The Informational Content of Geographical Indications

March 23, 2020

Abstract

Geographical indications (GIs) convey information about the place of production as a proxy for the attributes of agricultural products. We define the informational content of the GI proxy as its capacity to describe the tangible characteristics of production sites, instead of random noise or intangible factors from political bargaining about designation. We estimate econometrically the informational content of wine-related GIs for the *Côte d'Or* region of Burgundy, France. We show that GIs signal vineyard attributes with high precision, while we find some persistent bias from lobbying effects. We also study alternative classifications, from history and from simulations, which reveal a significant increase in the informational content of GIs over the last hundred years or so, and provide guidelines for better designated GIs in the future.

Keywords: Food certification, semi-parametric model, strategic quality disclosure, variance decomposition, wine economics.

J.E.L. Codes: C24, L15, Q13.

Running head: Informational content of geographical indications

Author Information and acknowledgements: Jean-Sauveur Ay, *Chargé de Recherche*, Umr Cesaer, AgroSup, Inrae, Université de Bourgogne Franche-Comté, 21000 Dijon (France). Associated data and R code are available in the Replication Material from the repository http://github.com/jsay/geoInd/ under copyright licence GNU GPL V3. I am grateful to Mohamed Hilal and Florian Humbert for providing geographical and historical data, to Julie Le Gallo, Stéphan Marette, and Emmanuel Paroissien for their comments on previous versions of the article, and to the editor Timothy Richards. This research received financial support from the French national institute for agriculture, food and the environment (Inrae).

It is controversial in market relations to use the place of production to signal the quality of agricultural products (Josling, 2006; USTR, 2017). However, distinguishing high quality products from those of poor quality is recognized as fundamental for consumers and producers when quality cannot be assessed before deciding to buy or sell (Akerlof, 1970; Nelson, 1970). Thus, one major point in the debate is the extent to which geographical indications (GIs) provide accurate information about product quality (Caswell and Mojduszka, 1996; Messer et al., 2017). We study this informational content through the econometric relationship between the tangible characteristics of vineyards and wine-related GIs of the *Côte d'Or* region (Burgundy, France).

We estimate the informational content of current, past, and simulated GI designation schemes for an exhaustive data set of about 60 000 vineyard plots producing the most expensive wines in the world.[1] We disentangle the tangible information related to the natural attributes of vineyards that are known to impact wine quality (topography, geology, and climate) from intangible information related to a pre-existing administrative delineation (the *communes* that correspond to the French municipalities). The major empirical challenge at hand is to separate these two sources of variations in the GI signal as, even with the fine-scale data we use here, all the natural variables are not (and probably never will be) actually observable for econometric or statistical analysis.[2]

We use semiparametric ordered generalized additive models (OGAMs) that exploit the precise location of vineyard plots to control for unobserved spatial heterogeneity through fine-scale smooth functions of geographic coordinates (Wood et al., 2016). The corresponding identification strategy is based on the structural difference between the spatial continuity of natural characteristics of vineyards and the discontinuity of administrative boundaries. OGAMs are shown to significantly outperform more classical approaches based on parametric (polynomial and/ or interaction) functions of geographic coordinates both in terms of goodness-of-fit and of causal inference validity.

We find that the GIs under study are highly informative, with four times higher conditional variance than noise variance (corresponding to an R^2 of $\sim 80\%$). The simulation of certain increases in the number of GI levels reveals moderate potential improvements on this point. Nevertheless, we find persistent lobbying bias in the GI signal that decreases its informational content. Through the reputation of previous landowners, their influence with decision-makers or their collective actions, some administrative units enjoyed privileged treatment placing their vineyard plots higher in the hierarchy than similar plots in other administrative units. The simulations show that better balanced schemes would significantly increase the informational content of GIs.

We combine this evidences with historical data about the GIs of 1936 (the year INAO, the French national institute in charge of geographical indications, was founded) to show that the lobbying bias has declined since that time. This is consistent with the theory developed by Benabou and Laroque (1992) regarding strategic information transmission. When information is not fully reliable (because of signal bias), the possibility of honest mistakes (because of signal noise) is confusing for consumers. Market incentives leave the consumers' learning process incomplete and allow information to be manipulated, which only fades out in the long run. The long history of GIs in Burgundy and their independent management appear as two important determinants of their current informational content. This result could be relevant in explaining the differential performance of GIs on wine markets between the regions of the "old world" and those of the "new world" in quality-driven international competition (Duncan and Greenaway, 2008).

Contribution to the Literature

There is a vast literature on the effects of GIs and, more generally, of labels on welfare and many kinds of market outcomes (see Bonroy and Constantatos, 2014; Messer et al., 2017, for reviews). Their social desirability appears theoretically ambiguous. On the one hand, they increase the information available for consumers (Zago and Pick, 2004; Bonroy and Constantatos, 2008), provide incentives to produce high-quality goods (Moschini et al., 2008; Yu et al., 2017), and allow to share reputation investments (Marette et al., 1999; Menapace and Moschini, 2012). On the other hand, they introduce supply control (Lence et al., 2007; Mérel and Sexton, 2012), confusion for consumers (Lohr, 1999; Brécard, 2014), and stigmatization of regular products (Liaukonyte et al., 2015; Ay et al., 2017). Empirically, numerous hedonic studies find positive price premiums for positively labelled products (see Combris et al., 2000; Carew and Florkowski, 2010; Sáenz-Navajas et al., 2013, for the GIs under consideration here). The policy implications of these evidences are not clear cut (Unwin, 1999; Oczkowski and Doucouliagos, 2015). They may be attributable to spurious correlation between quality and reputation (Ali and Nauges, 2007), sorting on unobserved consumer characteristics (Gustafson et al., 2016), or simply measurement errors (Oczkowski, 2001). The lack of knowledge to design more efficient GIs is also regularly recognized in the economic literature (Bonroy and Constantatos, 2014; Deconinck and Swinnen, 2014).

We propose a framework for quantifying the capacity of GIs[3] to transform the natural attributes of vineyards into searchable attributes of wines for consumers. Among the numerous other determinants of utility derived from wine consumption (vintage, producer, variety, status, or context), we argue that disclosing a reliable signal about the *terroir*[4] is a robust objective for GI management in the long run.[5] These natural variables are widely considered to impact the taste of wines (Bokulich et al., 2014; Roullier-Gall et al., 2014; van Leeuwen et al., 2018) while the lack of knowledge or technical expertise by most consumers precludes the ascertainment of the taste of *terroir* (i.e., it is a credence attribute). In Burgundy, there is a long history of ranking vineyards according to their quality for wine production, dating back to the Middle Ages (Jullien, 1816; Meloni and Swinnen, 2018). Through GIs, this accumulated empirical knowledge allows consumers to assert the potential quality of a wine in relation to the unchangeable characteristics of vineyard quality.[6]

We draw insights from information theory (Laffont, 1985; Vives, 2010) by considering GIs as an information structure, i.e., a joint distribution between vineyard quality and the signal given to consumers on wine labels. Formally, we define the informational content of the GI signal by its precision and its bias in describing vineyard characteristics. The precision criterion stems from the principle that more informative signals lead to greater variability of conditional expectations (Blackwell, 1951; Ganuza and Penalva, 2010). Intuitively, more precise signals enable greater Bayesian updating and more dispersed posterior expectations once the GI is known (Bowsher and Swain, 2012). Signal bias is defined as the probability that two identical vineyard plots will have different GIs (De and Nabar, 1991; Guerra, 2001). This measures the systematic deviation of GI designations from tangible and verifiable attributes (Grossman, 1981; Li, 2017) that could strategically distort the GI signal (Benabou and Laroque, 1992).

The GIs under study provide an empirically tractable case study of the economics of strategic quality disclosure (Dranove and Jin, 2010). In contrast with typical frameworks in which quality is strategically determined by producers (Shapiro, 1982; Albano and Lizzeri, 2001; Jin and Leslie, 2003), vineyard quality depends on predetermined *terroir* that facilitates the identification of the informational content of GIs. This allows a more comprehensive analysis of the role of private information and lobbying in the quality signal conveyed by GI designations. In addition, vineyard quality relies exclusively on the unchangeable location of production sites, which precludes spurious correlations from the assortative matching between quality and name as in Tadelis (1999).

Context and Data

First, we present the vineyards from the *Côte d'Or* region under study. Next, we present the historical evolution of GIs in this region, followed by a precise description of current GIs. Finally, we provide some summary statistics on the data used.

The Côte d'Or Region

The *Côte d'Or* (literally, slope of gold) is a northeastern French administrative unit (*département*) included in the larger wine-producing region of Burgundy (Figure 1). We studied a subset of the most famous vineyards in this region (named *Climats* locally), which was granted World Heritage Status by UNESCO in 2015 (`https://whc.unesco.org/fr/list/1425`). The area under consideration is a strip of approximately 65 km from the north to south and at most 5 km from east to west located between latitudes 46.9 and 47.3 and longitudes 4.7 and 5 (World Geodetic System 1984). The main tangible attributes of vineyards in the area are illustrated by the distribution of elevation in the left panel of Figure 1. The presence of *combes* (dry valley) results in some rounded patterns with fine-scale variations in the typical topographical variables (elevation, slope and exposition) that are known to have direct and indirect impacts on wine quality.

Firstly, elevation is expected to determine wine quality principally through its correlation with temperatures and atmospheric outcomes. Temperatures during the growing season and harvest are major determinants of the grape maturity cycle, sugar content, and the structure of aromas (van Leeuwen et al., 2004). The latitudinal position of vineyards is also correlated with temperature along the north-south gradient. Secondly, slope is expected to have both a direct effect through the drainage capacity of vineyard plots and an indirect effect through the correlated soil characteristics (steeper soils are generally older and thinner). The longitudinal position of vineyards indirectly correlates with precipitation in the area, as an escarpment to the west provides a protective barrier that limits rainfall and, consequently, soil moisture. Thirdly, exposition is expected to have a direct effect through sunshine cycles and an indirect effect through its correlation with the wind, which is known to be important in drying grapes and concentrating aromas (van Leeuwen et al., 2004).

