The informational content of geographical indications

Jean-Sauveur AY*

INRA UMR CESAER

Working Paper Version 1.1 : July 8, 2019

Abstract

Geographical indications (GIs) convey information about the place of production as a proxy for the quality of agricultural products. The quality of the GI proxy depends on its ability to follow the tangible characteristics of a production site instead of intangible factors, such as the influence of producers on the designation process. In this article, we disentangle the informational content of wine-related GIs for the *Côte d'Or* region of Burgundy, France. Due to their hierarchical and nested structure, GIs have high informational content with a signal to noise ratio of approximately 4. We also apply the signal decomposition to alternative wine classifications from history and counterfactual simulations to show significant improvements in GIs in the last century and potentially provide guidelines for better designated GIs in the future.

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

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

* Contact: jsay@inra.fr, UMR CESAER, AgroSup Dijon, INRA, Université de Bourgogne Franche-Comté. Adress: 25 boulevard Docteur Petitjean, 21000 Dijon (France). Data and R code are available in the Replication Material (RM) from the repository http://github.com/jsay/geoInd/ under copyleft licence GNU GPL V3. I acknowledge Mohamed Hilal, Jean-Marc Brayer, Pierre Curmi, and Florian Humbert for their help with the data.

1 Introduction

Using the place of production to signal the quality of agricultural products is not consensual in trade relations (Josling, 2006; USTR, 2017) but distinguishing good quality products from those of bad quality is recognized as being fundamental for consumers and producers when the quality cannot be assessed before deciding to buy or sell (Akerlof, 1970; Nelson, 1970). Thus, one obstacle in the debate is the extent to which geographical indications (GIs) provide information about product quality (Winfree and McCluskey, 2005; Yu et al., 2017). We studied this informational content of GIs through the econometric relationship between the natural and human characteristics of the vineyards and wine-related GIs of the *Côte d'Or* region (Burgundy, France).

Wine is an emblematic agricultural product, and its quality strongly depends on the natural conditions prevailing at production sites (Jackson and Lombard, 1993; Bokulich et al., 2014; Knight et al., 2015; van Leeuwen et al., 2018). Wine is also an experience good well-suited to studying the transmission of quality information between producers and consumers (Combris et al., 1997; Ali and Nauges, 2007; Ashenfelter, 2008; Storchmann, 2012). In Burgundy, the ranking of vineyards according to their quality potential for wine production has a long history dating back to the Middle Ages, with numerous modifications to the actual scheme (Jullien, 1816; Morelot, 1831; Lavalle, 1855; Danguy and Aubertin, 1892; Garcia, 2011; Wolikow and Jacquet, 2018). 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 items: *Côteaux Bourguignons* < *Bourgogne Régional* < *Bourgogne 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) and other wine-producing countries (Germany, United States and Italy, see Storchmann, 2005; Costanigro et al., 2010, 2019).

The main objective of this article was to estimate the informational content of actual, past, and simulated GI designation schemes for approximately 60 000 vineyard plots. The informational content is defined as the ability to describe the tangible characteristics of production sites. It is measured by the variance of a conditional quality index, according to the principle that more informative signals lead to greater variability of conditional expectation (Ganuza and Penalva, 2010). We propose disentangling tangible information from intangible information about production sites by decomposing a latent quality index estimated from actual GIs (Bowsher and Swain, 2012). The first set of tangible information relates to the natural attributes of vineyard plots that are known to impact wine quality: topography (elevation, slope, aspect), geology (subsoil material, soil depth, soil humidity), and climate (solar radiation, longitude, latitude). The second set of information relates to human characteristics that have historically impacted the GI designation process. Through the reputation of landowners, their influence with decision-makers or their collective actions, some administrative units have had differential treatment that could bias the signaled quality. Knowing the geographic co-variations between tangible and intangible characteristics, and the difficulty controlling for all tangible variables that impact vineyard quality (i.e., *terroir* variables), the major empirical challenge is to disentangle these two sources of variations. We propose a semiparametric approach that exploits the precise location of vineyard plots to control for the unobserved spatial heterogeneity through smooth functions of geographic coordinates (Wood et al., 2016). The empirical strategy is based on the difference between the spatial continuity of *terroir* and the discontinuity of administrative borders according to the axiom that nature does not make jumps.

This article is an empirical contribution to the literature about quality disclosure and strategic certification (see Bagwell, 2001; Dranove and Jin, 2010 for reviews). The vineyard quality index that we study is based exclusively on natural vineyard characteristics, in contrast to typical frameworks in which quality is strategically chosen by producers (Shapiro, 1982; Besanko et al., 1987; Albano and Lizzeri, 2001; Jin and Leslie, 2003; Desquilbet and Monier-Dilhan, 2014). The resulting exogeneity makes identification of the informational content of the quality signal easier and allows more transparent analysis of the role of history in the information conveyed by actual GIs put on the labels of wine bottles. We argue that the long history of GI designations allows us to neglect the role of actual wine producers and their undoubtedly tangible impact on wine quality. In effect, as generations of producers succeed each other with numerous vineyards bought and sold, the informational content of GIs is a predetermined collective reputation (Tirole, 1996) that is reasonably independent from the actual individual practices or skills of producers. In addition, the vineyard quality index 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). Because a GI name cannot be separated from its associated vineyard quality based on tangible characteristics, the GI information put on the label by producers is not related to their own characteristics, which is another source of tangible information not studied here.

A large body of literature about wine quality disclosure is concerned with expert reviews and the use of this information by consumers. Such ratings have been shown to have mainly short-term effects, both on the demand for (Friberg and Grönqvist, 2012) and the price of wines (Ali et al., 2008; Dubois and Nauges, 2010). The major problems with their aggregation (Ashenfelter and Quandt, 1999; Cardebat and Paroissien, 2015) and consistency (Cao and Stokes, 2010; Bodington, 2017) create some doubts about their own interest to consumers (Ashenfelter and Jones, 2013). Ratings by experts, judges, or websites have also been shown to be significantly divergent from historical GIs for Bordeaux wines (Thompson and Mutkoski, 2011), probably because of their fundamental differences. Ratings are exogenous year-to-year sources of information and not directly comparable to more stable public GIs voluntarily put on wine labels by producers. This observation introduces the tedious question of the endogenous adoption of quality disclosure for GIs, which is not a concern for expert review (Hollander et al., 1999). This could have unintended economic consequences, such as counter-signaling, in situations in which the certification is not adopted to signal the high quality of products (Bederson et al., 2018). This is *a priori* not the case for wine-related GIs in Burgundy, because their economic (Combris et al., 2000; Carew and Florkowski, 2010; Sáenz-Navajas et al., 2013) and historical (Meloni and Swinnen, 2018) importance is such that, to the best of our knowledge, wine producers and sellers in the region puts the GIs as the main informational message on wine labels.

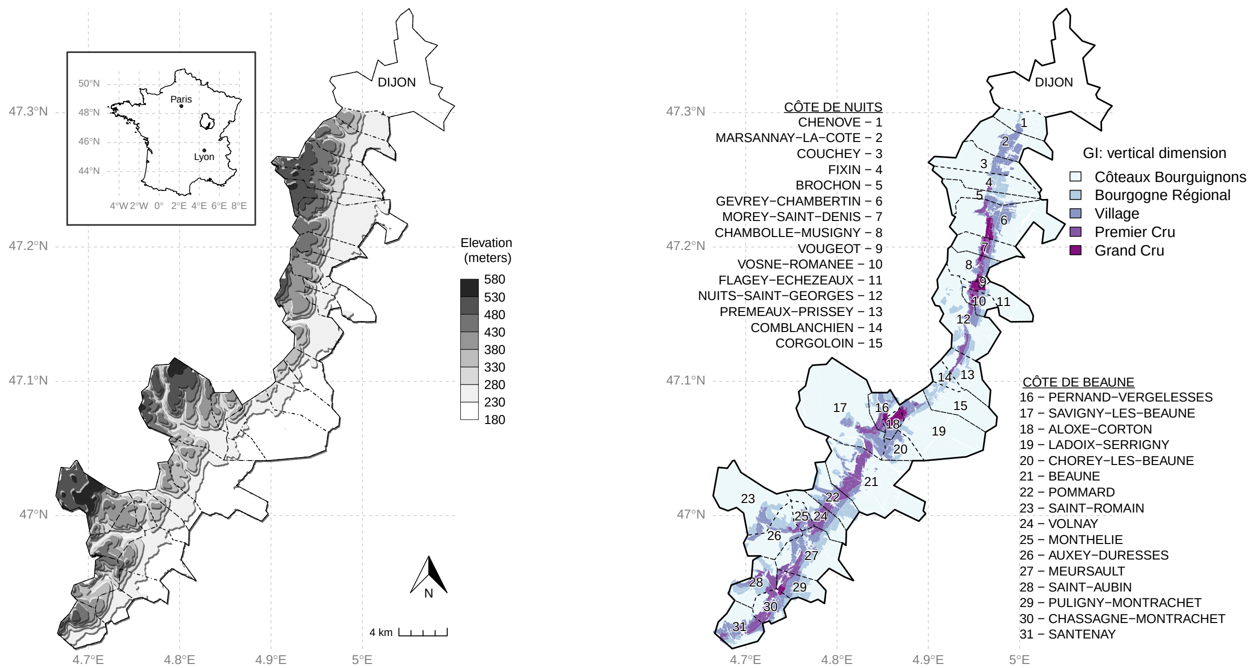
The analysis shows high informational content of actual GIs in terms of the underlying tangible vineyard characteristics, with 4-time times higher signal variance than noise variance (corresponding to an R^2 of ~ 80%). This high informational content agrees with the evidence of some signal bias due to intangible *commune* effects. This small administrative unit in France corresponds to the scale at which collective actions and lobbying have been performed historically, particularly the scale at which the syndicates (groups of wine producers) are structured (Jacquet, 2009). Nevertheless, this bias has decreased since the creation of GIs in 1936 through continuous evolution of GI designation schemes. This historical decrease in bias illustrates of the theory developed by Benabou and Laroque (1992) regarding strategic information transmission. We found that some producers or landowners could have profited from private information on vineyard quality manipulating the GI signal and extracted rents through undeserved high rating for vineyards in their administrative units. However, the hierarchical GI certification appears increasingly less biased, or more efficient in the sense of De and Nabar (1991); the probability that a vineyard plot is classified at least in its category is increasingly higher than the probability of lower quality plots being classified in that category. We also show that a monopolistic certifying party discloses useful information in the form of rank orderings, as predicted by Guerra (2001). This contrasts with models that found weak, if any, welfare gains associated with the information conveyed by a monopolistic certifying party (Shapiro, 1986; Lizzeri, 1999). These two results suggest that the high informational content of GIs and their actual economic importance in Burgundy come from their long history and independent management.

