wine economics

AMERICAN ASSOCIATION OF WINE ECONOMISTS

AAWE WORKING PAPER
No. 265
Economics

THE SIGNALING VALUES OF NESTED WINE NAMES

Jean-Sauveur Ay and Julie Le Gallo

May 2021

www.wine-economics.org

The Signaling Values of Nested Wine Names*

Jean-Sauveur AY Julie LE GALLO

UMR 1041 CESAER

AgroSup Dijon, INRAE, Université Bourgogne Franche-Comté

26 boulevard du Docteur Petitjean, 21000 Dijon, France

May 17, 2021

Abstract

Unobserved quality challenges the empirical content of signaling theory, and often precludes the valuation of quality signals such as wine names. This paper uses the location of vineyard plots to control for unobserved wine quality when estimating the causal value of wine names on vineyard prices. The identification tackles unobserved spatial heterogeneity by newly combining a multi-cutoff spatial regression discontinuity design with plausibly exogenous name variations. We deal with standard requirements of causal inference – unconfoundedness and overlap – with instrumental variables and high-dimensional propensity models in a double robust framework. For the Burgundy region of France, we then recover the full causal signaling scheme of nested wine names with both a horizontal and a vertical dimension. This typical structure of names is monotone and complementary, as the names are consistently ordered within each dimension (rank preservation) and they present spillovers between them (umbrella effect). We find a high importance for unobserved wine quality, which produces heterogeneous signaling values.

Keywords: Information disclosure ; unobserved *terroir* quality ; multi-dimensional signal ; spatial regression discontinuity ; double robustness ; overlap for causal inference.

J.E.L. Codes: C25,C26, C51, R33.

*Correspondance: jean-sauveur.ay@inrae.fr and julie.le-gallo@agrosupdijon.fr. The authors are grateful to the data providers: *SCAFR – Terre d'Europe* for vineyard sales, *Institut National de l'Origine et de la Qualité* for spatial delineations of wine names. This paper has benefited from discussions at the econometric seminars of AAWE (Mendoza), Beta (Strasbourg/ Nancy), Économie Publique (Paris), GREQAM (Marseille), SFER (Anger), VDQS (Colmar). We thank Andrew Plantinga for his encouragement in the initial stage of this work.

1 Introduction

When quality is not observable, two identical goods with different names can have systematically different prices. From the elementary signaling model (Spence, 2002),[1] the price difference is the signaling value of a name relatively to the other. This valuation is empirically challenging because it requires to control for unobserved quality, which is the typical reason for existing quality signals (Dranove and Jin, 2010). Ideal valuation design would consist in changing names of identical goods (Tyler et al., 2000; Clark and Martorell, 2014) in line with the counterfactual Rubin framework (Imbens and Rubin, 2015). The lack of such designs in the real world precludes the causal valuation of the various components of nested naming schemes, such as those used to designate wines.

We use the location of vineyard plots to estimate the causal signaling values of wine names on vineyard prices. The main identifying assumption is that wine quality impacts vineyard prices only from the natural attributes of vineyard plots. In line with *terroir* narratives,[2] these attributes vary smoothly between neighboring plots as nature does not make jumps (*natura non facit saltus*). This article provides new results both for causal inference with unobserved spatial heterogeneity and for the empirics of signaling theory. We show that exogenous name variations allow extrapolating spatial regression discontinuity design away from borders and that nested wine names are monotone and complement. As we will see, this indicates a consistent and relevant quality signaling.

The vineyard market is particularly suitable to the estimation of the signaling value of nested names. The strong dependence of vine cultivation on natural characteristics causes sharp variations in average wine quality over vintages, while this information cannot be assessed by consumers in most wine-purchasing situations. Whereas these natural sources of quality variations are exogenous by definition, they require precise biophysical data to be controlled adequately. The confounding influences of other inputs (such as labor and wine-making techniques) are implicit functions of *terroir*, as their prices are constant across vineyards (Cross et al., 2011). Consequently, the given supply of *terroir* impacts wine names and vineyard prices without endogenous quality feedback, from which we estimate the willingness to pay to use a wine name when buying a vineyard.

[1] Where quality does not depend on the level of investment in the signal, as it could be argued for our case study.

[2] We denote *terroir* the non-mobile, non-modifiable, and non-reproducible natural land attributes that are known to impact wine quality (Bokulich et al., 2014; Roullier-Gall et al., 2014; van Leeuwen et al., 2018).

The region under study also displays favorable characteristics for identification. Centuries of land divisions have produced a mosaic of vineyard plots of tiny size that are perfectly digitalized.[3] Intensive activity on the vineyard market during the last two decades generates fine-scale price variations for contiguous plots that are of similar *terroir* quality. Comparing vineyard prices on either side of a name border controls for unobserved wine quality and circumvents the traditional "ability bias" that restricts the empirical content of signaling theory.[4] Moreover, lobbying actions for the designation of wines during the 19th century provide plausibly exogenous name variations between administrative subdivisions. Used as instruments, they make distant plots with different names more overlapped in order to extrapolate heterogeneous signaling values away from borders. The constancy of name delineations through the 20th century solves potential simultaneity issues in relation to current reputations, preferences, and bargaining power (Malter, 2014). Coupled with the unidirectional *terroir* impacts previously described, this secures the identification of causal signaling values from vineyard price variations between similar plots with different names.

The names under study have a nested structure with both a horizontal and a vertical dimensions. They are related exclusively to the geographical origin of the grapes used to make the wines, according to two spatial delineations. The horizontal delineation is administrative, following the decree of 1789 after the French Revolution. This separation of *communes* (i.e., municipalities) based on the spatial distribution of churches built between the ninth and twelfth centuries, was not intended to signal wine quality (while it could be incidentally correlated with). The vertical delineation follows the classification of Professor Lavalle (1855) that ranked vineyards according to the quality of wines produced in the early 19th century. While the author used the same four-level classification for each *commune* in the region, he refused to produce a classification between them.[5] Despite this, the classification paved the way for lobbying efforts to modify the hierarchy. Jacquet (2009) and Ay (2021) show that this ended in systematic biases in favor of certain *communes* that have larger areas of high vertical names than warranted by their *terroir* endowment. This process was frozen in 1934 with the creation of the *Appellations d'Origine Contrôlée* nationally.

[3]We study about 60 000 plots with a wine name, for a total acreage of 115 km^2 (0.2 ha of average plot size).

[4]This term comes from seminal studies about estimating the returns to education (Griliches, 1977; Card, 2001). The exogeneity of the natural attributes behind *terroir* makes the identification easier, although the method we propose cannot be used to discriminate between endogenous quality and signaling theories (Huntington-Klein, 2020).

[5]Lavalle (1855) wrote: "I have studied the wines of each *commune* as if the others had not existed and the classification that I give is true only for each *commune* taken in isolation" (p.162, our translation).

Nowadays, wine producers have a legal obligation to reference their wines both by one of the 34 horizontal names from *commune* delineations and one of the five vertical names from the official hierarchical classification.[6] This typical nested structure allows us to investigate the value added by bi-dimensional multi-valued signals in a general causal framework. We are able to estimate and decompose the signaling value of each combination of names and to test two shape restrictions of general interest. We define monotonicity within each name dimension (i.e., does the hierarchy of names is preserved among the dimensions?) and show that it is equivalent to a rank condition (Chetverikov et al., 2018). As such, monotone names provide consistent quality signals whatever the combinations they lie in the nested scheme. We also define complementarity between name dimensions (i.e., does the value of one dimension spill over the value of the other dimension?) as a causal generalization of umbrella branding (Wernerfelt, 1988). Accordingly, complementary nested names are relevant to learn about quality, as they provide more information than each dimension of names taken in isolation. The price decomposition we propose allocates the total signaling values to the horizontal and vertical dimensions from counterfactual price comparisons.

Our first set of results is derived from a hedonic model under the assumption that fine-scale biophysical variables present in our data account fully for *terroir* quality. Within this framework, we find that the values of the biophysical variables decreases sharply when vertical names are included in the regressions. This suggests a strong dependence between these two groups of variables, while the vertical names ultimately present the highest price premiums. Conversely, the inclusion of horizontal names does not impact the value of biophysical variables. By including interactions between names, we find that the hierarchy of horizontal names is preserved between vertical names. This indicates a consistent quality disclosure in line with the monotonicity condition. We also find that horizontal name premiums are higher for *communes* with higher vertical names. This denotes the presence of umbrella effects, according to which names are complementary. Maintaining the hedonic assumption, we obtain a total signaling value of (horizontal and vertical) names for the Burgundy region of about €1.8 billions (in 2017), representing about €162 000 by hectare on average. These high values correspond to about six years of gross margin for the regional wine production and about 65% of the average per-ha vineyard price, respectively.

[6]The wine labels reported in Figure A.1 of the Online Appendix (OA) provide some examples about the use of these names by one Burgundian producer, more details about their structure are given in the data section.

A second part of evidence follows the presentation of a bi-dimensional multi-valued causal framework considering *terroir* variables as only partially observed in the data. The full signaling scheme is defined through more than 90 average treatment effects on the treated, from which we derive the two shape restrictions and the price decomposition. Under the assumption that *terroir* produces smooth geographical variations in wine quality (formally stated as spatial continuity), nested wine names offer a multi-cutoff spatial regression discontinuity (SRD) design (Keele and Titiunik, 2015; Cattaneo et al., 2016) on vineyard prices. This allows us to exploit the proximity between vineyard plots to control for *terroir* quality. We find that the first hedonic results, by misleadingly attributing the value of unobserved *terroir* to vertical names, overestimate their signaling values by a factor of two compared to what is found on each side of the borders between names. This result shows the high economic importance of wine quality that is unobserved from data by the econometrician. SRD estimations do not recover the monotonicity and complementarity of wine names, suggesting that they were due to unobserved *terroir* quality bias in the hedonic evidence. However, SRD estimates are local statistics that require strong restrictions to be extrapolated away from the borders (Angrist and Rokkanen, 2015; Bertanha and Imbens, 2019).

Our preferred third set of doubly robust (DR) results (Robins et al., 1994; Słoczyński and Wooldridge, 2018) take the best of the two previous approaches. We use the ordered model of name designations with lobbying effects presented by Ay (2021) for the same area. We estimate generalized propensity scores (Imbens, 2000) with a high-dimensional specification of geographical coordinates (Belloni et al., 2014; Wood et al., 2016). Still under the spatial continuity assumption, the high spatial density of the population of vineyard plots allows us to precisely account for *terroir* with penalized regressions in the propensity model. We derive formally the DR weights that exploit the historical name variations from lobbying as instrumental variables for the identification of the full signaling scheme. Estimated DR causal signaling values are similar to those from SRD at the borders of names, while they are significantly higher away from them. By combining the internal validity of the SRD approach (unconfoundedness) and the external validity of the hedonic approach (overlap), the DR approach produces signaling values that are between the two. We recover monotonicity and complementarity of wine names on the area, but with less significance than in the hedonic results. The total signaling value is revised downward but stayed high, at about €1.3 billions regionally, corresponding to €115 000 by ha (45% of average vineyard price).

The outline of the rest of the article is as follows. We highlight our contribution to the literature in the following Section 2. We present the context, data, and hedonic evidence in Section 3. Section 4 contains the causal framework used to define the signaling values. Section 5 presents the multi-cutoff spatial regression discontinuity evidence, and Section 6 reports our preferred doubly robust evidence from the historical variations. Section 7 concludes.

2 Contributions

This paper presents an observational case study that is conducive to estimate the signaling values of names, an issue that has given rise to a vast theoretical literature (Shapiro, 1983; Erdem and Swait, 1998; Tadelis, 1999; Neeman et al., 2019). This kind of empirical result is scarce because of the difficulty in controlling for unobserved quality (McDevitt, 2014; Bronnenberg et al., 2015 are two exceptions, see Bronnenberg et al., 2019 for a review). We estimate the signaling values of wine names as average treatment effects on the treated in the counter-factual framework. They are the difference between the value that plots of a given name have and the value they would have if they were named differently (see Graetz, 2017; Aryal et al., 2021 for alternative definitions in other contexts). The bi-dimensional multi-valued structure of our causal framework defines monotone quality signals that can be consistently ordered by their signaling values. This shape restriction is well know from signaling theory as a fundamental assumption (Milgrom, 1981; Athey and Levin, 2018). We verify the monotonicity from Spearman correlation coefficients for both the vertical and horizontal dimensions of names, showing their consistency as quality signals. This result is particularly relevant as some recent papers have shown that complex signals could be non monotone in some situations (Olszewski and Wolinsky, 2016; Currarini et al., 2020).

Burgundy's vineyards allow us to recover and decompose the full signaling schemes of nested names. This structure is common for consumption goods that nest a brand name and a series number (Dyson V8, Samsung Galaxy S20, and Windows 10 for instance), while it also concerns many other signals (a high-school diploma nests both an establishment and a grade, the former is horizontal and the latter is vertical). A vast literature on information theory has questioned the complementarity of nested signals (Börgers et al., 2013; Zhu and Dukes, 2017; Fong et al., 2019) although comprehensive empirical evidence is still lacking. We assess the complementarity of wine names by Kendall correlation coefficients and provide a decomposition of the total signaling value between both dimensions. We find the two dimensions of wine names are complementary, highlighting the relevance of this structure for information transmission (Krishna and Morgan, 2001; Liang and Mu, 2020). This result of complement names confirms the presence of umbrella effects between names from the causal framework (Wernerfelt, 1988; Hakenes and Peitz, 2009).[7]

We also contribute to the hedonic literature that aims to estimate collective values from private transactions on the land market (Starrett, 1981; Kanemoto, 1988; Bishop et al., 2020). The signaling values defined as average treatment effects on the treated are also capitalization measures (Kuminoff and Pope, 2014). We estimate them from cross-sectional price variations, thanks to the exogeneity of *terroir* quality and the spatial continuity assumption. Accounting for endogenous supply and unobserved heterogeneity are the two main challenges in turning hedonic theory to the data (see Brown and Rosen, 1982; Nerlove, 1995; Abbott and Klaiber, 2011; Kuminoff et al., 2013). We provide a detailed comparison between ordinary least squares, regression discontinuity design, and instrumented doubly robust methods. We are able to make sense of the different results thanks to a common framework. We show that classical hedonic estimations and regression discontinuity designs do not recover causal signaling values with unobserved heterogeneity. Our preferred approach is based on a DR method that exploits the spatial continuity more generally than in SRD. We show how to use exogenous variations of the treatments to predict more overlapped propensity scores under unconfoundedness. This strategy is generally relevant to exploit the spatial continuity assumption to deal with spatial heterogeneity for causal inference (Keele and Titiunik, 2016; Michalopoulos and Papaioannou, 2018).

This paper provides a new identification result, according to which an exclusion restriction increases overlap for DR estimates of heterogeneous treatment effects under unconfoundedness. While the credibility of unconfoundedness increases with the dimension of pre-treatment variables, this comes at a cost in terms of overlap between different treatment statuses (Khan and Tamer, 2010; D'Amour et al., 2021), especially for a large number of treatments (Flores and Mitnik,

[7]The principle of an umbrella effect is particularly relevant for Burgundy wines as, in the middle of twentieth century, 13 *communes* decided to join their most famous vertical wine name to their horizontal name. For instance, the *commune* of *Gevrey* changed its name to *Gevrey-Chambertin* in 1946. More anecdotally, Mullainathan et al. (2008) used Burgundy wines as an exemple of the link betwen coarse thinking and umbrella branding (p.601–604).