We do not use climate variables due to their typical coarse scale availability, which makes them unsuitable for the narrowness of our study area and the tiny size of vineyard plots. For example, historical climate data from *Météo-France* are usually available at 8 km resolution where the vineyard strip depicted in Figure 1 is at most 5 km wide. Moreover, topography (available at a 5 m resolution in our data) is regularly used to interpolate climate observations that artificially increase the resolution of climate variables. Using such interpolations to control for climate variables is not relevant to our econometric analysis as they are redundant with raw topographic variables.

Historical Context

Archaeological evidence locates the earliest vineyards in the region in antiquity (Garcia, 2014). The earliest written evidence dates from the 7th century, with abbey archives describing the donation of vineyards between groups of Benedictine monks whose names are still used in actual

GI classifications (e.g., *Abbayes de Bèze* or *de Saint-Vivant*). The origin of Burgundy's vineyard classification can be found in the work of the Cistercian monks who delineated plots of land that produced wine of distinct character (12th century according to Lavalle, 1855). However, the earliest exhaustive spatial delineation of the region was an administrative separation of *communes* following the decree of 1789 after the French Revolution. The delineation of *communes* was based on the spatial distribution of churches (usually built between the 9th and the 12th centuries), without the goal of signaling wine quality. This administrative subdivision represents the current horizontal dimension of GIs, as the *communes* are not explictely ordered in terms of vineyard quality.

The first exhaustive vertical classification scheme of vineyard quality was created by Lavalle (1855), a Professor of Natural and Medical History at Dijon University, inspired by the writings of other scientists, particularly Jullien (1816) and Morelot (1831). He provided a ranking of vineyards on four levels, from the best *Tête de Cuvée* to *Première*, *Deuxième* and *Troisième Cuvées*. The interaction between the horizontal and vertical dimensions is of particular importance in his work: "I have studied the wines of each of the *communes* of the *Côte* as if the other *communes* had not existed and the classification that I give is true only for each *commune* taken in isolation" (p.162, our translation).

These two spatial delineations were merged in an 1860 map by the *Comité d'Agriculture et de Viticulture de l'Arrondissement de Beaune*, the local organization of wine producers. This map contains small modifications from the initial 1789 and 1855 classifications (Wolikow and Jacquet, 2018) and was used extensively as a legal basis to regulate wine trade in the region. It paved the way for court trials, collective actions, and lobbying for the right to use the names of both dimensions that were not yet called GIs. The capacity of producers and owners to negotiate or influence judgments and delineations is determined by the reputation of the *commune* to which they belong (Jacquet, 2009). The author showed that there was unequal treatment between *communes* in terms of the vertical differentiation of vineyards, whereas the separation between advantaged and disadvantaged *communes* was not well established: "the reputation of the wine-growing *communes* of Burgundy is not an objectively measurable phenomenon" (Jacquet, 2009, p.189; our translation).

In 1936, a French national institute, INAO, was created to legally manage what became the GIs of all wine regions of the country in a common legal basis. In Burgundy, the first official GIs came from the map of 1860 and the jurisprudence occurring thereafter. Some modifications were then implemented during the 20th century with the creation of *Premiers Crus* in 1943 and the fine-scale digitization of plot-level delineation in a Geographical Information System after 2000. The GIs have been called *Appellation d'Origine Contrôlée* in France since 1936, corresponding to Protected Designation of Origin for the European Union (https://ec.europa.eu/agriculture/quality/schemes_en).

Current GI Designations

Thus, the GIs that we study are based on the fine-scale location of the vineyard plots, with both a vertical and horizontal dimension of differentiation. The vertical dimension is a quality ranking with five levels: *Côteaux Bourguignons* < *Bourgogne Régional* < *Village* < *Premier Cru* < *Grand Cru*. The horizontal dimension is 1 of the 31 *communes* (i.e., administrative municipalities) without an explicit hierarchy between them, such as *Beaune*, *Gevrey-Chambertin*, *Pommard*, or *Fixin*. Such a hierarchical and nested structure is common for wine-related GIs in France (Bordeaux, Rhône Valley, see Gergaud et al., 2017).

The highest quality vineyards are labeled *Grands Crus*, each of which has its own independent appellation name (e.g., "*Clos de la Roche*" or "*Chevalier-Montrachet*"). There are 32 *Grands Crus* in the area, eight in the *Côte de Beaune* (southern part) and 24 in the *Côte de Nuits* (northern part), with a total area of 472.6 ha (4.2% of acreage with GIs). In the hierarchy, it follows 404 *Premiers Crus* in the area that have to be associated with their *commune* names on wine labels (e.g., "*Les Chaumes*" from *Vosne-Romanée* or "*La Chapelle*" from *Auxey-Duresse*). There are 1619 ha of *Premiers Crus* in the *Côte de Beaune*, accounting for 20.5% of the sub-region and 433 ha in the *Côte de Nuits* (12.75%). The third vertical level corresponds to *Bourgogne Village* with or without a name (e.g., *Pommard Village* with name and *Côte de Nuits Village* without), accounting for 2500 ha

(31.75%) in the *Côte de Beaune* and 1563 ha (46%) in the *Côte de Nuits*. The vertical differentiation of GIs ends with *Bourgogne Régional* (2788 ha, 24.73% of the GI area) and *Coteaux Bourguignons* (1899 ha, 16.85%), which are sometimes grouped in the same *Régional* level.

The picture of current GIs in the area is not complete without mentioning the complexities between the vertical and horizontal dimensions. Note that the terms *commune* and *village* are often used synonymously for the administrative delineations in rural areas of France, whereas the first is related to the horizontal dimension and the second to the vertical dimension. In addition, the same verical level name from *Grand Cru*, *Premier Cru* or even *Villages* can be found in two different *communes*.[7] Furthermore, at the beginning of the 20th century, 10 *communes* added the name of their most famous *Grand Cru* to their administrative name, such as *Aloxe-Corton* or *Gevrey-Chambertin*. Consequently, the name of a *Grand Cru* is labeled in the horizontal information for wines that are not *Grand Cru*.[8] However, the legal obligation to mention the vertical level *Grand Cru*, *Premier Cru*, *Village*, *Régional* or *Coteaux Bourguignons* as the main information on wine labels suggests that this information is clearly apparent to consumers.

Summary Statistics

From this long-run history of GIs, we exploit the precise location of about 60 000 current vineyard plots to estimate the informational content of current GIs, controlling for the unobserved spatial heterogeneity from *terroir*. The precision of econometric estimations for disentangling the sources of variation in GIs depends on a balanced distribution of tangible variables and vertical levels between and within the horizontal *commune* items. The left-hand panel of Figure 1 shows that each *commune* contains approximately the whole range of elevation, slope, and exposition of the area, whereas the right-hand panel shows that administrative delineations of *communes* articulate with each other on the north-south gradient, which ensures sharp climatic differences between them. Figure AO1 in the online appendix presents the acreages and shares of each vertical level for each horizontal *commune*. Every *commune* has at least two of the five possible vertical levels. The majority of *communes* count three different vertical levels, with an average number of 3.87 levels per *commune*. Vineyards ranked as *Village*, *Premier Cru*, and *Grand Cru* are present in 28, 24, and 11 *communes* accounting for 90%, 77.4%, and 35.5% of all of them, respectively.

Table A1 in the online appendix presents the summary statistics about the exhaustive plot-level data that we use for the 31 *communes* of the region. For approximately 60 000 vineyard plots of a tiny average size of 0.2 ha (~ 0.5 acres), the elevation is distributed between 200 and 500 m, with an average of 286 m. Slopes are on average 5.73 degrees with high standard deviation (the coefficient of variation is ~ 100%). Solar radiation is distributed from 0.58 to 1.23 million Joules, with an average of 1.05 million J. To add flexibility to the econometric estimations, the exposition variable is discretized into eight dummy variables for different semi-quadrants, which shows that more than 50% of vineyard plots have a south-eastern exposition, between 90 and 180 degrees. Table A1 also shows the current distribution of the vertical dimensions of GIs and the distribution in 1936 when the INAO was created. We also use additional geological and pedological variables as fixed effects to control for sub-soil and soil characteristics. Because such variables are not central to the empirical strategy that we propose, we do not report them here.

Model of GI Designation

First, we present the structural model of GI designation that is assumed to be the data-generating process. Next, we discuss the empirical challenge of separating the *terroir* effects from the intangible influences and the specification procedure that we propose. Finally, we describe the decomposition of the vineyard quality signal from the GI information available to consumers.

Structure of GIs

The fine-scale variation of natural characteristics (i.e., *terroir*) between vineyard plots is the basis of the GI classification scheme. The vineyard quality index is an unknown function $q : \mathbb{R}^{K^*} \mapsto \mathbb{R}$ of the K^* natural characteristics $\mathbf{X}^*$ of each vineyard plot. From this scalar quality, GIs are designated through a continuous latent variable y^* defined as the sum of the vineyard quality index and idiosyncratic random designation noise noted ξ such that:

$$y^* = q(\mathbf{X}^*) + \xi. \tag{1}$$

The mapping between tangible *terroir* characteristics $\mathbf{X}^*$ and the vineyard quality index represents the cumulative knowledge from informed people who have contributed to the vineyard classification throughout history. At this stage, we consider the latent variable as an unbiased, though imperfect because of designation noise, signal of vineyard quality with $\mathbb{E}(\xi \mid \mathbf{X}^*) = 0$. Designation noise could be attributed to imperfect knowledge or anecdotal facts that cause random deviations around the quality signal. Designation noise is more generally due to the absence of a deterministic rule between vineyard characteristics and GIs; thus, the orthogonality between the designation noise and $\mathbf{X}^*$ is more a definition than an assumption.

