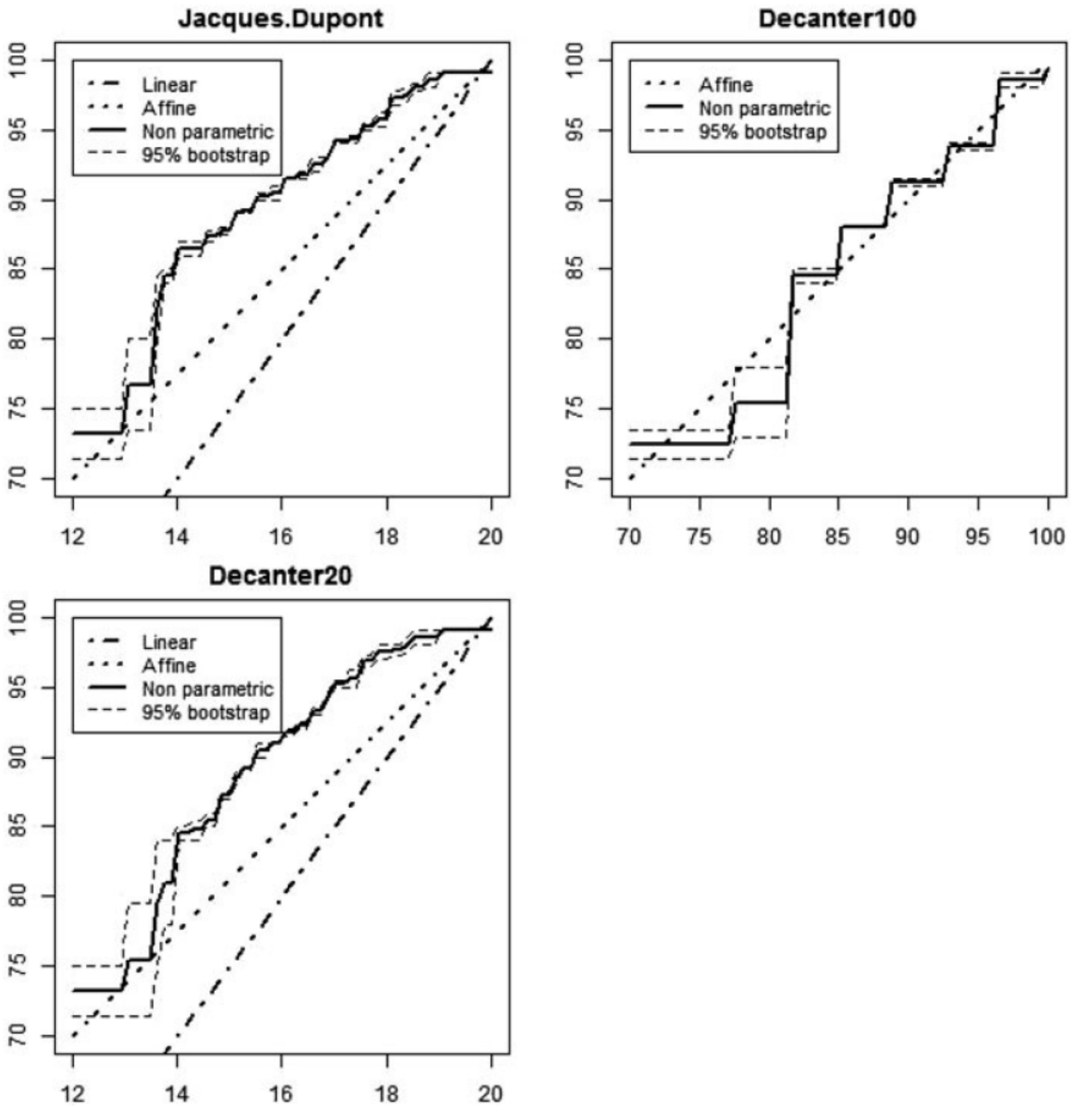


TABLE B.1 – Correlation Matrix after Conversion

	RP	NM	JR	WS	AG	BD	JD	JS	JC	JL	RVF	JMQ	RG	TA	D20	D100
RP		0.58	0.42	0.60	0.39	0.59	0.49	0.67	0.68	0.77	0.58	0.67	0.58	0.59	0.70	0.55
NM	0.58		0.49	0.60	0.57	0.57	0.49	0.67	0.58	0.74	0.57	0.65	0.57	0.59	0.66	0.59
JR	0.42	0.49		0.48	0.17	0.49	0.36	0.44	0.55	0.40	0.43	0.50	0.45	0.58	0.61	0.36
WS	0.60	0.60	0.48		0.54	0.59	0.46	0.68	0.70	0.69	0.56	0.62	0.60	0.65	0.67	0.55
AG	0.39	0.57	0.17	0.54		0.44	0.33	0.47		0.57	0.42	0.47	0.54	0.35	0.18	0.58
BD	0.59	0.57	0.49	0.59	0.44		0.48	0.62	0.71	0.65	0.62	0.66	0.55	0.64	0.74	0.70
JD	0.49	0.49	0.36	0.46	0.33	0.48		0.54	0.56	0.59	0.52	0.56	0.46	0.53	0.60	0.62
JS	0.67	0.67	0.44	0.68	0.47	0.62	0.54		0.66	0.72	0.64	0.65	0.63	0.54	0.67	0.55
JC	0.68	0.58	0.55	0.70		0.71	0.56	0.66		0.70	0.57	0.62	0.64	0.63	0.72	
JL	0.77	0.74	0.40	0.69	0.57	0.65	0.59	0.72	0.70		0.69	0.75	0.66	0.57	0.67	0.62
RVF	0.58	0.57	0.43	0.56	0.42	0.62	0.52	0.64	0.57	0.69		0.69	0.54	0.54	0.66	0.69
JMQ	0.67	0.65	0.50	0.62	0.47	0.66	0.56	0.65	0.62	0.75	0.69		0.63	0.67	0.71	0.68
RG	0.58	0.57	0.45	0.60	0.54	0.55	0.46	0.63	0.64	0.66	0.54	0.63		0.58	0.65	0.63
TA	0.59	0.59	0.58	0.65	0.35	0.64	0.53	0.54	0.63	0.57	0.54	0.67	0.58		0.66	0.62
D20	0.70	0.66	0.61	0.67	0.18	0.74	0.60	0.67	0.72	0.67	0.66	0.71	0.65	0.66		
D100	0.55	0.59	0.36	0.55	0.58	0.70	0.62	0.55		0.62	0.69	0.68	0.63	0.62		
Average	0.59	0.59	0.45	0.60	0.43	0.60	0.51	0.61	0.64	0.65	0.58	0.63	0.58	0.58	0.64	0.60
Abs. Difference	0.000	0.002	0.011	0.008	0.018	0.003	0.010	0.013	0.004	0.004	0.021	0.011	0.007	0.016	0.005	0.025

Appendix C

Appendix of Chapter 3

TABLE C.1 – Number of medals per wine before/after transaction

		After				
		0	1	2	3+	Total
Before	0	13,298	302	62	26	13,688
	1	1,517	129	32	10	1,688
	2	612	13	7	0	632
	3+	385	5	1	0	391
	Total	15,812	449	102	36	16,399

TABLE C.2 – Number of medals across competitions, before and after transaction date

Competition	Before transaction date				After transaction date			
	# Medals	# Bronze	# Silver	# Gold	# Medals	# Bronze	# Silver	# Gold
BOR	1,119	294	410	415	178	42	74	62
BRU	214	0	129	85	60	0	37	23
CHA	358	99	141	118	125	56	45	24
CVI	171	70	45	56	30	13	8	9
DEC	233	84	21	5	68	36	9	0
FEM	248	88	95	65	48	9	25	14
LYO	258	26	71	161	44	5	15	24
MAC	735	300	195	240	112	36	39	37
PAR	727	109	274	344	69	12	27	30
VIN	145	86	51	8	24	15	7	2
VII	30	0	28	2	11	0	10	1

TABLE C.3 – Estimates of α_M simultaneously by color and competition

Estimate	$M_{BORgold}$	$M_{BORsilver}$	$M_{BORbronze}$	$M_{PARgold}$	$M_{PARsilver}$	$M_{PARbronze}$
$\hat{\alpha}^{OLS}_{M_j}$	0.206	0.093	0.052	0.225	0.062	0.103
	(0.012)	(0.011)	(0.013)	(0.013)	(0.014)	(0.022)
$\hat{\alpha}^{DIF}_{M_j}$	0.134	0.073	-0.015	0.135	-0.029	0.113
	(0.028)	(0.031)	(0.04)	(0.05)	(0.046)	(0.048)
$\widehat{\alpha_{M_j} + \beta_{M_j}}$	0.206	0.092	0.054	0.221	0.057	0.104
	(0.012)	(0.011)	(0.013)	(0.013)	(0.013)	(0.022)
$\hat{\beta}_{F_j}$	0.072	0.019	0.068	0.086	0.086	-0.009
	(0.026)	(0.028)	(0.037)	(0.046)	(0.044)	(0.039)
Characteristics X	Yes					
Fixed effects	Yes					
N	16.399					
R^2	0.925					

TABLE C.4 – Distribution of expected profit in euro, including stickers costs

Competition	π	Mean	S.d.	Min	Max	p25	p75	$\%E(Profit) < 0$
Bordeaux	0.05	61	212	-166	5,034	-48	88	0.518
Challenge	0.05	50	220	-154	5,121	-66	81	0.563
Decanter	0.05	-43	315	-337	7,256	-209	-2	0.751
Paris	0.05	45	212	-189	5,029	-64	72	0.574
Bordeaux	0.10	208	423	-247	10,153	-11	261	0.294
Challenge	0.10	212	441	-196	10,353	-21	273	0.322
Decanter	0.10	160	630	-428	14,758	-172	241	0.548
Paris	0.10	193	423	-275	10,161	-25	247	0.339
Bordeaux	0.20	500	847	-409	20,392	62	606	0.135
Challenge	0.20	535	881	-280	20,818	70	658	0.14
Decanter	0.20	566	1,260	-610	29,762	-98	727	0.36
Paris	0.20	490	847	-446	20,425	53	597	0.154
Bordeaux	0.30	793	1,270	-571	30,630	135	952	0.079
Challenge	0.31	891	1,366	-373	32,329	170	1,081	0.075
Decanter	0.59	2,148	3,717	-1321	88,275	189	2,624	0.167
Paris	0.24	609	1,016	-515	24,530	84	737	0.122