2013). Precisely controlling for unobserved spatial heterogeneity through penalized regressions with high-dimensional spline transformations of geographical coordinates (Wood et al., 2016), we show that *commune* membership provides exogenous variations in vertical names. This makes plots with different vertical names to have closer predicted propensity scores. In other words, horizontal *communes* are used as instrumental variables and are assigned counter-factually to conduct causal inference (Holland, 1986; Ichimura and Taber, 2001). This result is of particular interest where high-dimensional methods with a high number of variables or transformations of variables are used for the propensity score. They often almost perfectly predict the treatment status without securing enough overlap (Belloni et al., 2014; Athey and Imbens, 2019). In particular, it contrasts with the usual recommendation of not including instrumental variables as pre-treatment variables without unconfoundedness (Bhattacharya and Vogt, 2007; Wooldridge, 2016).

Finally, this paper contributes to the wine economics literature by studying the long-standing question of the dependence of wine quality on natural conditions (Combris et al., 1997; Gergaud and Ginsburgh, 2008; White et al., 2009; Ashenfelter and Storchmann, 2010; Cross et al., 2011).[8] Separating natural from man-made determinants of wine quality is a recurrent concern in determining the effective supply constraints involved in wine production. This is central to disentangle the virtuous goal of signaling quality from the production of undeserved rents captured by landowners because of an artificially-reduced supply. This indeterminacy leads to intense debates in multilateral trade negotiations over the recognition of geographical indications (Josling, 2006; Meloni and Swinnen, 2018). Our results illustrate the high economic importance of *terroir* quality that is observed neither by the econometrician from the usual data nor by the consumer who does not taste a wine before buying it. We document the relevance of long-standing nested geographical indications from an economic point of view, as a reliable quality signal for wine markets (Costanigro et al., 2010; Yu et al., 2017; Mérel et al., 2020). Moreover, we find that wine quality revealed from vineyard prices for a given combination of names is positively related to the *terroir* quality that underlies the designation of the vertical names. As such, the accumulated historical knowledge about wine quality from names' designations is still relevant nowadays, despite evolving preferences, production technologies, and the globalization that characterized wine markets in the last decades.

[8] Adam Smith adressed this question in "An Inquiry into the Nature and Causes of the Wealth of Nations" 1776, Edwin Cannan (ed.), 5th ed. London 1904, Book 1, Chap XI, p. 219. Most of this literature, in social and natural sciences, assumes that *terroir* quality effect is observable from biophysical, chemical, or sensory wine characteritics.

3 Hedonic evidence

The area under study is a strip of about 5 km from West to East and 60 km from North to South, located between the cities of *Paris* and *Lyon* in the *Côte d'Or* region of France (Figure A.2 in the OA). Nested wine names indicate the vineyard plots from which the grapes come[9] by combining one horizontal *commune* name among 34 and one vertical name among five levels: *Coteaux*, *Région*, *Village*, *Premier cru*, and *Grand cru* (in increasing order of quality).

3.1 Data

We use both exhaustive population data about the 60 000 vineyard plots with an official wine name in the studied area and a quasi-exhaustive sample of 9 000 plots sold over the period 1992–2017.[10] We merge each vineyard plot in the population with the best disaggregated available data about land topography (5-meter resolution) and computed average elevation, slope, solar radiation, and exposition for each plot. Other data about climate, subsoil, and soil characteristics are available for the area, but are not used due to their coarse spatial resolution.[11]

Table A.1 in the OA reports summary statistics for both the population of plots and the sample of sales. Vineyards are sold at an average price of 350 000 euros/ha, which is about 50 times more than the average price of nearby farmland (Ay et al., 2012). Importantly, the distributions of biophysical variables are very similar in the two data sets. Unreported analysis find that the sample of sales is representative of the population for conditional vineyard prices.[12] Each combination

[9] To use an official wine name, a wine producer must also comply with a set of constraints on yields, alcohool content, and plant varieties (for the most significant ones). These contraints vary only slightly over the studied area. For instance, there is only one red grape variety *Pinot Noir* and one white grape variety *Chardonnay* allowed (except for the lowest level, the *Coteaux* where *Gamay* and some other negligible varieties are allowed). Maximum yields for a *Régional* red wine are 60 hl/ha with a minimum alcohol content 10.5%, but 50 hl/ha and 11% for a red wine *Village* of *Gevrey-Chambertin*. Aggregate statistics show these constraints have not been binding in the last decades and we therefore conjectured that they do not cause much vineyard price variations and focused on names only.

[10] In France, there is a legal obligation to declare each land sale to the *Safer* office from which we obtained the data. Due to a preliminary period at the beginning of the 90's, the vineyard sales data becomes exhaustive only around 1997.

[11] The best disaggregated historical climate data from *Météo-France* are available at an 8-km resolution, which is not relevant given the tiny size of vineyard plots and the shape of the studied area. Soil and subsoil variables are available from coarse punctual observations interpolated with topography variables already controlled in econometric models.

[12] Like Cross et al. (2017), we estimate various sample selection models without obtaining significant Mill's ratios. Therefore, we consider in the rest of the article that sale data are a random sample drawn from the plot population.

of horizontal and vertical names that exists in the population is present in the sample, with an average sampling intensity of 15%. Some combinations of names do not exist in the population nor, consequently, in the sample: two *communes* do not have a *Premier Cru* level (*Marsannay* and *Côte-de-Nuits-Village*) and only seven *communes* have a *Grand Cru* level (*Gevrey-Chambertin*, *Morey-Saint-Denis*, *Chambolle-Musigny*, *Vosne-Romanée-Vougeot*, *Aloxe-Corton-Ladoix*, *Puligny-Montrachet*, *Chassagne-Montrachet*, from North to South).[13] It is worth mentioning that only the wines from the three highest vertical levels are allowed to put the corresponding horizontal name on wine labels. Although all the plots in the two lowest levels also belong to an administrative *commune*, it is forbidden to put its name on wine labels (see Figure A.1 in OA).

3.2 Hedonic regressions

We first run a series of OLS regressions by exploiting the availability of biophysical variables at a fine spatial resolution. We noted y_i the natural logarithm of per-square-meter vineyard prices for the whole sample of plots sold $i = 1, 2, \dots, N$.[14] For a given observation i, we bind in the vector $\mathbf{z}_i$ an intercept and the reported sale characteristics (year of sale, acreage, tenure status, occupation, type of seller, and type of buyer). Another vector $\mathbf{x}_i$ bind third-order polynomials of each biophysical variable (elevation, slope, solar radiation, and exposition). The vectors $\mathbf{d}_i$ and $\mathbf{c}_i$ code for the vertical and horizontal names through respectively 4 and 18 dummy variables. A last vector $\mathbf{a}_i \equiv \mathbf{d}_i \otimes \mathbf{c}_i$ of dimension $4 \times 18 = 72$ is the Kronecker product of the two name dimensions to include the interaction between names in the following linear model:

$$y_i = \mathbf{z}_i^\top \boldsymbol{\theta} + \mathbf{x}_i^\top \boldsymbol{\beta} + \mathbf{d}_i^\top \boldsymbol{\delta} + \mathbf{c}_i^\top \gamma + \mathbf{a}_i^\top \boldsymbol{\lambda} + \varepsilon_i. \tag{1}$$

The matrix formed by stacking the row vectors $\mathbf{a}_i^\top$ is not full rank. Indeed, 14 interaction variables take a constant zero value because the corresponding combination of names do not exist in the data

[13] As detailed in the notes below Table A.1 in the OA, we grouped some contiguous *communes* together to simplify the borders used in the identification strategy, without loss of generality. We obtained 19 horizontal wine names from the 31 initial administrative *communes* by merging contiguous *communes* with similar horizontal name premiums.

[14] To deal with multi-plot sales, each price observation was inversely weighted by the number of plot in the sale and, in all the regression results presented in this article, standard errors were clustered at the sale level. This leads to very similar coefficients and standard errors to those obtained by limiting the sample to the 2 329 sales with only one vineyard plot (this latter sub-sample accounts for 57% of the 4 054 total vineyard sales).

(see bottom panel of Table A.1 in OA). These terms are dropped from the explanatory variables to reach identification. We also estimate some restricted versions of this model, by setting some other coefficients to zero. We consider models without $\mathbf{x}_i$, $\mathbf{d}_i$, $\mathbf{c}_i$, and $\mathbf{a}_i$ respectively to evaluate the marginal contribution of each group of variables to explaining vineyard prices. Finally, we estimate a model with more parsimonious interactions, by limiting $\mathbf{a}_i$ to be equal to $[d_{2i}, d_{3i}] \otimes \mathbf{c}_i$. This choice is justified by the legal prohibition on *Coteaux* and *Région* wines ($j = 0, 1$) putting the horizontal names on labels, and by the small number of observations for *Grand Cru* ($j = 4$).

3.3 Hedonic results

Table 1 displays partial R^2 statistics on the relative importance of each group of variables in the various versions of Equation 1. The full R^2 at the bottom of columns (0), (1), and (2) show that, individually, vertical names are the most significant, followed by biophysical characteristics, and horizontal names. However, this hierarchy is considerably modified when the variables are introduced jointly. Model (3) shows that including vertical names and biophysical variables jointly decreases both marginal contributions, where the latter are more impacted (however, they stay significant at 5% from a joint F test). This indicates a high correlation between these two groups of variables, while the four dummies for vertical names remain the largest part of the explanatory power. Model (4) shows a different pattern for horizontal names, as both partial R^2 increase with the inclusion of biophysical variables. The partial R^2 for horizontal names is also stable with the inclusion of vertical names in model (5), whereas the marginal explanatory power of biophysical variables decreases again (again, they stay significant at 5% from a F test). The last column (6) shows that interactions between names have a partial R^2 of 0.12, for a full R^2 of more than 0.80.

The estimated coefficients for control variables, biophysical variables, and name values are reported in the OA. Table A.2 shows that vineyard price negatively depends on plot acreages (with an elasticity of about −0.1), occupied plots at the moment of sale are 25% more expensive, tenured plots are less expensive, and buyers' characteristics are generally more significant than those of the sellers.[15] Figure A.3 shows a sharp increase in average wine prices over the last decade, a

[15]We have tested other sets of buyers' and sellers' characteristics (age, numbers of persons, municipality of origin) without obtaining significant coefficients or significant interactions with names premiums.

Table 1: **Partial R^2 from Hedonic Models of Vineyard Prices** ($N = 8\,987$)**.**

	Outcome variable: natural logarithm of per-hectare vineyard price						
Groups of variables	(0)	(1)	(2)	(3)	(4)	(5)	(6)
Controls	0.23	0.19	0.24	0.19	0.22	0.18	0.15
Biophysical variables	0.21			0.06	0.22	0.03	0.03
Vertical names		0.48		0.25		0.21	0.00
Horizontal names			0.11		0.16	0.17	0.00
Interaction btw names							0.12
Full R^2	0.46	0.71	0.37	0.73	0.56	0.78	0.81

Notes: The partial R^2 statistics represent the marginal increase in explanatory power following the introduction of each group of variables in succession, with other reported groups already included. The statistic of 0.21 in the second row is the R^2 obtained by regressing the residuals from a model with only control variables on biophysical variables. These partial statistics do not sum to the full R^2 in columns because of partial correlations between groups of variables. The main effects for horizontal and vertical names are redundant with their interactions, so their partial R^2 are zero by definition in column (6). The details of the estimated coefficients are reported in Table A.3 of the OA.

coarse linear pattern for the vertical values (on average, the per-ha vineyard price is multiplied by two from a vertical level to the following), and a less predictable pattern for the nevertheless significant horizontal values. Figure A.4 shows a U-inverted marginal effect for elevation, while the effects of slope and solar radiation are more monotonically increasing. This non-linear structure of observed *terroir* effects is consistent with those obtained for other wine regions in other countries (Ashenfelter and Storchmann, 2010; Cross et al., 2011).

The full set of estimated name values is reported in the Table A.3 of the OA. The vertical and horizontal values are stable between specifications. Figure A.5 in the OA displays the value of wine names from the hedonic model with parsimonious interactions (i.e., with $\mathbf{a}_i = [d_{2i}, d_{3i}] \otimes \mathbf{c}_i$ in Equation 1). An average plot from the lowest vertical *Coteaux* level has a predicted per-ha price of €20 000, which is less than half the predicted price for the second-lowest *Région* level. More strikingly, an average plot from the highest *Grand Cru* level has a predicted per-ha price 80 times higher than the lowest *Coteaux* level. Within the intermediate *Village* level, horizontal names provide price variations up to a factor of 10 (between a *Marsannay Village* predicted at €75 000 and *Puligny-Montrachet Village* predicted at €750 000). Horizontal values appear closely correlated between *Village* and *Premier cru* levels, which suggests similar horizontal values that are included within the signaling values of vertical names. Figure A.6a and Figure A.6b in the OA provide similar evidence for the hedonic model with full interactions.

4 Causal framework

Previous hedonic evidences has been based on the assumption that the observed biophysical variables fully control for *terroir* effects, a critical assumption regarding the complex processes underlying this concept.[16] However, two important points stem from this prior analysis: (i) strong correlations exist between biophysical variables and vertical names, and (ii) strong interactions exist between horizontal and vertical signaling values. These are the two main motivations for the bi-dimensional multi-valued causal framework with partially-observed *terroir* that is now presented. This allows us to define causal signaling values, shape restrictions, and a price decomposition.

4.1 Price equation

We consider observed vineyard prices to be generated according to a counter-factual mapping between the three main groups of variables considered above.[17] We note y the price variable, X^* the full set of partially-observed *terroir* variables, $d \in \{0, 1, \ldots, J\}$ the vertical names from the lowest to the highest level, $c \in \{0, 1, \ldots, K\}$ the horizontal *commune* names from the northernmost to the southernmost (see Figure A.15 of OA), and ε the errors with:[18]

$$y = h(X^*, d, c) + \varepsilon, \quad \text{with} \quad \mathbb{E}(\varepsilon \mid X^*, d, c) = 0. \tag{2}$$

We assume additive errors with conditional mean independence to focus on the three determinants of vineyard price inside the unspecified price function. This model is fairly general as the full set of *terroir* variables X^* is not assumed to be observed in the data. We consider that we observe only a subset $X \subset X^*$ of them, indicating that some aspects of the *terroir* that impact vineyard price through wine quality are not available to the econometrician.[19] The effects of vertical and horizontal names d and c on price are unrestricted, they can be arbitrarily correlated with the *terroir* X^*.

[16] The precise linear proxy assumptions that are sufficient to interpret previous hedonic evidence as causal according to the framework of this section are presented in Section A1.1 in OA.

[17] In the rest of this section, we omit control variables Z of sale characteristics without loss of generality. They are reintroduced additively in the empirical part of the paper.

[18] Henceforth, all these terms are considered as random variables. Univariate variables are noted in lower case, multivariate variables are noted in upper case. Realizations of these variables are indexed by i.

[19] These *terroir* variables are nevertheless observed in the field by informed persons, such as buyers and sellers.

We assume that the causal hedonic function h is twice continuously differentiable in X^* for each value of d and c, and that y has finite moments. These are the classical regularity conditions of the semiparametric setting (Robinson, 1988; Li and Racine, 2007; Cattaneo, 2010). In contrast, the discrete nature of the random variables d and c produces discontinuous jumps of the hedonic price function that are taken into account in the identification strategy we propose. A more implicit restriction in Equation 2 is the stable unit treatment value assumption, according to which the price of one plot is not affected by the names of the other plots (Holland, 1986).