The hierarchical structure of GIs is modeled through the multi-valued scalar $y \in \{1, \dots, 5\}$ that represents the vertical differentiation of GIs: *Côteaux Bourguignons* < *Bourgogne Régional* < *Village* < *Premier Cru* < *Grand Cru*. The GI of a given vineyard plot is a crude measurement of the underlying latent variable through a threshold-crossing relationship:

$$y = j \quad \Leftrightarrow \quad \alpha^c_{j-1} < y^* < \alpha^c_j, \quad \text{for} \quad j = 1, \dots, 5, \tag{2}$$

where $\alpha^c_0 = -\infty < \alpha^c_1 < \cdots < \alpha^c_5 = +\infty$ for every *commune* $c \in \{1, \dots, 31\}$ by construction. The superscript c on the thresholds indicates the *commune* in which the vineyard is located among the 31 *communes* of the area under consideration, and represents the horizontal dimension of GIs by municipality-specific thresholds. The variation in the thresholds between *communes* corresponds to the differential treatments that have been documented by historians and presented above. For example, a *commune* c_1 receives preferential treatment in terms of *Premier Cru* ($j = 4$) if its corresponding thresholds are lower than those of another given *commune* c_2: $\alpha^{c_1}_3 < \alpha^{c_2}_3$ and $\alpha^{c_1}_4 < \alpha^{c_2}_4$.

This means that the quality requirements for *Premier Cru* of the *commune* c_1 are less stringent and, consequently, the average quality is lower: $\mathbb{E}(y^* \mid y = 4, c = c_1) < \mathbb{E}(y^* \mid y = 4, c = c_2)$.[9]

Within a given *commune*, the ordered structure of GIs provides an efficient (i.e., unbiased) certification process as defined by De and Nabar (1991); the probability that a vineyard is classified in at least its own quality category is higher than the probability that another lower-quality vineyard will be classified in at least that category. For two vineyard plots, 1 and 2, with differentiated tangible characteristics such that $q(\mathbf{X}_1^*) > q(\mathbf{X}_2^*)$ and located within the same *commune* c_0, one can show that $\text{Prob}(y_1 \geqslant j) > \text{Prob}(y_2 \geqslant j)$ for all j because:

$$\text{(3)} \qquad \text{Prob}(y_i \geqslant j) = F\left[q(\mathbf{X}_i^*) - \alpha_{j-1}^{c_0}\right], \quad \text{for} \quad i = 1, 2.$$

where F is the strictly increasing cumulative distribution function of $-\xi$. The efficiency of the GI designation scheme is also verified in the absence of threshold variations between *communes* (i.e., if α_j^c is constant among c for each j), which is equivalent to lack of bias in the GI signal.

The efficiency property (or absence of bias) is no longer true for vineyard plots located in different *communes*, say c_1 and c_2 to continue with the same example. The lesser quality vineyard plot 2 has a higher probability of being classified at least j_1 (the GI quality level of vineyard 1) if $\alpha_{j_1}^{c_2} - \alpha_{j_1}^{c_1} > q(\mathbf{X}_1^*) - q(\mathbf{X}_2^*)$. In this case, the preferential treatment given to *commune* c_2 is a source of bias in the GI classification that contradicts the efficiency of the vertical GI differentiation ($\alpha_{j_1}^{c_2} > \alpha_{j_1}^{c_1}$ is a necessary condition to have a higher probability for the vineyard plot 2 compared to 1). In particular, the probability that another plot from another *commune* (e.g., plot 3 from *commune* c_3) of the same quality as plot 1 but higher in the GI classification is equal to the ordinal superiority measure defined by Agresti and Kateri (2017):

$$\text{(4)} \qquad \gamma_{3|1}^{j_1} \equiv \text{Prob}(y_3 > y_1 \mid \mathbf{X}_1^*) \approx F\left(\frac{\alpha_{j_1}^{c_3} - \alpha_{j_1}^{c_1}}{\sqrt{2}}\right).$$

We use the approximation that the cdf of the normalized difference between designation noises is equal to the marginal cdf; this approximation is exact in the case of a Gaussian distribution. This measure of ordinal superiority determines the bias in the GI designation independently of the conditioning tangible characteristics X_1^* of vineyard plots. This allows a direct comparison between the horizontal dimension c of GIs for each vertical level j. For a given *commune* of reference (e.g., c_1 in Equation 4), this implies $30 \times 5 = 150$ measures of ordinal superiority. Therefore, we assume an additive separability between the horizontal and vertical intercepts to simplify the comparison, $\alpha_j^c = \alpha_j - \mu_c$. The ordinal superiority measure between two identical plots located in given *communes* A and B becomes $\gamma_{A|B} = F\left[(\mu_{c_B} - \mu_{c_A})/\sqrt{2}\right]$ regardless of j, which allows the number of ordinal superiority measures to be divided by 5. The resulting 30 statistics provide objective measures of the differential treatments that have been applied between *communes* according to the GI vertical classification of their vineyards. The presence of significant ordinal superiority measures indicates some significant bias in the GI signal, for which the ordinal superiority measures estimate the size.

Ordered Generalized Additive Model

The estimation of the unknown function $q(\cdot)$ that relates tangible attributes of vineyards to the vineyard quality index is subject to two empirical challenges that we consider jointly: the specification of the functional form for the effect of a given tangible variable x_k and the presence of unobserved *terroir* variables that impact vineyard quality. These unobserved effects for the econometrician are taken into account in GI designations by observations in the field because they are known to people involved in GI designations. This is a serious econometric concern due to the potential confounding effect that such variables could have through their spurious correlations with *commune* delineations that group together adjacent vineyard plots. Identifying the information conveyed by GIs about tangible variables requires that all of these *terroir* variables be observable, which is unfortunately not the case, and probably never will be.

Instead, we propose to estimate an Ordered Generalized Additive Model (OGAM, Wood et al., 2016; Wood, 2017) that allows a semiparametric specification of the effect of each observed tangible variable and enables us to control for omitted *terroir* variables through bivariate smoothing of geographic coordinates. This identification strategy is based on the definition of *terroir* as the full set of natural variables that impact the vineyard quality index. As they originate from natural processes (Pickett and Cadenasso, 1995), we consider them as spatially continuous according to the axiom that nature makes no jumps, in contrast to the discontinuities introduced by administrative delineations of *communes* related to lobbying and political bargaining over the GI designations.

Consider that we only observe the realizations of a subset $\mathbf{X}_i \subset \mathbf{X}_i^*$ of all *terroir* variables that are taken into account in the GI designation scheme for a given vineyard plot $i = 1, \ldots, N$. These observed tangible variables are elevation, slope, exposition, solar radiation, geology, pedology, and geographic coordinates in our data. By noting C_i the row vector of dimension 31, with the typical element c_{ih} equal to 1 if vineyard i is located in *commune* h and zero otherwise, the specification of a logistic distribution for the reduced-form errors leads to a classical parametric ordered logit model that can be estimated by maximum likelihood:

$$\text{Prob}(y_i > j \mid \mathbf{X}_i, \mathbf{C}_i) = \Lambda[\mathbf{B}(\mathbf{X}_i)^\top \boldsymbol{\beta} + \mathbf{C}_i^\top \boldsymbol{\mu} - \alpha_j], \tag{5}$$

where Λ is the logistic cdf. The intangible determinants that impact GIs through varying designation thresholds, noted μ_c previously, are taken into account by the dummy variables $\mathbf{C}_i$ which work as *commune* fixed effects. In the absence of theoretical priors for the effects of all observed tangible variables $\mathbf{X}_i$, we specify them through a series of functional transformations noted as $\mathbf{B}(\cdot)$ with an associated vector of coefficients $\boldsymbol{\beta}$. From an initial set of K observed tangible variables (with $K < K^*$), the series and vector of coefficients are of dimension $\widetilde{K} = \sum_k L_k$, where L_k is the number of transformations used to specify the effect of each variable x_k. For example, a second-order polynomial specification for all observed tangible variables is noted $\mathbf{B}(\mathbf{X}_i) = [x_{1i}\ x_{1i}^2\ x_{2i}\ x_{2i}^2\ \cdots\ x_{Ki}\ x_{Ki}^2]$ with a set of $\widetilde{K} = 2 \times K$ coefficients to estimate.

The results presented below will show that polynomial specifications have limited performance in accounting for the complex interactions between natural characteristics of vineyards and the continuous quality index used in GI designations. Thus, we turn to semiparametric thin plate regression splines that have optimal smooth approximation properties (Wood, 2017). The matrix $\mathbf{B}(\mathbf{X})$ is specified through additive low rank isotropic smoothers of the individual tangible variables x_k. The cost of this additional flexibility is the need to estimate jointly a smoothing parameter that controls the penalization of the overfit. Accordingly, the complexity of the spline transformations is determined endogenously for a given maximum basis reduction for each variable through a quadratic penalty. The penalized deviance is minimized by penalized iterated weighted least squares and the smoothing parameter is estimated using a separate criterion from the restricted maximum likelihood framework. The computational details are given in Wood et al. (2016).

The complexity of the effect of a given variable or of the whole model can be assessed by the effective degrees of freedom that account for the endogenous penalization of any given dimension reduction (Wood, 2017, p.273). The most sensitive point is the estimation of the smoothing parameter which is a source of additional uncertainty, whereas Wood et al. (2016) provide some corrections for inference and traditional goodness of fit measures, such as Akaike Information Criteria (AIC). Unfortunately, goodness of fit measures provide little guidance about the causal inference of *commune* effects that measure the intangible effects that bias the GI signal. We propose to determine the sufficient level of spatial smoothing with a heuristic procedure based on auxiliary regressions and surrogate residuals recently defined by Liu and Zhang (2018), online appendix presents more details about this specification procedure.