2 Context and data

2.1 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, 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 some direct and indirect impacts on wine quality. Firstly, elevation is expected to determine wine quality principally through its indirect 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. The latitude position of vineyards is also indirectly 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 longitude position of vineyards indirectly correlates with precipitation on the area, as a hill at the west provides a protective barrier that limits rains and, consequently, soil moisture. Thirdly, the 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 have strong importance for dry grapes and to concentrate aromas. Recognizing the indirect effects of these observed topographical and position variables is important, as fine climate data do not exist for consistent use on the vineyard plot scale.

Figure 1: **Topography and geographical indications (GIs) of the vineyards of the *Côte d'Or***
Notes: The elevation on the left side map is decretized in 8 classes of 50 m intervals. From the east to the west, the elevation is first convex then concave, which means that the highest slopes are for average elevations. GIs on the right side map are located on these highest 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 panel also reports the names of the 31 *communes* of the area, considered as the horizontal dimension of GIs.



2.2 Historical context

Some archaeological evidence located the first vineyards in the region in antiquity (Garcia, 2014). The first written evidence dated 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 first exhaustive spatial delineation of the region was an administrative separation of *communes* following the decree of 1789 after the French revolution. What we consider the horizontal dimension developed before the vertical dimension of actual GIs (Garcia, 2011, p.40). 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. 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 providesd 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, translation from the author).

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 are 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; translation by the author). In 1936, a French national institute, INAO, was created to legally manage what became the GIs of all wine regions of the country on 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).

2.3 Actual GI designations

Actual GIs are a nest between a vertical quality ranking of five items and a horizontal differentiation scheme through 31 administrative municipalities (*communes*, Figure 1, right panel). The highest quality vineyards are labeled *Grands Crus*, each of which has its own independent appellation title (e.g., "*Clos de la Roche*" or "*Chevalier-Montrachet*"). There are 32 *Grands Crus* in the area, 8 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 difference between these last two GIs was justified initially in terms of grape varieties (*Pinot noir* or *Chardonnay* for *Bourgogne Régional* and *Gamay noir* or *Aligoté* for *Côteaux Bourguignons*), but this distinction is less and less relevant as *Pinot noir* and *Chardonnay* become the main varieties.

The picture of actual GIs in the area is not complete without mentioning of the complexities that exist between the vertical and horizontal dimensions, which could lead to difficulties for consumers distinguishing their respective informational content. 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 name as a vertical item from *Grand Cru*, *Premier Cru* or even *Villages* can be found in two different *communes*: 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*. 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*. 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. However, the legal obligation or prohibition to mention the vertical item *Grand Cru*, *Premier Cru*, *Village*, *Régional* or *Coteaux Bourguignons* as the main information on wine bottle labels suggests that this information is clearly identifiable to consumers.

2.4 Summary Statistics

The precision of econometric estimations for disentangling the sources of variation in GIs depends on a balanced distribution of tangible variables and vertical items between and within the horizontal *commune* items. The left panel of Figure 1 shows that each *commune* approximately contains the whole range of elevation, slope, and exposition of the area, whereas the right panel shows that administrative delineations of *communes* articulate with each other on the north-south gradient, which ensures sharp climatic differences between them. Figure 3 in the Appendix presents the acreages and shares of each vertical item for each horizontal *commune* item. Every *commune* has at least two of the five possible vertical items. The majority of *communes* count three different vertical items, with an average number of 3.87 items 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 4 in the Appendix presents some summary statistics about the exhaustive plot-level data that we use on 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 an average 5.73 degrees wit high standard deviation (the coefficient of variation is ∼ 100%). The 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 aspect variable is discretized in 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 4 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 in the empirical strategy that we propose, we do not report them here. Interested readers can access these variables through the Replication Material (RM) from the link on the title page of this article.

3 Model of GI designation

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

3.1 Structure of GIs

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

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

The mapping between tangible *terroir* characteristics X^* and the objective quality index represents the cumulative knowledge from informed people that have contributed to the vineyard classification throughout history. At this stage, we consider the latent variable as an unbiased, through imperfect, signal of the quality of vineyards with $\mathbb{E}(\xi \mid X^*) = 0$. The designation noise could be attributed to imperfect knowledge or anecdotal facts that cause random deviations around the signal. The presence of designation noise is more generally due to the absence of a deterministic rule between vineyard natural characteristics and GIs; thus, the orthogonality of the designation noise is more a definition than an assumption. Determining the correlation between this quality signal and consumer preferences for the taste of wines and the related question of the value of GI information would require economic data about wine prices or consumer' surveys that we did not use here. Instead, we evaluated the relevance of the GI information according to this long-term quality signal which is different than evaluating the relevance of the quality signal itself. The ordered structure of the vertical dimension of GIs explains our reference to objective quality.

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* < *Bourgogne 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:

$$y = j \quad \Leftrightarrow \quad \alpha_{j-1}^c < y^* < \alpha_j^c, \quad \text{for} \quad j = 1, \dots, 5, \tag{2}$$

where $\alpha_0^c = -\infty < \alpha_1^c < \cdots < \alpha_5^c = +\infty$ for all *commune* $c \in \{1, \dots, 31\}$ by construction. The exponent c on the thresholds marks the *commune* in which the vineyard is located among the 31 *communes* of the area under consideration and represents the horizontal dimension of GIs. 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 smaller than those of another given *commune* c_2: $\alpha_3^{c_1} < \alpha_3^{c_2}$ and $\alpha_4^{c_1} < \alpha_4^{c_2}$. This means that the quality requirements for *Premier Cru* of the *commune* c_1 are less stringent and, consequently, the average quality is smaller: $\mathbb{E}(y^* \mid y = 4, c = c_1) < \mathbb{E}(y^* \mid y = 4, c = c_2)$.[1]

Within a given *commune*, the ordered structure of GIs provides an efficient certification process as defined by De and Nabar (1991); the probability that a vineyard is classified into at least its own quality category is higher than the probability that another vineyard with lower quality will be classified into at least that category. For two vineyard plots, 1 and 2, with differentiated tangible characteristics such that $q(X_1^*) > q(X_2^*)$ 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:

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

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

The efficiency property and absence of bias are 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 rank of vineyard 1) if $\alpha_{j_1}^{c_2} - \alpha_{j_1}^{c_1} > q(X_1^*) - q(X_2^*)$. 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 given 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 scheme is equal to the ordinal

[1] The link with average quality from this last inequality requires the addtional assumption that $\mathbb{E}(\xi \mid X^*, c) = 0$, 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 implies it by the law of iterated expectations: $\mathbb{E}(\xi \mid X^*) = \mathbb{E}[\mathbb{E}(\xi \mid X^*, c) \mid X^*] = 0$.

superiority measure defined by Agresti and Kateri (2017):

$$\gamma_{3|1}^{j_1} \equiv \text{Prob}(y_3 > y_1 \mid X_1^*) = F\left(\frac{\alpha_{j_1}^{c_3} - \alpha_{j_1}^{c_1}}{\sqrt{2}}\right). \tag{4}$$

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 bias in the GI signal, and the ordinal superiority measures are used to estimate the size of the bias.

3.2 Informational content

The formal analysis 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). Thus, we consider GIs as an information structure, i.e., a joint distribution between the states of the world (long-term vineyard qualities index) and the GIs (y and c represent the vertical and horizontal dimensions, respectively). We propose evaluating the extend to which the observation of y and c from wine labels allows consumers to recover wine quality, assuming that a more informative signal leads to a more dispersed distribution of conditional expectations. Contrary to Ganuza and Penalva (2010), we measure the dispersion through conditional variance of the signals as it is allowed by the work of Bowsher and Swain (2012). This leads to four nested variance decomposition:

$$\textit{Total decomposition}: \mathbb{V}(y^*) = \mathbb{V}[q(X^*)] + \mathbb{V}[\xi] \tag{5}$$

$$\textit{Joint decomposition}: \mathbb{V}[q(X^*)] = \mathbb{V}\{\,\mathbb{E}[q(X^*) \mid y, c]\,\} + \mathbb{E}\{\,\mathbb{V}[q(X^*) \mid y, c]\,\} \tag{6}$$

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

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

The *total decomposition* in Equation 5 comes from the law of total variance, the law of iterated expectations, and the definition of designation errors by $\mathbb{E}(\xi \mid X^*) = 0$. 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 $\mathbb{V}[q(X^*)]/\mathbb{V}[\xi]$ gives the proportion of relevant information conveyed by the continuous quality grade $q(X^*)$ 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 the wine label.