4.2 Incremental name premiums

This causal model of vineyard price is closely related to the potential outcome framework used in program impact evaluation (Wooldridge, 2010). The only differences concern the additivity of mean-independent errors ε and the specification of individual heterogeneity exclusively through the partially-observed *terroir* variables X^*. In particular, this framework presents a bi-dimensional multi-valued treatment structure with unspecified heterogeneous treatment effects. For a given vineyard plot characterized by (X_i^*, ε_i), its counter-factual price for any vertical name $j \neq d_i$ and any horizontal name $k \neq c_i$ is $h(X_i^*, j, k) + \varepsilon_i$. This implies $(J+1) \times (K+1) - 1$ counter-factual prices (94 in our case) for each vineyard plot, one for each alternative combination of wine names.

A high number of causal statistics can be defined from this counterfactual framework, but some are of more particular interest. Firstly, we note $\delta_{jk}(X^*) \equiv h(X^*; j, k) - h(X^*; j-1, k)$ the individual incremental vertical premiums and $\gamma_{jk}(X^*) \equiv h(X^*; j, k) - h(X^*; j, k-1)$ the individual incremental horizontal premiums, each defined for $j = 1, \ldots, J$ and $k = 1, \ldots, K$ for a given X^*. They represent respectively the causal value added by changing the vertical name from $j-1$ to j and by changing its horizontal name from $k-1$ to k.[20] Secondly, we defined two aggregate measures of these individual incremental premiums by averaging them across plots according to the current designation of names. These latter statistics are well-known average treatment effects on the treated:

$$\delta_{jk} \equiv \mathbb{E}[\delta_{jk}(X^*) \mid d = j, c = k] \quad \text{and} \quad \gamma_{jk} \equiv \mathbb{E}[\gamma_{jk}(X^*) \mid d = j, c = k]. \tag{3}$$

[20]This incremental structure is more immediate for the vertical names indexed by j than for the horizontal names indexed by k. Nevertheless, Figure A.15 in the OA shows that, because of the spatial contiguity, all the bilateral horizontal premiums can be recovered additively from the individual incremental horizontal premiums defined here.

The expectation operators are taken from the distribution of X^* conditionally on the particular values of d and c. There is a slight abuse of notations in Equation 3, as the indices j and k of the statistics represent both the individual incremental causal premium of interest (following the expectation operators) and the vineyard plots that are averaged (within the conditioning sets). This notation significantly decrease the number of treatment effects under consideration, while keeping them consistent with the capitalization literature by focusing on treatment effects on the treated (Kuminoff and Pope, 2014). As $\mathbb{E}[h(X^*; j, k) \mid d = j, c = k]$ is directly estimable from a random sample of sales, the critical counterfactual estimands are $\mathbb{E}[h(X^*; j-1, k) \mid d = j, c = k]$ and $\mathbb{E}[h(X^*; j, k-1) \mid d = j, c = k]$. Their estimation requires the computation of the average price that vineyards of vertical name j within the *commune* k would have if they were named as $j-1$ without changing k, and the average price that the same plots would have is they were named as $k-1$ without changing j. The spatial proximity between $j-1$ and j, and $k-1$ and k (Figure A.2 in the OA) suggests that, among the full set of average treatment effects on the treated defined, these incremental causal estimands δ_{jk} and γ_{jk} compare neighboring plots, which are expected to have the most similar *terroir* under spatial continuity of the partially-observed variables X^*.

Thirdly, we aggregate the latter statistics by averaging them respectively on all horizontal and vertical names. This give J average incremental vertical premiums $AIVP_j$ and K average incremental horizontal premiums $AIHP_k$ on which we base our main identification results.

$$AIVP_j \equiv \mathbb{E}[h(X^*; j, c) - h(X^*; j-1, c) \mid d = j] = w_{j\cdot}^{-1} \textstyle\sum_{k=0}^{K} w_{jk} \cdot \delta_{jk}, \tag{4}$$

$$AIHP_k \equiv \mathbb{E}[h(X^*; d, k) - h(X^*; d, k-1) \mid c = k] = w_{\cdot k}^{-1} \textstyle\sum_{j=0}^{J} w_{jk} \cdot \gamma_{jk}. \tag{5}$$

The expectation operator in $AIVP_j$ is taken from the joint distribution of X^* and c, conditionally on $d = j$ (and symmetrically for $AIHP_k$). The term w_{jk} is the population acreage of plots with both the vertical name j and the horizontal name k, $w_{j\cdot}$ is for plots with the vertical name j (whatever the horizontal name), and $w_{\cdot k}$ for plots with the horizontal name k (whatever the vertical name). For all plots currently designated as j, the $AIVP_j$ statistic is the average signaling value they have from being designated j instead of $j-1$, for their given distribution of horizontal names (and symmetrically for $AIHP_k$). Because they average individual signaling values according to the joint distribution of d and c, these statistics include the interaction between wine names.

4.3 Shape restrictions

This rich set of treatment effects allows us to define and test shape restrictions of general economic interest, in order to characterize the distribution of signaling values between horizontal and vertical names more extensively. Due to the partially-observed nature of X^*, we focused on aggregate shape restrictions that are also conditional moment inequalities (Chetverikov et al., 2018).

Monotonicity. The first shape restriction concerns the consistency of quality signaling, according to which the ordering of names is stable across dimensions. We define monotonicity separately for horizontal and vertical dimensions, and begin with the vertical one that is more intuitive given its hierarchical structure. The vertical name j is monotone relatively to $j' \neq j$ if:

$$\mathbb{E}[h(X^*; j, k) - h(X^*; j', k) \mid d = j, c = k] \leqslant 0, \quad \text{for all} \quad k = 0, \ldots, K. \tag{6}$$

Vertical monotonicity implies that a higher name has more value within each *commune*, while a lower name could be more valued between *communes*. For each couple (j', j), the average value added by j is either positive or negative for all horizontal names, while we expect it is positive if $j' < j$ and negative otherwise. With $J = 4$, there are $J(J + 1)/2 = 10$ bilateral comparisons between all (j', j) pairs for each $K + 1 = 19$ horizontal names. Hence, the vertical monotonicity is assessed from a total of 190 bilateral comparisons. Figure A.6a in OA shows that vertical monotonicity is verified from the hedonic evidence (we observe only one gap between the *Coteaux* and *Région* levels in *Auxey-Duresse-Saint-Romain*, see Section A.1.1 in the OA for the formal derivation of monotonicity for the hedonic model).

Symmetrically, the horizontal name k is monotone relatively to $k' \neq k$ if:

$$\mathbb{E}[h(X^*; j, k) - h(X^*; j, k') \mid d = j, c = k] \leqslant 0, \quad \text{for all} \quad j = 0, \ldots, J. \tag{7}$$

Monotone horizontal names present the same ordering of the bilateral signaling values across all the vertical names. For our application with $K = 18$, there were $K(K+1)/2 = 171$ bilateral comparisons between k and k' for each of the five vertical names, which give a total of 855 comparisons. Figure A.6b in the OA shows that, because of the high number of horizontal names, monotonicity is less easily assessed visually. We test it by bilateral Spearman's rank correlation coefficients between the horizontal signaling values of each vertical level, as reported in Table 2. Horizontal monotonicity has a strong support from the hedonic evidence, more particularly between the vertical levels *Village* and *Premier cru* and between the *Coteaux* and *Region*. Horizontal names present an implicit ordering that is generally preserved between all the vertical names.

Table 2: **Spearman's Rank Correlation between Horizontal Values for all Vertical Levels.**

	Coteaux	*Region*	*Village*	*Premier cru*
Region	0.868			
Village	0.688	0.833		
Premier cru	0.587	0.637	0.902	
Grand cru	0.699	0.773	0.806	0.746

Notes: Spearman's rank correlation coefficients are statistical measures of monotonicity, they are computed between the horizontal name values reported in Figure A.6b in the OA, for each pair of vertical levels. The Spearman's coefficients are all significant at 95%. The details of their computations are reported in Section A1.1 of the OA.

Complementarity. The second set of shape restriction concerns the presence of spillovers between the two dimensions of names. Accordingly, vertical premiums of complementary names are higher for higher horizontal premiums (and inversely). The combination (j, k) of names is complementary relatively to the combination (j', k') with $j' < j$ and $k' < k$ if:[21]

$$\mathbb{E}[h(X_i^*, j, k) - h(X_i^*, j', k) \mid d = j, c = k] > \mathbb{E}[h(X_i^*, j, k') - h(X_i^*, j', k') \mid d = j, c = k]. \tag{8}$$

For the $(J + 1) \times (K + 1) = 95$ combinations of vertical and horizontal names in our case study, Equation 9 consists in 4 465 bilateral combinations that cannot be summarized easily. We assessed the complementarity of wine names by the Kendall rank correlation coefficients between vertical and horizontal name premiums. This is a classical measure for the concordance of signs between the two sides of Equation 9. We obtained a significant value of 0.67 from the hedonic results, which suggests that wine names are complementary in two out of three cases. Figure A.7 in the OA shows more intuitively that vertical premiums spill over the horizontal names. Horizontal names that have more acreages with high vertical names have higher signaling values. This can be interpreted as the presence of umbrella effects under the linear proxy assumptions of A.1.1 in the OA.

[21] There is a slight abuse of notation in $k' < k$, where we implicitely supposed a monotone ranking of horizontal names. If the names are not monotone (vertically or horizontally), this definition of complementarity is not relevant.

4.4 Price decomposition

Finally, we decompose the total signaling value of the wine names into a vertical part, a horizontal part, and an interaction part. As such, we estimate the respective contributions of each name dimension in the full signaling scheme. Such detailed decompositions are not unique because they are based on the arbitrary choice of the reference modality (Fortin et al., 2011).[22] We define this reference modality as the least valuable combination (j^*, k^*) of names without loss of generality. By noting Π the total signaling value of all wine names on the area, we have:

$$\begin{aligned}
\Pi &\equiv \Sigma_{j=0}^{J}\Sigma_{k=0}^{K}\, w_{jk}\ \mathbb{E}[h(X^*_; j,k) - h(X^*_; j^*,k^*) \mid d=j, c=k] \\
&= \Sigma_{j=0}^{J}\, w_{j\cdot}\ \mathbb{E}[h(X^*_; j,c) - h(X^*_; j^*,c) \mid d=j] \\
&\quad + \Sigma_{k=0}^{K}\, w_{\cdot k}\ \mathbb{E}[h(X^*_; d,k) - h(X^*_; d,k^*) \mid c=k] \\
&\quad + \Sigma_{j=0}^{J}\,\Sigma_{k=0}^{K}\, w_{jk}\ \mathbb{E}[h(X^*_; j^*,k) + h(X^*_; j,k^*) - h(X^*_; j,k) - h(X^*_; j^*,k^*) \mid d=j, c=k].
\end{aligned} \tag{9}$$

The first equality is definitional, the second equality contains three additive terms: the part of the value attributed to vertical names, the part attributed to horizontal names, and the interaction between them. One can verify that all terms of the vertical (horizontal) part are positive if the vertical (horizontal) names are monotone, and the third part is negative for complementary names. Because the vertical and horizontal parts are averaged according to the joint distribution of c and d, they count twice the interactions between names that are subtracted in the third term. The two first terms can be decomposed once more by separating the value of each in summed terms. From hedonic results, the total signaling value is about €1.8 billions, representing about €160 000 per hectare on average (65% of the average price). The vertical part represents 80% (€130 000 per ha), the horizontal part 35% (€55 000 per ha), and the interaction part −15% (€25 000 per ha). The *Village* level is the greatest contributor to the vertical part, followed by *Premier cru*, *Région*, and *Grand cru* (see A.1.1 of OA for the details). The relative contributions of the horizontal names are more balanced because the heterogeneity of horizontal values is less marked. The interaction part is negative because the wine names are both monotonic and complementary.

[22] Using the terminology of Fortin et al. (2011), we are interested in the detailed decomposition of the structural (or unexplained) effects, which are shown to be average treatment effects on the treated by the authors.

5 Regression discontinuity evidence

This section presents the multi-cutoff spatial regression discontinuity (SRD) framework applied to the geographic delineations of wine names. The normalized distances between vineyard plots and borders are used as running variables in order to identify local incremental signaling values.

5.1 Spatial structure of delineations

The spatial delineations of horizontal and vertical wine names produce a complex web of borders (see Figure A.8 in the OA). We count about 300 km of borders between the five vertical levels and about 50 km of borders between the 19 horizontal names. The spatial structure of wine names is such that horizontal borders are globally West-East oriented and vertical borders are mostly North-South oriented (see also Figure 1 below and Figure A.9 in the OA).

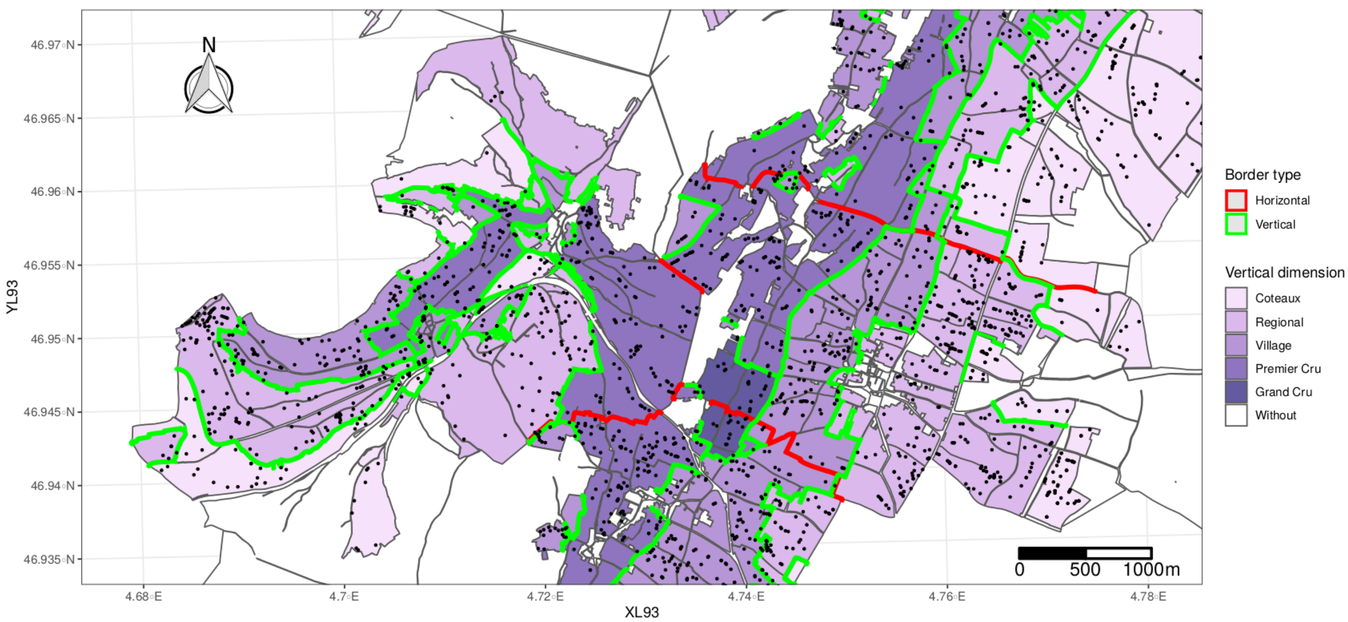
We noted V the bi-dimensional random vector of geographic coordinates (i.e., longitude and latitude). $\ell_{j|k}(V)$ is the euclidean distance between the plot at V and the closest point of the border that separates the vertical names $j-1$ and j in the *commune* k.[23] We note $\ell'_{k|j}(V)$ the distance between the plot at V and the closest point of the border between the horizontal names $k-1$ and k within the vertical level j. As the vertical levels j and $j-1$ are overwhelmingly contiguous and the k index is ordered from North to South, these incremental borders represent more the 98% of all borders (see Figure 1 for one of the few counter-examples). Hence, the centroid V of each plot with a given combination of names (j, k) is uniquely defined by the four distances to its closest incremental borders: $\ell_{j-1|k}(V)$, $\ell_{j|k}(V)$, $\ell'_{k-1|j}(V)$, and $\ell'_{k|j}(V)$.