Informational Content

The formal analysis that we develop about the precision part of the informational content of GIs is based on the framework of Ganuza and Penalva (2010) for information signal ordering, in addition to the variance decomposition formulas provided by Bowsher and Swain (2012). We consider GIs as an information structure, i.e., a joint distribution between the states of the world (vineyard quality index) and the GIs. We propose to evaluate the extend to which the observation of y and c allows consumers to recover vineyard quality, assuming that a more informative signal leads to a more dispersed distribution of conditional expectations. We measure the dispersion through conditional variance of the signals. This leads to four nested variance decomposition:

$$\text{(6)} \qquad \textit{Total decomposition}: \mathbb{V}(y^*) = \mathbb{V}[q(X^*)] + \mathbb{V}[\xi]$$

$$\text{(7)} \qquad \textit{Joint decomposition}: \mathbb{V}[q(X^*)] = \mathbb{V}\{\,\mathbb{E}[q(X^*) \mid y, c]\,\} + \mathbb{E}\{\,\mathbb{V}[q(X^*) \mid y, c]\,\}$$

$$\text{(8)} \qquad \textit{Vertical decomposition}: \mathbb{V}\{\,\mathbb{E}[q(X^*) \mid y, c]\,\} = \mathbb{V}\{\,\mathbb{E}[q(X^*) \mid y]\,\} + \mathbb{E}\{\,\mathbb{V}[\mathbb{E}(q(X^*) \mid y, c) \mid y]\,\}$$

$$\text{(9)} \qquad \textit{Horizontal decomposition}: \mathbb{V}\{\,\mathbb{E}[q(X^*) \mid y, c]\,\} = \mathbb{V}\{\,\mathbb{E}[q(X^*) \mid c]\,\} + \mathbb{E}\{\,\mathbb{V}[\mathbb{E}(q(X^*) \mid y, c) \mid c]\,\}$$

The *total decomposition* in Equation 6 comes from the law of total variance, the law of iterated expectations, and the definition of designation errors by $\mathbb{E}(\xi \mid \mathbf{X}^*) = 0$. It presents the variance of the latent variable as the sum of a *signal variance* and a *noise variance* defined from the data-generating process. The signal to noise ratio $\mathbb{V}[q(\mathbf{X}^*)]/\mathbb{V}[\xi]$ gives the proportion of relevant information conveyed by the continuous quality grade $q(\mathbf{X}^*)$ in terms of the irrelevant information from noise ξ. This decomposition represents the maximum informational content that any GI signal can achieve for the data-generating process under consideration. This corresponds to the case in which the continuous quality grade is conveyed to consumers as a continuous score on wine labels.

The *joint decomposition* in Equation 7 comes from the law of total variance applied to the continuous quality grade (Bowsher and Swain, 2012). It disentangles the part of the signal that is conveyed jointly by the vertical and horizontal dimensions of GIs (the *joint signal*, which is the variance of the expectation) and the part that is lost due to the discretization of the continuous quality information (the *joint noise*, which is the expectation of the variance). If the continuous quality grade $q(\mathbf{X}^*)$ was observable, the share of the *joint signal* in the *total signal* would be the R^2 of the regression of $q(\mathbf{X}^*)$ on the full set of dummy variables from y and c.

The *vertical decomposition* in Equation 8 separates the *joint signal* into the part that is conveyed through the vertical dimension of GIs (the *vertical signal*, the variance of the expectation) and the residual part that remains for the horizontal dimension (the *vertical residual*). The first term represents the variance of the quality information that can be assessed by consumers only through the vertical dimension y of GIs. Consumers may choose to favor this dimension by choice based on their experience. An important point is that, in the absence of preferential treatment between *communes* in the GI designation scheme, the residual part of this decomposition would be zero. In such a case, the vertical dimension would be unbiased and provide all of the relevant information about quality available to consumers. The only loss in information would be due to the discretization of the continuous quality index and the *joint signal* would be equal to the *vertical signal*.

The *horizontal decomposition* in Equation 9 is symmetric to vertical decomposition, as it defines a *horizontal signal* and a *horizontal residual*. This means that decomposition of the *joint signal* between a *vertical* and *horizontal* part is non-unique, depending on the GI dimension that is privileged by consumers. The first *horizontal signal* measures the dispersion of the expectation of vineyard quality conditionally on the *commune* of the vineyards. This informational content is due both to the incidental spatial correlation between vineyard quality and *commune* delineations, and to the historical factors that have made GI thresholds dependent on the *communes*. In the absence of any preferential treatment of certain *communes*, this signal is reliable, as it indicates that some *communes* have better tangible conditions to make wines of better quality. Thus, the residual part of the decomposition is the marginal gain of using the vertical dimension of GIs for consumers who rely only on the horizontal dimension.

Results

We first present the estimation of the ordered models of GI designations, and interpret the varying thresholds as ordinal superiority measures between *communes*. Next, we discuss the precision criteria of the informational content of GIs with variance decomposition formula, turn to the ordered models of 1936 GIs to finally study the informational content of simulated GIs.

Models of GI Designation

The first column (0) of Table 1 reports the joint significance statistics from a standard ordered logit model with quadratic effects for the three topographic variables, third-order polynomials with full interactions for spatial coordinates, and pedology, geology, exposition, and *commune* fixed effects. The reported χ^2 statistics are equivalent to F-statistics for models with discrete outcomes. The tests indicate that all variables are significant at the 1% level, for an overall pseudo-R^2 of 36.7%. The most significant series of variables is the set of 31 *commune* dummies that represent the intangible lobbying effects on GI designations. This set of variables is closely followed by the pedology fixed effects and the polynomial transformation of spatial coordinates that controls for the effects of the longitude and latitude of vineyards. Elevation, solar radiation, geology, exposition, and slope variables follow in decreasing order of joint significance, for an overall significance that is slightly higher for tangible variables than intangible variables.

The non-linear effects of the three topographical variables on the latent quality index are reported in Figure A2 of the online appendix. Elevation and slope variables have inverted-U effects, with the highest vineyard quality at about 290 meters and 10 degrees. The effect of solar radiation increases linearly and southern exposition provides the highest marginal probability of a high GI classification. The top-left panel of Figure 0A3 in the online appendix shows the marginal effects of spatial coordinates on the latent index. The third-order parametric specification with full interactions produces some smooth ellipsoidal patterns with two central kernels that describe a core-periphery structure.

Columns (I) to (V) in Table 1 report the same significance statistics from OGAMs with increasing complexity in the spatial smoothing terms from left to right, as appears from the effective degrees of freedom for spatial coordinates. The semiparametric structure of these models maintains the same degrees of freedom for pedology, geology, exposition and *commune* fixed effects with 13, 14, 7, and 31 degrees, respectively. Increasing the complexity of the spline series of spatial coordinates increases the pseudo-R^2 to 75% and the percentage of good predictions to 90% in the most complex OGAM reported in the last column (V).

Simultaneously, the joint significance of spatial coordinates increases and the significance of all other explanatory variables decreases, except slope and exposition variables, for which the decrease of significance is not monotone. As expected, the spatial patterns of GI designations are increasingly grasped by spatial coordinates at the expense of other explanatory variables. Figure A2 in the online appendix shows the comparative advantage of OGAMs over the parametric model (0) in estimating the marginal effects of each explanatory variable. Panel A of Figure A2 shows that the strong effect of elevation in the 0–300 m range is not found in the parametric model, Panel B shows the same result for slope on the 0–5 degree range. These results are particularly stringent as these ranges concentrate the majority of vineyard plots. In terms of spatial smooth effects reported in Figure A3 of online appendix, OGAMs produce more detailed spatial variations than the broad ellipsoid pattern from the parametric model (0). This suggests some fine-scale spatial variations in the latent quality index according to the GI designation scheme. The significance of *commune* fixed effects decreases sharply by increasing the complexity of spatial smoothing, whereas it remains the second most important set of variable in the model (V).

Ordinal Superiority of Communes

The last row of Table 1 reports the bootstrapped F-statistics for the joint significance of *commune* dummies on surrogate residuals from auxiliary models that do not account for such fixed effects (see online appendix about causal inference). Figure A4 in the online appendix shows the relevance of smoothing spatial coordinates to control for unobserved *terroir* variables and improve the causal inference about *commune* effects. Initially, it appears that OGAMs allow to decreasing significantly the correlation between auxiliary residuals and *commune* effects, compared to the parametric ordered logistic model (0). A maximum effective degrees of freedom of approximately 700, which corresponds to model (IV) in Table 1, is a sufficient complexity level to rule out potentially correlated omitted *terroir* effects, as the insignificance of *commune* dummies on the surrogate residuals from the auxiliary regressions cannot be rejected according to the median of the bootstrapped F-statistics. The fact that *commune* effects remain significant in models (IV) and (V) of Table 1 indicates persistent effects of intangible lobbying effects from political bargaining about the GI designation scheme, even for precisely controlled *terroir* effects. Similar vineyard plots from one side or another of administrative boundaries have significantly different probabilities of being in different vertical levels of GIs.

Ordinal superiority measures computed from OGAMs with 700, 800, and 900 maximum edf are reported in Figure 2. A positive measure indicates that the *commune* is advantaged relatively to the average *commune* (Agresti and Kateri, 2017). We see that only vineyard plots from four *communes* are not differently designated from the average *commune* of the area. *Communes* from the *Côte de Nuits* in the north of the region are, on average, more advantaged than those of the *Côte de Beaune* in the south, as eight *communes* from this part of the region are among the 12 most advantaged. The proximity to Dijon, where trials of the use of vineyard names occurred between 1860 and 1936, is one potential explanation for this result, as well as the fact that it was usual that influential people living in Dijon owned vineyards in the *Côte de Nuits*, which is closer to Dijon than *Côte de Beaune* (Wolikow and Jacquet, 2018, see Figure 1 for the location of Dijon).

The *communes* that have a syndicate engaged in collective action appear to be privileged, but the separation is not clear-cut.[10] This hierarchy of advantaged and disadvantaged *communes* from current GI designation scheme is not significantly correlated with the average vertical level of the vineyards, as some advantaged *communes* do not have vineyards of high level on average (*Ladoix-Serrigny* and *Chorey-les-Beaune*), and some *communes* with high level vineyards are disadvantaged by the designation scheme (*Flagey-Echezeaux* and *Pommard*). We found that the ordinal superiority measures are only weakly positively correlated with average levels of current GIs ($R^2 = 0.06$, $t = 1.27$, see Figure 0A5 in the online appendix).