TABLE C.5 – Distribution of expected profit in euro, without stickers costs

Competition	π	Mean	S.d.	Min	Max	p25	p75	$\%E(Profit) < 0$
Bordeaux	0.05	122	257	-84	5,555	-21	167	0.345
Challenge	0.05	95	257	-110	5,528	-48	140	0.449
Decanter	0.05	54	372	-244	7,935	-153	120	0.61
Paris	0.05	103	257	-102	5,537	-39	149	0.421
Bordeaux	0.10	328	513	-82	11,194	43	419	0.145
Challenge	0.10	302	513	-109	11,168	16	392	0.212
Decanter	0.10	354	745	-242	16,115	-60	485	0.344
Paris	0.10	310	513	-101	11,176	24	400	0.191
Bordeaux	0.20	741	1,027	-80	22,474	170	922	0.048
Challenge	0.20	715	1,027	-106	22,447	144	895	0.077
Decanter	0.20	953	1,490	-238	32,475	125	1,215	0.145
Paris	0.20	723	1,027	-98	22,456	152	904	0.068
Bordeaux	0.30	1,154	1,540	-77	33,753	298	1,426	0.017
Challenge	0.31	1,169	1,592	-104	34,855	284	1,449	0.033
Decanter	0.59	3,290	4,394	-223	96,281	847	4,064	0.018
Paris	0.24	888	1,232	-97	26,967	203	1,105	0.049

TABLE C.6 – Estimates of α_{M_j} with M_j being the number of gold, silver and bronze medals

Estimate	M_{gold}	M_{silver}	M_{bronze}
$\hat{\alpha}_M^{OLS}$	0.152	0.064	0.057
	(0.007)	(0.006)	(0.007)
$\hat{\alpha}_M^{DIF}$	0.087	0.007	-0.003
	(0.009)	(0.009)	(0.016)
$\widehat{\alpha_M + \beta_M}$	0.152	0.064	0.058
	(0.007)	(0.006)	(0.007)
$\hat{\beta}_F$	0.064	0.058	0.061
	(0.006)	(0.007)	(0.015)
Characteristics X	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
N	16399	16399	16399
R^2_{OLS}	0.924	0.924	0.924
R^2_{DIF}	0.924	0.924	0.924

Appendix D

Appendix of Chapter 4

Expectation of the exponential of a truncated normal variable (section 4.4.1)

Let $\epsilon \sim \mathcal{N}(\mu, \sigma^2)$, Φ and ϕ be respectively the distribution and the density functions of a normal $\mathcal{N}(0,1)$, and $\phi_{\mu,\sigma}$ be the density of the law $\mathcal{N}(\mu, \sigma^2)$. Recall that:

$$\Phi(y) = \int_{-\infty}^{y} \phi(x)\mathrm{d}x \qquad ; \qquad \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad ; \qquad \phi_{\mu,\sigma}(x) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)$$

The computation of the expectation of e^{ϵ} given $\epsilon < 0$ is elementary:

$$\begin{aligned}
E(e^{\epsilon}|\epsilon < 0) &= \frac{1}{\mathbb{P}(\epsilon < 0)} \int_{-\infty}^{0} e^{x}\phi_{\mu,\sigma}(x)\mathrm{d}x \\
&= \frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{0} e^{x}\phi_{\mu,\sigma}(x)\mathrm{d}x \\
&= \frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{0} e^{x}\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x \\
&= \frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}+x}\mathrm{d}x \\
&= \frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[x-(\mu+\sigma^2)]^2}{2\sigma^2}+(\mu+\frac{1}{2}\sigma^2)}\mathrm{d}x \\
&= \frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[x-(\mu+\sigma^2)]^2}{2\sigma^2}} e^{(\mu+\frac{1}{2}\sigma^2)}\mathrm{d}x \\
&= e^{(\mu+\frac{1}{2}\sigma^2)}\frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{[x-(\mu+\sigma^2)]^2}{2\sigma^2}}\mathrm{d}x \\
&= e^{(\mu+\frac{1}{2}\sigma^2)}\frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{-\frac{\mu+\sigma^2}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\mathrm{d}x \\
&= e^{(\mu+\frac{1}{2}\sigma^2)}\frac{1}{\Phi(-\frac{\mu}{\sigma})} \int_{-\infty}^{-\frac{\mu+\sigma^2}{\sigma}} \phi(x)\mathrm{d}x \\
E(e^{\epsilon}|\epsilon < 0) &= e^{(\mu+\frac{1}{2}\sigma^2)}\frac{\Phi(-\frac{\mu+\sigma^2}{\sigma})}{\Phi(-\frac{\mu}{\sigma})}
\end{aligned}$$

TABLE D.1 – Acronyms and abbreviations

Short name	Full name
Models	
ADL	Auto-Distributed Lags model
DADL	Differenced Auto-Distributed Lags model
VAR	Vector Auto-Regressive model
ECM	Error-Correction Model
UCM	Unobserved Component Model
Estimation method	
OLS	Ordinary Least Squares
2SLS	Two-Stage Least Squares
Statistical tests	
ADF	Augmented Dickey-Fuller test
ERS	Elliot-Rothenberg-Stock test
KPSS	Kwiatkowski–Phillips–Schmidt–Shin test
Acronyms	
AMIS	Agricultural Market Information System
AOC	*Appellation d'Origine Contrôlée*
BCF	*Banque de France*
CIVB	*Conseil Interprofessionnel du Vin de Bordeaux*
GDP	Gross Domestic Production
HUR	Harvest-to-Use Ratio
INAO	*Institut National de l'Origine et de la Qualité*
INSEE	*Institut National de la Statistique et des Etudes Economiques*
MAPE	Mean Average Percentage Error
MRAE	Mean Relative Absolute Error
OECD	Organisation for Economic Co-operation and Development
OIV	*Organisation Internationale de la Vigne et du Vin*
SUR	Stock-to-Use Ratio
UK	United Kingdom
USA	United States of America
Abbreviations	
ha	hectare
hl	hectolitre

Time series econometrics reminder

Autoregressive Distributive Lag models

Let y_t be a stationary process following an ADL(p,q), with $(x_{1t}, ..., x_{nt})$ being the set of the explanatory variables. The variables x_{it} are also stationary. The dynamics of y_t are given by the following equation.

$$y_t = \alpha_0 + \sum_{k=1}^{p} \alpha_k y_{t-k} + \sum_{i=1}^{n} \sum_{l=0}^{q} \beta_{il} x_{it-k} + \epsilon_t \tag{D.1}$$

where ϵ_t is a zero-mean, exogenous and time-independent error term with a constant variance σ_ϵ^2.

The process y_t is said to follow a differentiated ADL(p,q), noted DADL(p,q), if the differentiated series Δy_t follows an ADL(p,q).

Cointegration

Let $X_t = (x_{1t}, ..., x_{nt})$ be a set of non-stationary variables for which ΔX_t is stationary. X_t is cointegrated if there exists a vector $\gamma = (\gamma_1, ..., \gamma_n)^t$ such that $X_t \gamma = \sum_{i=1}^{n} \gamma_i x_{it}$ is a stationary process.

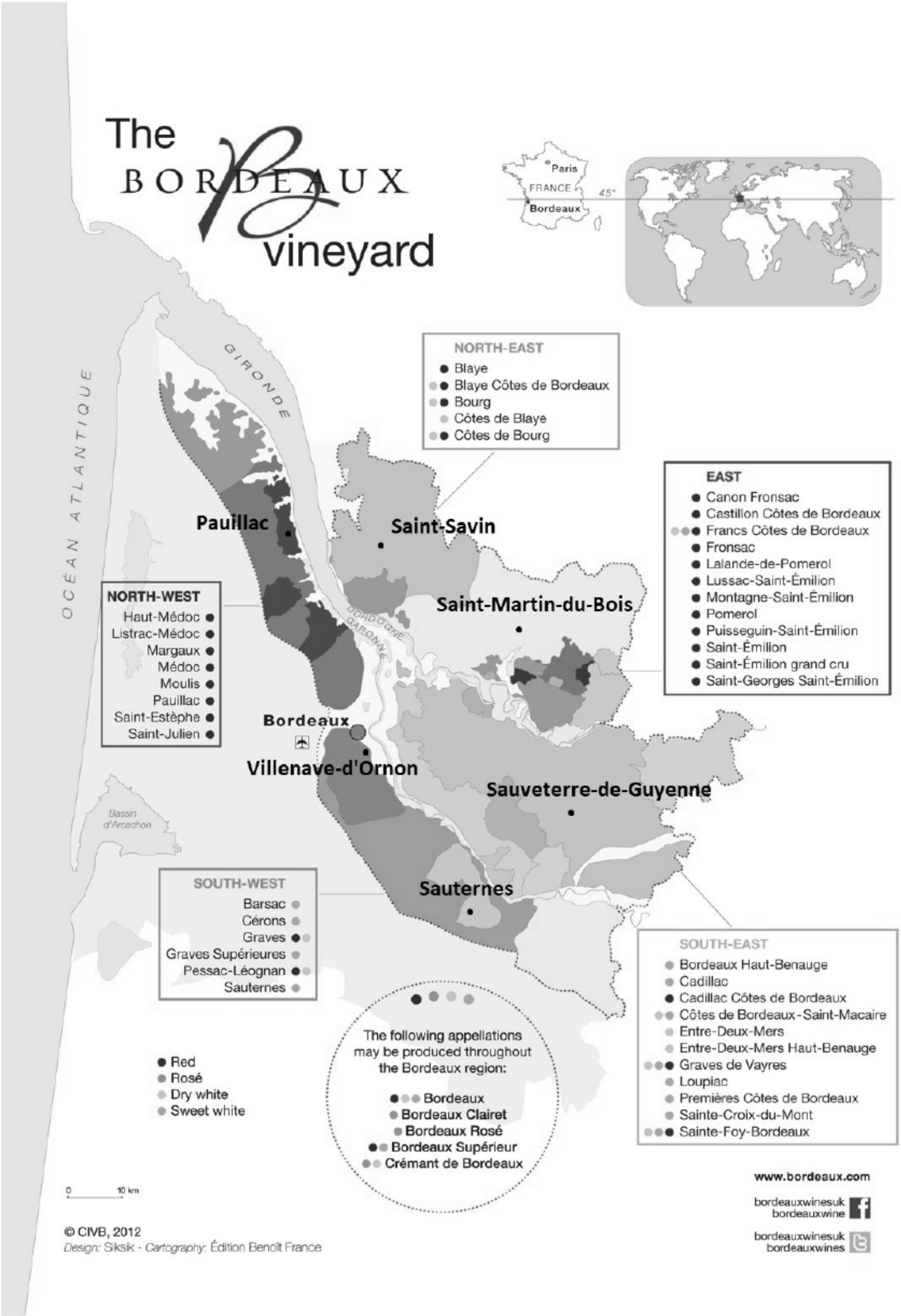
Error-Correction representation in the Engle and Granger (1987) framework

Let y_t be a non-stationary process and $X_t = (x_{1t}, ..., x_{nt})$ be a set of non-stationary variables such that (y_t, x_t) is cointegrated, and $\gamma = (\gamma_1, ..., \gamma_n)^t$ be one cointegration vector. y_t is said to follow an ECM(p,q) if its dynamics are given by the following equation.

$$\Delta y_t = \alpha_0 + \sum_{k=1}^{p} \alpha_k \Delta y_{t-k} + \sum_{i=1}^{n} \sum_{l=0}^{q} \beta_{il} \Delta x_{it-l} + \delta(y_{t-1} - \sum_{i=1}^{n} \gamma_i x_{it-1}) + \epsilon_t \tag{D.2}$$

where ϵ_t is a zero-mean, exogenous and time-independent error term with a constant variance σ_ϵ^2. In the 2SLS procedure of Engle and Granger (1987), the cointegration vector γ is estimated by OLS in a first stage regression of y_t on X_t. This cointegration equation is said to represent the long-run equilibrium around which prices fluctuate. The other coefficients are also estimated by OLS in second stage regression of Δy_t on the Δy_{t-k}, Δx_{it-1}, and the lagged error term of the long-run equation $(y_{t-1} - \sum_{i=1}^{n} \hat{\gamma}_i x_{it-1})$. Johansen (1991) proposed an extended framework where all variables are endogenous and which may include several cointegrating relationships. If the dynamics of each endogenous variable are still given by equation (D.2), the estimation procedure is a maximum likelihood assuming normal errors and a number r of non-collinear cointegrating relationships.