To derive the incremental vertical premiums of *commune* k, we considered the subset of plots within the horizontal name $c = k$ with a vertical name $d \in \{j-1, j\}$. These plots present two different vertical names for a same horizontal name, and are geographically separated by only one vertical incremental border where $\ell_{j|k}(V)$ is the distance to it. Plots designated as $d = j-1$ are on one side and plots designated as $d = j$ are on the other side, within the same *commune* (see Figure 1). The SRD approach consists in normalizing these distances to the incremental border

[23]For a given plot $V_i = [v_{1i}, v_{2i}]$, we have $\ell_{j|k}(V_i) = \sqrt{(v_{1i} - v_{1(j|k)})^2 + (v_{2i} - v_{2(j|k)})^2}$ where $\mathbf{v}_{j|k} \equiv [v_{1(j|k)}, v_{2(j|k)}]$ represents the geographic coordinates of the closest point of the border between $j-1$ and j within the *commune* k.

Figure 1: **Focus on the *communes* of *Puligny-Montrachet* and *Saint-Aubin*.**

Notes: The map displays two southern *communes* of *Puligny-Montrachet* (in the East) and *Saint-Aubin* (in the West, see Figure A.2 of the OA). Parts of the *communes* of *Meursault* and *Chassagne-Montrachet* are respectively located in the North and South of the map. The centroids of plots sold over the period 1992–2017 are market by black dots. Horizontal borders always delineate contiguous *communes* (i.e., between $k-1$ and k) and vertical borders delineate overwhelmingly incremental names (i.e., between $j-1$ and j). Two exceptions appear on this map, the East of the *Grand cru* ($j = 4$) is contiguous to a *Village* ($j = 2$), and a part at the North-West border between a *Village* name ($j = 2$) a *Coteaux* name ($j = 0$). These non-incremental borders are not considered in the SRD analysis. The white areas without an official names are developed lands at the center of the *communes* and cropland or forest at the periphery.



through the following running function:

$$r_{j|k}(V) \equiv \mathbf{1}[d = j] \cdot \ell_{j|k}(V) - \mathbf{1}[d = j-1] \cdot \ell_{j|k}(V) = \ell_{j|k}(V) \cdot [d_j - d_{j-1}], \tag{10}$$

where $d_j \equiv \mathbf{1}[d = j]$ is a dummy variable defined from the indicator function $\mathbf{1}[\cdot]$. This class of functions $r_{j|k}(\cdot)$ for $j > 0\ \forall k$ normalizes the distances to give negative values for plots designated as $j-1$ and positive values for plots designated as j. We defined a symmetric class for horizontal names $k > 0$ within a same vertical name j: $r'_{k|j}(V) = \ell'_{k|j}(V) \cdot [c_k - c_{k-1}]$ with $c_k \equiv \mathbf{1}[c = k]$.

These two classes of running functions are used to estimate local values of the incremental signaling premiums δ_{jk} and γ_{jk} in a SRD procedure that is presented below. Following the common practice in multi-cutoff SRD (Cattaneo et al., 2016), the estimations of aggregated incremental premiums $AIVP_j$ and $AIHP_k$ are obtained by pooling each horizontal and vertical names respectively, in addition to the normalization of distances. Considering the subset of plots with a vertical name

$d \in \{j-1, j\}$ for $j > 0$ regardless of the *commune* k they belong, we define:

$$r_j(V) \equiv \Sigma_{k=0}^{K} c_k \cdot r_{j|k}(V) = (d_j - d_{j-1}) \cdot \Sigma_{k=0}^{K} c_k \cdot \ell_{j|k}(V). \tag{11}$$

These functions return the distance to the closest incremental vertical border within the *commune* k to which plot V belong, positively for $d = j$ and negatively for $d = j-1$. They are used to pool horizontal names for each couple of contiguous vertical names in order to estimate average incremental vertical premiums $AIVP_j$. The symmetric class $r'_k(V) = (c_k - c_{k-1}) \cdot \Sigma_{j=0}^{J} d_j \cdot \ell'_{k|j}(V)$ for $k > 0$ is defined from the pooled sample of vertical names to estimate incremental horizontal premiums $AIHP_k$. The four classes of running functions $r_{j|k}(V)$, $r'_{k|j}(V)$, $r_j(V)$, and $r'_k(V)$ produce a sharp SRD design for the wine names, as they depend deterministically on V.

5.2 Spatial assumptions for *terroir*

Under the two following assumptions, we show that SRD identifies local incremental signaling values, from which we derive the implicit weights of the pooled SRD estimates.

SRD Assumptions.

- SRD.1 (Spatial continuity): $\mathbb{E}[h(X^*, j, k) \mid V, d, c]$ is continuous in V for all j, k.
- SRD.2 (Spatial ignorability): $\mathbb{E}[\varepsilon \mid X^*, V, d, c] = \mathbb{E}[\varepsilon \mid X^*, d, c]$.

The first assumption SRD.1 considers that *terroir* variables X^* produce spatially continuous counterfactual price variations. The expected price for any given combination of names (j, k) is continuous conditionally on V, and whatever d, c. This corresponds to the combination of Assumptions 2 (continuity) and 5 (cutoff ignorability) of Cattaneo et al. (2016), and implies that the designation of wine names is ignorable conditionally on V: $\mathbb{E}[h(X^*, j, k) \mid V, d, c] = \mathbb{E}[h(X^*, j, k) \mid V]$. This spatial continuity assumption SRD.1 concerns the full support of V on the studied area. It is not limited to a particular cutoff point or to a particular set of points as in classical and multi-cutoff RD (Imbens and Lemieux, 2008; Abadie and Cattaneo, 2018). This difference is not restrictive, as situations where continuity holds at some arbitrary cutoff points but not for all the other points are peculiar in applied settings (Lee and Lemieux, 2010). SRD.1 stems from our definition of *terroir* as coming exclusively from the natural characteristics of land, jointly with the axiom that nature does not make jumps. This assumption is not rejected by a placebo analysis of the absence of significant discontinuities for the conditional expectations of observed biophysical variables at borders.

The second assumption SRD.2 considers that the errors from observed vineyard prices (already assumed to be linearly independent from *terroir* and wine names by the causal model) do not present any spatial pattern. Wooldridge (2010) in chapter four (p.138), describes the ignorability assumption as rarely controversial because a proxy variable is generally irrelevant to explain the outcome in a conditional mean sense. In our case, SRD.2 amounts to using geographical coordinates V as proxy variables for the partially-observed *terroir*. They are prevented from having a direct effect on the errors from vineyard prices. This restriction that V would not matter if X^* are known is not testable, but it is made virtually in all empirical applications that use spatial continuity for identification (Keele and Titiunik, 2016; Michalopoulos and Papaioannou, 2018).

5.3 Identification from spatial discontinuities

The SRD framework is quite general as it requires neither the full observation of *terroir*, nor the presence of constant signaling premiums. For the population of vineyard plots designated as $(j-1, j, k)$ with $j > 0$, the usual SRD approach identifies a local average of the individual incremental vertical values presented in Equation 3 (see A.1.2 in OA for the complete proof):

$$\begin{aligned}
\delta_{jk}^{SRD} &\equiv \lim_{\epsilon \to 0^+} \mathbb{E}[y \mid r_{j|k}(V) = \epsilon, c = k] - \lim_{\epsilon \to 0^-} \mathbb{E}[y \mid r_{j|k}(V) = \epsilon, c = k] \\
&= \lim_{\epsilon \to 0^+} \mathbb{E}[h(X^*; j, k) \mid \ell_{j|k}(V) = \epsilon] - \lim_{\epsilon \to 0^+} \mathbb{E}[h(X^*; j-1, k) \mid \ell_{j|k}(V) = -\epsilon] \\
&= \mathbb{E}[\delta_{jk}(X^*) \mid V = \mathbf{v}_{j|k}].
\end{aligned} \tag{12}$$

The first line of Equation 12 is definitional, the following is implied by spatial continuity (SRD.1), by spatial ignorability (SRD.2), and by the definition of the normalized running functions. The last equality comes from the continuity of the hedonic function in X^* for given $d = j$ and $c = k$, $\forall j, k$, as defined from the causal model. The two-dimensional vector $\mathbf{v}_{j|k}$ corresponds to the geographical coordinates of the centroid of the border that delineates the vertical levels $j-1$ and j within the horizontal name k.[24] It is straightforward to obtain the symmetric result for the local signaling values of horizontal names $\gamma_{jk}^{SRD} = \mathbb{E}[\gamma_{jk}(X^*) \mid V = \mathbf{v}_{k|j}]$ from each triplet $(j, k-1, k)$.

These results show that local incremental name premiums are identified from classical SRD procedures at each incremental vertical and horizontal border. These local statistics do not represent average treatment effects on the treated if the incremental values depend on unobserved *terroir* variables, in which case the plots at $V = \mathbf{v}_{j|k}$ are not representative of the population of treated vineyards. Nevertheless, these values could be policy relevant to estimate the partial equilibrium causal effect of a marginal change of name delineation (i.e., a small movement of the border).

Consider now the SRD estimation from pooling vineyard plots designated as $j-1$ or j regardless of the horizontal *commune* name k they belong to. Using previously defined $r_j(V)$ as the running variable, this leads to (see, again, Section A.1.2 in OA for the complete proof):

$$\begin{aligned} \delta_j^{SRD} &\equiv \lim_{\epsilon \to 0^+} \mathbb{E}[y \mid r_j(V) = \epsilon] - \lim_{\epsilon \to 0^-} \mathbb{E}[y \mid r_j(V) = \epsilon] \\ &= \Sigma_{k=0}^{K}\, \omega_{jk}^{SRD} \cdot \mathbb{E}[\delta_{jk}(X^*) \mid V = \mathbf{v}_{j|k}], \\ \text{with } \; & \omega_{jk}^{SRD} \equiv \mathbb{P}(c = k) \times f_{V|c=k}(\mathbf{v}_{j|k} \mid c = k)/f_V(\mathbf{v}_j). \end{aligned} \tag{13}$$

This result closely follows Lemma 1 of Cattaneo et al. (2016) for the general multi-cutoff RD. We note $f_V(\mathbf{v}_j) \equiv \Sigma_{k=0}^{K}\, f_{V|c=k}(\mathbf{v}_{j|k} \mid c = k) \cdot \mathbb{P}(c = k)$ the density of plots at the proximity of the incremental vertical border of j regardless k. The proof requires the spatial continuity of this density, which is assessed and not rejected by the data. In Equation 13, pooling horizontal names puts an implicit weighting scheme on local incremental signaling values, the weights sum to one for each couple of incremental vertical names across *communes*. The pooled estimators have the same local meaning as previous δ_{jk}^{SRD} and γ_{jk}^{SRD} so that, in general, $\delta_j^{SRD} \neq AIVP_j$ and $\gamma_k^{SRD} \neq AIHP_k$.

As shown by Cattaneo et al. (2016), pooled SRD estimators are double averages: they are the weighted averages between borders of local averages within borders. If the incremental values do not depend on *terroir* variables (e.g., if they are constant for the treated with $\delta_{jk}(X^*) = \delta_{jk}$), the local SRD premiums identified from Equation 12 can be interpreted causally ($\delta_{jk}^{SRD} = \delta_{jk}$). In this

[24] Contrary to Keele and Titiunik (2016), we did not consider heterogeneous treatment effects within a given border. We represented a border by its centroid to simplify the analysis, this is not a strong assumption given the small length of the borders defined for each triplet $(j-1, j, k)$ as shown by Figure 1 in text and Figure A.9 in the OA.

case, the pooled premiums from Equation 13 identify $AIVP_j$ under the harmless condition that the distribution of population acreages $w_{jk}/w_{j\cdot}$ is equal to the distribution of incremental borders ω_{jk}^{SRD} (i.e., the size of a region with a given name is proportional to the length of its incremental borders). Figure A.16 in the OA provides some support for this restriction with the data at hand, so the pooled SRD weights are not expected to produce a high bias for $AIVP_j$ and $AIHP_k$. The local nature of individual and pooled SRD estimates (for $V = \mathbf{v}_{j|k}$) is probably the most important source of bias for vineyard plots far from borders. Given the small size of incremental borders and the small number of vineyard sales close to some of them, the pooled SRD estimates are determinant to obtain precise estimates of average vertical and horizontal name signaling values. Recognizing the rich set of causal information that can be derived from the multi-cutoff SRD also allows us to test several shape restrictions (but not all) about the full signaling scheme of nested wine names.

5.4 Results from spatial discontinuities

Table 3 presents the average incremental vertical premiums δ_j^{SRD} from pooled SRD with third-order polynomials and standard errors clustered within sales (Cattaneo et al., 2019, see also Figures A.11, A.12, A.13, and A.14 in the OA for the graphical analysis). The top panel of the table shows that SRD premiums (expressed as a percent of vineyard prices) are much smaller than the hedonic premiums from Section 2. Hedonic WLS estimations present upward biases distributed from 150% to 190%, where the highest biases are found for the highest vertical names. Table 3 also reports alternative SRD estimators with *commune* fixed effects and with vineyard prices demeaned by *commune*. This evaluates the robustness of pooled estimations when controlling for unobserved *commune* effects (see Section A.1.2 in OA for details). These alternative SRD estimators are not significantly different from the raw SRD estimations but they are more precise, in line with Calonico et al. (2019). The ranking of vertical incremental values is the same between SRD and hedonic models. The biggest premiums are found between *Région* and *Village* levels, followed by *Premier cru* and *Grand cru*, *Coteaux* and *Région*, and finally *Village* and *Premier cru*. The bottom panel of Table 3 reports the placebo analysis of the spatial continuity of observed biophysical variables and plot size (see also Figures A.11, A.12, A.13, and A.14 in the OA). The borders between wine names do not produce other discontinuities than in vineyard prices, in line with assumption SRD.1.