The *joint decomposition* in Equation 6 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 $q(X^*)$ was observable, the share of the *joint signal* in the *total signal* would be the R^2 of the regression of $q(X^*)$ on the full set of dummy variables from y and c. According to the nested structure of the *total* and *joint* decomposition, we define the *joint informational content* of horizontal and vertical dimensions of GIs as $\mathbb{V}\{\,\mathbb{E}[q(X^*) \mid y, c]\,\}/(\mathbb{E}\{\,\mathbb{V}[q(X^*) \mid y, c]\,\} + \mathbb{V}[\xi])$. This statistic measures the share of the quality information that is conveyed to consumers through both dimensions of GIs.

The *vertical decomposition* in Equation 7 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 or they can have a bounded rationality due to limited cognitive ability in understanding the full complexities of GIs. 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*. We also propose defining *vertical noise* as the sum of the *vertical residual* and the *joint noise*. This corresponds to the information loss of using only the vertical dimension:

$$\textit{Vertical noise}: \mathbb{E}\{\, \mathbb{V}[q(X^*) \mid y]\,\} = \mathbb{E}\{\, \mathbb{V}[q(X^*) \mid y, c]\,\} + \mathbb{E}\{\, \mathbb{V}[\mathbb{E}(q(X^*) \mid y, c) \mid y]\} \tag{9}$$

The *horizontal decomposition* in Equation 8 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. 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 that rely only on the horizontal dimension. Finally, we also define the *horizontal noise* as the sum of the *joint noise* and the *horizontal residual*, and it corresponds to the loss in GI signal when using only the horizontal dimension of GIs:

$$\textit{Horizontal noise :}\ \mathbb{E}\{\,\mathbb{V}[q(X^*)\mid c]\,\} = \mathbb{E}\{\,\mathbb{V}[\mathbb{E}(q(X^*)\mid y,c)\mid c] + \mathbb{E}\{\,\mathbb{V}[q(X^*)\mid y,c]\,\} \tag{10}$$

3.3 Ordered Generalized Additive Model

The estimation of the unknown function $q(\cdot)$ that relates tangible attributes of vineyards to the long-term 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. For example, the data set that we use does not contain fine-scale climate variables, such as temperature or precipitation, but one can also consider other unobserved *terroir* variables such as local variations in soil quality. These unobserved effects for the econometrician are taken into account in GI designations by observations in the field because they are known by 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 estimating an Ordered Generalized Additive Model (OGAM, Wood et al., 2016; Wood, 2017; Kammann and Wand, 2003; Lausted Veie and Panduro, 2015 for econometric applications) that allows to semiparametric specification of the effect of each observed tangible variable and 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 long-term vineyard quality. As they originate from natural processes, 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 intangible human determinants of GIs.

Consider that we only observe the realizations of a subset $X_i \subset X_i^*$ 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 that are described as having both direct and indirect effects on vineyard quality. 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 parametric ordered logit model that can be estimated by maximum likelihood:

$$\text{Prob}(y_i > j \mid X_i, C_i) = \Lambda[B(X_i)^\top \beta + C_i^\top \mu - \alpha_j], \tag{11}$$

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 C_i which work as *commune* fixed effects. In the absence of theoretical priors for the effects of all observed tangible variables X_i, we specify them through a series of functional transformations noted as $B(\cdot)$ with an associated vector of coefficients β. 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 $B(X_i) = [x_{1i}\ x_{1i}^2\ x_{2i}\ x_{2i}^2\ \cdots\ x_{Ki}\ x_{Ki}^2]$ with a set of $\widetilde{K} = 2 \times K$ coefficients to estimate.

Polynomial specifications will be shown empirically to 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 turned to semiparametric thin plate regression splines that have optimal smooth approximation properties (Wood, 2017). The matrix $B(X)$ 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 superfluous wiggliness. 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 sensible 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).

Goodness of fit measures provide information about the predictive abilities of estimated parameters but give little guidance on identifying the individual effects of the explanatory variables that are impacted by the degree of smoothing of geographic coordinates. To determine the sufficient complexity that allows control of unobserved spatial heterogeneity that correlates with *commune* delineations, we used the surrogate residuals recently defined by Liu and Zhang (2018) from auxiliary regressions that do not take into account *commune* fixed effects. Using residuals for specification purposes has a long history in econometrics, complemented by generalized residuals for non-linear outcomes (Pagan and Hall, 1983; Gourieroux et al., 1987; Chesher and Irish, 1987).

Define a surrogate variable $S \mid X, y \sim \lambda\left[\, B(X)^{\top}\beta - \alpha_y \mid y \,\right]$ that follows a truncated logistic distribution conditionally on the observed distribution of the vertical dimension of GIs. The principle of using the observed values of y to estimate the residuals is shared by generalized residuals, as the originality of the surrogate approach is to randomly draw the realizations rather than compute them analytically. This allows an estimation of the full distribution of model errors instead of only their first moments and sensibly extends the potential applications in regression diagnostics (Liu and Zhang, 2018). We obtained the residuals from N random draws of the surrogate variable S_i with:

$$R_i = S_i - \mathbb{E}(S_i) = S_i + \alpha_{y_i} - B(X_i)^{\top}\beta \tag{12}$$

and regress them on the *commune* fixed effects. This allows us to test the presence of correlated residual patterns after accounting only for tangible variables in the auxiliary regressions. By increasing the complexity of $B(X_i)$ through increasing spline base dimensions of the smooth functions of geographical coordinates, the joint significance of *commune* fixed effect decreases as the unobserved spatial patterns are increasingly accounted for. Failing to reject the null hypothesis of a Fisher test of joint significance of *commune* fixed effects is expected to determine that the sufficient complexity is attained by the auxiliary model. Once this is attained, we jointly estimate the effect of tangible and intangible GI determinants in a full OGAM with the obtained degree of spatial smoothing as in any parametric regression framework. In the absence of residual spatial effects correlating with *commune* dummies, the estimated ordinal superiority measures are unbiased. Note that the F-statistics are bootstrapped to take into account the additional uncertainty attributable to the random draws of surrogate residuals.

4 Results

4.1 Models of GI designation

The first column (0) of Table 1 reports the joint significance statistics from a standard ordered logit model with additive 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 (i.e., with the highest significance statistic) is the set of 31 *commune* dummies that represent the intangible determinants of human influence on GI designations. This set 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 positions 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 (data not shown). The non-linear effects of the three topographical variables on the latent quality index are reported in Figure 4 of the 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 linearly increases and southern exposition provides the highest marginal probability of a high GI classification (see the Replication Material, RM, from the link on the title page of this article). The top-left panel of Figure 5 in the 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. The coefficients estimated from *communes* fixed effects are interpreted in the next subsection in terms of ordinal superiority.

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 it appears from the effective degrees of freedom for spatial coordinates. The semiparametric structure of these models keeps 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 percent 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 4 in the Appendix shows the comparative advantage of OGAMs over parametric model (0) in estimating the marginal effects of each explanatory variable. Panel A of Figure 4 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 degrees range. These results are particularly stringent as these ranges concentrate the majority of vineyard plots. In terms of spatial smooth effects reported in Figure 5 of Appendix, OGAMs produce more detailed spatial

Table 1: **Joint Variable Significance for Ordered Models of 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

Notes: ** accounts for joint significance at 1% 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-orders 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, shrinked 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 surrogate residuals from auxiliary regressions without *commune* dummies.

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 GI designation scheme. The significance of *commune* fixed effects is affected by increasing the complexity of spatial smoothing, whereas it remains the second most important set of variable in model (V).

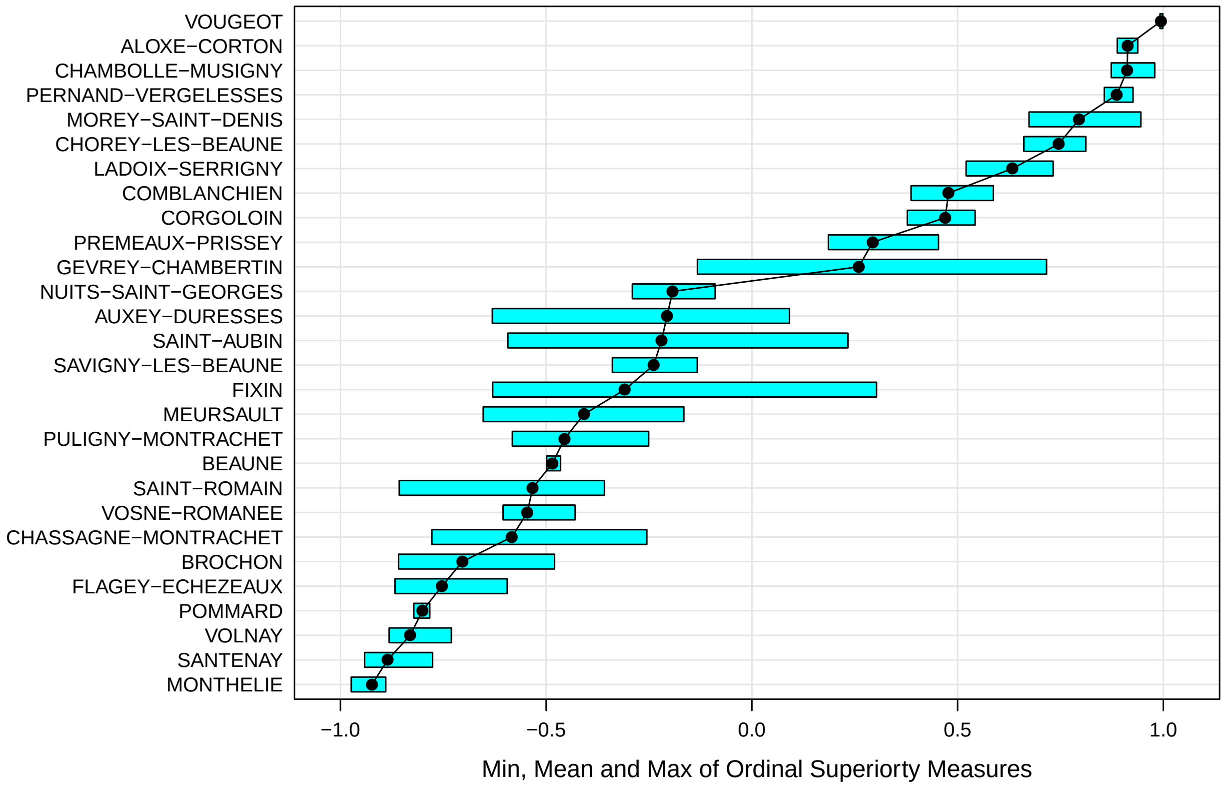
4.2 Ordinal superiority of *communes*

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. Figure 6 in the Appendix presents in more detail the relevance of smoothing spatial coordinates to control for unobserved *terroir* variables. Initially, it appears that OGAMs provide some important progress compared to the parametric ordered logistic model (0) from which surrogate residuals are highly correlated between *communes*. 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 statistics. This indicates persistent effects of intangible human-related characteristics on the GI designation scheme, even for precisely controlled *terroir* effects. Similar vineyard plots from one side or another of administrative borders have significantly different probabilities of being in different vertical levels of GIs.