Informational Content of Current GIs

Table 2 reports the decomposition computed from equations (6) to (9) with $q(X_i^*)$ predicted from the five OGAMs (I) to (V) reported in Table 1. The empirical formulas, and the R code used to compute the different terms are available on the replication material file mentionned in the aknowledgements. As expected, the total signal shares reported in the first row of Table 2 increase from left to right and the total noise decreases.[11]

In contrast to this monotonic relationship between the total signal and the complexity of the spatial smoothing terms, the results from joint, vertical, and horizontal decomposition are more stable between specifications. For all models, the vertical and horizontal dimensions of GIs have high joint information content. From the last column of Table 2, the joint signal of approximately 78% is four times higher than the joint noise of 19%. The vertical dimension has higher informational content than the horizontal dimension, with a signal to noise ratio of 2 (65/32) compared to 0.33 (24/73). The horizontal residual, which represents the marginal informational content of the vertical dimension after the horizontal dimension is fully taken into account, is higher than the horizontal signal when using the horizontal dimension alone.

This result reinforces the superiority of the vertical dimension for conveying quality information, though the dimension of the signal is lower (5 levels instead of 31 items). From the vertical residual terms, we see that the vertical dimension of GIs has approximately 20% (13/65) bias in conveying information about vineyard quality. We find that the marginal and residual contributions of the horizontal dimension are quite low to inform the vertical dimension of current GIs.

Models of 1936 GIs

We estimate the same set of ordered models with the vertical GIs of 1936 as the outcome variable. At that point of history, the vertical dimension of GIs counted only three levels, as reported in the summary statistics in Table A1 in the online appendix: *Régional* < *Village* < *Grand Cru* with respectively 57%, 41%, and 3% of current vineyard plots.[12] The joint significance, and the marginal effects of *terroir* variables are reported in Table A2, and Figure A6, respectively, in the online appendix. The hierarchy of the joint significance of explanatory variables is very close to what is obtained for current GIs. The *commune*, pedology fixed effects, and geographic coordinates have the highest significance, followed by elevation, geology, solar radiation, slope, and exposition.

The marginal effects of elevation and slope also have an inverted-U pattern with similar maximum values, and the spatial smoothed patterns are also very close to what is found for current GIs. For the 1936 GI designation scheme, the control for omitted *terroir* variables is reached for smaller maximum edf of spatial coordinates (bootstrapped F-statistics are reported at the bottom of Table A2). In contrast, Figure A7 in the online appendix shows the ordinal superiority measures were more marked between *communes* in 1936. The *communes* from the *Côte de Nuits* already appeared as relatively advantaged (seven *communes* among the 11 most advantaged) and the effect of the syndicates of producers appear more clearly (see footnote 3). Figure 3 shows that the *commune* where a vineyard is located was a more important determinant of GI designations in the middle of the 20th century than it is currently. For 18 *communes* out of 25 (72%) the ordinal superiority measure decreases in absolute values in the last decades. This indicates that the GI designation scheme is increasingly efficient in the sens of De and Nabar (1991).

The first column of Table 3 reports the decomposition of the latent quality index according to the 1936 GIs. The GIs of 1936 have lower joint informational content than current GIs with a joint signal to noise ratio close to 1 (48/49.5), which indicates an increase of GI precision over the last hundred years or so. The bias from the vertical dimension doubles to 40% (14/35). The informational content of the vertical and horizontal dimensions are more balanced, whereas the vertical dimension remains more informative. The vertical dimension of 1936 GIs has a signal to noise ratio of 0.54 (34.4/63.1) compared to 0.32 (23.8/73.7) for the other horizontal dimension.

These improvements in the informational content of GIs over the last hundred years or so provide a concrete illustration of the ways some other GI schemes could be revised. Because the *commune* delineations have not changed since 1936, this evolution is exclusively attributable to the increase in the number of vertical levels and the allocation of plots between these levels. Since 1936, about 57% of current vineyard plots have had the same level. The two new levels *Coteaux bourguignons* and *Premiers crus* are composed of respectively 98% and 93.5% of vineyard plots that were on the closest current levels (i.e., respectively *Bourgogne régional* and *Village*). Importantly, the increases in the hierarchy are more frequent for vineyards from disadvantaged *communes* (39% for the 15 most disadvantaged *communes* of Figure 2) than for advantaged *communes* (27% for the 15 most advantaged *communes* in Figure 2). Following the management by INAO, more balanced GI designation schemes between administrative units have significantly increased their informational content.

Informational Content of Simulated GIs

In order to provide guidelines for better designated GIs, we performed different simulations of GI designation schemes as reported in columns S.I to S.VI in Table 3. These simulations consist in changing the GI schemes and evaluating the consequences on their informational content. The six vertical designation schemes under consideration were simulated by changing the predictions of the latent quality index that is mapped to GI levels by thresholds (in columns S.I, S.II, and S.III), and by changing the number of vertical levels from five to six (in S.IV, S.V and S.VI). We did not consider changing the horizontal dimension of GIs, because changing the administrative boundaries of *communes* in order to improve wine quality signaling is not policy-relevant.

Scheme S.0 is a benchmark scheme that tries to reproduce actual GI designations by adding simulated designation noises from surrogate residuals to the predictions of the latent quality index. This stochastic index is mapped to the vertical dimension of the simulated GIs with estimated thresholds and *commune* fixed effects. The second column, S.0, in Table 3 shows that the decomposition terms are similar to those obtained in the last column of Table 2 from current GIs. Next, we drop the designation errors from surrogate residuals in S.I, we drop the intangible *commune* effects in S.II, and we drop both designation errors and *commune* effects in S.III. In the last simulated designation schemes S.IV, S.V, and S.VI, *Bourgogne*, *Village* and *Premier Cru* levels are respectively divided into two different levels by adding a threshold, fixed at the mean of the estimated thresholds used for S.0. Each of these schemes corresponds to the creation of an additional level (e.g., *Bourgogne supérieur*, *Village supérieur* ou *Premier Cru supérieur*) enabling consumers to distinguish them.

The decomposition reported in Table 3 show that dropping the lobbying effects associated with *commune* effects is the most important policy for increasing the informational content of the vertical dimension of GIs. Conversely, reducing the designation noise is more important for increasing the joint signal, which corresponds to the assumption that consumers use the information of both GI dimensions. These two policy changes for GIs seem to be additively cumulative for increasing the informational content of both the vertical and joint signals. In particular, the marginal gain of dropping the *commune* effects is about the same with and without designation noise. Table 3 also shows that dropping the *commune* effect in the designation scheme increases the joint signal more than adding a sixth vertical level as in S.III, S.IV or S.V. Among these latter alternative schemes, we find that splitting the intermediate level *Village* is more efficient, but the differences are small.

Conclusion

We present a framework for modeling geographical indications (GIs) and disentangling their informational content, i.e., their capacity to describe the tangible characteristics of production sites. Applied to the wine-producing region of *Côte d'Or* (Burgundy, France), we find simultaneously a high precision of the vertical levels of GIs and a persistent bias from differential treatments between administrative units. This latter horizontal dimension corresponds to the spatial scale at which collective action and lobbying was historically made by wine professionals for GI designations. This indicates that improving the informational content of GIs is more a political than a technical matter (i.e., reducing the bias matters more than improving the precision).

The informational content of GIs that we study is complementary to the value of this information for consumers. Because it does not depend on taste, knowledge, perception, or fashion, increasing the informational content could appear as a more robust policy objective in the long run. Nevertheless, some important aspects of the value of the GI information are omitted. Firstly, the informational content of GIs generally increases with the complexity of the signal, while a complex signal could be too difficult (so, less valuable) to interpret for consumers. Secondly, the informational content treats symmetrically high and low levels of GIs, while the informational content of high levels would be more valuable than that of low levels (consumers who chose high levels would supposedly have a greater preference for quality). Accordingly, more research is needed to convert these results in terms of the value of GI information, which would require economic data about wine prices, vineyard prices, or surveys of consumers' preferences.

Our empirical strategy is based on the difference between the assumed spatial continuity of *terroir* and the discontinuity of administrative boundaries, from which we disentangle the tangible and intangible determinants of GIs. Due to the small size of vineyard plots in the region, the smooth functions of geographic coordinates allow us to control for the fine-scale variations of unobserved heterogeneity from the *terroir*. The estimated spatial patterns of the latent quality index grasped by these functions are not exclusively related to tangible characteristics that matter for wine quality. In particular, they can grasp some spatial interactions of reputation or influence between vineyard plots on both sides of a *commune* boundary. Nevertheless, we find that the main decomposition results about the average magnitude for the signal-to-noise ratios are robust to the degree of spatial smoothing used in ordered regressions. Taking into account fine-scale variations in *terroir* is important when estimating the bias for the GI signal, but not decisive for signal precision.

The benefits of the long-term history for the informational content of GIs would require some second thoughts about GI flexibility, which is sometimes required to keep pace with changing consumers' preferences and changing determinants of wine quality (particularly in the face of climate change). As a human institution, which requires political bargaining and the involvement of producers with private information, the unbiased nature of the GI signal would probably not be attained spontaneously. Moreover, the regular modifications that would be required to keep pace with the changing preferences or changing environment would increase the confusing correlation between tangible and intangible characteristics and, consequently, decrease the informational content of GIs. The stability of GIs and their third-party management probably account for a large proportion of their informational content and their value that is currently observed on wine markets.

Notes

[1]More than half of the 50 most expensive wines in the world are from the *Côte de Nuits* and *Côte de Beaune* studied here, according to `https://www.wine-searcher.com/most-expensive-wines`, accessed March 23, 2020.

[2]One may think of the need to control for local climate patterns (Labbé et al., 2019), soil microbes (Bokulich et al., 2014; Gilbert et al., 2014), or any other form of spatial heterogeneity (Pickett and Cadenasso, 1995) that could confound the lobbying effect associated with the administrative location of a given vineyard plot.