Table 3: **Average Incremental Vertical Premiums from SRD Estimations.**

	Région (N = 3 090)	*Village* (N = 6 297)	*Premier cru* (N = 4 671)	*Grand cru* (N = 750)
Estimator: WLS	0.956^{***}	1.907^{***}	0.700^{***}	1.279^{***}
	(0.075)	(0.041)	(0.033)	(0.119)
Estimator: SRD pooled	0.584^{***}	1.265^{***}	0.390^{**}	0.686^{**}
	(0.201)	(0.291)	(0.180)	(0.350)
Estimator: SRD with fixed effects	0.384^{***}	0.962^{***}	0.658^{***}	0.744^{***}
	(0.143)	(0.233)	(0.155)	(0.280)
Estimator: SRD within *commune*	0.302^{*}	1.013^{***}	0.390^{***}	0.696^{***}
	(0.174)	(0.218)	(0.131)	(0.260)
Placebo: Elevation	0.010	0.260	0.287	0.797
	(1.039)	(1.149)	(0.463)	(1.265)
Placebo: Slope	−0.385	0.408	0.850	−1.956
	(0.706)	(1.012)	(0.830)	(1.283)
Placebo: Solar Radiation	−0.002	0.417	0.120	−1.571
	(0.151)	(0.273)	(0.130)	(1.125)
Placebo: Exposition	−0.212	0.026	-0.167^{*}	-0.321^{*}
	(0.152)	(0.106)	(0.090)	(0.195)
Placebo: Plot Size	0.182	−0.005	−0.305	−0.120
	(0.585)	(0.265)	(0.258)	(0.499)

Standard errors (reported in parentheses) are clustered by vineyard sales: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
Notes: The first row of the Table is a benchmark estimation of the hedonic model of Equation 1 by weighted least squares (see Section A1.1 of the OA for details). Each remaining cell in the top panel is a pooled spatial regression discontinuity estimator of incremental vertical premiums with the log of per-ha vineyard price as the outcome variable. The first column reports the causal effect of having a *Région* name relatively to a *Coteaux* name (in % of vineyard price), the second column is the effect of having a *Village* name relatively to a *Région* name, and so on. Figures A.11a, A.12a, A.13a, and A.14a of the OA represent the results graphically. In the following rows of the top panel, *commune* fixed effects are introduced as covariates and a within *commune* transformation of the outcome is applied before SRD. The bottom panel reports the pooled SRD estimations for observed biophysical variables and plot size, as a heuristics placebo analysis for SRD.1. These latter results are displayed in Figures A.11, A.12, A.13, and A.14 in the OA.

Table A.4 in the OA reports pooled horizontal SRD premiums γ_k^{SRD} for the 18 sub-samples of contiguous horizontal names (see also Figure A.15a in the OA). In the majority of cases, both hedonic and SRD premiums are not significantly different from zero because the narrowness of horizontal borders decreases the precision of the estimates. We do not observe a systematic upward bias between hedonic and SRD estimations (i.e., between column 1 and columns 2 to 4 in Table A.4 of OA). This confirms that controlling for unobserved *terroir* effects is not as critical for horizontal names as for vertical names, as suggested by the partial R^2 of Table 1. Table A.5 in the OA reports both hedonic and SRD estimations for the incremental values of vertical levels, separately for each horizontal *commune* name (see also Figure A.15b of the OA). These individual incremental values are rarely significant between *Coteaux* and *Région* names. This result can be interpreted both as the absence of differentiated signaling values and as the lack of statistical power due to small sample sizes close to the individual incremental borders. This is not the case for signaling premiums between *Région* and *Village* vertical names. They are almost all significant with a similar hierarchy of horizontal names to that in hedonic models.

We observe in Table A.5 of the OA that vertical names are still monotone from SRD results (only one significant negative value exists between *Coteaux* and *Région* levels for the *commune* of *Chassagne-Montrachet*). Spearman rank correlations between the horizontal values recovered from the same SRD estimations are not significantly different from zero. This does not validate horizontal monotonicity. Under the additional assumptions reported in Section A.1.2 of OA (about additive *terroir* effects and parsimonious interactions) pooled vertical SRD estimations δ_j^{SRD} reported in the second row of Table 3 identify $AIVP_j$, and the SRD estimations δ_{2k} between *Région* and *Village* levels in the fourth column of Table A.5 in the OA identify $AIHP_k$. This shows that the complementarity of wine names fades out in SRD estimations (Figure A.16 of the OA) and suggests that the complementarity of names from hedonic estimation is an artifact due to unobserved *terroir* heterogeneity. According to the price decomposition of Equation 9 (still under the assumptions of A.1.2 in OA), vertical names have a total value of €618 millions (average of €55 000 per ha, 42% of which is obtained from hedonic estimations) and horizontal names have a total value of €724 million (average of €63 000 per ha, 115% of which is obtained from hedonic estimations). The full decomposition of the total signaling value cannot be computed because neither the interaction part, nor the total signaling value are recovered from SRD.

6 Doubly robust evidence

This section presents our preferred doubly robust (DR) evidence that combines the control for unobserved *terroir* quality and the extrapolation of causal signaling values away from the borders. We show that name designations depend on historical lobbying actions that increase the overlap between vineyard plots, therefore providing precise estimates for the full signaling scheme.

6.1 Specification assumptions

The DR approach is based on an ordered model with *commune*-specific thresholds, where vertical name designations are assumed to depend on partially-observed *terroir* variables. Following Ay (2021), the latent variable $d^* \equiv b(X^*) + \xi$ specifies a vineyard quality index that underlies vertical designations. The unknown function $b(\cdot)$ describes the relation between *terroir* and vertical names, according to historically accumulated knowledge and repeated wine tasting that are not observable to the econometrician. The random part ξ accounts for unobserved variables not related to the *terroir*, which enter incidentally into the designation under $\mathbb{E}(\xi \mid X^*) = 0$. The designation model combines the latent vineyard quality to additively separable thresholds $\alpha_j + \mu_k$. The vertical name of the plot i in the *commune* $c_i = k$ satisfies (with $\alpha_0 \equiv -\infty < \alpha_1 < \cdots < \alpha_J \equiv +\infty$):

$$d_i = j \quad \Leftrightarrow \quad \alpha_{j-1} + \mu_k < d_i^* < \alpha_j + \mu_k. \tag{14}$$

The designation of vertical names is specified through varying thresholds μ_k as the horizontal delineations k correspond to the administrative scale at which lobbying took place during the second part of the ninetieth century and the beginning of the twentieth century (Jacquet, 2009; Ay, 2021). Through the reputation of past landowners, their influence with decision makers or their collective actions, some *communes* enjoyed privileged treatment. Their vineyard plots are placed higher in the vertical hierarchy than similar plots in other administrative units. Recall that the horizontal *commune* delineations pre-existed the vertical ones and were not initially intended to signal wine quality. Some plots with the same *terroir* quality $b(X^*)$ but in different *communes* face different probabilities of having a given vertical name, without any tangible justification. The name variations from μ_k do not depend on potential *terroir* quality and will serve our identification strategy.

At first glance, the estimation of the designation model is subject to the same limitation as the vineyard price equation. The identifications of the latent index $b(X^*)$ and both terms of the thresholds $\alpha_j + \mu_k$ rely on the control of *terroir* that is only partially observed from the data. However, the crucial difference lies in the possibility of estimating the designation model on the *exhaustive* population of vineyard plots, whereas the price model can only be estimated from the sample of sales. This multiplies by more than six the spatial density of observations (60 000 instead of

9 000 in the same area). This also allows to precisely control for unobserved spatial heterogeneity through high-dimensional methods (Belloni et al., 2014; Athey and Imbens, 2019). With the same population data, Ay (2021) found that spline transformations of biophysical variables and geographic coordinates by penalized maximum likelihood (Wood et al., 2016) accounted for unobserved *terroir* effects in estimating *commune* coefficients from the delineation of vertical names.

In line with this prior result, we specify partially-observed vineyard quality $b(X^*)$ through additive spline transformations of covariates noted $B(X, V)^\top \psi$. The high-dimensional vector ψ of unknown coefficients is associated with the series transformation of the biophysical variables and the geographical coordinates (see Section A.1.3 in OA for details). This specification gives the first *sufficient* DR.1 assumption presented below. The second *sufficient* DR.2 assumption concerns the specification of counter-factual hedonic functions, which are expected to depend on the vineyard quality index of the designation model. As is well known (Robins and Rotnitzky, 1995; Słoczyński and Wooldridge, 2018), the DR identification is reached if at least the propensity score or the outcome is well specified, which amounts to one of the following DR assumptions in our case:

DR Assumptions.

- DR.1 (Designation specification): $\mathbb{P}(d \leqslant j \mid X, V, c) = \Phi[\alpha_j + \Sigma_k \, \mu_k \cdot c_k - B(X, V)^\top \psi]$,
- DR.2 (Hedonic specification): $\mathbb{E}[h(X^*; j, k) \mid X, V] = \kappa_{jk} + \rho_{jk} \cdot B(X, V)^\top \psi$.

In DR.1, Φ denotes the cumulative distribution function of errors ξ, from which generalized propensity scores are specified to make causal inference about multi-valued treatments under unconfoundedness (Imbens, 2000; Cattaneo, 2010; Uysal, 2015). As the probability of having a low vertical name d increases with μ_k, a *commune* k with a high coefficient μ_k is systematically disadvantaged by designations. Because the range of a cumulative function is bounded by the unit interval, DR.1 implies a vertical overlap of plots with $0 < \mathbb{P}(d = j \mid X, V, c) < 1$ for the full range of X, V, c (i.e., where c is included as a pre-treatment variable). In DR.2, we use the dimension reduction property of the propensity score (Rosenbaum and Rubin, 1983) for the specification of the hedonic function, where ρ_{jk} measures the partial correlation between the latent vineyard quality and expected vineyard prices. This linear dependence is allowed to change between j and k, whereas the saturated form of wine name effects κ_{jk} is included without loss of generality.

6.2 Doubly robust identification

In a direct application of Lemma 3.2 from Słoczyński and Wooldridge (2018) to our framework, we show in Section A.1.3 of the OA that Assumptions SRD and DR.1 are sufficient to identify the full set of signaling premiums defined from the causal model of Section 4. We obtain:

$$\mathbb{E}[h(X^*_i, j, k) \mid d = l, c = m] = w_{lm}^{-1} \cdot \mathbb{E}[\omega^{DR}_{jk|lm}(X, V) \cdot d_j \cdot c_k \cdot y], \tag{15}$$

$$\text{with} \quad \omega^{DR}_{jk|lm}(X, V) \equiv \frac{\Phi[\alpha_l + \mu_m - B(X, V)^\top \psi] - \Phi[\alpha_{l-1} + \mu_m - B(X, V)^\top \psi]}{\Phi[\alpha_j + \mu_k - B(X, V)^\top \psi] - \Phi[\alpha_{j-1} + \mu_k - B(X, V)^\top \psi]} \times \frac{\mathbb{P}(c = m \mid X, V)}{\mathbb{P}(c = k \mid X, V)}.$$

From the first line of Equation 15, the average price that vineyards designated as $(d = l, c = m)$ would have if they were designated as $(d = j, c = k)$ is identified by a weighted expectation of the observed prices y for the vineyards currently designated as $(d = j, c = k)$. The weighting scheme is based on the ratio of two generalized propensity scores from DR.1. They are multiplied by the ratio of two conditional probabilities about horizontal designations (at the end of the second line of Equation 15). While the vertical overlap is implied by DR.1, the horizontal overlap $0 < \mathbb{P}(c = k \mid X, V) < 1, \forall k$, is more critical, and could generate an irregular identification (Khan and Tamer, 2010).

This first DR result is quite general, as it identifies average counterfactual prices for each combination of names (j, k, l, m) with $j, l = 0, \dots, J$ and $k, m = 0, \dots, K$. The second part of the DR proof about the sufficiency of DR.2 is more direct and is only reported in Section A.1.3 of OA for completeness. Accordingly, the full causal signaling scheme is identified from Equation 15 and the shape restrictions can be tested for each combination of names in the data (i.e., for $w_{jk}, w_{lm} > 0$). Average individual incremental premiums of Equation 3 are recovered from:

$$\begin{aligned} \delta_{jk} &= (w_{\cdot k}/w_{jk}) \cdot \mathbb{E}\Big\{[d_j - \omega^{DR}_{(j-1)k|jk}(X, V) \cdot d_{j-1}] \cdot y \mid c = k\Big\}, \\ \gamma_{jk} &= (w_{j\cdot}/w_{jk}) \cdot \mathbb{E}\Big\{[c_k - \omega^{DR}_{j(k-1)|jk}(X, V) \cdot c_{k-1}] \cdot y \mid d = j\Big\}. \end{aligned} \tag{16}$$

The ratio of the conditional probabilities $\mathbb{P}(c = m \mid X, V)/\mathbb{P}(c = k \mid X, V)$ from $\omega^{DR}_{jk|lm}(X, V)$ in Equation 15 that threaten identification simplifies in $\omega^{DR}_{(j-1)k|jk}(X, V)$. So, horizontal overlap is not required to identify the individual incremental vertical values of the first line of Equation 16 because only vineyard plots with $c = k$ are used. As a side effect, the name variations between *communes* from historical lobbying are not used for identification because μ_k appears both at the numerator and the denominator of the weights. In the second line of Equation 16, the identification of incremental horizontal values relies on $\mathbb{P}(c = k \mid X, V)/\mathbb{P}(c = k-1 \mid X, V)$ in $\omega^{DR}_{j(k-1)|jk}(X, V)$ and consequently requires horizontal overlap. Because the *terroir* bias is not as critical for horizontal names, we put this point aside and come back to it empirically in Section 6.3 below.

The pooled average incremental vertical premiums can be recovered from Equation 16, by using $AIVP_j = w_{j\cdot}^{-1} \, \Sigma_{k=0}^{K} \, w_{jk} \cdot \delta_{jk}$ from Equation 4. This identification relies on using *communes* as pre-treatment variables, which is not required to have vertical unconfoundedness from SRD.1.[25] In fact, $AIVP_j$ can be identified from generalized propensity scores evaluated at any arbitrary *commune* m without loss of generality. Under the sufficient assumptions SRD and DR.1, we can identify the pooled premiums of vertical names from (see Section A.1.3 in the OA for the complete proof):

$$\mathbb{E}[h(X^*_i, j, c) \mid d = l] = w_{l\cdot}^{-1} \, \mathbb{E}[\omega^{DR}_{jm|lm}(X, V) \cdot d_j \cdot y], \qquad \forall m = 0, \ldots, K. \tag{17}$$

This original result shows that average treatment effects on the treated can be identified from generalized propensity scores with a counterfactual assignation of vineyard plots between the *communes*. The proof of Section A.1.3 in OA presents the formula for averaging propensity scores among the set of *communes*, which produces the highest overlap in our empirical application. As first recognized by Ichimura and Taber (2001), an unconfoundedness assumption is an exclusion restriction for the treatment that, in our case, is equivalent to considering *commune* dummies as instrumental variables. The exclusion restriction for c in SRD.1 is not necessary (as $AIVP_j$ can be recovered from equations 4 and 15) but allows us to increase overlap between vineyard plots with different vertical names. From the chain rule, $0 < \mathbb{P}(d_j = 1 \mid X, V, c = m) < 1$ is implied by $0 < \mathbb{P}(d_j \cdot c_m = 1 \mid X, V) < 1$ $(\forall j, m)$. The overlap condition for Equation 15 is more restrictive than for Equation 17 on the full support of (X, V) $\forall m$. Alternative identification strategies have been developed recently to deal with limited overlap, though distributional assumption (Rothe, 2017), resampling (Ma and Wang, 2020) or bias correction (Sasaki and Ura, 2018). To the best of our knowledge, the use of exogenous variations of the treatment assignment is new.

[25] SRD.1 implies that $\mathbb{E}[h(X^*_i, j, k) \mid X, V, d] = \mathbb{E}[h(X^*_i, j, k) \mid X, V]$ by the law of iterated expectations, so vineyard prices are conditionally mean-independent from d even without controlling for c as a pre-treatment variable.