FIGURE D.1 – Geographic locations of the six weather stations



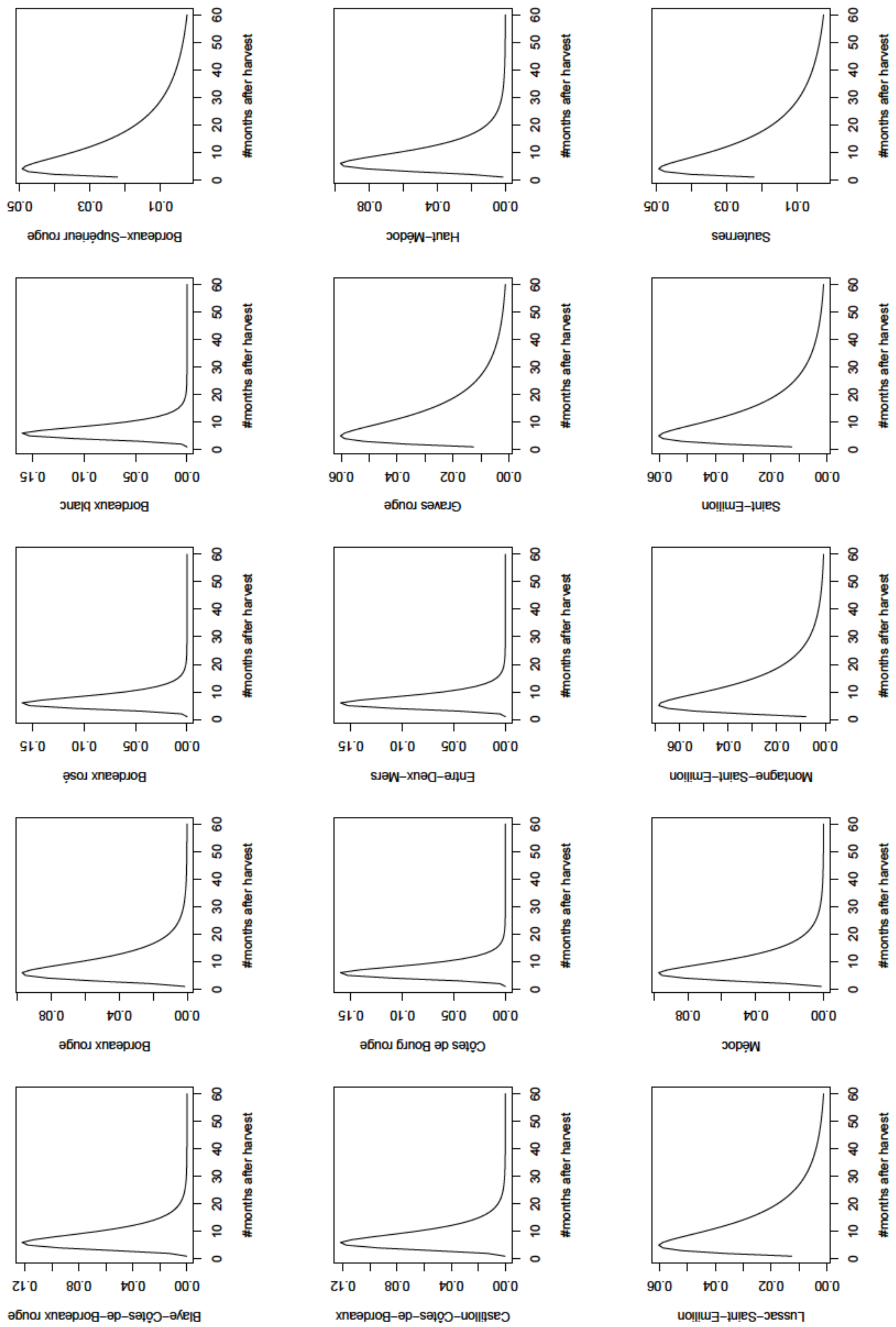


FIGURE D.2 – Flow functions for the weights of each vintage in the indicator of average transacted quality

TABLE D.2 – Attribution and sample means of the weather variables

Station	AOC	$\overline{F4}$	$\overline{TX46}$	$\overline{RR57}$	$\overline{RR89}$
Pauillac	HME, MED	5.19	21.07	163.92	116.73
Saint-Martin-du-Bois	CAS, LU, MSE, SE	2.1	21.19	178.71	117.84
Saint-Savin	BLA, CBO	1.82	21.44	168.97	120.21
Sauternes	SAU	7.56	21.35	177.36	125.99
Sauveterre-de-Guyenne	BB, BRO, BR, BSR, EDM	10.42	21.03	171.35	125.22
Villenave-d'Ornon	GRA	1.87	21.46	180.03	130.53

TABLE D.3 – Estimated variances of the parameters of the UCM filter for prices

AOC	$\hat{\sigma}^2_{\epsilon_i^p}$	$\hat{\sigma}^2_{\nu_i^p}$	$\hat{\sigma}^2_{\xi_i^p}$	$\hat{\sigma}^2_{\omega_i^p}$
BLA	3.0e-03	4.5e-04	4.1e-07	1.7e-08
BSR	2.1e-03	5.0e-04	3.9e-07	3.2e-05
BR	1.5e-03	1.4e-03	7.1e-06	8.7e-09
BRO	2.3e-03	1.2e-03	1.4e-07	1.5e-08
BW	8.9e-05	1.0e-03	2.2e-07	1.5e-08
CAS	2.1e-02	1.5e-04	1.1e-06	2.1e-06
CBO	2.7e-03	5.9e-04	3.0e-07	1.0e-05
EDM	4.0e-03	1.1e-03	1.8e-09	1.2e-06
GRA	3.8e-03	1.2e-03	3.7e-07	5.9e-05
HME	1.4e-02	1.0e-03	7.7e-07	7.4e-06
LU	2.0e-03	1.1e-03	1.1e-10	7.5e-09
MED	4.0e-03	1.4e-03	2.7e-07	3.0e-05
MSE	2.4e-03	1.3e-03	2.8e-10	6.7e-08
SE	9.8e-04	1.2e-03	5.1e-09	4.2e-09
SAU	2.1e-03	1.1e-03	1.8e-10	1.2e-07

FIGURE D.3 – Observed monthly real prices (black) and the estimated underlying levels (grey) by AOC