Ordinal superiority measures from models with 700, 800 and 900 maximum edf are distributed inside the -1 and 1 interval in a way that a positive value indicates an advantage to the average and a negative value indicates a disadvantage (Agresti and Kateri, 2017). From Figure 2, only vineyard plots from four *communes* are not different from the average *commune* in terms of the designation of their vineyards along the vertical dimension of GIs. *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 11 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 own some vineyards in the *Côte de Nuits*, closer to Dijon than *Côte de Beaune* (Wolikow and Jacquet, 2018). The *communes* that have a syndicate engaged in collective action appear to be privileged,

Figure 2: **Ordinal superiorty measures for the *communes* in the actual GI designation scheme**

Notes: For a given *commune* c 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 all *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, whereas relatively disadvantged *communes* appear at the bottom.



but the separation is not clear-cut.[2] This hierarchy of advantaged and disadvantaged *communes* does not strictly follow their past or actual reputations, as some advantaged *communes* are not reputed (*Ladoix-Serrigny* and *Chorey-les-Beaune*), and some reputed *communes* are disadvantaged (*Flagey-Echezeaux* and *Pommard*). We found that the ordinal superiority measures only weakly positively correlated with average levels of actual GIs (R^2= 0.06, see Figure 7 in the Appendix).

[2] 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*.

4.3 Informational content of GIs

Table 2 reports the decomposition computed from equations (5) to (8) with $q(X_i^*)$ predicted from the five OGAMs reported in Table 1 with increasing complexity of spatial coordinates (the empirical formulas used are reported in RM Appendix A, jointly with the R code to compute them). As expected, the total signal reported in the first row of Table 2 increases from left to right and the total noise decreases.[3] 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 stationary between specifications. For all models, the vertical and horizontal dimensions of GIs have high information content. From the last column of Table 2, the joint signal of approximately 78% is 4-times higher than the average 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 taken into account, is higher than the horizontal signal when only using the horizontal dimension. This result reinforces the superiority of the vertical dimension to convey quality information, though the content is fewer items (5 instead of 31). From the vertical residual terms (i.e., 6th row), we see that the vertical dimension of GIs has approximately 20% (13/65) bias in conveying information about vineyard quality, and this bias due to preferential treatment between *communes* can be assessed by consumers through the horizontal dimension.

4.4 Alternative GI designation schemes

We estimate the same ordered models with the GIs of 1936 as outcomes. This year corresponds to the creation of the INAO. At this time, the vertical dimension of GIs counted only three levels, as reported in the summary statistics in Table 4 in the Appendix: *Régional* < *Village* < *Grand Cru*

[3]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.

Table 2: **Signal Decompositions from OGAMs with Spatial Coordinates**

		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

Notes: The effective degrees of freedom for spatial smoothing terms in parentheses show that the columns correspond to model (I) to (V) from Table 1. Decomposition terms are expressed in percent of variance of the latent variable y^* according to equations (5) to (8) 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*.

with respectively 57%, 41%, and 3% of actual vineyard plots.[4] The joint significance, the plots of the marginal effects and the spatial smooth effects are reported in Table 5, Figure 8 and Figure 10, respectively, in the Appendix. For these older GIs, the control for omitted *terroir* variables is reached for smaller maximum edf of spatial coordinates (bootstrapped F-statistics are reported in the bottom of Table 5 in Appendix, the violin plot is only reported in the RM, p.XX). Surprisingly, the hierarchy of the joint significance of explanatory variables is comparable to what is obtained for actual GIs. The *commune* and 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 close maximum values, and the spatial smoothed patterns are also very close to what is found with actual GIs. In contrast, the ordinal superiority measures are more varied between *communes* (see Figure 9 in the Appendix and Figure XX in RM, p.XX). For given *terroir* characteristics, the *commune* where a vineyard is

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

located was a more important determinant of GIs designations in the middle of the 20th century than currently. This is equivalent to claim 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 actual GIs with a joint signal to noise ratio close to 1 (48/49.5). However, the informational content of the vertical and horizontal dimensions are more balanced, whereas the vertical dimension stays 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 horizontal dimension. Because the *commune* delineations have not changed since 1936, the informational content of the horizontal dimension does not change. Note that the lower informational level of the vertical dimension of the GIs of 1936 is not due exclusively to the higher importance of *communes* in GI delineation (i.e., intangible determinants), an immeasurable part of the loss could be attributable to the lower number of vertical items (3 instead of 5 in current GIs). These results indicate significant improvements in the informational content of GIs in the last century in combination with the decreased bias from intangible determinants.

We performed different simulations of counter-factual GI designation schemes as reported in columns S.0 to S.VI in Table 3. The six vertical designation schemes were simulated by changing the underlying latent predictions of the quality index (in S.I, S.II and S.III), and by changing the number of vertical items in the GI schemes (in S.IV, S.V and S.VI). The detailed formula and R codes used to simulate alternative GI designation schemes are reported in RM (p.XX). We did not consider changing the horizontal dimension of GIs, because changing the administrative boundaries of *communes* 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 noised 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 very close to those obtained in the last column of Table 2. Next, we drop the designation errors from surrogate residuals in S.I, we drop the intangible

Table 3: **Signal Decompositions from Alternative GI Designations**

		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

Notes: 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 to mimic actual GIs. S.I drops the random idiosynchratic 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.

commune effects in S.II, and we drop both designation errors and *commune* effects in S.III. The last three designation schemes S.IV, S.V and S.VI represent an increase in the number of vertical items in GIs from the actual 5 to 6 items. The *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 item (like, e.g., *Bourgogne supérieur*, *Village supérieur* ou *Premier Cru supérieur*) that allows consumer to distinguish them from the wine labels.

The numbers reported in Table 3 show that dropping the intangible effects associated with *commune* effects is the most important policy to increase the informational content of the vertical dimension of GIs. Conversely, reducing the designation noise is more important for increasing the joint signal, which corresponds with 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 6th vertical level as in S.III, S.IV or S.V. Among these latter alternative schemes, we found that splitting the intermediate level *Village* is more efficient, but the differences are small. Note that the measure of the informational content that we propose is not directly related to the value of the GI information, as it treats symmetrically high and low levels of GIs. More research is needed to convert these results in terms of the efficient amount of information to give consumers, as information about high levels of GIs would be more valuable as wine becomes more expensive. In all cases, these potential improvements from the vertical dimension of GIs have no impact on the informational content of the horizontal dimension, which maintains the same order of magnitude among simulations.

5 Conclusion

We present a framework for modeling the geographical indications (GIs) and disentangle their informational content, i.e., their ability to describe the tangible characteristics of production sites. Applied to the wine-producing region of *Côte d'Or* (Burgundy, France), we found significant effects of both tangible and intangible characteristics of vineyard plots, described as stemming from natural and human outcomes, respectively. In particular, the presence of intangible effects is robust to the precise control of the omitted *terroir* variables under the assumption that, as a natural pattern, they vary smoothly in space. This implies that historical elements, such as the reputation of landowners or producers, their influence with the decision makers, or their collective actions, have some persistent effects on the GI designation scheme. Nevertheless, this differential treatment between administrative units decreases over the long-term with the continuous changes made in the scheme under the control of the INAO.

We interpret this result in regards to the dynamic analysis of strategic transmission information developed by Benabou and Laroque (1992). The authors show that, when information is not fully reliable (here, because of the human influence on the GI designation process), the possibility of honest mistakes (here, because of designation noise) produces some confusion for consumers. Consequently, market incentives keep the consumers' learning process incomplete and allow market manipulation. According to the authors, this negative economic outcome "is limited only in the long run by the public's constant reassessment of their credibility" (Benabou and Laroque, 1992, p.947). This analysis is particularly relevant for signaling wine quality, as we studied it here, knowing the difficulties defining and observing the notion of *terroir* and agreeing about the quality of wines, which worsens the problem of consumer verification of the relevance of the informational content of GIs.

These benefits of the long-term history of the informational content of GIs creates some doubt about the flexibility required to follow changing consumers' preferences and changing determinants of wine quality (particularly in the face of climate change, as argued by White et al., 2009). As a human institution, which requires the involvement of producers with private information about vineyard and wine quality, the unbiased nature of the GI signal would probably not be reached spontaneously following changes. Moreover, the regular modifications that would be required to follow the changing preferences or changing environment would increase the confusing correlation between tangible and intangible characteristics and, consequently, decrease the informational content of GIs for consumers. The stability of GIs and their third-party management probably account for a great portion of their value that is actually observed in the wine markets.