6.3 Doubly robust estimation

Słoczyński and Wooldridge (2018) showed that weighted least squares WLS on sub-samples defined from the treatment status recovered causal statistics under DR assumptions very close to ours (see Cattaneo, 2010; Uysal, 2015 for alternative estimation procedures). For each couple (j, k) of names, we can consistently estimate the κ and ρ coefficients by using the weights from Equation 15. Because we estimate average signaling values, the coefficients depend on the population of treated plots (through the weighting scheme), so using the identification result from Equation 16 for each combination (j, k, l, m) necessitates the estimation of 5 184 models (from 72 weighting schemes for each of the 72 combinations of names). From them, we can then compute (we omit indexation of the weights on coefficients for the sake of clarity):

$$\hat{\delta}_{jk}^{DR} = (\hat{\kappa}_{jk} - \hat{\kappa}_{(j-1)k}) + (\hat{\rho}_{jk} - \hat{\rho}_{(j-1)k}) \cdot w_{jk}^{-1} \cdot \Sigma_{i=1}^{N} \mathbf{1}[d_i = j, c_i = k] \cdot B(\mathbf{x}_i, \mathbf{v}_i)^\top \hat{\psi}, \tag{18}$$
$$\hat{\gamma}_{jk}^{DR} = (\hat{\kappa}_{jk} - \hat{\kappa}_{j(k-1)}) + (\hat{\rho}_{jk} - \hat{\rho}_{j(k-1)}) \cdot w_{jk}^{-1} \cdot \Sigma_{i=1}^{N} \mathbf{1}[d_i = j, c_i = k] \cdot B(\mathbf{x}_i, \mathbf{v}_i)^\top \hat{\psi}.$$

Under SRD assumptions, only one of assumptions DR.1 and DR.2 is sufficient to have $\hat{\delta}_{jk}^{DR} = \delta_{jk}$ and $\hat{\gamma}_{jk}^{DR} = \gamma_{jk}$. The possibility to finely control for V from the full population data with a high spatial density of plots allows us to use the spatial continuity of *terroir* as in SRD estimation of Section 5, without being limited to estimate local causal statistics. We can compare DR and SRD results by computing individual average incremental premiums from Equation 18 at $\mathbf{v}_i = \mathbf{v}_{j|k}$.

Using the identification result of Equation 17 consists in pooling vineyard sales with a different horizontal name c for each level of vertical name j. This leads to the model:

$$y_i = \mathbf{c}_i^\top \boldsymbol{\kappa}_j + \rho_j \cdot B(\mathbf{x}_i, \mathbf{v}_i)^\top \hat{\psi} + \epsilon_i. \tag{19}$$

The unknown coefficients $\boldsymbol{\kappa}_j$ and ρ_j are estimated separately for each $J + 1 = 5$ sub-samples of vertical names and for each 5 different weighting schemes, with $l = 0, \ldots, J$ in $\tilde{\omega}_{j|l}^{DR}(\mathbf{x}_i, \mathbf{v}_i) \equiv \Sigma_{m=0}^{K} \omega_{jm|lm}^{DR}(\mathbf{x}_i, \mathbf{v}_i)/(K+1)$. These 25 WLS estimations give 25 sets of coefficients noted $\hat{\boldsymbol{\kappa}}_j^{(l)}$ for the $K + 1$ horizontal premiums and $\hat{\rho}_j^{(l)}$ for the coefficients associated with the predicted latent vineyard quality (the exponents in parenthesis mark the vertical level l used for the weights). Note that index

k is dropped from DR.2 in Equation 19 because this restriction is not rejected by the data (the interaction between $\mathbf{c}_i$ and $B(\mathbf{x}_i, \mathbf{v}_i)^\top \hat{\psi}$ is not significant from a F test on WLS models). Finally, all incremental signaling values, all shape restrictions, and the full price decomposition from Section 4 can be recovered from these sub-sample WLS estimations (see Section A.1.3 in OA for details).

6.4 Doubly robust results

We first present the results from name designation models on the whole population data with a logistic distribution of errors (see the left panel of Table A.1 in the OA for the summary statistics). Table 4 reports the relative importance of each variable or group of variables in explaining the designation of vertical names with different degrees of spatial smoothing of geographical coordinates (as measured by χ^2 statistics about joint significance). As a whole, the three models have high explanatory and predictive performances with pseudo-R^2 greater than 0.7 and more than 85% of correct predictions. Following the bivariate smoothing function of geographical coordinates, the *commune* dummies are the second most important group of variables in terms of joint significance, followed by elevation, solar radiation, slope, and exposition. As expected, increasing the complexity of the smoothing functions of geographical coordinates (i.e., decreasing the spatial smoothing of the latent quality variable) decreases the joint significance of other included exogenous variables, while the *commune* dummies about lobbying effects stay the second most significant group.

The similarity between the marginal effects of biophysical variables in the designation of vertical wine names (Figure A.17 in OA) and their marginal effects in the hedonic models of vineyard prices (Figure A.4 in OA) is striking. The spline transformations of elevation produce a U-inverted effect with a maximum probability of a high vertical name at less than about 300 meters. Slope and solar radiation also have very similar effects as in hedonic models. This illustrates once more the strong statistical relationships between the designation of vertical wine names, biophysical variables, and vineyard prices. The unequal treatments between *communes* in terms of vertical name designations are displayed by Figure A.18 in OA. It shows highly significant differences between *communes* and a stable ranking of them for the different spatial smoothing specifications of Table 4. The rankings of the *communes* are similar to those obtained by Ay (2021), who provides additional interpretations related to the local context and the historical justifications of lobbying

Table 4: **Joint Significance χ^2 Statistics of Exogenous Variables in Designation Models.**

	Outcome variable: Ordered Vertical Wine Names		
Variable	(df= 500)	(df= 700)	(df= 900)
Elevation	2 378.3***	1 994.7***	1 974.6***
	[8.84]	[8.83]	[8.841]
Slope	273.94***	183.69***	167.74***
	[8.368]	[8.25]	[8.051]
Solar Radiation	1 245.2***	1 155.3***	811.18***
	[8.035]	[8.181]	[7.938]
Spatial Coordinates	104 224***	107 858***	114 093***
	[484.4]	[662.1]	[841.2]
Exposition	69.604***	47.667***	15.037***
	[4]	[4]	[4]
Communes	4 988.8***	3 255.6***	2 578***
	[19]	[19]	[19]
Number of Observations	59 838	59 838	59 838
McFadden R^2	69	72.24	75.15
Percent of correct predictions	84.31	86.81	88.83
Akaike Information Criteria	53.72	48.6	44.04

Effective degrees of freedom [reported in brakets] determine significance: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
Note: Each column of the Table reports the χ^2 statistics for the joint significance of each group of spline transformations related to the variables in rows (Wood et al., 2016). The models are of increasing spatial complexity (i.e., decreasing spatial smoothing) from left to right. The first column restricts the maximal degree of freedom from spline transformations of geographical coordinates to be less than 500, to obtain a penalized effective degree of freedom of 484.4 (eighth line). The two following columns put the maximal degrees of freedom to respectively 700 and 900.

effects. Figure A.19 in OA shows the spatial patterns of the predicted vineyard quality that underlies the designation of vertical wine names for df = 900. The spatial precision allowed by the small size and high density of vineyard plots from population data produces very fine spatial variations of the vineyard quality index as revealed by the designation of vertical names. It is important to note that horizontal delineations from administrative *commune* borders do not produce any discontinuity in these predictions that rely only on the geographical coordinates of vineyard plots.

Figure A.20a and Figure A.20b in the OA show the gain of overlap from counter-factually assigning vineyard plots between *communes*. The raw predictions of propensity scores (without *commune* averaging in the panels A) only produce enough overlap to estimate incremental premiums between contiguous vertical names. For instance, vineyards named as *Coteaux* present an overlap only with *Région* vineyards, as the overlap areas with *Village*, *Premier cru*, and *Grand cru* are thin or even empty. As displayed in the bottom panels B of the Figures, averaging propensity scores between counter-factual assignment of *communes* membership increases the overlap between plots with different vertical names sharply. *Coteaux* vineyards are then overlapped with *Région*, *Village*, and *Premier cru* vineyards. However, the *commune* instruments are not sufficiently strong to obtain an overlap between *Coteaux* and *Grand cru* vineyards, so the corresponding results should be interpreted with caution. The Figure also shows that overlap gains are higher for df = 500 than for df = 900, i.e., where the *commune* coefficients are more significant (as reported in Table 4).

Table 5 gives the pooled vertical values from DR estimations, where average incremental vertical values $AIVP_j$ appear just above the empty diagonal. The pooled vertical DR values correspond to incremental signaling values (per hectare in 2017) of €19 300, €114 000, €150 000, and €400 000, for respectively *Région*, *Village*, *Premier cru*, and *Grand cru* (as the respective average prices are €36 400, €87 100, €325 500, and €688 400). Vertical DR premiums are much lower than vertical hedonic premiums of Table A.3 in the OA and slightly higher than vertical SRD premiums of Table 3. Incremental DR vertical premiums are not significantly different from SRD premiums when generalized propensity scores are evaluated at the borders, i.e., for $\mathbf{v} = \mathbf{v}_{j|k}$ in Equation 19 (unreported results). The similarity between SRD and DR decreases significantly when moving away from the diagonal of the Table (i.e., when we compare plots further from geographical borders). Symmetrically, average premiums reported in the first line of Table 5 are closer to the hedonic values. The DR approach combines the internal validity of the SRD and the external validity of the hedonic approach, and produces results that are between the two.

Table A.6 in OA displays additional diagnostics for the 25 WLS regressions used to obtain the vertical signaling values. *Commune* fixed effects from Equation 19 are more significant for *Village* and *Premier cru* levels. This is in accordance with the impossibility for lower levels *Coteaux* and *Région* to combine their horizontal names and with the small number of *Grand cru* observations. The vineyard quality index from the designation model has a globally significant positive effect on price, especially for *Village* names. The full R^2 are significantly lower than in the hedonic regressions because of both sub-sampling and the approach chosen to account for control variables $\mathbf{z}$ from Equation 1.[26] The results are obtained without trimming extreme predicted propensity scores,

[26] To control for sale characteristics, we regress the logarithm of per-ha vineyard price on the variables to obtain the vector $\boldsymbol{\theta}$ of coefficients, then the WLS sub-sample estimations are made on the residuals.

Table 5: **Pooled Average Vertical Premiums from Double Robust Estimations.**

	Coteaux	*Régional*	*Village*	*P. cru*	*G. cru*
Coteaux		0.53^{***}	2.01^{***}	2.73^{***}	4.06^{***}
		(0.10)	(0.15)	(0.16)	(0.29)
Régional	-0.60^{***}		1.31^{***}	2.37^{***}	3.68^{***}
	(0.14)		(0.08)	(0.22)	(0.38)
Village	-2.12^{***}	-1.45^{***}		0.46^{***}	1.58^{***}
	(0.39)	(0.07)		(0.05)	(0.19)
Premier cru	-3.25^{***}	-2.28^{***}	-0.47^{***}		0.58^{***}
	(0.79)	(0.53)	(0.09)		(0.13)
Grand cru	-6.18^{***}	-3.90^{***}	-1.89^{***}	-0.91^{*}	
	(1.32)	(0.62)	(0.38)	(0.40)	

Standard errors (reported in parentheses) are from stratified bootstrap: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
Notes: Reported numbers correspond to pooled average vertical premiums (as a percent of vineyard price) between the names in columns and the names in rows for vineyards actually designated as the name in columns (i.e., average treatment effects on the treated). In the second cell of the first row, *Région* vineyards are 53% more expensive that they would be if designated as *Coteaux*, according to their actual distribution among horizontal names. Because of the estimated weights and latent quality index, standard errors are bootstrapped with a stratification of vineyard sales. The details of the 25 WLS estimations from which these numbers are computed are reported in Table A.6 in the OA.

so the number of observation is constant in each sub-sample. Table A.7 in the OA reports the average incremental horizontal premiums $AIHP_k$, showing that they are a little weaker than the hedonic ones reported in Table A.3 in OA. Again, controlling for unobserved *terroir* heterogeneity appears less important for horizontal premiums than for vertical ones.

From Table 5, vertical signaling values are pairwise monotone, as we find positive increasing signaling values at the top right corner and negative decreasing ones in the bottom left corner. Table 6 reports Spearman's rank correlation coefficients from *commune* fixed effects of Table A.7 in the OA. It shows that horizontal monotonicity is verified from DR results, although with a smaller significance than the hedonic results from Table 2. We find that the complementarity of horizontal and vertical names is also verified, as displayed by Figure A.21 in OA. An increase of 1% of the average vertical value for a given *commune* is associated with an increase of 0.66 points in its horizontal value. Nevertheless, this evidence is only significant at 15% ($t = 1.58$). According to the price decomposition of Equation 9, we find a total DR signaling value of about €1.3 billions (in 2017), representing about €115 000 by hectare on average. The vertical part represents 61% (€70 150 per ha), the horizontal part 53% (€60 950 per ha), and the interaction part −12% (€13 800 per ha), instead of respectively 80%, 35% and −15% from the hedonic results. The total signaling value of nested wine names is revised downward to about 40% compared to hedonic results. The part of horizontal names increases (both in share of total value and absolute value) and the share of the interaction part remains roughly stable, while it is divided by two in monetary units.

Table 6: **Spearman's rank correlation between horizontal premiums for each vertical level.**

	Coteaux	*Region*	*Village*	*Premier cru*
Region	0.707			
Village	0.317	0.410		
Premier cru	0.612	0.612	0.897	
Grand cru	0.143	0.257	0.086	0.429

Notes: Spearman's rank correlation coefficients are statistical measures of pairwise monotonicity. We computed them from horizontal fixed effects from Equation 19 in each sub-sample of vertical names. The designation model used to predict vineyard quality index is with df = 900, see also Table A.6 for the estimated values of raw coefficients.

7 Conclusion

This article provides a causal valuation of nested wine names, controlling for unobserved *terroir* quality that impacts both name designation and vineyard price. Our preferred identification strategy exploited a historical bias in name designation, from lobbying actions that took place one century ago. We show that averaging counter-factual propensity scores along this exogenous dimension increases the overlap between observations, which improves the causal inference.

By providing a hierarchy of quality to consumers, the vertical dimension of names has the highest signaling value, although it was over-estimated by a hedonic model (by a factor of about two) and under-estimated by a spatial regression discontinuity design (by a factor of about 0.85). The differences between the methods are respectively explained by unobserved *terroir* heterogeneity and by unobserved heterogeneity of signaling values, which both increase for vineyard plots away from the borders between names. Conversely, the signaling values of horizontal names are not significantly biased by unobserved *terroir* heterogeneity, and benefit from spillovers of vertical names. For a given *terroir* quality, a horizontal name is more valuable when it is nested with high-valued vertical names. This provides a first causal evidence of complementary quality disclosure from bi-dimensional multi-valued signals with both a horizontal and a vertical dimension.

The complementarity of nested names accounts for about 15% of their total signaling value. This could explain the frequency of this nested structure for consumption goods in particular and for many other quality signals in general. For several centuries in Burgundy, wine production was segmented according to this nested structure, and we show that it is still relevant today despite changes in production techniques, tastes, and market structures. Moreover, the monotonicity of wine names within their respective horizontal and vertical dimensions indicates a consistent quality signaling that gives a stable ordering of names. The relative signaling value of a given horizontal or vertical name is preserved whatever the combination to which it belongs. This empirical finding provides a causal support for the monotonicity condition often used in the information theory.