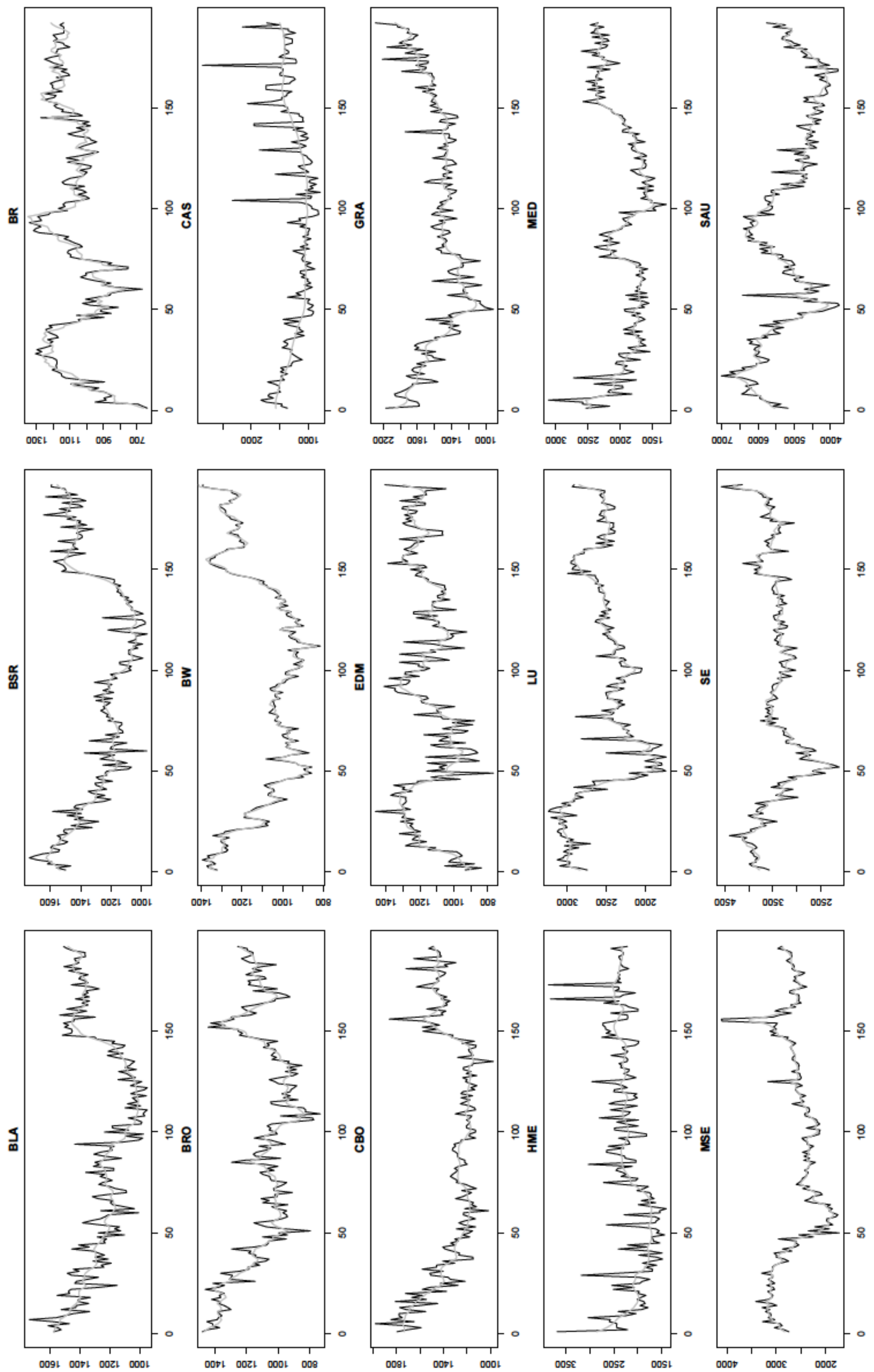


TABLE D.4 – Estimated variances of the parameters of the UCM filter for stocks

AOC	$\hat{\sigma}^2_{\epsilon^s_i}$	$\hat{\sigma}^2_{\nu^s_i}$	$\hat{\sigma}^2_{\xi^s_i}$	$\hat{\sigma}^2_{\omega^s_i}$
BLA	4.4e-09	5.4e-04	1.1e-08	1.2e-05
BSR	3.0e-08	4.5e-04	3.4e-13	2.3e-09
BR	1.2e-09	6.3e-03	2.9e-11	8.2e-09
BRO	1.1e-10	7.1e-03	9.1e-10	2.7e-09
BW	2.5e-09	1.0e-03	8.2e-09	5.1e-05
CAS	6.5e-10	7.1e-04	2.6e-09	1.9e-05
CBO	3.5e-05	5.0e-04	8.9e-10	7.0e-10
EDM	1.2e-10	3.4e-03	2.9e-10	4.9e-05
GRA	2.2e-14	4.1e-04	7.6e-08	1.1e-05
HME	1.5e-08	1.4e-04	5.5e-11	5.2e-07
LU	1.6e-11	6.5e-04	9.0e-12	1.2e-09
MED	7.1e-08	1.4e-04	2.6e-06	4.9e-06
MSE	1.5e-13	4.8e-04	4.7e-12	7.1e-10
SE	1.5e-09	3.8e-04	2.5e-10	7.0e-06
SAU	1.4e-10	5.0e-04	4.0e-10	3.5e-06

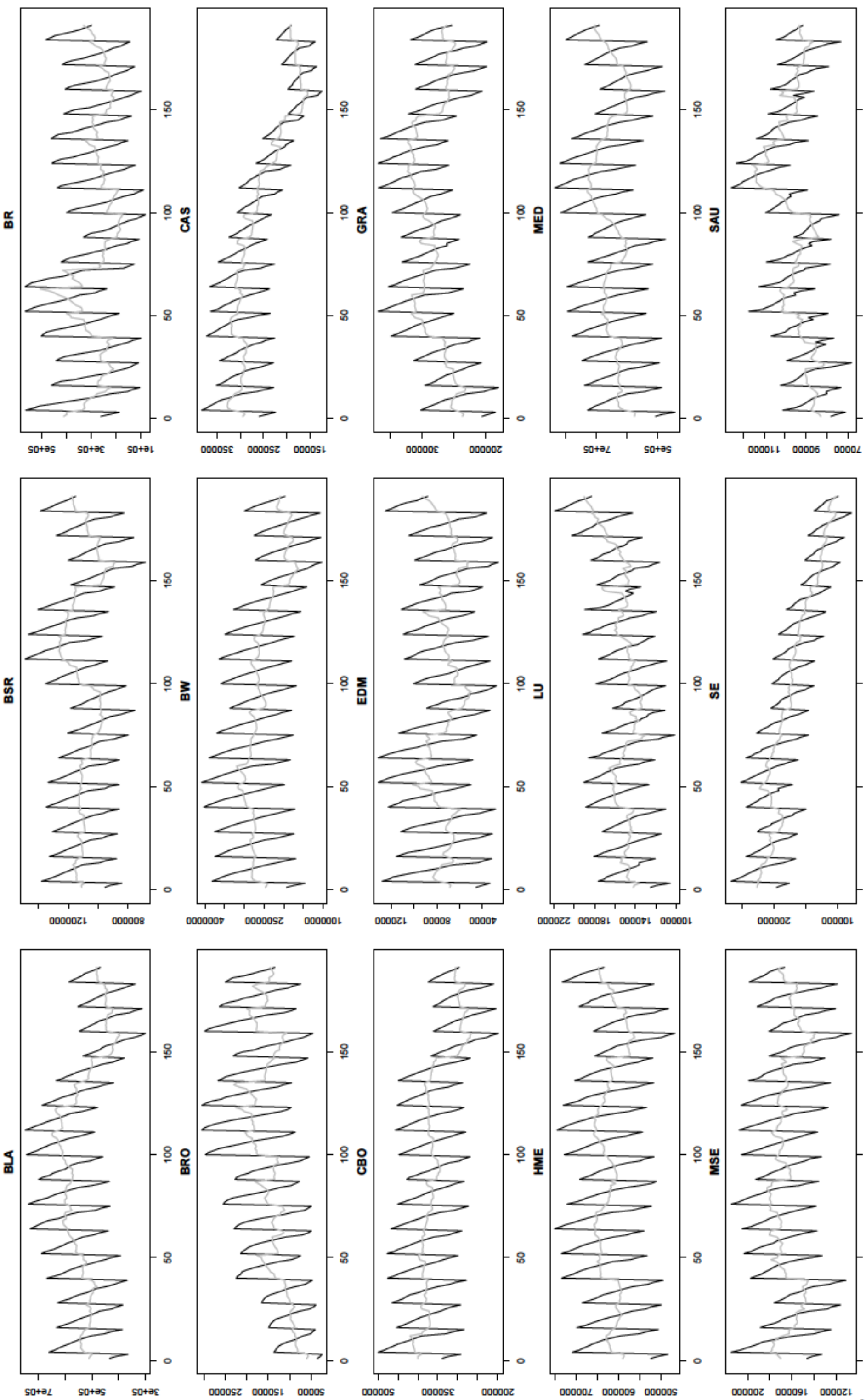


FIGURE D.4 – Observed monthly stocks (black) and the estimated underlying levels (grey) by AOC

TABLE D.5 – Estimated variances of the parameters of the UCM filter for deliveries

AOC	$\hat{\sigma}^2_{\epsilon^d_i}$	$\hat{\sigma}^2_{\nu^d_i}$	$\hat{\sigma}^2_{\xi^d_i}$	$\hat{\sigma}^2_{\omega^d_i}$
BLA	4.0e-02	2.1e-04	7.6e-16	1.5e-04
BSR	8.7e-03	4.8e-04	2.4e-10	1.3e-03
BR	3.9e-02	1.8e-03	1.4e-08	5.8e-08
BRO	5.0e-02	3.3e-07	5.9e-07	5.5e-03
BW	6.7e-03	7.0e-04	2.3e-10	7.3e-04
CAS	4.4e-02	1.3e-03	3.4e-10	9.2e-04
CBO	3.0e-02	1.4e-04	2.7e-10	7.3e-04
EDM	4.7e-02	4.3e-06	2.4e-11	6.9e-04
GRA	2.7e-02	3.8e-04	6.3e-11	3.8e-07
HME	2.9e-02	7.5e-04	1.2e-10	3.3e-05
LU	7.9e-02	3.9e-08	2.7e-10	1.6e-03
MED	2.0e-02	1.6e-03	5.8e-12	1.0e-07
MSE	6.7e-02	7.5e-04	4.8e-12	8.4e-07
SE	4.4e-02	9.1e-04	2.6e-10	2.4e-03
SAU	6.4e-02	1.2e-03	1.8e-10	3.7e-04

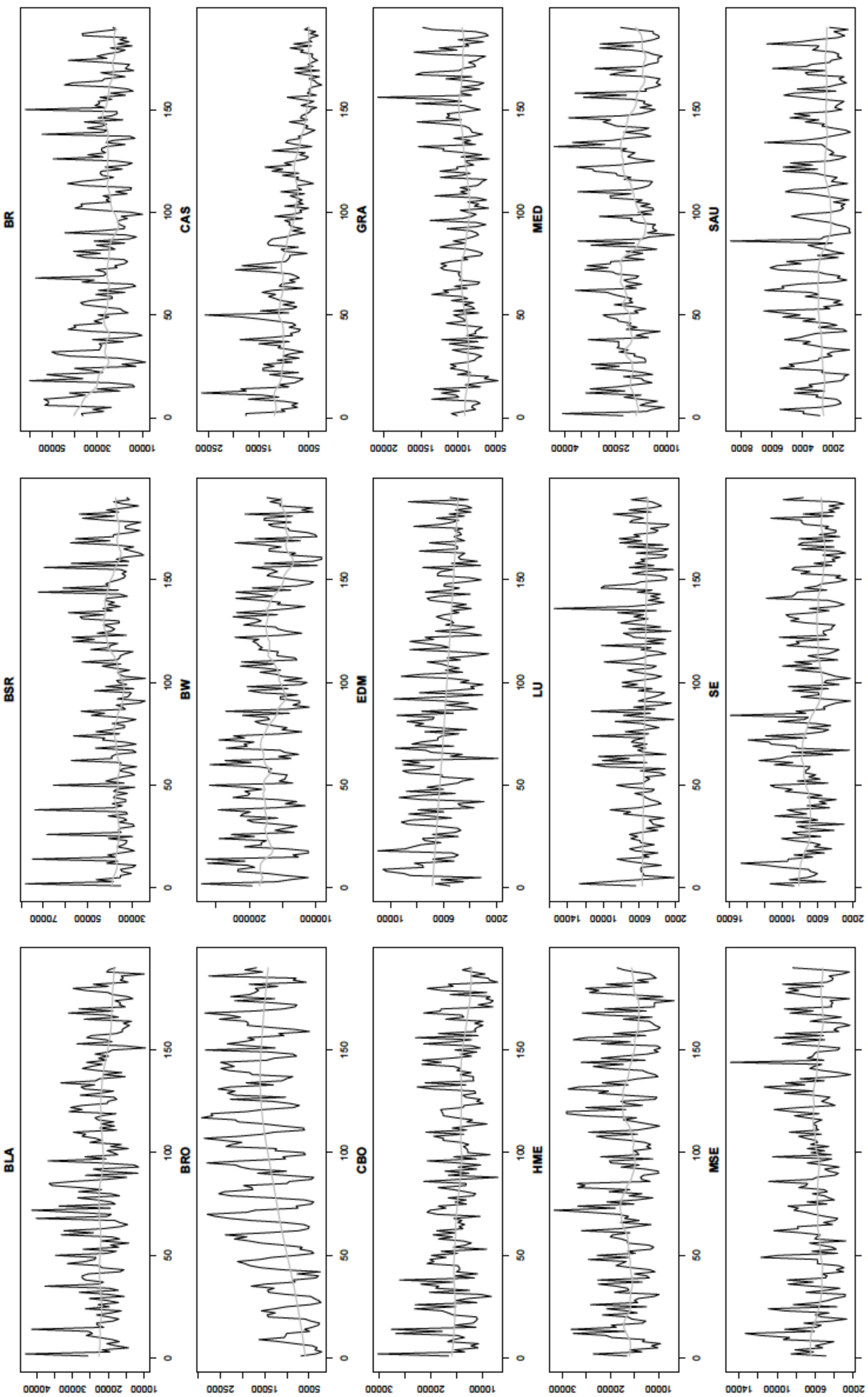


FIGURE D.5 – Observed monthly deliveries (black) and the estimated underlying levels (green) by AOC