[3] The GIs that we consider combine a spatial delineation of vineyards and a set of production constrains about maximum yields, plant varieties, and minimum alcohol content from the official *cahier des charges*. We focus on the informational content of spatial delineations, the most common attribute of GIs around the world.

[4] We group in *terroir* the immobile and non-reproducible characteristics of vineyards, i.e., those that best justify the reference to production sites in the long run. This focus on natural determinants is usual in economics (Gergaud and Ginsburgh, 2008; Cross et al., 2011), while this definition is narrower than others that include human determinants (know-how, usual practices, cultural heritage, Barham, 2003.)

[5] Another pivotal concept for evaluating a GI is the value of information (Foster and Just, 1989; Rousu et al., 2014), which depends more on short-run determinants such as perception, fashion, preference, choice set, or cognitive ability (Klain et al., 2014; Liaukonyte et al., 2015). It represents a less robust objective for GIs in the long run.

[6] We use the term vineyard quality in reference to a hierarchical structure of GIs that is modeled from a continuous latent variable crossing ordered thresholds (Storchmann, 2005; Ashenfelter and Storchmann, 2010).

[7] The *Grand Cru Bonnes Mares* is shared between the *communes* of *Chambolle-Musigny* and *Morey-Saint-Denis*, the *Fixin Premier Cru Clos de la Perrière* is shared between the *communes* of *Brochon* and *Fixin*, and the *Vosnes-Romanée Village* is shared between the *communes* of *Vosnes-Romanée* and *Flagey-Echézeaux*.

[8]This complexity reaches its maximum in the two *communes* of *Chassagne-Montrachet* and *Puligny-Montrachet*, which share *Grand Cru Montrachet* and have chosen to add it to their administrative names.

[9]The link with average quality from this last inequality requires the addtional assumption that $\mathbb{E}(\xi \mid \mathbf{X}^{*}, \mathbf{C}) = 0$, i.e., that the random part of the latent variable is unrelated between *communes*. We make this assumption in the rest of the article, which has the same rationale as the orthogonality of designation noise in regard to *terroir* variables presented above and even implies it by the law of iterated expectations: $\mathbb{E}(\xi \mid \mathbf{X}^{*}) = \mathbb{E}[\mathbb{E}(\xi \mid \mathbf{X}^{*}, \mathbf{C}) \mid X^{*}] = 0$.

[10]Jacquet 2009 (p.189, 211) reports that the *communes* of *Vougeot*, *Aloxe-Corton*, *Ladoix-Serrigny*, *Gevrey-Chambertin*, *Vosne-Romanée*, and *Santenay* had the first syndicates, with some internal conflicts for *Santenay*.

[11]As the variance of errors is normalized to identify ordered models and the variance of y^* from the data-generating process is constant between models, the increase in the total signal and the decrease in total noise are two sides of the same coin, as they come from the increase in the variance of the latent quality index predicted by tangible variables.

[12]We drop the *communes* of *Chenôve*, *Marsannay-la-Côte*, *Couchey*, *Comblanchien*, *Corgoloin*, and *Saint-Romain* because they contained only one vertical GI level in 1936, so their fixed effects are not identified.

References

Agresti, A. and Kateri, M. (2017). Ordinal probability effect measures for group comparisons in multinomial cumulative link models. *Biometrics* 73: 214–219.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics* : 488–500.

Albano, G. L. and Lizzeri, A. (2001). Strategic certification and provision of quality. *International Economic Review* 42: 267–283.

Ali, H. H. and Nauges, C. (2007). The pricing of experience goods: The example of *en primeur* wine. *American Journal of Agricultural Economics* 89: 91–103.

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. *The Review of Economics and Statistics* 92: 333–349.

Ay, J.-S., Chakir, R. and Marette, S. (2017). Distance decay in the willingness to pay for wine: Disentangling local and organic attributes. *Environmental and Resource Economics* 68: 997–1019.

Barham, E. (2003). Translating terroir: The global challenge of French AOC labeling. *Journal of Rural Studies* 19: 127–138.

Benabou, R. and Laroque, G. (1992). Using privileged information to manipulate markets: Insiders, gurus, and credibility. *Quarterly Journal of Economics* 107: 921–958.

Blackwell, D. (1951). Comparison of experiments. *proceedings of the second berkeley symposium on mathematical statistics and probability* ed. by J. Neyman, University of California Press, Berkeley: 93–102.

Bokulich, N. A., Thorngate, J. H., Richardson, P. M. and Mills, D. A. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proceedings of the National Academy of Sciences* 111: E139–E148.

Bonroy, O. and Constantatos, C. (2008). On the use of labels in credence goods markets. *Journal of Regulatory Economics* 33: 237–252.

Bonroy, O. and Constantatos, C. (2014). On the economics of labels: How their introduction affects the functioning of markets and the welfare of all participants. *American Journal of Agricultural Economics* 97: 239–259.

Bowsher, C. G. and Swain, P. S. (2012). Identifying sources of variation and the flow of information in biochemical networks. *Proceedings of the National Academy of Sciences* 109: E1320–E1328.

Brécard, D. (2014). Consumer confusion over the profusion of eco-labels: Lessons from a double differentiation model. *Resource and energy economics* 37: 64–84.

Carew, R. and Florkowski, W. J. (2010). The importance of geographic wine appellations: Hedonic pricing of Burgundy wines in the British Columbia wine market. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 58: 93–108.

Caswell, J. A. and Mojduszka, E. M. (1996). Using informational labeling to influence the market for quality in food products. *American Journal of Agricultural Economics* 78: 1248–1253.

Combris, P., Lecocq, S. and Visser, M. (2000). Estimation of a hedonic price equation for Burgundy wine. *Applied Economics* 32: 961–967.

Cross, R., Plantinga, A. J. and Stavins, R. N. (2011). What is the value of terroir? *American Economic Review* 101: 152.

De, S. and Nabar, P. (1991). Economic implications of imperfect quality certification. *Economics Letters* 37: 333–337.

Deconinck, K. and Swinnen, J. (2014). The political economy of geographical indications. *AAWE working paper No 174* .

Dranove, D. and Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature* 48: 935–63.

Duncan, A. and Greenaway, D. (2008). The economics of wine. *The Economic Journal* 118: F137–F141.

Foster, W. and Just, R. E. (1989). Measuring welfare effects of product contamination with consumer uncertainty. *Journal of Environmental Economics and Management* 17: 266–283.

Ganuza, J.-J. and Penalva, J. S. (2010). Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica* 78: 1007–1030.

Garcia, J.-P. (2014). La construction des climats viticoles en Bourgogne, la relation du vin au lieu au Moyen Âge. *Atelier du Centre de Recherches Historiques* 12: 22 p.

Gergaud, O. and Ginsburgh, V. (2008). Natural endowments, production technologies and the quality of wines in bordeaux. does terroir matter? *Economic Journal* 118: F142–F157.

Gergaud, O., Livat, F., Rickard, B. and Warzynski, F. (2017). Evaluating the net benefits of collective reputation: The case of Bordeaux wine. *Food Policy* 71: 8–16.

Gilbert, J. A., Lelie, D. van der and Zarraonaindia, I. (2014). Microbial terroir for wine grapes. *Proceedings of the National Academy of Sciences* 111: 5–6.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics* 24: 461–483.

Guerra, G. A. (2001). Certification disclosure and informational efficiency: A case for ordered ranking of levels. *University of Oxford, Department of Economics, Discussion Paper 64* ISSN 1471-0498.

Gustafson, C. R., Lybbert, T. J. and Sumner, D. A. (2016). Consumer sorting and hedonic valuation of wine attributes: exploiting data from a field experiment. *Agricultural economics* 47: 91–103.

Jacquet, O. (2009). *Un siècle de construction du vignoble bourguignon. Les organisations vitivinicoles de 1884 aux AOC*. Editions Universitaires de Dijon.

Jin, G. Z. and Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *Quarterly Journal of Economics* 118: 409–451.

Josling, T. (2006). The war on terroir: geographical indications as a transatlantic trade conflict. *Journal of Agricultural Economics* 57: 337–363.

Jullien, A. (1816). *Topographie de tous les vignobles connus*. Colas, Paris.

Klain, T. J., Lusk, J. L., Tonsor, G. T. and Schroeder, T. C. (2014). An experimental approach to valuing information. *Agricultural economics* 45: 635–648.

Labbé, T., Pfister, C., Brönnimann, S., Rousseau, D., Franke, J. and Bois, B. (2019). The longest homogeneous series of grape harvest dates, beaune 1354–2018, and its significance for the understanding of past and present climate. *Climate of the Past Discussions* 2019: 1–33, doi: 10.5194/cp-2018-179.

Laffont, J.-J. (1985). *Économie de l'incertain et de l'information*. Economica.

Lavalle, J. (1855). *Histoire et statistique de la vigne et des grands vins de la Côte d'Or*. Daumier, Dijon.

Lence, S. H., Marette, S., Hayes, D. J. and Foster, W. (2007). Collective marketing arrangements for geographically differentiated agricultural products: Welfare impacts and policy implications. *American Journal of Agricultural Economics* 89: 947–963.

Li, D. (2017). Expertise versus bias in evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics* 9: 60–92.

Liaukonyte, J., Streletskaya, N. A. and Kaiser, H. M. (2015). Noisy information signals and endogenous preferences for labeled attributes. *Journal of Agricultural and Resource Economics* : 179–202.

Liu, D. and Zhang, H. (2018). Residuals and diagnostics for ordinal regression models: A surrogate approach. *Journal of the American Statistical Association* : 1–10.

Lohr, L. (1999). Welfare Effects of Eco-label Proliferation: Too Much of a Good Thing? Tech. rep., Food and Agricultural Marketing Consortium (FAMC).

Marette, S., Crespi, J. M. and Schiavina, A. (1999). The role of common labelling in a context of asymmetric information. *European Review of Agricultural Economics* 26: 167–178.

Meloni, G. and Swinnen, J. (2018). Trade and terroir. the political economy of the world's first geographical indications. *Food Policy* 81: 1–20.