Our empirical strategy is based on the difference between the assumed spatial continuity of *terroir* and the discontinuity of administrative borders, 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 centroid geographic coordinates (longitude and latitude) allow the fine-scale variations of unobserved heterogeneity from the *terroir* to be controlled. The estimated spatial patterns of the latent quality index grasped by these functions are not exclusively related *a priori* to tangible characteristics that matter for wine quality, particularly, spatial interactions of reputation or influence between vineyard plots in close proximity on both sides of a *commune* border. Nevertheless, we found that the main decomposition results for the informational content of GIs and the general result of a signal-to-noise ratio of 4 are robust in regards to the degree of spatial smoothing used in regressions. Taking into account fine-scale variations in *terroir* is important to estimate the ordinal superiority measures but not determinant of the informational content of GIs.

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., Lecocq, S. and Visser, M. (2008). The impact of gurus: Parker grades and en primeur wine prices. *Economic Journal* 118: F158–F173.

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. (2008). Predicting the quality and prices of bordeaux wine. *Economic Journal* 118: F174–F184.

Ashenfelter, O. and Jones, G. V. (2013). The demand for expert opinion: Bordeaux wine. *Journal of Wine Economics* 8: 285–293.

Ashenfelter, O. and Quandt, R. (1999). Analyzing a wine tasting statistically. *Chance* 12: 16–20.

Bagwell, K. (2001). *The Economic Analysis of Advertising*. Edward Elgar Publishing.

Bederson, B. B., Jin, G. Z., Leslie, P., Quinn, A. J. and Zou, B. (2018). Incomplete disclosure: Evidence of signaling and countersignaling. *American Economic Journal: Microeconomics* 10: 41–66.

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

Besanko, D., Donnenfeld, S. and White, L. J. (1987). Monopoly and quality distortion: Effects and remedies. *Quarterly Journal of Economics* 102: 743–767.

Bodington, J. (2017). Disentangling wine judges' consensus, idiosyncratic, and random expressions of quality or preference. *Journal of Wine Economics* 12: 267–281.

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.

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.

Cao, J. and Stokes, L. (2010). Evaluation of wine judge performance through three characteristics: Bias, discrimination, and variation. *Journal of Wine Economics* 5: 132–142.

Cardebat, J.-M. and Paroissien, E. (2015). Standardizing expert wine scores: An application for bordeaux en primeur. *Journal of Wine Economics* 10: 329–348.

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.

Chesher, A. and Irish, M. (1987). Residual analysis in the grouped and censored normal linear model. *Journal of Econometrics* 34: 33–61.

Combris, P., Lecocq, S. and Visser, M. (1997). Estimation of a hedonic price equation for Bordeaux wine: Does quality matter? *Economic Journal* 107: 390–402.

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

Costanigro, M., McCluskey, J. J. and Goemans, C. (2010). The economics of nested names: Name specificity, reputations, and price premia. *American Journal of Agricultural Economics* 92: 1339–1350.

Costanigro, M., Scozzafava, G. and Casini, L. (2019). Vertical differentiation via multi-tier geographical indications and the consumer perception of quality: The case of Chianti wines. *Food Policy* .

Danguy, R. and Aubertin, C. (1892). *Les grands vins de Bourgogne (la Côte d'Or)*. Librairie H. Armand, Dijon.

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

Desquilbet, M. and Monier-Dilhan, S. (2014). Are geographical indications a worthy quality label? a framework with endogenous quality choice. *European Review of Agricultural Economics* 42: 129–150.

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

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: 205–212.

Friberg, R. and Grönqvist, E. (2012). Do expert reviews affect the demand for wine? *American Economic Journal: Applied Economics* 4: 193–211.

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. (2011). *Les* climats *du vignoble de Bourgogne comme patrimoine mondial de l'humanité*. Ed. Universitaires de Dijon.

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., 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.

Gourieroux, C., Monfort, A., Renault, E. and Trognon, A. (1987). Generalised residuals. *Journal of Econometrics* 34: 5–32.

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.

Hollander, A., Monier-Dilhan, S. and Ossard, H. (1999). Pleasures of cockaigne: Quality gaps, market structure, and the amount of grading. *American Journal of Agricultural Economics* 81: 501–511.

Jackson, D. and Lombard, P. (1993). Environmental and management practices affecting grape composition and wine quality-a review. *American Journal of Enology and Viticulture* 44: 409–430.

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.

Kammann, E. and Wand, M. P. (2003). Geoadditive models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 52: 1–18.

Knight, S., Klaere, S., Fedrizzi, B. and Goddard, M. R. (2015). Regional microbial signatures positively correlate with differential wine phenotypes: Evidence for a microbial aspect to terroir. *Scientific reports* 5.

Lausted Veie, K. and Panduro, T. E. (2015). An alternative to the standard spatial econometric approaches in hedonic house price models. *Land Economics* 91: 386–409.

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

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

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

Lizzeri, A. (1999). Information revelation and certification intermediaries. *The RAND Journal of Economics* : 214–231.

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

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

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

Pagan, A. R. and Hall, A. D. (1983). Diagnostic tests as residual analysis. *Econometric Reviews* 2: 159–218.

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.

Shapiro, C. (1986). Investment, moral hazard, and occupational licensing. *The Review of Economic Studies* 53: 843–862.

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

Storchmann, K. (2012). Wine economics. *Journal of Wine Economics* 7: 1–33.

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

Thompson, G. M. and Mutkoski, S. A. (2011). Reconsidering the 1855 bordeaux classification of the medoc and graves using wine ratings from 1970–2005. *Journal of Wine Economics* 6: 15–36.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies* 63: 1–22.

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

White, M. A., Whalen, P. and Jones, G. V. (2009). Land and wine. *Nature Geoscience* 2: 82.

Winfree, J. A. and McCluskey, J. J. (2005). Collective reputation and quality. *American Journal of Agricultural Economics* 87: 206–213.

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.

A Appendix

Table 4: **Descriptive Statistics for the Variables used in the Econometric Analysis**

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Acreage [1000 m^2]	59113	0.002	0.003	0.000	0.001	0.002	0.177
Elevation [1000 m]	59113	0.286	0.056	0.210	0.241	0.319	0.505
Slope [degree]	59113	5.772	5.478	0.000	1.556	8.747	36.970
Solar radiation [millions J]	59113	1.060	0.049	0.581	1.048	1.076	1.230
Longitude [degree]	59113	4.837	0.104	4.665	4.740	4.955	5.003
Latitude [degree]	59113	47.060	0.110	46.900	46.980	47.170	47.300
Actual GI [*Coteaux*]	59113	0.164	0.370	0	0	0	1
Actual GI [*Régional*]	59113	0.229	0.420	0	0	0	1
Actual GI [*Village*]	59113	0.428	0.495	0	0	1	1
Actual GI [*Premier Cru*]	59113	0.147	0.354	0	0	0	1
Actual GI [*Grand Cru*]	59113	0.032	0.177	0	0	0	1
1936 GI [*Régional*]	59113	0.565	0.496	0	0	1	1
1936 GI [*Village*]	59113	0.407	0.491	0	0	1	1
1936 GI [*Grand Cru*]	59113	0.027	0.163	0	0	0	1
Aspect [0 – 45]	59113	0.046	0.210	0	0	0	1
Aspect [45 – 90]	59113	0.186	0.389	0	0	0	1
Aspect [90 – 135]	59113	0.362	0.481	0	0	1	1
Aspect [135 – 180]	59113	0.212	0.409	0	0	0	1
Aspect [180 – 225]	59113	0.100	0.300	0	0	0	1
Aspect [225 – 270]	59113	0.044	0.206	0	0	0	1
Aspect [270 – 315]	59113	0.030	0.170	0	0	0	1
Aspect [315 – 360]	59113	0.021	0.142	0	0	0	1

Notes: Topographic variables were computed by a Geographical Information System from a Digital Elevation Model with 5 m resolution. Longitude and latitude correspond to the center of each vineyard plot. Current GI are dummy variables that account for the vertical dimension in 2018, and Past GI comes from the 1936 map. Aspect is discretized according to radian range reported in brakets.

Figure 3: **Distribution of the Vertical Dimension within and between the Horizontal Dimension**

Notes: For each *commune* on the y-axis (the horizontal dimension of GIs), the bars represent the cumulative vineyard area designated in each item of the vertical dimension represented with different colors. The number reported is the percentage the each vertical item represent in each *commune*.