TABLE D.6 – Annual long-run equilibrium OLS estimation by AOC using only 2001-2016 data

AOC	All	BLA	BSR	BR	BRO	BW	CAS	CBO	EDM	GRA	LSE	LU	MED	MSE	SE	SAU
H	-0.08		-0.45	-0.23	-0.35	-0.33			-0.18	-0.54	-0.38	-0.69	-0.83	-0.44		
	(0.04)		(0.16)	(0.15)	(0.05)	(0.13)			(0.12)	(0.15)	(0.2)	(0.27)	(0.18)	(0.28)		
SP	-0.25	-0.69	-0.61		-0.29	-0.29			-0.35	-1.2	-1.4	-1.08	-1.82	-1.33		-1.06
	(0.04)	(0.15)	(0.25)		(0.04)	(0.07)			(0.1)	(0.14)	(0.45)	(0.44)	(0.28)	(0.43)		(0.3)
E	1.17		2.69	2.22			3.38	1.72				2.8	0.66		0.77	
	(0.18)		(0.56)	(0.58)			(0.65)	(0.6)				(0.88)	(0.36)		(0.55)	
Y					1.06								0.7			
					(0.25)								(0.25)			
C	-0.18										-0.89		-0.3		-0.51	
	(0.1)										(0.29)		(0.18)		(0.31)	
Q										1.02		0.87	1.12	1.22		
										(0.38)		(0.69)	(0.41)	(0.82)		
r					-2.66					-2.88						
					(0.78)					(0.94)						
R^2	0.96	0.59	0.71	0.55	0.94	0.72	0.66	0.37	0.51	0.88	0.7	0.5	0.95	0.45	0.28	0.47

TABLE D.7 – MRAE of the annual forecasts where naive forecasts are July prices

AOC	2012	2013	2014	2015	2016	Average
BLA	1.23	1.01	0.303	1.048	0.452	0.708
BSR	0.698	0.735	0.205	3.123	0.736	0.753
BR	0.578	0.922	0.253	0.21	1.476	0.53
BRO	0.736	0.778	0.214	0.114	0.1	0.268
BW	0.28	1.236	3.85	1.418	1.066	1.15
CAS	1.363	0.451	4.878	10	0.614	1.791
CBO	1.327	1.202	0.204	2.4	0.202	0.691
EDM	0.299	1.136	0.412	0.668	0.861	0.604
GRA	0.924	0.105	1.383	1.018	0.519	0.589
HME	0.64	0.583	0.405	0.1	10	0.685
LU	0.284	1.219	0.86	1.196	0.1	0.513
MED	10	0.781	10	3.411	0.402	2.546
MSE	0.323	0.705	0.188	1.713	0.543	0.524
SE	0.473	0.567	0.263	0.553	0.537	0.461
SAU	3.214	4.034	0.135	0.246	10	1.34
Average	0.814	0.804	0.561	0.889	0.685	0.741

FIGURE D.6 – 2016 prices density forecasts (black), observed prices (light grey) and past prices (dark grey)

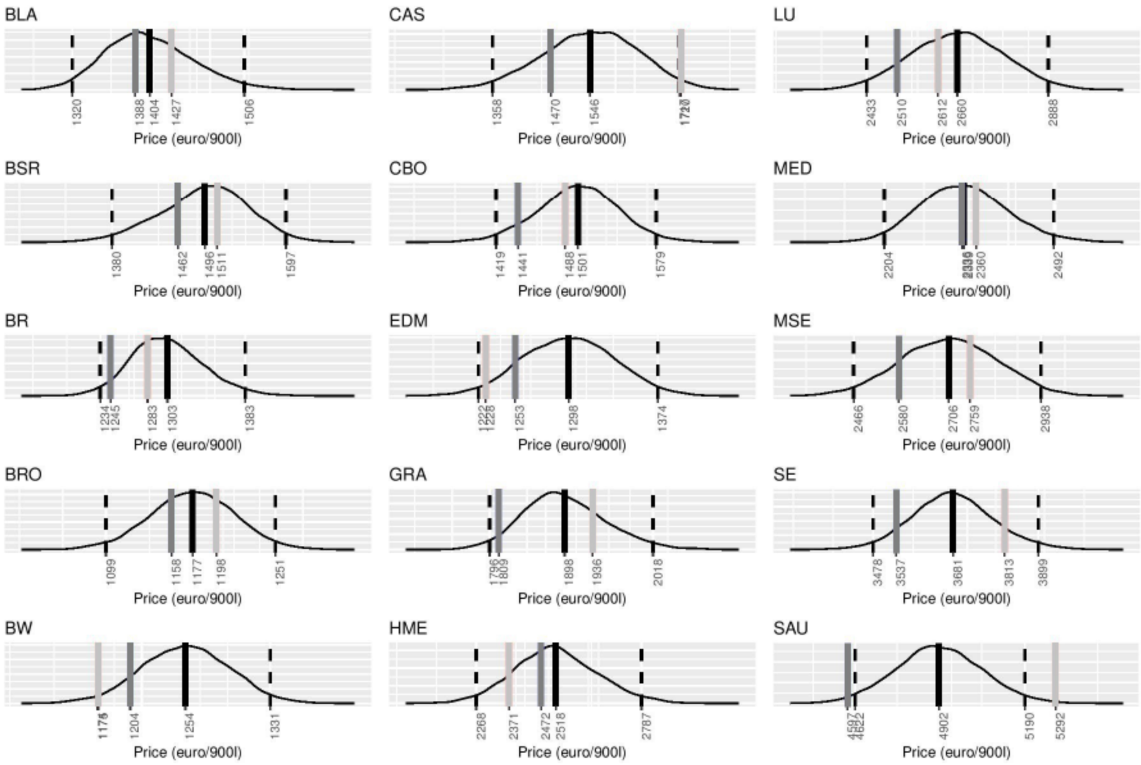


TABLE D.8 – Comparison of forecasts across different specifications for the monthly models

Criterion	BIC		AIC	
Prices	Real	Nominal	Real	Nominal
Overall MAPE	0.051	0.051	0.052	0.051
Overall MRAE	1.056	1.039	1.073	1.059

TABLE D.9 – MAPE of the monthly forecasts accuracy by AOC and year

MAPE	2012	2013	2014	2015	2016	Average
BLA	0.068	0.076	0.079	0.061	0.042	0.064
BSR	0.051	0.066	0.094	0.050	0.050	0.060
BW	0.030	0.054	0.050	0.023	0.031	0.036
BRO	0.038	0.076	0.111	0.061	0.038	0.059
BR	0.035	0.045	0.069	0.032	0.024	0.038
CAS	0.065	0.082	0.066	0.139	0.108	0.088
CBO	0.030	0.077	0.098	0.065	0.045	0.058
EDM	0.063	0.059	0.048	0.051	0.048	0.054
GRA	0.042	0.045	0.033	0.054	0.081	0.049
HME	0.051	0.073	0.103	0.086	0.038	0.066
LU	0.065	0.023	0.044	0.085	0.018	0.040
MED	0.043	0.034	0.038	0.048	0.059	0.044
MSE	0.052	0.032	0.071	0.084	0.036	0.051
SE	0.027	0.022	0.044	0.039	0.020	0.029
SAU	0.035	0.053	0.043	0.038	0.106	0.050
Average	0.044	0.050	0.061	0.056	0.043	0.051

TABLE D.10 – MAPE of the monthly univariate forecasts by AOC and horizon

MAPE	1	2	3	4	5	6	7	8	9	10	11	12	Average
BLA	0.032	0.037	0.041	0.046	0.053	0.070	0.081	0.089	0.100	0.099	0.103	0.099	0.065
BSR	0.038	0.045	0.051	0.063	0.072	0.073	0.075	0.090	0.090	0.097	0.111	0.108	0.072
BW	0.025	0.029	0.036	0.037	0.038	0.042	0.040	0.041	0.050	0.053	0.059	0.043	0.040
BRO	0.038	0.045	0.058	0.070	0.068	0.077	0.073	0.082	0.095	0.089	0.091	0.108	0.072
BR	0.021	0.024	0.036	0.045	0.050	0.057	0.061	0.063	0.066	0.062	0.061	0.064	0.048
CAS	0.071	0.080	0.073	0.079	0.075	0.074	0.087	0.102	0.113	0.113	0.113	0.112	0.089
CBO	0.036	0.042	0.048	0.059	0.071	0.071	0.067	0.060	0.065	0.064	0.065	0.071	0.059
EDM	0.036	0.046	0.055	0.057	0.054	0.056	0.055	0.073	0.071	0.075	0.068	0.056	0.057
GRA	0.048	0.051	0.045	0.054	0.055	0.054	0.057	0.056	0.040	0.032	0.042	0.037	0.047
HME	0.072	0.070	0.068	0.076	0.078	0.071	0.067	0.073	0.088	0.082	0.076	0.072	0.074
LU	0.023	0.022	0.022	0.028	0.027	0.031	0.028	0.033	0.039	0.031	0.042	0.042	0.030
MED	0.036	0.040	0.046	0.054	0.048	0.050	0.068	0.066	0.072	0.080	0.071	0.086	0.058
MSE	0.028	0.042	0.043	0.039	0.047	0.043	0.055	0.065	0.068	0.068	0.065	0.071	0.051
SE	0.019	0.014	0.022	0.026	0.031	0.024	0.028	0.034	0.033	0.044	0.044	0.047	0.029
SAU	0.036	0.040	0.038	0.037	0.043	0.045	0.049	0.050	0.052	0.045	0.058	0.057	0.045
Average	0.035	0.038	0.043	0.049	0.052	0.053	0.057	0.062	0.065	0.064	0.068	0.067	0.053