Menapace, L. and Moschini, G. (2012). Quality certification by geographical indications, trademarks and firm reputation. *European Review of Agricultural Economics* 39: 539–566.

Mérel, P. and Sexton, R. J. (2012). Will geographical indications supply excessive quality? *European Review of Agricultural Economics* 39: 567–587.

Messer, K. D., Costanigro, M. and Kaiser, H. M. (2017). Labeling food processes: The good, the bad and the ugly. *Applied Economic Perspectives and Policy* 39: 407–427.

Morelot, D. (1831). *Statistique de la vigne dans le département de la Côte d'Or*. Lagier, Dijon.

Moschini, G., Menapace, L. and Pick, D. (2008). Geographical indications and the competitive provision of quality in agricultural markets. *American Journal of Agricultural Economics* 90: 794–812.

Nelson, P. (1970). Information and consumer behavior. *The Journal of Political Economy* 78: 311–329.

Oczkowski, E. (2001). Hedonic wine price functions and measurement error. *Economic record* 77: 374–382.

Oczkowski, E. and Doucouliagos, H. (2015). Wine prices and quality ratings: A meta-regression analysis. *American Journal of Agricultural Economics* 97: 103–121.

Pickett, S. T. and Cadenasso, M. L. (1995). Landscape ecology: spatial heterogeneity in ecological systems. *Science* 269: 331–333.

Roullier-Gall, C., Boutegrabet, L., Gougeon, R. D. and Schmitt-Kopplin, P. (2014). A grape and wine chemodiversity comparison of different appellations in Burgundy: Vintage vs terroir effects. *Food chemistry* 152: 100–107.

Rousu, M. C., Marette, S., Thrasher, J. F. and Lusk, J. L. (2014). The economic value to smokers of graphic warning labels on cigarettes: Evidence from combining market and experimental auction data. *Journal of Economic Behavior & Organization* 108: 123–134.

Sáenz-Navajas, M.-P., Campo, E., Sutan, A., Ballester, J. and Valentin, D. (2013). Perception of wine quality according to extrinsic cues: The case of Burgundy wine consumers. *Food Quality and Preference* 27: 44–53.

Shapiro, C. (1982). Consumer information, product quality, and seller reputation. *The Bell Journal of Economics* : 20–35.

Storchmann, K. (2005). English weather and Rhine wine quality: An ordered probit model. *Journal of Wine Research* 16: 105–120.

Tadelis, S. (1999). What's in a name? Reputation as a tradeable asset. *American Economic Review* 89: 548–563.

Unwin, T. (1999). Hedonic price indexes and the qualities of wines. *Journal of Wine Research* 10: 95–104.

USTR (2017). 2017 Special 301 Report. *Office of the United States Trade Representative* 81 p.: https://ustr.gov/sites/default/files/301/2017Special301ReportFINAL.PDF.

van Leeuwen, C., Friant, P., Chone, X., Tregoat, O., Koundouras, S. and Dubourdieu, D. (2004). Influence of climate, soil, and cultivar on terroir. *American Journal of Enology and Viticulture* 55: 207–217.

van Leeuwen, C., Roby, J.-P. and Rességuier, L. de (2018). Soil-related *terroir* factors: A review. *OENO One* 52: 173–188.

Vives, X. (2010). *Information and learning in markets: the impact of market microstructure*. Princeton University Press.

Wolikow, S. and Jacquet, O. (2018). *Bourgogne(s) viticole(s) : Enjeux et perspectives historiques d'un territoire*. Éditions Universitaires de Dijon.

Wood, S. N. (2017). *Generalized additive models: An introduction with R*. Chapman and Hall/CRC, second edition.

Wood, S. N., Pya, N. and Säfken, B. (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association* 111: 1548–1563.

Yu, J., Bouamra-Mechemache, Z. and Zago, A. (2017). What is in a name? Information, heterogeneity, and quality in a theory of nested names. *American Journal of Agricultural Economics* 100: 286–310.

Zago, A. M. and Pick, D. (2004). Labeling policies in food markets: Private incentives, public intervention, and welfare effects. *Journal of Agricultural and Resource Economics* : 150–165.

Tables

Table 1: **Joint Variable Significance for Ordered Logit Models of Current GI Designations**

Variable	(0)	(I)	(II)	(III)	(IV)	(V)
Elevation	4 029.6**	4 123.2**	1 793.1**	1 189.9**	1 014.1**	867.04**
	[2]	[8.913]	[8.882]	[8.85]	[8.79]	[8.81]
Slope	531.9**	922.46**	343.61**	168.47**	155.46**	190.06**
	[2]	[8.3]	[8.241]	[8.331]	[8.173]	[7.722]
Solar Radiation	1 885.2**	2 091.3**	981.64**	797.71**	646.51**	530.96**
	[2]	[8.1]	[8.052]	[8.283]	[7.977]	[7.331]
Spatial Coords	7 602.7**	32 524**	59 294**	74 154**	78 445**	86 597**
	[15]	[98.59]	[295]	[483.2]	[666.6]	[841.4]
Pedology	8 810.7**	2 447.2**	713.07**	450.42**	408.64**	387.9**
	[13]	[13]	[13]	[13]	[13]	[13]
Geology	1 715.6**	977.42**	557.45**	500.46**	406.43**	440.86**
	[14]	[14]	[14]	[14]	[14]	[14]
Exposition	743.48**	61.043**	81.266**	171.5**	158.98**	130.52**
	[7]	[7]	[7]	[7]	[7]	[7]
Commune	9 767.6**	3 007.9**	2 295.2**	2 353.7**	1 721.6**	1 363.5**
	[31]	[31]	[31]	[31]	[31]	[31]
Nb Observ.	59 113	59 113	59 113	59 113	59 113	59 113
McFadden R^2	36.7	53.23	63.1	68.4	72.48	75.65
Pc good pred.	63.69	74.85	80.38	84.35	87.25	89.47
Akaike IC	104	77.22	61.4	53.09	46.76	41.93
Surrogate F	156.35	17.7	5.64	3.94	1.98	1.82

Note: $^{**}p < 0.001$ for joint significance tests from the reported chi-square statistics, effective degrees of freedom are in brackets. Column (0) corresponds to an ordered logit model with quadratic effects for elevation, slope, and solar radiation (df= 2) with a full interaction between third-order polynomials for longitude and latitude (df= $3 + 3 + 3 \times 3 = 15$) and with 13, 14, 7, and 31 dummy variables for pedology, geology, exposition, and *communes* fixed effects, respectively. Models (I) to (V) are OGAMs with elevation, slope and solar radiation additively specified with a maximum of 9 edf, shrunk endogenously by a quadratic penalization. Spatial coordinates are specified in increasing order of complexity with the maximum edf of 100, 300, 500, 700, and 900. The last row reports the bootstraped F-statistics for the joint nullity of *commune* effects on residuals from auxiliary regressions without *commune* dummies.

Table 2: **Signal Decompositions of the Informational Content of GIs**

		Effective degrees of freedom for spatial smoothing				
Decomp.	Term	(99)	(295)	(483)	(667)	(841)
Total	Signal	85.30	94.47	96.03	97.31	97.49
	Noise	14.70	5.53	3.97	2.69	2.51
Joint	Signal	69.73	70.15	76.71	75.19	78.62
	Noise	15.60	24.35	19.35	22.15	18.90
Vertical	Signal	54.05	48.77	51.68	56.25	65.18
	Residual	15.68	21.38	25.03	18.94	13.44
	Noise	31.25	45.70	44.36	41.07	32.31
Horizontal	Signal	18.34	16.61	25.60	22.62	23.82
	Residual	51.41	53.56	51.14	52.59	54.83
	Noise	66.99	77.88	70.46	74.72	73.70

Note: The effective degrees of freedom for spatial smoothing terms in parentheses show that the columns correspond to models (I) to (V) in Table 1. Decomposition terms are expressed as a percentage of variance of the latent variable y^* according to equations (6) to (9) in the text. For each column, the sum of *vertical signal* and *vertical residual* equals the *joint signal*, as does the sum of *horizontal signal* and *horizontal residual*. The *vertical noise* equals the sum of the *vertical residual* and the *joint noise*, and the *horizontal noise* equals the sum of *horizontal residual* and *joint noise*.

Table 3: **Signal Decompositions from Alternative GI Designation Schemes**

		Alternative scenarios of GI designations							
Decomp.	Term	1936	S.0	S.I	S.II	S.III	S.IV	S.V	S.VI
Total	Signal	97.49	97.49	97.49	97.49	97.49	97.49	97.49	97.49
	Noise	2.51	2.51	2.51	2.51	2.51	2.51	2.51	2.51
Joint	Signal	48.00	78.21	80.96	79.47	81.52	79.02	79.48	78.87
	Noise	49.52	19.31	16.55	18.05	15.99	18.50	18.03	18.64
Vertical	Signal	34.41	64.60	68.16	69.74	72.59	65.62	66.12	65.48
	Residual	13.59	13.61	12.80	9.73	8.94	13.40	13.36	13.40
	Noise	63.08	32.89	29.33	27.75	24.90	31.87	31.37	32.01
Horizontal	Signal	23.82	23.82	23.82	23.82	23.82	23.82	23.82	23.82
	Residual	24.19	54.42	57.17	55.67	57.73	55.22	55.69	55.08
	Noise	73.70	73.70	73.70	73.70	73.70	73.70	73.70	73.70

Note: Latent quality index used to simulate GI designation schemes is predicted from model (V) of Table 1, which provides the best fit of current GIs. The first column reports the informational content of the GIs of 1936. Scheme S.0 is a benchmark simulation that adds surrogate residuals to the latent quality index in order to mimic current GIs. S.I drops the random idiosyncratic terms, S.II drops the intangible determinants through averaging *commune* effects, and S.III drops both random terms and intangible determinants of GIs. Schemes S.IV, S.V, and S.VI add a vertical level on actual GIs for *Bourgogne*, *Village*, and *Premier Cru*, respectively, by an additional threshold fixed at the mean.