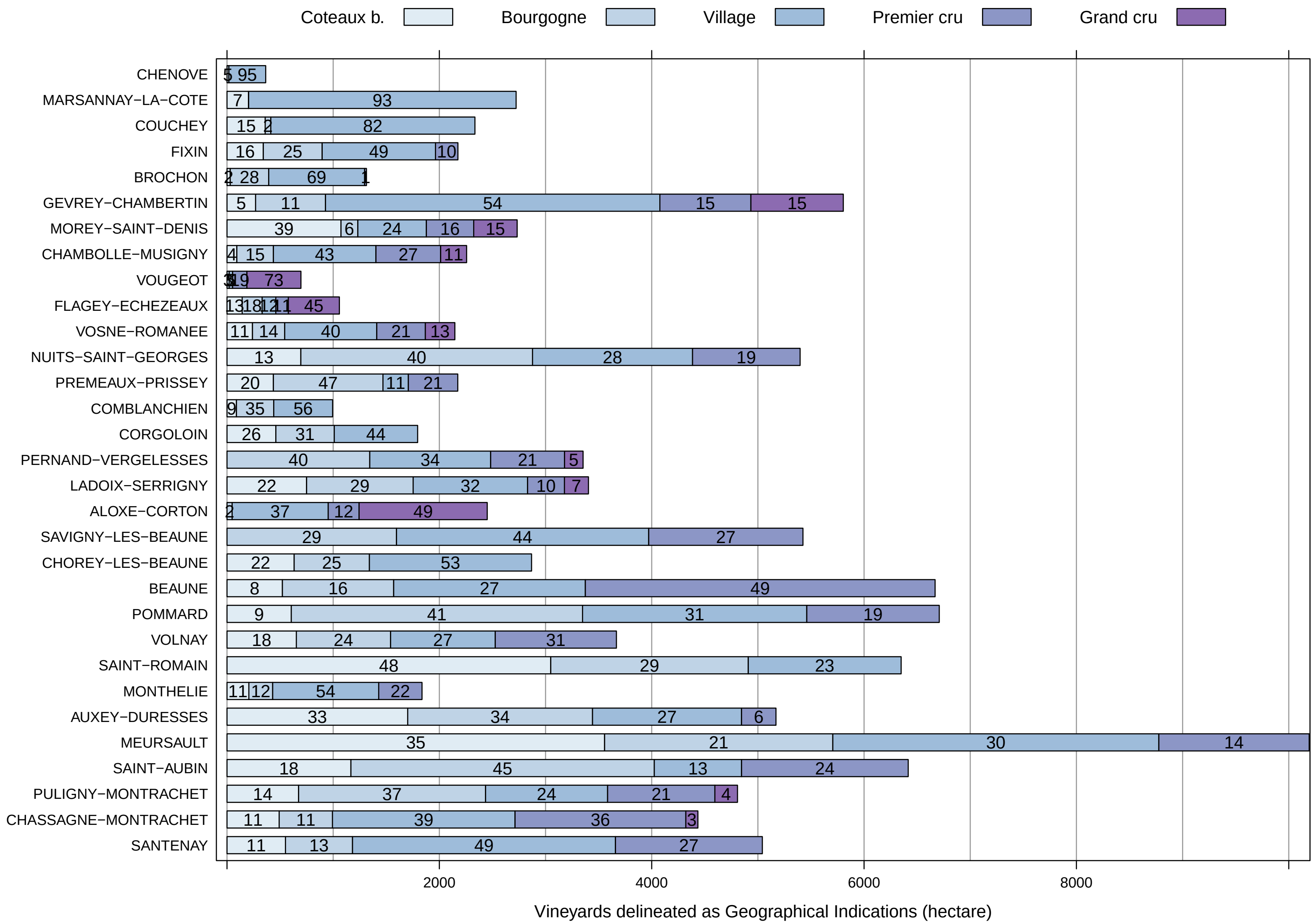


Figure 4: **Nonlinear Effects of Topographic Variables on GI Designations**

Notes: Dotted lines represent the quadratic effects from model (0) in Table 1, centered at zero with all other explanatory variables at their sample means. Continuous lines represent the centered effects from 10 OGAMs with increasing darkening for increasing effective degrees of freedom for spatial smoothing terms. Models (I) to (V) in Table 1 are a subset of these OGAMs with maximum effective degrees of freedom uniformly distributed between 100 and 1000. The histograms at the bottom of each plot represent the marginal distributions of each explanatory variable in the region.

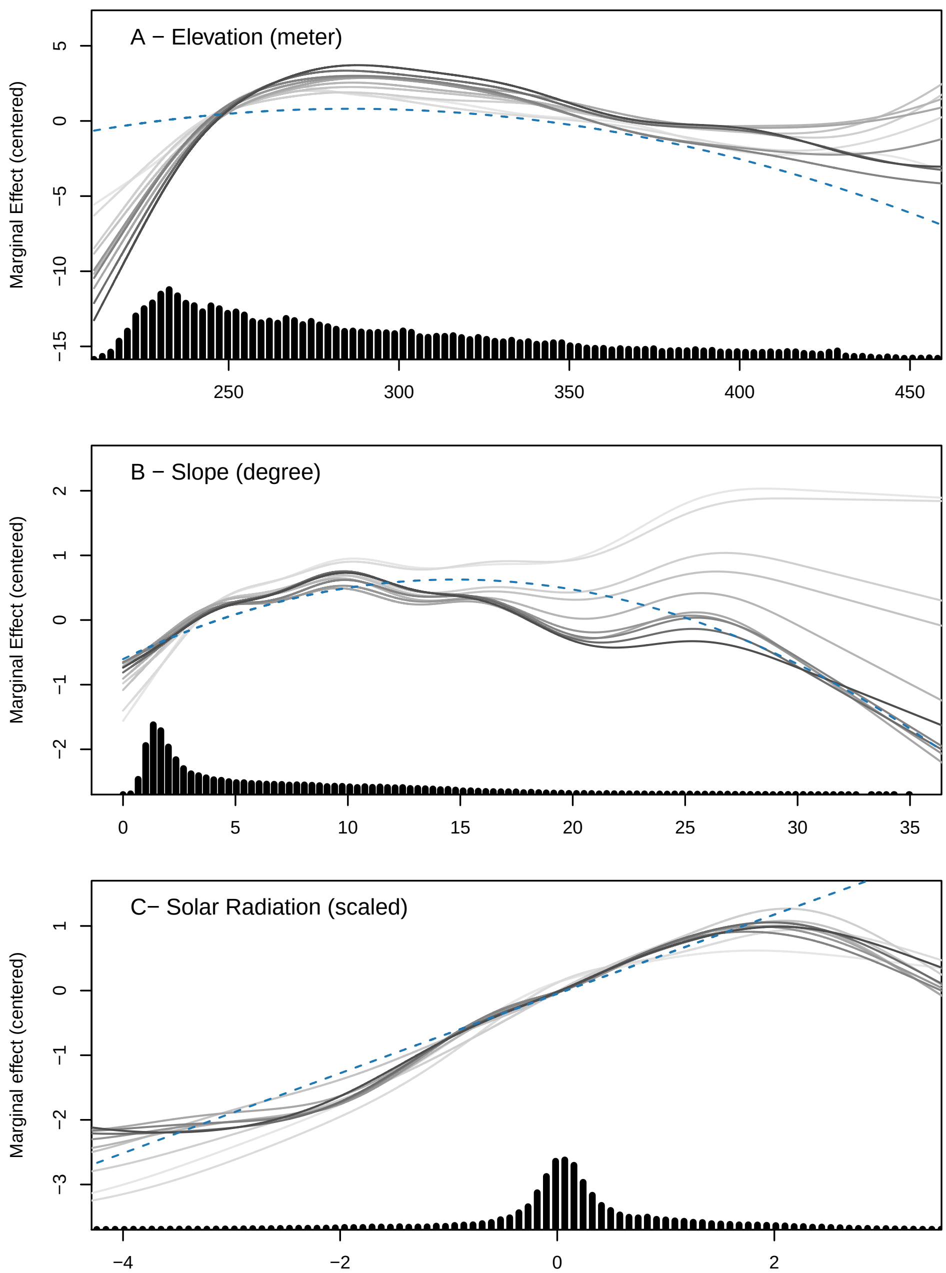


Figure 5: **Spatial Smoothed Effects from Ordered GI Designation Models**

Notes: The smooth surfaces are predicted from spatial coordinates with other explanatory variables at their sample means, with a uniform normalization to be inside the unit interval. A. The smooth prediction from parametric ordered logistic model (0) in Table 1. B to F. the prediction of OGAMs (I) to (V) with increasing effective degrees of freedom as reported at the top of each plot.