TABLE D.11 – MAPE of the monthly univariate forecasts by AOC and year

MAPE	2012	2013	2014	2015	2016	Average
BLA	0.072	0.081	0.074	0.074	0.036	0.065
BSR	0.059	0.103	0.111	0.075	0.039	0.072
BW	0.048	0.043	0.064	0.029	0.027	0.040
BRO	0.043	0.108	0.105	0.091	0.043	0.072
BR	0.052	0.077	0.061	0.039	0.025	0.048
CAS	0.065	0.076	0.062	0.150	0.124	0.089
CBO	0.027	0.076	0.114	0.063	0.048	0.059
EDM	0.069	0.080	0.044	0.048	0.053	0.057
GRA	0.032	0.057	0.036	0.060	0.056	0.047
HME	0.050	0.066	0.105	0.096	0.069	0.074
LU	0.022	0.031	0.041	0.051	0.017	0.030
MED	0.058	0.072	0.071	0.050	0.042	0.058
MSE	0.039	0.036	0.096	0.077	0.033	0.051
SE	0.018	0.028	0.050	0.029	0.027	0.029
SAU	0.035	0.030	0.042	0.045	0.095	0.045
Average	0.043	0.059	0.067	0.059	0.043	0.053

TABLE D.12 – MRAE of the monthly forecasts by AOC and Year, estimation on real prices

MRAE	2012	2013	2014	2015	2016	Average
BLA	0.994	0.707	1.604	1.891	1.219	1.211
BSR	0.881	1.113	1.696	1.692	1.043	1.240
BW	0.959	0.704	1.482	0.886	1.204	1.013
BRO	0.648	1.038	1.248	0.865	1.008	0.939
BR	1.142	0.821	1.412	1.017	0.953	1.051
CAS	0.680	0.606	0.625	1.547	1.551	0.908
CBO	0.857	1.058	1.191	1.457	1.142	1.124
EDM	1.088	0.753	1.212	0.984	1.039	1.003
GRA	1.118	0.814	1.074	1.402	1.061	1.078
HME	1.022	1.122	1.088	0.930	1.153	1.060
LU	2.286	0.890	0.935	1.770	1.021	1.280
MED	0.987	0.644	0.660	1.233	1.194	0.908
MSE	1.926	0.995	0.868	1.857	0.862	1.216
SE	0.978	0.607	0.778	1.066	0.741	0.817
SAU	1.014	1.519	0.884	1.049	1.184	1.111
Average	1.045	0.863	1.070	1.262	1.077	1.056

TABLE D.13 – MRAE of the monthly univariate forecasts by AOC and year

MRAE	2012	2013	2014	2015	2016	Average
BLA	1.271	0.934	1.521	2.237	1.104	1.349
BSR	0.941	1.420	1.776	2.294	0.931	1.384
BW	1.114	0.744	1.566	1.107	1.202	1.115
BRO	0.763	1.314	1.152	1.164	0.893	1.037
BR	1.290	1.017	1.128	1.445	1.061	1.178
CAS	0.709	0.626	0.583	1.580	1.677	0.927
CBO	0.768	1.019	1.301	1.443	1.358	1.148
EDM	1.299	1.403	0.998	1.071	1.223	1.189
GRA	0.922	1.024	0.970	1.402	0.856	1.019
HME	0.936	1.025	1.093	1.066	2.092	1.185
LU	0.983	0.989	0.927	1.246	0.904	1.003
MED	1.241	1.055	1.118	1.324	0.846	1.104
MSE	1.371	1.007	1.362	1.763	0.863	1.234
SE	0.749	0.697	0.974	0.805	0.759	0.792
SAU	0.853	0.860	0.855	1.169	1.059	0.951
Average	0.990	0.984	1.116	1.355	1.078	1.097

TABLE D.14 – MRAE of the monthly univariate forecasts by AOC and horizon

MRAE	1	2	3	4	5	6	7	8	9	10	11	12	Average
BLA	0.885	0.951	1.072	1.407	1.110	1.462	1.692	1.583	1.615	1.647	1.634	1.508	1.349
BSR	1.265	1.287	1.313	1.380	1.597	1.374	1.352	1.391	1.372	1.322	1.458	1.529	1.384
BW	0.884	1.031	0.982	1.068	1.001	1.088	1.195	1.165	1.190	1.239	1.434	1.209	1.115
BRO	1.053	1.076	0.967	1.032	0.990	0.954	0.942	0.973	1.036	1.006	1.079	1.410	1.037
BR	1.364	1.126	1.239	1.124	1.158	1.253	1.369	1.260	1.213	1.125	0.990	0.988	1.178
CAS	1.556	0.837	0.667	0.698	0.674	0.832	0.891	0.979	1.079	1.125	1.100	1.022	0.927
CBO	0.996	1.134	1.020	1.320	1.515	1.249	1.222	0.933	1.247	0.959	1.037	1.286	1.148
EDM	1.453	1.221	1.415	1.241	1.042	1.021	1.035	1.268	1.223	1.215	1.102	1.125	1.189
GRA	1.349	1.151	1.071	1.104	1.115	1.381	1.206	0.982	0.864	0.707	0.855	0.717	1.019
HME	1.571	1.220	1.000	1.183	1.249	1.156	1.081	1.058	1.379	1.229	1.087	1.112	1.185
LU	1.675	1.188	0.903	1.128	0.887	0.941	0.829	0.962	1.086	0.832	0.907	0.938	1.003
MED	1.114	1.034	1.063	1.252	0.948	1.038	1.159	0.907	1.232	1.282	1.154	1.136	1.104
MSE	1.256	1.298	1.392	1.213	1.334	1.105	1.369	1.218	1.301	1.110	1.120	1.139	1.234
SE	0.958	0.596	0.776	0.840	0.881	0.723	0.750	0.686	0.778	0.796	0.881	0.909	0.792
SAU	1.011	1.009	0.903	0.855	1.004	0.851	0.947	0.988	1.096	0.756	1.044	0.997	0.951
Average	1.200	1.060	1.032	1.104	1.076	1.075	1.111	1.069	1.163	1.062	1.107	1.114	1.097

Synthèse en Français

Symbole de vie, de renouveau et de prospérité dans les civilisations Méditerranéennes antiques, le vin est un attribut essentiel de l'identité culturelle de nombreux peuples de part le monde. Dans certains pays européens comme la France, l'Italie ou le Portugal, le vin est un produit de consommation courante[1] et la vigne une partie intégrante du paysage rural. En France, le secteur vitivinicole contribue à hauteur de 1.2% au PIB et emploie plus de 500 000 personnes. Au total, plus de 100 000 vins différents sont produits en France et leurs prix de détail varient d'environ cinq euros à plusieurs milliers d'euros pour une bouteille standard de 75 centilitres[2]. Cette diversité des prix sur le marché du vin fascine la communauté des économistes en ce qu'elle permet de révéler sur le rôle des signaux de qualité dans la formation des prix. En particulier, la question de la pertinence et l'influence des critiques est cruciale dans la mesure où ces critiques sont censés permettre une meilleure adéquation entre les prix et la qualité. Du propre aveu des critiques professionnels[3], il existe une part de hasard et de subjectivité dans l'évaluation de la qualité des vins par la dégustation. Cette situation où des signaux de qualité imparfaits influencent la détermination des prix a été jusqu'ici l'un des principaux sujets d'intérêt des économistes dans le marché du vin.

Si cet enjeu est essentiel au niveau microéconomique pour le segment des vins "super-premiums" (au-delà de 15 euros la bouteille), les mécanismes macroéconomiques d'équilibre offre-demande sont prépondérants sur le marché des vins de consommation de masse. En particulier, les rôles du stockage et des anticipations des producteurs dans la formation des prix sont des enjeux importants dans la littérature en économie agricole, mais ont été jusqu'à présent largement négligés dans la littérature spécifique au vin. Cette thèse réconcilie ces deux littératures en analysant à la fois les influences des signaux de qualité au niveau des produits et des déterminants macroéconomiques dans la formation des prix des vins. Le premier chapitre propose une nouvelle méthode pour dissocier l'influence des critiques et des déterminants observables de la qualité (météo, producteur) sur les prix de détail des vins haut de gamme. Le second chapitre introduit une méthode de d'échelonnage des notes données par plusieurs critiques pour faciliter la comparaison et l'agrégation de ces notes. Ces deux premiers chapitres concernent essentiellement le marché des grands crus et en cela s'inscrivent dans la tradition de l'économie du vin. En revanche, les deux seconds chapitres de cette thèse donnent à voir des aspects inexplorés du marché des vins destinés à la consommation de masse (en dessous de 15 euros la bouteille). Le troisième chapitre estime l'influence des concours vinicoles sur les prix payés aux producteurs, ainsi que leurs incitations à participer à ces concours. L'influence des déterminants macroéconomiques est traitée dans le quatrième et dernier chapitre de cette

[1] La consommation individuelle de vin des Français est d'environ 40 litres par an, ce qui équivaut à un verre par jour.

[2] Le prix de détail indiqué sur le site www.winedecider.com pour un Château Pétrus du millésime 2000 est d'environ 5 000 euros pour une bouteille standard de 75 centilitres, d'après une consultation en octobre 2017.

[3] Le fameux critique américain Robert Parker confesse dans un entretien donné au Naples Daily News en 2007: "Je pense vraiment que la différence entre les notes 96, 97, 98, 99 et 100 sur 100 se joue probablement dans l'émotion de l'instant" (Mobley-Martinez 2007, traduction libre). Jancis Robinson, critique anglaise renommée, écrit sur son site: "Je sais bien qu'il serait plus pratique pour tout le monde que l'on puisse évaluer tous les vins sur une échelle unique, mais je ne crois pas que cela soit réaliste étant donnée la myriade des styles et des types de vins qui, fort heureusement, existe encore." (traduction libre).

thèse, où je développe, estime et évalue un modèle de prévision des cours du vin. Chaque innovation méthodologique présentée dans cette thèse est illustrée par une application sur données originales de la région bordelaise.

1 Avis d'experts et prix des vins

Les vins les plus cotés sont généralement évalués par plusieurs notes de dégustation, celles-ci étant constituées d'un commentaire et d'une note sur 20 ou sur 100 selon les critiques. Ces scores de qualité présentent plusieurs défauts intrinsèques, souvent rappelées par les critiques elles-mêmes (cf. note de bas de page 3). Ils sont néanmoins scrutés et amplement commentés par les acteurs du marché. Dans la mesure où elles rendent compte d'un avis au moins partiellement subjectif, leur influence sur les prix est un sujet d'inquiétude récurrent. Cette controverse habite la communauté des économistes s'intéressant au vin depuis les années 1990 (Ashenfelter et al., 1995; Hadj Ali and Nauges, 2007; Hadj Ali et al., 2008; Ashenfelter and Jones, 2013). Ces articles concluent généralement à un effet mineur des experts sur les prix par rapport aux attributs "objectifs" tels que le nom du producteur (effet de réputation) ou le millésime. Dans ce chapitre, issu d'une collaboration avec Jean-Marie Cardebat et Jean-Marc Figuet[4], une nouvelle méthode est proposée pour évaluer les influences respectives des attributs subjectifs (notes) et objectifs (producteur, millésime) sur les prix de détail.