Figures

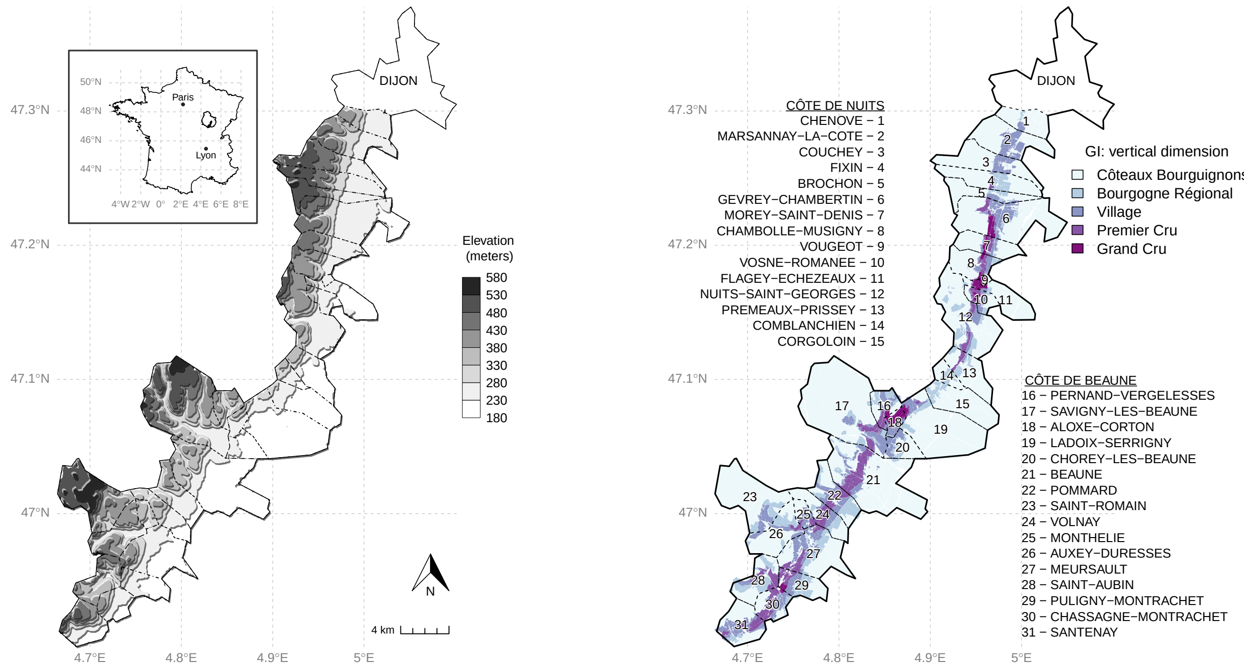

Figure 1: **Topography and geographical indications of the vineyards of the *Côte d'Or***

Note: The elevation on the left-hand map is discretized in 8 classes of 50 m intervals. From east to west, the elevation is first convex then concave, which means that the steepest slopes are for average elevations. GIs on the right-hand map are located on these steepest slopes. The spatial precision of the vertical dimension of GIs is such that best vineyards, classified as *Grands Crus*, are not visually well-separated from just below *Premiers Crus*. The right-hand panel also reports the names of the 31 administrative *communes* of the area, used to identify lobbying effects.

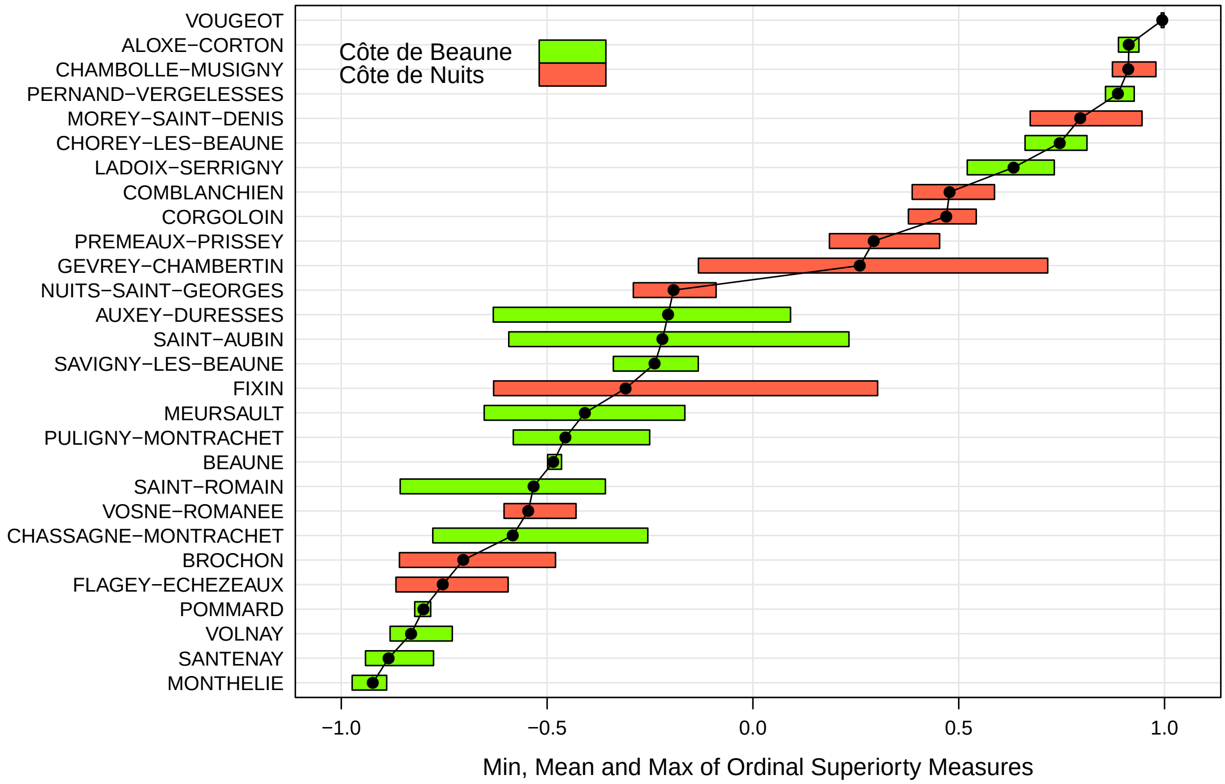

Figure 2: **Ordinal superiorty measures for the current GI designation scheme**

Note: For a given *commune* on the y-axis, ordinal superiority measures are computed as the difference between the estimated fixed effect μ_c and the average fixed effect $\overline{\mu}$ of every *commune* according to: $\Delta_c = 2 \times \Lambda[(\mu_c - \overline{\mu})/\sqrt{2}] - 1$. The horizontal bars represent the range of measures according to the OGAMs with 700, 800, and 900 maximum edf for the effects of spatial coordinates. Black dots represent the average of these measures. Relatively privilegied *communes* appear at the top of the y-axis, whereas relatively disadvantged *communes* appear at the bottom.

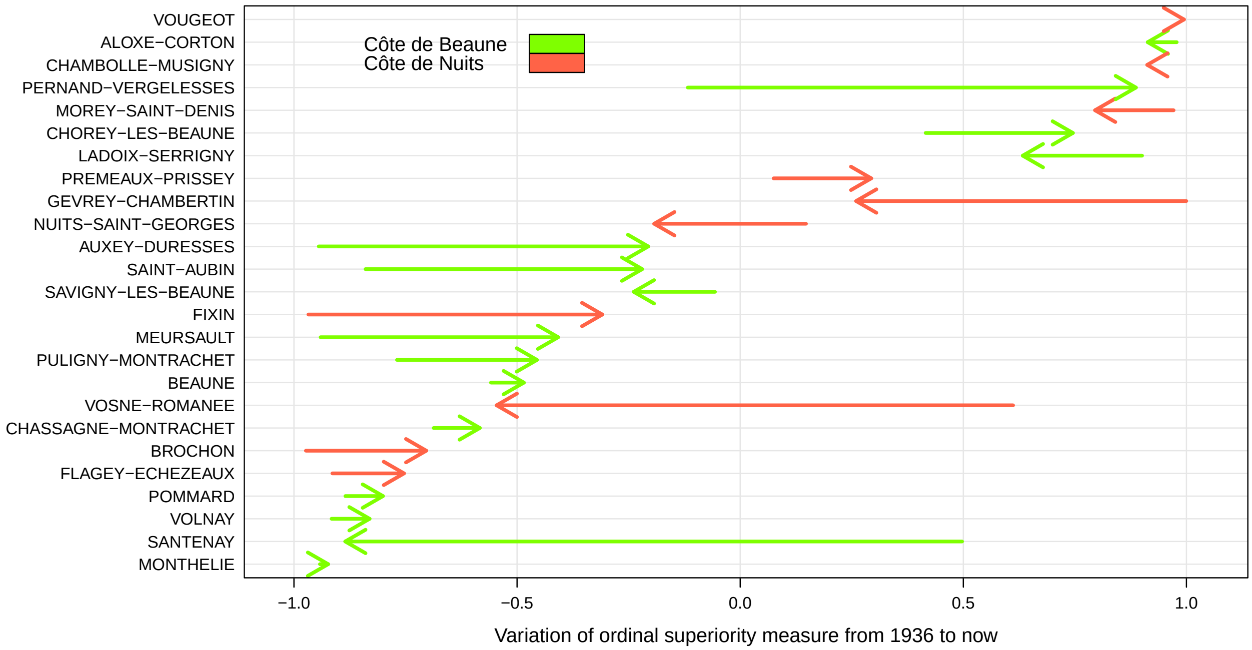

Figure 3: **Variations of ordinal superiorty measures between 1936 and today**

Note: For a given *commune* on the y-axis, ordinal superiority measures are computed as the difference between the estimated fixed effect μ_c and the average fixed effect $\overline{\mu}$ of every *commune* according to: $\Delta_c = 2 \times \Lambda[(\mu_c - \overline{\mu})/\sqrt{2}] - 1$. The arrows represent the change in the measures between the creation of INAO in 1936 and current GIs.