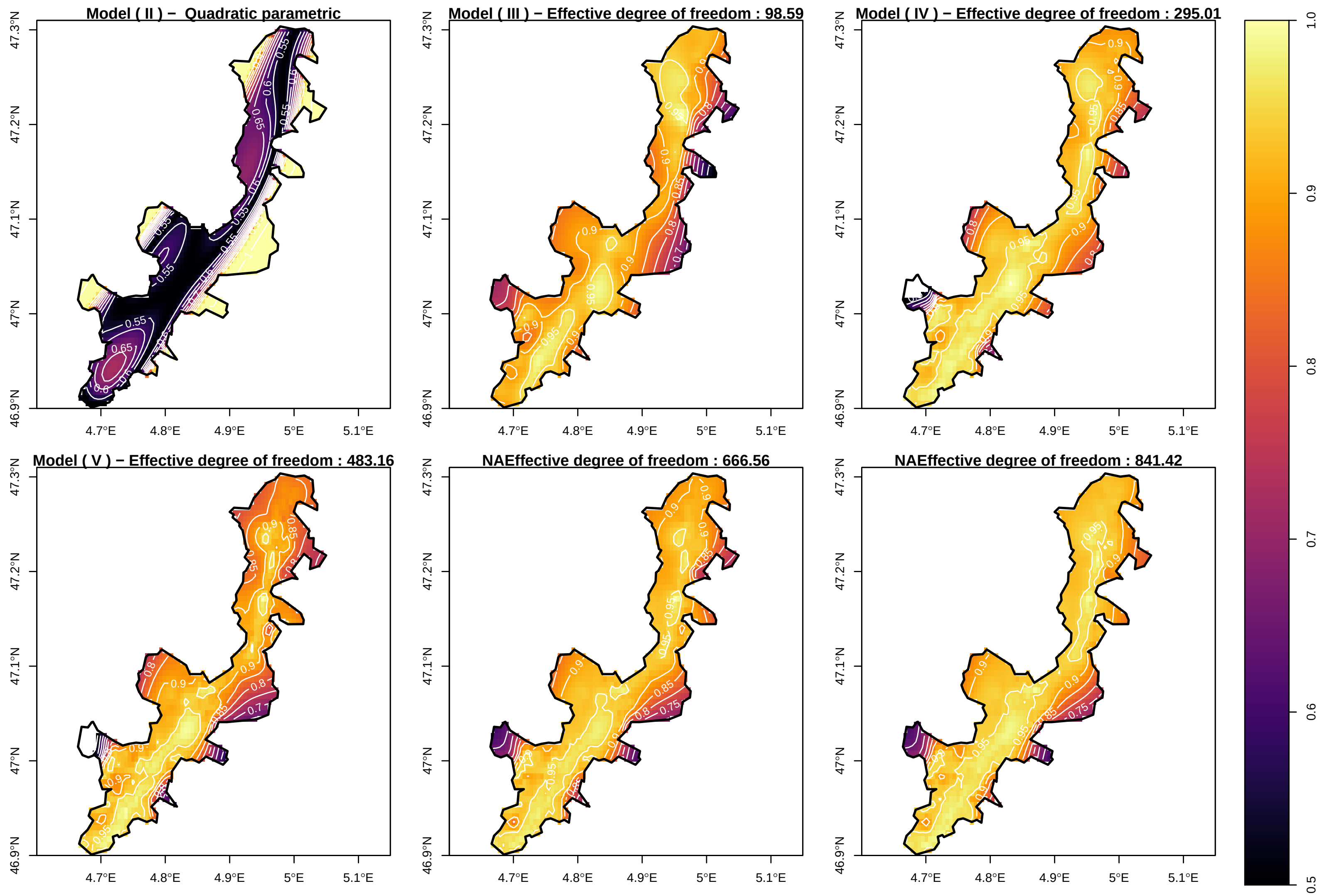


Figure 6: **F-statistics for Correlated Residual Effects**

Notes: For each model on the x-axis, the Figure reports the distribution of the bootstrapped F-statistics (log scale). Increasing the complexity of spatial effects from left to right decreases the significance of the *commune* effects on the surrogate residuals from auxiliary regressions.

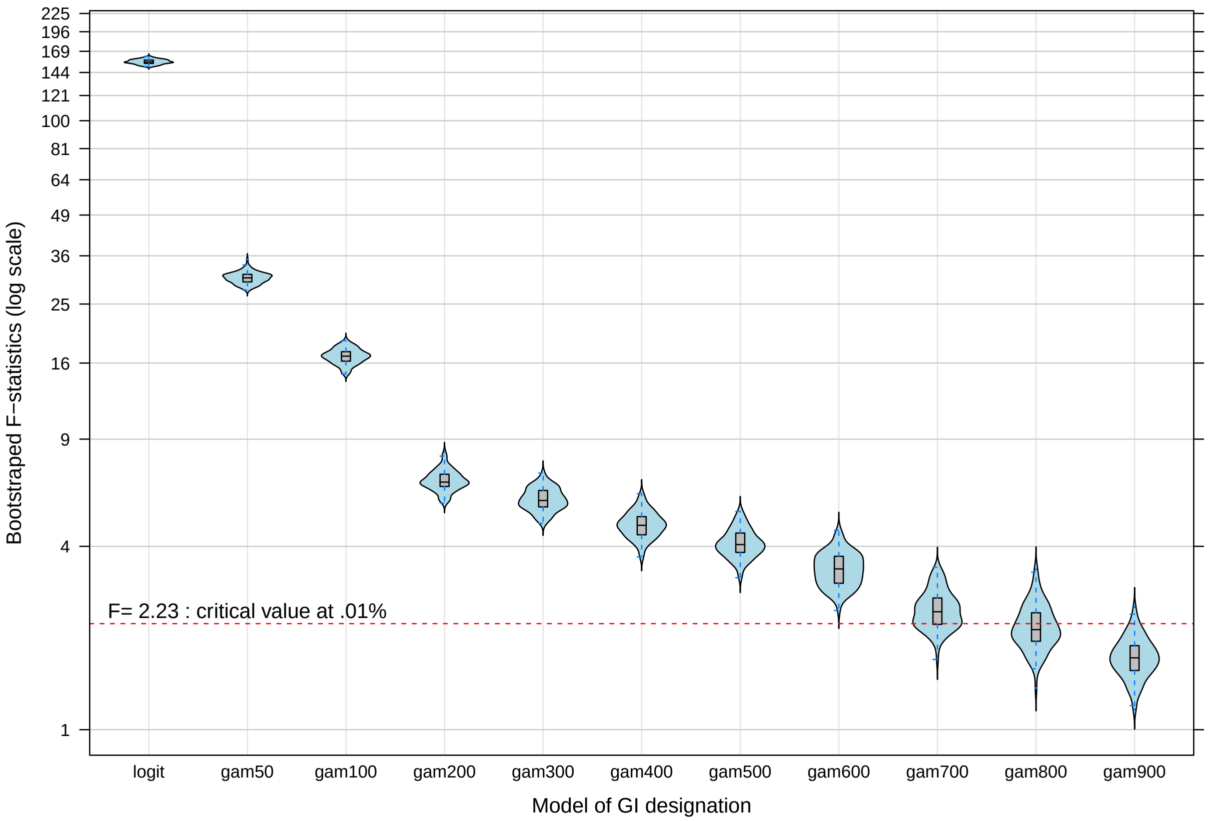


Figure 7: **Correlation between GI Grade and Ordinal Superiority Measures**

Notes: The ordinal superiority measures come from means in Figure 2. The average GI grade for each *commune* is the area-weighted mean of GIs coded from 1 to 5. Privileged *communes* (according to ordinal superiority measures) do not appear to have systematically more high GIs than average.

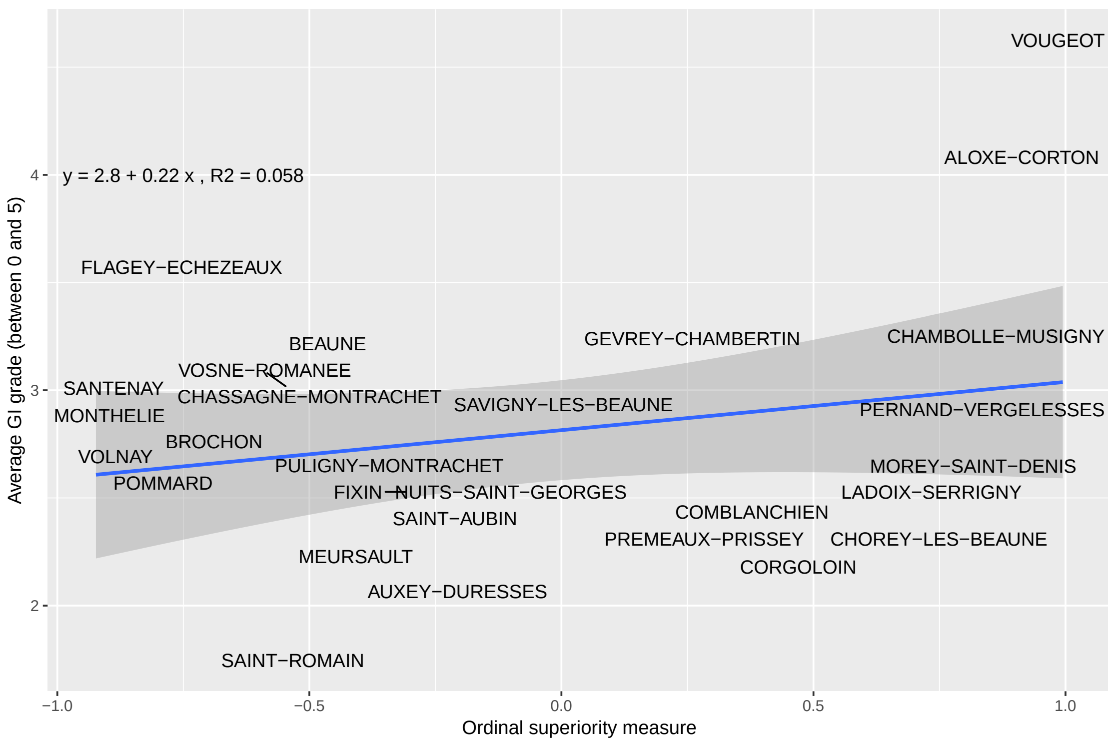


Table 5: **Joint Variable Significance for Ordered Models of the 1936 GI Designations**

Variable	(0)	(I)	(II)	(III)	(IV)	(V)
Elevation	982.42**	1 196.2**	197.72**	144.79**	265.02**	253.01**
	[2]	[8.826]	[7.628]	[8.232]	[8.659]	[7.42]
Slope	409.2**	478.13**	466.46**	297.06**	190.45**	169.07**
	[2]	[8.754]	[8.729]	[8.743]	[8.774]	[7.493]
Solar Radiation	859.1**	208.81**	139.42**	99.245**	87.676**	142.83**
	[2]	[8.04]	[1.082]	[8.114]	[7.419]	[7.425]
Spatial Coords	5 814.5**	6 760**	14 559**	17 285**	18 979**	20 906**
	[15]	[48.73]	[97.95]	[147.1]	[194.3]	[235.3]
Pedology	4 099.2**	2 820.6**	898.79**	599.37**	537.03**	539.28**
	[13]	[12]	[12]	[12]	[12]	[12]
Geology	982.42**	1 047**	692.13**	710.2**	585.81**	509.32**
	[14]	[14]	[14]	[14]	[14]	[14]
Exposition	287.18**	177.45**	131.87**	58.532**	43.002**	64.03**
	[7]	[7]	[7]	[7]	[7]	[7]
Commune	8 600.1**	3 720.9**	2 639.2**	2 177.2**	1 831.7**	1 264.7**
	[25]	[25]	[25]	[25]	[25]	[25]
Nb Observ.	50 000	50 000	50 000	50 000	50 000	50 000
McFadden R^2	44.63	49.68	61.32	66.06	69.82	72.36
Pc good pred.	81.86	83.74	87.88	89.84	91.35	92.21
Akaike IC	45	41.21	31.82	28.09	25.12	23.12
Surrogate F	92.72	8.45	5.4	3.43	2.75	2.03

Notes: ** accounts for joint significance at 1% from the reported chi-squared statistics. The 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 3-orders polynomials for longitude and latitude (edf= $3 + 3 + 3 \times 3 = 15$), with 7 and 25 dummy variables for exposition and *communes*, respectively. Five *communes* were dropped because they contained only one GI in 1935. Models (I) to (V) are OGAMs with elevation, slope and solar radiation additively specified with a maximum of 9 edf, shrinked endogenously by a quadratic penalization. Spatial coordinates are specified in increasing order of complexity with the maximum edf of 100, 150, 200, 250 and 300. The last row reports the average of bootstrapped Fisher statistics for the joint nullity of *commune* dummies on surrogate residuals.