Nous avons extrait les prix renseignés sur le site www.winedecider.com durant la dernière semaine d'Avril 2011 pour 137 producteurs du Bordelais et pour les millésimes 2000 à 2010. Sur ce même site, nous avons également collecté les notes de quatre experts, chacun d'eux ayant noté l'intégralité des vins de la base. Afin de dissocier les influences spécifiques des notes de l'influence de la qualité du millésime, nous utilisons des données météorologiques de températures et de précipitations pour chacune des années 2000 à 2010. Dans une première étape, nous estimons par la méthode des Moindres Carrés Ordinaires (MCO) une équation où les notes sont expliquées par les variables météorologiques et des effets fixes relatifs à chaque producteur. Les valeurs prévues par cette équation sont donc uniquement déterminées par le savoir-faire spécifique à chaque producteur ainsi que par les conditions climatiques. En revanche, les résidus sont orthogonaux à ces informations, et sont à ce titre interprétés comme les opinions subjectives des experts. Les deux composantes des scores sont ensuite utilisées pour expliquer les prix dans une seconde équation, également estimée par les MCO.

Dans l'estimation de la première étape, les opinions ne représentent que 2% à 7% de la composante objective, ce qui indique que l'hétérogénéité des scores est relativement bien expliquée par les effets fixes producteurs et les variables météorologiques. L'estimation de l'équation de prix révèle que ceux-ci sont plus fortement impactés par une hausse d'un point de la composante objective (+13.7%) que par une hausse d'un point d'une opinion subjective (de +0.4% à +4.5% selon l'expert, et un total de +10.9% pour une hausse d'un point de toutes les opinions). En cela, nous consolidons les résultats précédents de la littérature. Dans une extension incluant l'écart-type des notes dans l'équation de prix, nous mettons en évidence la relation positive entre le prix d'un vin et la dispersion de ses notes. Ce résultat est contre-intuitif dans la mesure où les consommateurs sont généralement considérés averses au risque et donc devraient préférer les vins faisant consensus auprès des critiques. Nous proposons une explication à ce phénomène en montrant que la note maximale obtenue par chaque vin est également la note la plus corrélée au prix. Cette observation suggère l'existence d'un biais de médiatisation accrue pour la note maximale. De façon générale, ce résultat illustre la différence souvent

[4]Cette collaboration a abouti à la publication d'un article au *Journal of Wine Economics* en 2014 (Cardebat et al., 2014). La version de ce chapitre intègre des éléments répondant aux commentaires reçus après la publication.

négligée entre information disponible (les notes sont disponibles sur internet) et information effectivement utilisée par les consommateurs.

2 Uniformisation des scores de qualité

Le second chapitre est issu d'une collaboration avec Jean-Marie Cardebat[5], et répond à la problématique de l'agrégation de l'information disponible pour le consommateur. Nous utilisons une méthode dérivée de travaux en psychométrie et en économie de l'éducation pour évaluer les résultats de tests (Braun and Holland, 1982; Kolen and Brennan, 2014) afin de comparer et d'agréger les scores de qualité donnés par plusieurs sources. Dans cette étude, nous avons utilisé les notes de quinze critiques donnés à 4 333 vins de Bordeaux entre 2000 à 2014.

Un problème majeur dans la comparaison entre les notes est que les critiques européens notent sur 20, et les critiques américains sur 100. La méthode présentée dans ce chapitre, appelée *equipercentile equating*, consiste à égaler les quantiles des distributions des notes données par chaque expert. Par exemple, 15/20 est le quantile d'ordre 0.092 au sein de la distribution des scores donnés par Jancis Robinson (ce qui signifie que 9.2% de ses notes sont inférieures à 15/20). Dans la distribution des scores donnés par Robert Parker, ce quantile est 86/100. Notre méthode consiste donc à estimer qu'un 15/20 de Jancis Robinson équivaut à un 86/100 de Robert Parker. En convertissant ainsi toutes les notes des critiques dans l'échelle d'un expert de référence, par exemple Robert Parker, nous obtenons une base de donnée où toutes les notes sont directement comparables entre elles. Le critère d'optimalité retenu est celui de la rareté: deux notes données par deux sources différentes sont équivalentes si et seulement si elles sont également rares.

Notre principale contribution est méthodologique, cette méthode étant jusque là absente de la littérature en économie du vin. La conversion des notes que nous proposons permet de faciliter la comparaison et l'agrégation des notes pour différents experts, vins, ou millésimes. Dans notre application aux vins de Bordeaux, nous retrouvons le consensus selon lequel Robert Parker est plus généreux que ses pairs avec les vins de Bordeaux. Nous estimons en revanche qu'il a été le critique le plus sévère avec les vins du millésime 2013. Notre méthode permet enfin de calculer des écart-types pertinents entre les notes des critiques, et donc de mieux mesurer le degré de consensus pour chaque vin.

3 L'impact des médailles sur les prix payés aux producteurs

Si la plupart des études sur les liens entre qualité et prix des vins se concentre sur le très haut de gamme et les notes des critiques, encore peu d'études existent sur les certifications de qualité des vins plus courants (en-dessous de 15 euros la bouteille) qui constituent pourtant plus de 95% de la consommation mondiale. Sur ce marché, le moyen le plus répandu de se différencier de ses concurrents est de gagner une médaille à l'un des nombreux concours vinicoles existant. En France, plus de 130 de ces concours sont organisés chaque année. Ils permettent d'identifier qualitativement une partie significative des producteurs. Par exemple, environ 20% des vins

[5]Nous avons publié un article au *Journal of Wine Economics* en 2015 (Cardebat and Paroissien, 2015). Comme le premier chapitre, la version présentée dans cette thèse précise certains éléments suite aux commentaires reçus après la publication.

bordelais obtiennent une ou plusieurs médailles dans ces concours[6]. Ces concours communiquent sur leur impact positif sur les ventes afin d'attirer un maximum de participants[7], mais aucun article scientifique n'a encore estimé leur effet. En France, ils récompensent entre 25 et 30% des participants, le plafond réglementaire étant de 33%[8]. A l'étranger, la réglementation des concours est plus souple, voire inexistante. Le concours organisé par le journal *Decanter* récompense par exemple jusqu'à 60% des participants. Pour les producteurs, participer aux concours apparaît donc très attractif, quand bien même ils doivent s'acquitter d'un coût fixe de participation. Comme dans le cas des critiques sur le marché des grands crus, l'influence de ces concours sur les prix peut se révéler problématique s'ils sont peu fiables pour signaler la qualité des vins. Dans la littérature en économie, deux études américaines (Hodgson, 2008, 2009) ont justement pointé du doigt le manque de fiabilité de certains concours Californien. D'autres articles questionnent le fonctionnement des concours et la méthode d'agrégation des avis des juges pour désigner les gagnants (Ashenfelter and Quandt, 1999; Ashenfelter, 2006; Ginsburgh and Zang, 2012; Balinski and Laraki, 2013).

Ce troisième chapitre est issu d'un document de travail co-rédigé avec Michael Visser. Nous conduisons la première étude d'impact des concours européens sur les prix payés aux producteurs. En outre, nous fournissons des indicateurs indirects de fiabilité pour ces concours. Pour ce faire, nous utilisons les transactions d'un des plus importants bureaux de courtage en vin de la région bordelaise sur la période courant de 2006 à 2016, pour un total de 16 399 transactions après nettoyage de la base. Les vins concernés par ces transactions sont des vins de consommation courante, 99% des prix observés étant inférieurs à 8.6€/75cL. Nous avons également collecté tous les palmarès des onze concours principaux pour les vins de Bordeaux, de manière à identifier précisément tous les vins médaillés selon le type de médaille (bronze, argent ou or), le concours, la date précise de la récompense (au jour près) et celle de la transaction. Cette opération a révélé qu'une partie substantielle des vins récompensés (18%) ont été vendus avant d'obtenir leurs médailles[9]. Sous certaines conditions explicitées dans le chapitre, cette information nous permet d'identifier l'effet causal de l'obtention d'une médaille sur le prix en corrigeant l'effet de la qualité inobservée[10].

Nos estimations révèlent un effet causal important des médailles, principalement généré par les médailles d'or dont l'effet (+13%) dépasse largement celui des médailles d'argent (4.4%) et de bronze (4.2%), ces deux derniers étant statistiquement indifférenciables. De plus, notre méthode révèle une relation positive et statistiquement significative entre le prix de transaction et les médailles obtenues seulement après la transaction. Nous en déduisons que dans une certaine mesure, les concours sont capables d'identifier les vins les meilleurs - et donc les plus chers même en l'absence de médaille. Un effet spécifique à chaque concours est estimé mais l'impact causal n'est statistiquement non nul que pour six concours. Le lien suggéré par nos estimations entre médaille et qualité n'est statistiquement significatif que pour quatre concours, tous présentant une ancienneté d'au moins 40 ans. En outre, ces quatre concours sont

[6]Libre ensuite aux gagnants d'arborer ou non ces médailles sur leurs bouteilles, sachant que les macarons sont généralement payants et coûtent autour de 3 centimes d'euro l'unité.

[7]Françoise Harrewyn, responsable du concours de Bordeaux, a notamment déclaré: "Dans le cadre d'une médaille d'or, la vente du vin médaillé se conclut plus vite et il peut se commercialiser jusqu'à 30% plus cher.", propos rapportés sur la page https://www.lenouveleconomiste.fr/lesdossiers/les-concours-14338/ (consultée le 24 octobre 2017).

[8]Cette réglementation n'a été mise en place qu'en 2013, mais cette contrainte était déjà respectée par les principaux concours.

[9]Ceci peut être dû au simple fait que le producteur ait vendu une partie ou la totalité de sa production sans attendre les résultats, ou encore au fait que le négociant acheteur du lot ait participé à un concours après la transaction. Certains concours interdisent cette dernière pratique.

[10]Un vin médaillé est a priori de meilleure qualité donc aurait été un peu plus cher quand bien même il n'aurait pas reçu de médaille. Il ne suffit donc pas de comparer les prix des vins médaillés et des vins non-médaillés pour mesurer l'effet causal.