The extent of *terroir*'s significance for vine cultivation and for the production of high-quality wines is regularly debated in the natural science literature and in the wine industry. We provide evidence of a high economic importance of fine-scale *terroir* quality that is neither unobserved by the researcher from data, nor by the typical consumer from bottled wines that cannot taste it. These *terroir* variables are only observable from the field and probably have to be combined with external knowledge from experiences of wine production and wine tasting to reach their fullest extent in value. By focusing on the willingness to pay on the land market (by buyers and sellers, but also real estate and consulting agencies), we reveal the importance of *terroir* quality from the agents that hold this knowledge and use it in bidding for a wine name when buying a vineyard. This estimation relies on their perceptions from the land market and it would be interesting to objectify the underlying natural processes behind the *terroir* concept in future researches.

Our results were obtained from a spatial continuity assumption about *terroir* quality. This is justified by the well-recognized smooth geographical variations of natural processes and the tiny size of vineyard plots for the area under study. We have to highlight that the causal signaling values that we define and estimate are partial equilibrium statistics under the current market equilibrium of Burgundy's wines. Their external validity for other areas or other products needs to be addressed adequately. Additional research is also needed to study the general equilibrium determinants of the signaling values of other names, in line with changing consumers' knowledge, learning, fashion, or taste. This research agenda would inform better policies about better quality disclosures, by endogenizing both consumers' and producers' reactions.

8 Data availability

The population data about the official wine names and biophysical variables for the 60 000 vineyard plots of the area are freely available at the dataverse https://data.inrae.fr/dataset.xhtml?persistentId=doi:10.15454/ZZWQMN

References

Abadie, A. and Cattaneo, M. D. (2018). Econometric methods for program evaluation. *Annual Review of Economics* 10: 465–503.

Abbott, J. K. and Klaiber, H. A. (2011). An embarrassment of riches: Confronting omitted variable bias and multi-scale capitalization in hedonic price models. *Review of Economics and Statistics* 93: 1331–1342.

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

Angrist, J. D. and Krueger, A. B. (1999). Empirical strategies in labor economics. In *Handbook of Labor Economics*. Elsevier, *3*, 1277–1366.

Angrist, J. D. and Rokkanen, M. (2015). Wanna get away? Regression discontinuity estimation of exam school effects away from the cutoff. *Journal of the American Statistical Association* 110: 1331–1344.

Aryal, G., Bhuller, M. and Lange, F. (2021). Signaling and employer learning with instruments. *National Bureau of Economic Research* .

Ashenfelter, O. and Storchmann, K. (2010). Using hedonic models of solar radiation and weather to assess the economic effect of climate change: the case of Mosel valley vineyards. *Review of Economics and Statistics* 92: 333–349.

Athey, S. and Imbens, G. W. (2019). Machine learning methods economists should know about. *Annual Review of Economics* 11: 685–725.

Athey, S. and Levin, J. (2018). The value of information in monotone decision problems. *Research in Economics* 72: 101–116.

Ay, J.-S. (2021). The informational content of geographical indications. *American Journal of Agricultural Economics* 103: 523–542.

Ay, J.-S., Brayer, J.-M., Cavailhès, J., Curmi, P., Hilal, M. and Ubertosi, M. (2012). La valeur des attributs naturels des terres agricoles de Côte-d'Or. *UMR CESAER Working Paper* .

Belloni, A., Chernozhukov, V. and Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives* 28: 29–50.

Bertanha, M. and Imbens, G. W. (2019). External validity in fuzzy regression discontinuity designs. *Journal of Business & Economic Statistics* : 1–39.

Bhattacharya, J. and Vogt, W. B. (2007). Do instrumental variables belong in propensity scores? *National Bureau of Economic Research* .

Bishop, K. C., Kuminoff, N. V., Banzhaf, H. S., Boyle, K. J., von Gravenitz, K., Pope, J. C., Smith, V. K. and Timmins, C. D. (2020). Best practices in using hedonic property value models for welfare measurement. *Review of Environmental Economics and Policy* 14: 260–281.

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.

Börgers, T., Hernando-Veciana, A. and Krähmer, D. (2013). When are signals complements or substitutes? *Journal of Economic Theory* 148: 165–195.

Bronnenberg, B. J., Dubé, J.-P., Gentzkow, M. and Shapiro, J. M. (2015). Do pharmacists buy bayer? informed shoppers and the brand premium. *Quarterly Journal of Economics* 130: 1669–1726.

Bronnenberg, B. J., Dubé, J.-P. and Moorthy, S. (2019). The economics of brands and branding. In *Handbook of the Economics of Marketing*. Elsevier, *1*, 291–358.

Brown, J. N. and Rosen, H. S. (1982). On the estimation of structural hedonic price models. *Econometrica* : 765–768.

Calonico, S., Cattaneo, M. D., Farrell, M. H. and Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics* 101: 442–451.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica* 69: 1127–1160.

Cattaneo, M. D. (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics* 155: 138–154.

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2019). *A practical introduction to regression discontinuity designs*. Cambridge University Press.

Cattaneo, M. D., Titiunik, R., Vazquez-Bare, G. and Keele, L. (2016). Interpreting regression discontinuity designs with multiple cutoffs. *Journal of Politics* 78: 1229–1248.

Chetverikov, D., Santos, A. and Shaikh, A. M. (2018). The econometrics of shape restrictions. *Annual Review of Economics* 10: 31–63.

Clark, D. and Martorell, P. (2014). The signaling value of a high school diploma. *Journal of Political Economy* 122: 282–318.

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.

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.

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

Cross, R., Plantinga, A. J. and Stavins, R. N. (2017). Terroir in the new world: Hedonic estimation of vineyard sale prices in California. *Journal of Wine Economics* 12: 282–301.

Currarini, S., Ursino, G. and Chand, A. (2020). Strategic transmission of correlated information. *Economic Journal* 130: 2175–2206.

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

D'Amour, A., Ding, P., Feller, A., Lei, L. and Sekhon, J. (2021). Overlap in observational studies with high-dimensional covariates. *Journal of Econometrics* 221: 644–654.

Erdem, T. and Swait, J. (1998). Brand equity as a signaling phenomenon. *Journal of Consumer Psychology* 7: 131–157.

Flores, C. A. and Mitnik, O. A. (2013). Comparing treatments across labor markets: an assessment of nonexperimental multiple-treatment strategies. *Review of Economics and Statistics* 95: 1691–1707.

Fong, Y.-F., Liu, T., Meng, X. and Zhang, Y. (2019). Consumers' trust of multi-attribute experience-goods seller. *Available at SSRN 3247708* .

Fortin, N., Lemieux, T. and Firpo, S. (2011). Decomposition methods in economics. In *Handbook of Labor Economics*. Elsevier, *4*, 1–102.

Friendly, M., Monette, G. and Fox, J. (2013). Elliptical insights: Understanding statistical methods through elliptical geometry. *Statistical Science* 28: 1–39.

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.

Gibbons, C. E., Serrato, J. C. S. and Urbancic, M. B. (2018). Broken or fixed effects? *Journal of Econometric Methods* 8.

Graetz, G. (2017). Human capital, signaling, and employer learning. what insights do we gain from regression discontinuity designs? *CESifo Working Paper Series* .

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica* : 1–22.

Hakenes, H. and Peitz, M. (2009). Umbrella branding and external certification. *European Economic Review* 53: 186–196.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association* 81: 945–960.

Huntington-Klein, N. (2020). Human capital versus signaling is empirically unresolvable. *Empirical Economics* : 1–33.

Ichimura, H. and Taber, C. (2001). Propensity-score matching with instrumental variables. *American Economic Review* 91: 119–124.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika* 87: 706–710.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142: 615–635.

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

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

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

Kanemoto, Y. (1988). Hedonic prices and the benefits of public projects. *Econometrica* : 981–989.

Keele, L. and Titiunik, R. (2016). Natural experiments based on geography. *Political Science Research and Methods* 4: 65.

Keele, L. J. and Titiunik, R. (2015). Geographic boundaries as regression discontinuities. *Political Analysis* 23: 127–155.

Khan, S. and Tamer, E. (2010). Irregular identification, support conditions, and inverse weight estimation. *Econometrica* 78: 2021–2042.

Krishna, V. and Morgan, J. (2001). A model of expertise. *Quarterly Journal of Economics* 116: 747–775.

Kuminoff, N. V. and Pope, J. C. (2014). Do “capitalization effects” for public goods reveal the public’s willingness to pay? *International Economic Review* 55: 1227–1250.

Kuminoff, N. V., Smith, V. K. and Timmins, C. (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature* 51: 1007–62.

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

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48: 281–355.

Li, Q. and Racine, J. S. (2007). *Nonparametric econometrics: Theory and practice*. Princeton University Press.

Liang, A. and Mu, X. (2020). Complementary information and learning traps. *Quarterly Journal of Economics* 135: 389–448.

Ma, X. and Wang, J. (2020). Robust inference using inverse probability weighting. *Journal of the American Statistical Association* 115: 1851–1860.

Malter, D. (2014). On the causality and cause of returns to organizational status: Evidence from the grands crus classés of the Médoc. *Administrative Science Quarterly* 59: 271–300.

McDevitt, R. C. (2014). "A" business by any other name: Firm name choice as a signal of firm quality. *Journal of Political Economy* 122: 909–944.

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

Mérel, P., Ortiz-Bobea, A. and Paroissien, E. (2020). How big is the "lemons" problem? historical evidence from French wines. *SMART working papers* .

Michalopoulos, S. and Papaioannou, E. (2018). Spatial patterns of development: A meso approach. *Annual Review of Economics* 10: 383–410.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics* : 380–391.

Mullainathan, S., Schwartzstein, J. and Shleifer, A. (2008). Coarse thinking and persuasion. *Quarterly Journal of Economics* 123: 577–619.

Neeman, Z., Öry, A. and Yu, J. (2019). The benefit of collective reputation. *RAND Journal of Economics* 50: 787–821.

Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers. *European Economic Review* 39: 1697–1716.

Olszewski, W. and Wolinsky, A. (2016). Search for an object with two attributes. *Journal of Economic Theory* 161: 145–160.

Robins, J. M. and Rotnitzky, A. (1995). Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association* 90: 122–129.

Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association* 89: 846–866.

Robinson, P. M. (1988). Root-n-consistent semiparametric regression. *Econometrica* : 931–954.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.

Rothe, C. (2017). Robust confidence intervals for average treatment effects under limited overlap. *Econometrica* 85: 645–660.

Roullier-Gall, C., Lucio, M., Noret, L., Schmitt-Kopplin, P. and Gougeon, R. D. (2014). How subtle is the *terroir* effect? chemistry-related signatures of two *climats de Bourgogne*. *PLoS One* 9: e97615.

Sasaki, Y. and Ura, T. (2018). Estimation and inference for moments of ratios with robustness against large trimming bias. *arXiv preprint arXiv:1709.00981* .

Shapiro, C. (1983). Premiums for high quality products as returns to reputations. *Quarterly Journal of Economics* 98: 659–679.

Słoczyński, T. and Wooldridge, J. M. (2018). A general double robustness result for estimating average treatment effects. *Econometric Theory* 34: 112–133.

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review* 92: 434–459.

Starrett, D. A. (1981). Land value capitalization in local public finance. *Journal of Political Economy* 89: 306–327.

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

Tyler, J. H., Murnane, R. J. and Willett, J. B. (2000). Estimating the labor market signaling value of the GED. *Quarterly Journal of Economics* 115: 431–468.

Uysal, S. D. (2015). Doubly robust estimation of causal effects with multivalued treatments: An application to the returns to schooling. *Journal of Applied Econometrics* 30: 763–786.

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

Wernerfelt, B. (1988). Umbrella branding as a signal of new product quality: An example of signalling by posting a bond. *The RAND Journal of Economics* : 458–466.

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

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.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2016). Should instrumental variables be used as matching variables? *Research in Economics* 70: 232–237.

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.

Zhu, Y. and Dukes, A. (2017). Prominent attributes under limited attention. *Marketing Science* 36: 683–698.

A Online Appendix

A.1 Complete proofs

This section presents the complete proofs for, respectively, the causal interpretation of hedonic evidence (Sections 3 and 4 of the main paper), the regression discontinuity evidence (Section 5 of the paper), and the doubly robust evidence (Section 6 of the paper).

A.1.1 Causal interpretation of hedonic evidence

Under the linear proxy **Assumptions A1.1**, we show that hedonic estimates of Section 3 are asymptotically consistent for the causal statistics described in Section 4.

$$\textbf{A1.1}\ (\forall\, j,k) \quad \mathbb{E}[h(X^*_; j,k) \mid X,d,c] = \mathbb{E}[h(X^*_; j,k) \mid X] = \eta_{jk} + X^\top \boldsymbol{\beta}. \tag{20}$$

The first equality assumes that conditioning on the proxy $X \subset X^*$ is sufficient to ignore the distribution of horizontal and vertical names in the conditional expectations of hedonic functions. The second equality specifies these conditional expectations as a vector of homogeneous coefficients $\boldsymbol{\beta}$ and the saturated form of homogeneous names premiums η_{jk} (Angrist and Krueger, 1999). Under **A1.1** and the law of iterated expectations (LIE), the incremental causal statistics from Equation 3 of the main text can be written as linear forms with:

$$\begin{aligned}
\delta_{jk} &= \mathbb{E}\Big\{\mathbb{E}\Big[h(X^*_; j,k) - h(X^*_; j-1,k) \mid X\Big] \mid d=j, c=k\Big\} = \eta_{jk} - \eta_{(j-1)k}\\
\gamma_{jk} &= \mathbb{E}\Big\{\mathbb{E}\Big[h(X^*_; j,k) - h(X^*_; j,k-1) \mid X\Big] \mid d=j, c=k\Big\} = \eta_{jk} - \eta_{j(k-1)}.
\end{aligned}$$

By substituting these expressions respectively in Equation 4 and Equation 5 of Section 4, we have:

$$AIVP_j = w_{j\cdot}^{-1}\Sigma_{k=0}^{K}\, w_{jk}\,[\eta_{jk} - \eta_{(j-1)k}] \quad \text{and} \quad AIHP_k = w_{\cdot k}^{-1}\Sigma_{j=0}^{J}\, w_{jk}\,[\eta_{jk} - \eta_{j(k-1)}],$$

where w_{jk}, $w_{j\cdot}$, and $w_{\cdot k}$ are the population acreages for respectively plots with names j and k, plots with name j (whatever k), and plots with name k (whatever j).