Figure 8: **Non-linear effects of tangible variables in the 1936 GI designations**

Notes: Dotted lines represent the quadratic centered effects of model (0) presented in Table 5. Continuous lines represent the centered effects from OGAMs (I) to (V) with increasing darkening with increasing effective degrees of freedom. The histograms at the bottom of the plots represent the marginal distributions of each explanatory variable.

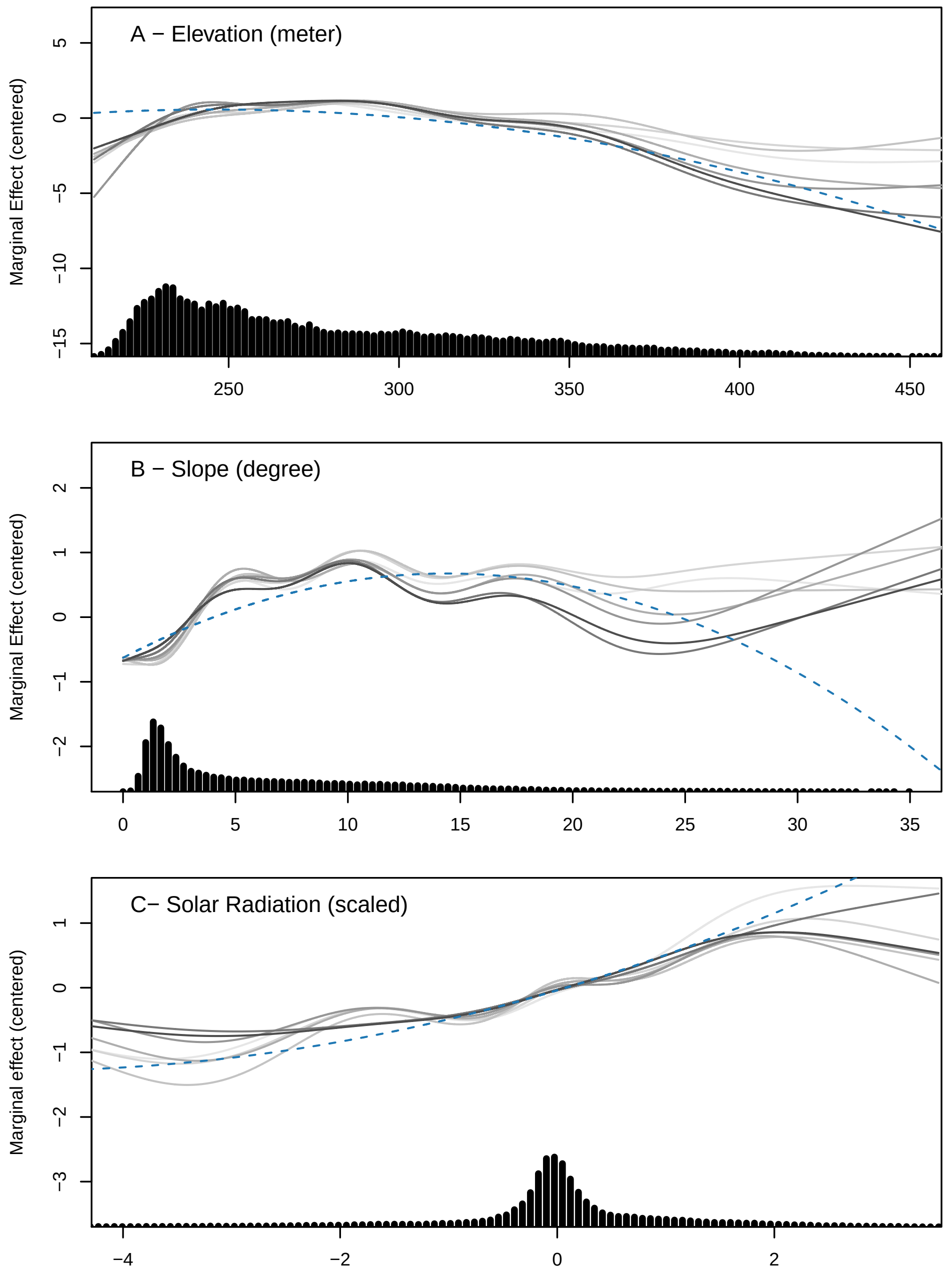


Figure 9: **Spatial Smoothed Effects from 1936 GI Designation Models**

Notes: The smooth surfaces are predicted from spatial coordinates with others explanatory variables at their sample means, with a uniform normalization to be inside the unit interval. Panel A displays the smooth prediction from parametric ordered logistic model (0) of Table 5. Panels B to F display the prediction from the OGAMs (I) to (V) with increasing effective degrees of freedom as reported at the top of each plot.

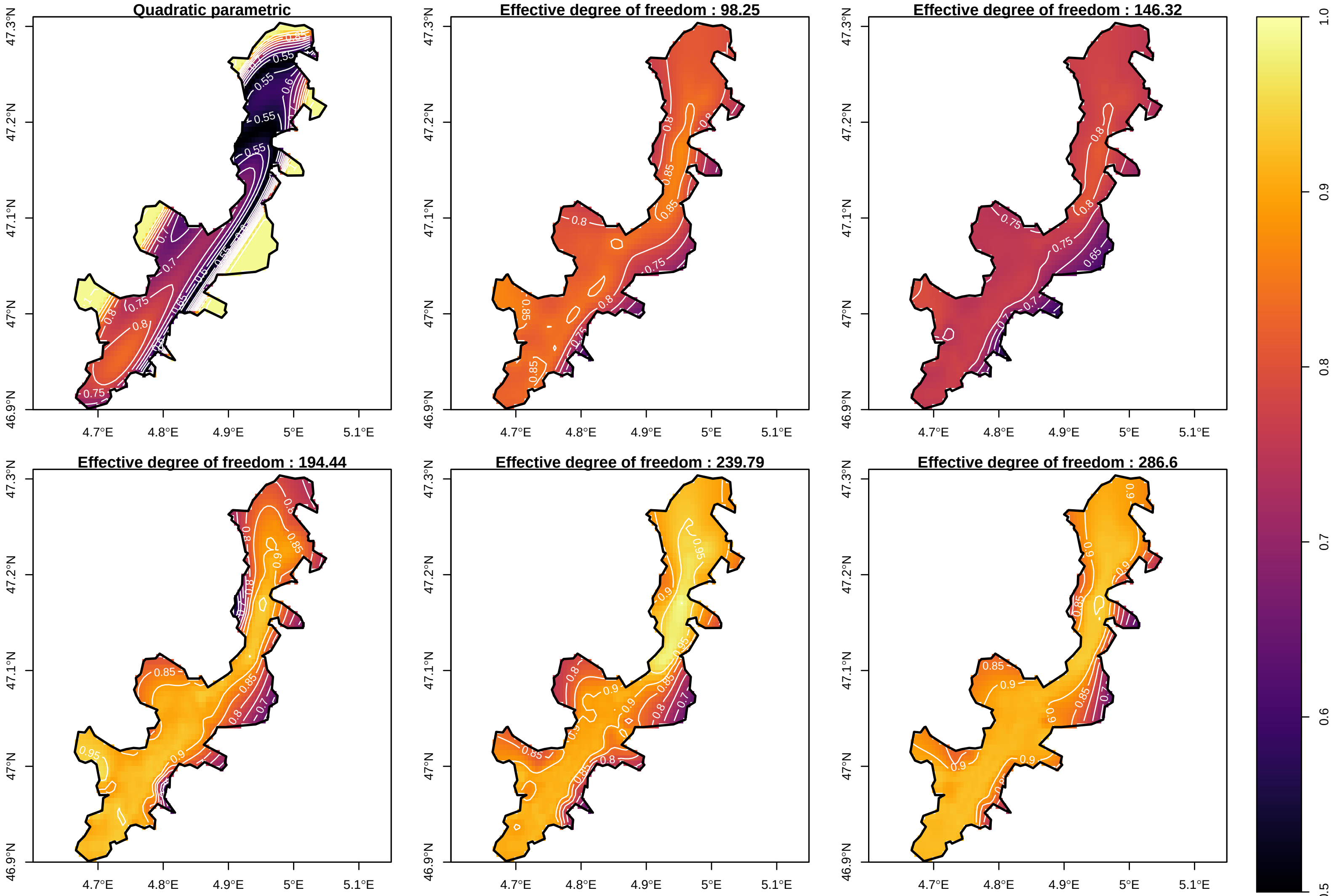


Figure 10: **Ordinal Superiority Measures for the *communes* in the 1936 GI designation scheme**

Notes: For a given *commune* c, ordinal superiority measures are computed from the difference between the own estimated fixed effect μ_c and the average fixed effect $\overline{\mu}$ according to: $\Delta_c = 2 \times \Lambda[(\mu_c - \overline{\mu})/\sqrt{2}] - 1$ as in the main text. The horizontal bars represent the range of measures according to different OGAMs with varying complexity for the effects of spatial coordinates, black dots represent the average of these measures. Relatively privilegied *communes* appear at the top of the Figure, whereas relatively disadvantged *communes* appear at the bottom.