également ceux pour lesquels les juges ont relativement peu de vins à évaluer. Bien que ces concours comportent une certaine part d'aléatoire, nous reconnaissons à ces concours la capacité de différencier une grande partie de la production en attirant une majorité de producteurs. Cependant, les méthodes de jugement sont assujetties à caution. Une réglementation standard pour les concours a été proposée par l'Organisation Internationale de la vigne et du vin, mais reste jusqu'ici peu suivie, que ce soit en France ou à l'étranger. Le fort impact sur les prix révélé par notre étude plaide pour un renforcement du contrôle des processus d'évaluation au sein des concours, afin de limiter la circulation de médailles peu représentatives de la qualité des vins.

4 Prévision des cours des vins

La relation entre prix et signaux de qualité étudiée dans les trois premiers chapitres est une problématique centrale dans la littérature en économie sur le marché du vin. Cette littérature s'est jusqu'ici plutôt concentrée sur l'étude des choix des consommateurs, peu de travaux concernent les choix des producteurs[11]. En économie agricole pourtant, le programme du producteur, les anticipations de prix et de récolte ainsi que la gestion des stocks sont des problématiques fondamentales. Depuis qu'Ezekiel (1938) a énoncé son *cobweb theorem* révélant les liens entre volatilité des prix et erreurs d'anticipation, les économistes de l'agriculture ont cherché améliorer les anticipations des producteurs. En particulier, une branche de la littérature s'est spécialisée dans la prévision des prix agricole[12], en appliquant les innovations techniques dans le traitement des séries temporelles, notamment la méthode Box-Jenkins (Box et al., 1970), les modèles VAR (Sims, 1980) et les modèles à correction d'erreur (Engle and Granger, 1987; Johansen, 1991). Plusieurs études ont finalement montré que lorsque des prix à terme étaient disponibles, ceux-ci constituaient des prévisions de prix satisfaisantes et aisément accessibles[13]. La prévision des prix agricole a donc été peu à peu désertée par les économistes, quand bien même tous les marchés agricoles ne sont pas équipés d'un marché à terme. Le marché du vin en est par exemple dépourvu, et les producteurs manquent de visibilité sur les prix. Il est en outre particulièrement difficile de concevoir un marché à terme dédié au vin du fait de la grande variété des produits[14].

Ce dernier chapitre renoue avec les problématiques historiques de l'économie agricole en développant plusieurs modèles de prévision des cours du vin. Ces modèles sont appliqués aux prix moyens des vins vendus en vrac par les producteurs bordelais pour chacune des quinze principales Appellations d'Origine Contrôlée (AOC). Contrairement à la littérature existante en économie du vin, ce dernier chapitre met l'accent sur le rôle des effets de volume dans la formation des prix[15]. Pour chaque AOC, je dispose du stock total en début d'année, de la récolte totale et des volumes retirés mensuellement des chais des producteurs. Toutes ces données de marché ont été fournies par le Conseil Interprofessionnel des Vins de Bordeaux, l'institution locale en charge de la collecte et de la diffusion des informations de marché. Le rôle des anticipations de récolte est également pris en compte dans les prévision de prix. J'utilise pour cela un historique de données quotidiennes de température et de précipitation pour six stations de la région de Bordeaux fourni par MétéoFrance. Enfin, les déterminants macroéconomiques

[11]On peut tout de même citer deux articles Alston et al. (2015a) et Anderson (2014), qui étudient les évolutions dans les choix des cépages dans les principaux pays producteurs de vin.

[12]Voir Allen 1994 pour une revue de cette littérature.

[13]Voir notamment Just and Rausser (1981), Tomek (1996), Kastens et al. (1998) et Ahumada and Cornejo (2016).

[14]La mise en place d'un marché à terme suppose d'identifier un produit standard pour lequel de nombreuses transactions sont observées. La diversité de l'offre de vin complique la définition d'un tel standard, qui risque d'être soit trop mal défini soit trop peu représentatif.

[15]Haeger and Storchmann (2006) est une exception notable.

sont intégrés aux différents modèles, dont les taux de change, les taux d'intérêt, la croissance économique des pays consommateurs, les productions de vin par pays et les flux détaillés de commerce international. Pour chacun de ces déterminants, j'ai collecté des historiques de données régulièrement mis à jour, de sorte que les modèles puisse permettre des prévisions de prix opérationnelles[16].

Dans un premier temps, chacun de ces jeux de données (météo, taux de change, etc.) a été agrégé en un indicateur global par AOC. L'information des données météorologiques a notamment été résumée au moyen d'un modèle de récolte ad hoc. Une étape essentielle du travail méthodologique a ensuite été de définir des critères adaptés pour la sélection des variables à inclure aux modèles dans le but d'optimiser la précision des prévisions. En effet, ajouter une variable n'est utile que dans la mesure où les données disponibles permettent une estimation suffisamment précise de son influence sur les prix[17]. Différents types de modèles sont utilisés dans les prévisions, notamment les modèles autorégressifs à retards distribués (ADL), les modèles à correction d'erreur (ECM) de Engle and Granger (1987), ainsi que les modèle à composantes inobservées (UCM) dus à Harvey (1989). Ces modèles sont tous relativement standards en économétrie des séries temporelles et permettent une certaine souplesse dans la sélection des variables. A partir de ces modèles classiques, je développe un modèle de prix spécifique à chaque AOC pour des prévisions en fréquence annuelle et mensuelle, et en adaptant dans chaque cas les choix de modélisation aux données disponibles. Les prévisions des modèles sont évaluées sur les cinq dernières années en les comparant à des prévisions naïves faisant l'hypothèse que les prix resteront inchangés. Malgré sa simplicité, cette hypothèse de référence pour la prévision des prix s'est révélée difficile à battre dans les travaux publiés[18].

Le principal résultat de ce chapitre est que les prévisions des modèles se révèlent plus précises en moyenne que les prévisions naïves, que ce soit à l'échelle annuelle ou mensuelle, et en particulier pour l'AOC régionale Bordeaux. De plus, les différents modèles sont plus efficaces en cas de choc de production important. Les prévisions mensuelles sont essentiellement efficaces entre Août et Décembre, période durant laquelle les données disponibles offrent une assez bonne visibilité sur le volume disponible pour les mois suivants. L'estimation des modèles de prévision permet en outre de commenter les influences respectives des déterminants. En particulier, mes estimations soulignent le rôle prépondérant du stockage dans la fluctuation des prix du vin, phénomène déjà mis en évidence dans le cas des céréales notamment[19]. En revanche, les variations de la qualité des millésimes ont peu d'influence sur les cours moyens, à l'exception des AOC les plus qualitatives comme Saint-Emilion. Au sein des déterminants macroéconomiques, les taux de change ont un impact important sur les prix car une part considérable de la production est exportée hors de la zone euro. Les conditions météorologiques ont également une influence significative sur les prix via les anticipations de récolte. Ce panorama nuance les enseignements de la littérature sur les prix du vin, jusqu'ici focalisée sur les grands crus. Pour la majorité de la production bordelaise, la volatilité des prix est moins guidée par la qualité des millésimes que par les variations des volumes disponibles et des conditions macroéconomiques. En outre, les estimations du modèle de rendement fournissent une évaluation des conditions optimales de température à Bordeaux. Il apparaît que les

[16]Les données de taux de change sont fournies par le Banque Centrale Européenne et collectées sur le site fx-top.com. Les données de taux d'intérêt sont collectées sur le site de la Banque de France. Les chiffres des PIB par pays sont extraits du site de l'Organisation de Coopération et de Développement Économiques (OCDE). Enfin, les productions de vin par pays ainsi que les flux de commerces sont donnés sur le site de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO).

[17]Ce genre de problème est récurrent en statistique, et porte le nom d'arbitrage biais-variance. Voir Shmueli (2010) pour une discussion de cet arbitrage dans le cas de la prévision.

[18]Voir Allen 1994 pour une revue de la littérature sur prévision des prix.

[19]Voir Wright (2011) et Gouel (2012) pour une revue de littérature sur le rôle du stockage sur la volatilité des prix agricoles.

conditions actuelles sont proches de l'optimum pour les rendements[20]. La tendance actuellement croissante des températures, si elle se poursuit, menace ainsi de faire diminuer les volumes produits à Bordeaux. Évidemment, les producteurs bordelais ont de nombreuses possibilités pour faire face au changement climatique, notamment en adaptant la conduite de la vigne, les procédés de vinification, les rendements maximums autorisés ou même les choix des cépages.

[20]Jones et al. (2005) arrivent à une conclusion similaire pour la qualité.

Nunc est bibendum

Résumé & Abstract

Résumé L'extraordinaire variété des prix et des produits sur le marché du vin suscite un intérêt grandissant au sein de la communauté des économistes. Justement, les acteurs de ce marché sont en demande de nouveaux outils économiques pour s'adapter aux évolutions récentes, dont la multiplication des experts influents et l'accélération de l'intégration des régions vinicoles à l'économie mondiale. Cette thèse se place au croisement des intérêts des économistes et des professionnels du vin en développant de nouvelles méthodes statistiques pour l'étude et la mesure de l'influence des déterminants des prix du vin. Chacun des quatre chapitres qui la compose présente une innovation méthodologique dont l'intérêt est illustré par une application empirique sur des données originales du marché des vins de Bordeaux. Le premier chapitre s'appuie sur des données météorologiques pour isoler la composante subjective des notes de dégustation et évaluer l'influence des critiques sur les prix de détail. Le deuxième chapitre introduit une méthode d'échelonnage pour comparer les scores de qualité estimés par plusieurs sources. Le troisième chapitre estime l'impact des médailles obtenues aux concours vinicoles sur les prix payés aux producteurs. Le quatrième et dernier chapitre compile une base de données exhaustive sur les déterminants macroéconomiques des fluctuations du marché Bordelais pour établir un modèle opérationnel de prévision des cours des vins par appellation d'origine contrôlée.
Mots-clés: vin, prix, qualité, experts, certification, prévision

Abstract The extraordinary variety of prices and products on the wine market has attracted an increasing interest from the economists community. On the other end, the agents of this market require new economic tools to adapt to the recent evolutions, such as the multiplication of influential experts and the accelerating integration of the wine regions in the global economy. This dissertation aligns the interests of economists and wine professionals by developing new methods for the study and the impact measurement of wine prices determinants. Each of the four chapters constituting this thesis introduces a specific methodological innovation and illustrates its benefits with an empirical application using novel data on the Bordeaux wine market. The first chapter builds on weather data to identify the subjective component of tasting grades and assess the influence of experts on retail prices. The second chapter proposes a scaling method to compare quality scores among different sources. The third chapter estimates the causal impact of the obtention of a medal at wine competitions on the prices paid to producers. The fourth and last chapter assembles a comprehensive database on macro-level determinants of the fluctuations of the Bordeaux wine market to build an operational forecasting model of the average prices by protected appellation of origin.
Keywords: wine, price, quality, experts, certification, forecasting