From Equation 2 with $X \subset X^*$ and LIE, we have $\mathbb{E}(\varepsilon \mid X,d,c) = 0$. As dummies for $d = 0$ and $c = 0$ are omitted and the design matrix is full rank, the OLS estimation of Equation 1 gives:

$$\begin{aligned}
\hat{\delta}_j^{OLS} &\xrightarrow{p} \mathbb{E}\Big[\mathbb{E}(y \mid X, d=j, c=0) - \mathbb{E}(y \mid X, d=0, c=0)\Big] = \eta_{j0} - \eta_{00} = \Sigma_{0<l\leqslant j}\, \delta_{l0}\\
\hat{\gamma}_k^{OLS} &\xrightarrow{p} \mathbb{E}\Big[\mathbb{E}(y \mid X, d=0, c=k) - \mathbb{E}(y \mid X, d=0, c=0)\Big] = \eta_{0k} - \eta_{00} = \Sigma_{0<m\leqslant k}\, \gamma_{0m}\\
\hat{\lambda}_{jk}^{OLS} &\xrightarrow{p} (\eta_{jk} + \eta_{00}) - (\eta_{j0} + \eta_{0k}) = \Sigma_{0<l\leqslant j}\,(\delta_{lk} - \delta_{l0}) = \Sigma_{0<m\leqslant k}\,(\gamma_{jm} - \gamma_{0m}).
\end{aligned}$$

The parameters δ_{l0} and γ_{0m} in the two first lines are the average incremental effects on the treated from Equation 3 in the text. The last line follows from $\hat{\delta}_j^{OLS} + \hat{\gamma}_k^{OLS} + \hat{\lambda}_{jk}^{OLS} \xrightarrow{p} \eta_{jk} - \eta_{00}$. The last two equality show that coefficients of interactions are double differences. In addition, we have $\hat{\delta}_j^{OLS} - \hat{\delta}_{j-1}^{OLS} + \hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{(j-1)k}^{OLS} \xrightarrow{p} \eta_{jk} - \eta_{(j-1)k}$ and $\hat{\gamma}_k^{OLS} - \hat{\gamma}_{k-1}^{OLS} + \hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j(k-1)}^{OLS} \xrightarrow{p} \eta_{jk} - \eta_{j(k-1)}$.

Hence, the average incremental vertical and horizontal premiums of Equation 4 and Equation 5 of the text can be consistently estimated from OLS estimation and population data by:

$$\widehat{AIVP}_j^{ols} = \hat{\delta}_j^{OLS} - \hat{\delta}_{j-1}^{OLS} + w_{j\cdot}^{-1}\Sigma_{k=0}^{K} w_{jk} \, [\hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{(j-1)k}^{OLS}]$$
$$\widehat{AIHP}_k^{ols} = \hat{\gamma}_k^{OLS} - \hat{\gamma}_{k-1}^{OLS} + w_{\cdot k}^{-1}\Sigma_{j=0}^{J} w_{jk} \, [\hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j(k-1)}^{OLS}]$$

This result allows us to compare the descriptive evidences with the causal results presented in Sections 5 and 6 of the main text. More precisely, the coefficients reported in the WLS row of Table 3 and in the WLS column of Table A.4 in this OA are estimated from weighted least square estimation of Equation 1 with the weights of Gibbons et al. (2018) [equation (3), page 4]:

$$\omega_{i|j}^{OLS} = \left[\hat{\mathbb{V}}(\tilde{d}_l \mid l = j)\right]^{-1/2} \quad \text{and} \quad \omega_{i|k}^{OLS} = \left[\hat{\mathbb{V}}(\tilde{c}_m \mid m = k)\right]^{-1/2}.$$

With this procedure, we obtain $\widehat{AIVP}_j^{ols}$ and $\widehat{AIHP}_k^{ols}$ directly from the OLS estimation and we cluster standard errors more easily according to the formula presented in Gibbons et al. (2018).

Under previous Assumptions **A1.1**, shape restrictions and decomposition from Section 4 can be assessed from the hedonic OLS coefficients with full interactions (Equation 1). The following inequalities are obtained by substituting previous OLS estimators from the definitions of the text.

The **vertical monotonicity** of Equation 6 implies $\forall j, j', k$:

$$\hat{\delta}_j^{OLS} - \hat{\delta}_{j'}^{OLS} \lesseqgtr -(\hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j'k}^{OLS}).$$

The **horizontal monotonicity** of Equation 7 implies $\forall j, k, k'$ such that:

$$\hat{\gamma}_k^{OLS} - \hat{\gamma}_{k'}^{OLS} \lesseqgtr -(\hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{jk'}^{OLS}).$$

The **complementarity** of Equation 9 implies $\forall j, j', k, k'$:

$$\hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j'k}^{OLS} \geqslant \hat{\lambda}_{jk'}^{OLS} - \hat{\lambda}_{j'k'}^{OLS} \quad \text{or} \quad \hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{jk'}^{OLS} \geqslant \hat{\lambda}_{j'k}^{OLS} - \hat{\lambda}_{j'k'}^{OLS}.$$

The estimated **decomposition** of the full capitalized value of vineyards is then:

$$\begin{aligned}\hat{\Pi}^{OLS} =& \Sigma_{j=0}^{J} \Sigma_{k=0}^{K} w_{jk}(\hat{\delta}_j^{OLS} - \hat{\delta}_{j^*}^{OLS} + \hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j^*k}^{OLS}) \\ &+ \Sigma_{j=0}^{J} \Sigma_{k=0}^{K} w_{jk}(\hat{\gamma}_k^{OLS} - \hat{\gamma}_{k^*}^{OLS} + \hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{jk^*}^{OLS}) \Leftrightarrow \\ &+ \Sigma_{j=0}^{J} \Sigma_{k=0}^{K} w_{jk}(\hat{\lambda}_{j^*k}^{OLS} + \hat{\lambda}_{jk^*}^{OLS} - \hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j^*k^*}^{OLS})\end{aligned} \qquad \begin{aligned}\hat{\Pi}^{OLS} =& \Sigma_{j=0}^{J} w_{j\cdot} \, (\hat{\delta}_j^{OLS} - \hat{\delta}_{j^*}^{OLS}) \, \Sigma_{k=0}^{K} w_{jk}(\hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j^*k}^{OLS}) \\ &+ \Sigma_{k=0}^{K} w_{\cdot k}(\hat{\gamma}_k^{OLS} - \hat{\gamma}_{k^*}^{OLS}) \, \Sigma_{j=0}^{J} w_{jk}(\hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{jk^*}^{OLS}) \\ &+ \Sigma_{j=0}^{J} \Sigma_{k=0}^{K} w_{jk}(\hat{\lambda}_{j^*k}^{OLS} + \hat{\lambda}_{jk^*}^{OLS} - \hat{\lambda}_{jk}^{OLS} - \hat{\lambda}_{j^*k^*}^{OLS}).\end{aligned}$$

By comparing this last decomposition with shape restrictions for $j' = j^*$ and $k' = k^*$, it appears that each term of the vertical (resp. horizontal) part of the decomposition is positive if the vertical (resp. horizontal) monotonicity is verified. Moreover, the third interaction term of the decomposition is negative if the names are complementary.

A.1.2 Details for Spatial Regression Discontinuity evidence

The observed vineyard price y from Equation 2 of the main text can be rewritten as the sum of counter-factual hedonic functions (with $d_l \equiv \mathbf{1}[d = l]$ and $c_m \equiv \mathbf{1}[c = m]$):

$$y = \Sigma_{l=0}^{J} \Sigma_{m=0}^{K} d_l \cdot c_m \cdot h(X^*, l, m) + \varepsilon,$$

which can be substituted in each SRD limit of Equation 12 of the main text:

$$\begin{aligned}
&\lim_{\epsilon \to 0^+} \mathbb{E}[y \mid r_{j|k}(V) = \epsilon, c = k] \\
&\quad = \lim_{\epsilon \to 0^+} \mathbb{E}[\Sigma_{l=0}^{J} \Sigma_{m=0}^{K} d_l \cdot c_m \cdot h(X^*, l, m) \mid r_{j|k}(V) = \epsilon, c = k] + \lim_{\epsilon \to 0^+} \mathbb{E}[\varepsilon \mid r_{j|k}(V) = \epsilon, c = k].
\end{aligned}$$

The second term is zero by LIE and spatial ignorability (SRD.2):

$$\mathbb{E}\Big\{ \underbrace{\mathbb{E}[\varepsilon \mid X^*, V, d, c]}_{=0} \mid r_{j|k}(V) = \epsilon, c = k] \Big\}.$$

This leaves the first term rewritten as:

$$\begin{aligned}
\lim_{\epsilon \to 0^+} \mathbb{E}[y \mid r_{j|k}(V) = \epsilon, c = k] &= \lim_{\epsilon \to 0^+} \Sigma_{l=0}^{J} \Sigma_{m=0}^{K} \mathbb{E}[d_l \cdot c_m \cdot h(X^*, l, m) \mid r_{j|k}(V) = \epsilon, c = k] \\
&= \lim_{\epsilon \to 0^+} \mathbb{E}[h(X^*, j, k) \mid r_{j|k}(V) = \epsilon, c = k] \\
&= \lim_{\epsilon \to 0^+} \mathbb{E}[h(X^*, j, k) \mid \ell_{j|k}(V) = \epsilon, c = k] \\
&= \lim_{\epsilon \to 0^+} \mathbb{E}[h(X^*, j, k) \mid \ell_{j|k}(V) = \epsilon] = \mathbb{E}[h(X^*, j, k) \mid V = \mathbf{v}_{j|k}].
\end{aligned}$$

In the right hand side of the first equality, $d_l \cdot c_m$ is different from zero only when $l = j$ and $m = k$, which is implicitly the case when $r_{j|k}(V) > 0$. The second equality simplifies with the definition of the running functions (Equation 10 of the main text). Moving from the third equality to the fourth requires spatial continuity of SRD.1 and the last equality comes from the fact that plot $\ell_{j|k}(V) = 0$ are such that $V = \mathbf{v}_{j|k}$. The proof of $\lim_{\epsilon \to 0^-} \mathbb{E}[y \mid r_{j|k}(V) = \epsilon, c = k] = \mathbb{E}[h(X^*, j-1, k) \mid V = \mathbf{v}_{j|k}]$ is symmetric and allows obtaining Equation 12: $\delta_{j|k}^{SRD} = \mathbb{E}[h(X^*, j, k) - h(X^*, j-1, k) \mid V = \mathbf{v}_{j|k}]$.

The proof of Equation 13 of the main text follows Lemma 1 of Cattaneo et al. (2016):

$$\begin{aligned}
&\lim_{\epsilon \to 0^+} \mathbb{E}[y \mid r_j(V) = \epsilon] = \lim_{\epsilon \to 0^+} \Sigma_{k=0}^{K} \mathbb{E}[c_k \cdot h(X^*, j, k) \mid \Sigma_{m=0}^{K} c_m \ell_{j|m}(V) = \epsilon] \\
&\quad = \lim_{\epsilon \to 0^+} \Sigma_{k=0}^{K} \mathbb{E}[h(X^*, j, k) \mid \ell_{j|k}(V) = \epsilon, c = k] \cdot \mathbb{P}(c = k \mid \Sigma_{m=0}^{K} c_m \ell_{j|m}(V) = \epsilon,).
\end{aligned}$$

From the chain rule, we have:

$$\mathbb{P}(c = k \mid \Sigma_{m=0}^{K} c_m \ell_{j|m}(V) = \epsilon) = \frac{\mathbb{P}(\Sigma_{m=0}^{K} c_m \ell_{j|m}(V) = \epsilon \mid c = k)}{\mathbb{P}(\Sigma_{m=0}^{K} c_m \ell_{j|m}(V) = \epsilon)} \cdot \mathbb{P}(c = k)$$

Hence, we swap the limit and the sum operators, and use our previous result to obtain:

$$\begin{aligned}\lim_{\epsilon\to 0^+}\mathbb{E}[y \mid r_j(V)=\epsilon] &= \sum_{k=0}^{K}\frac{\mathbb{P}(V=\mathbf{v}_{j|k}\mid c=k)}{\mathbb{P}(V=\mathbf{v}_j)}\cdot\mathbb{E}[h(X^*_i j,k)\mid V=\mathbf{v}_{j|k}]\cdot\mathbb{P}(c=k)\\ &= \sum_{k=0}^{K}\frac{f_{V|c}(\mathbf{v}_{j|k}\mid c=k)}{f_V(\mathbf{v}_j)}\cdot\mathbb{E}[h(X^*_i j,k)\mid V=\mathbf{v}_{j|k}]\cdot\mathbb{P}(c=k)\end{aligned}$$

from which we recover the weighting scheme noted ω_{jk}^{SRD} in the main text. Symmetrically, the continuity of the distribution of plots on each side of borders concludes the proof:

$$\lim_{\epsilon\to 0^+}\mathbb{E}[y \mid r_j(V)=\epsilon] = \Sigma_{k=0}^{K}\,\omega_{jk}^{SRD}\cdot\mathbb{E}[h(X^*_i j-1,k)\mid V=\mathbf{v}_{j|k}]$$

Vertically pooled SRD estimates δ_j^{SRD} give a local estimation of $AIVP_j$ for $j>0$ where *communes* are averaged according the the size of their vertical borders and the densities of vineyards sales around these borders instead of population acreages. We also report in Table 3 of the main paper pooled vertical SRD estimations with *commune* fixed effects as covariates (Calonico et al., 2019) and pooled vertical SRD on price deviations from *commune* averages using the log of $\tilde{y}\equiv y-\overline{y}_c$ as the outcome variable ($\overline{y}_c$ is the average vineyard price for the *commune* c). These estimators allow to control for unobserved effects to verify the robustness of raw pooled estimators.

Assessing shape restrictions from SRD estimates requires additional **Assumptions A1.2** in addition to SRD.1 and SRD.2 of the main text:

$$\textbf{A1.2}\ (\forall j,k)\quad \mathbb{E}[h(X^*_i j,k)\mid V]=\tau_{jk}+g(V), \tag{21}$$

with parsimonious interactions such that $\tau_{0k}=\tau_0$, $\tau_{1k}=\tau_1$, $\tau_{4k}=\tau_4$, and $\tau_{2k}=\tau_{3k}=\tau'_k$. This set of assumptions corresponds to the additive separation between names premiums and the un-specified spatial patterns of *terroir*. Under assumption SRD from the main text and **A1.2**, we have:

$$\begin{aligned}\delta_{jk}^{SRD} &= \tau_{jk}-\tau_{(j-1)k}=\delta_{jk} \quad\text{and}\quad \gamma_{kj}^{SRD}=\tau_{jk}-\tau_{j(k-1)}=\gamma_{jk}\\ \delta_j^{SRD} &= \Sigma_{k=0}^{K}\omega_{jk}^{SRD}\delta_{jk}\approx AIVP_j \quad\text{and}\quad \gamma_k^{SRD}=\Sigma_{j=0}^{J}\omega_{jk}^{SRD}\gamma_{jk}\approx AIHP_k.\end{aligned}$$

The last two equality are some approximation under the equivalence between area and border weighting schemes (see Figure A.16 of this Online Appendix).

Hence, the vertical and the horizontal parts of the decomposition are respectively:

$$\begin{aligned}&w_1.\delta_1^{SRD}+w_2.(\delta_1^{SRD}+\delta_2^{SRD})+w_3.(\delta_1^{SRD}+\delta_2^{SRD}+\delta_3^{SRD})+w_4.(\delta_1^{SRD}+\delta_2^{SRD}+\delta_3^{SRD}+\delta_4^{SRD})\\ &\text{and}\quad \Sigma_{k=0}^{K}(w_{2k}+w_{3k})\times(\delta_{2k}^{SRD}-\delta_{2k^*}^{SRD}).\end{aligned}$$

Note that neither the interaction, nor the total value of wine names can be computed from SRD estimations, so that the estimation of the full price decomposition cannot be recovered.

A.1.3 Details for Doubly Robust evidence

Firstly, consider the penalized estimation of the ordered generalized additive model about the designation of vertical names (DR.1), which allows us to compute generalized propensity scores in the DR approach. The matrix $B(X, V)$ of the main text is specified through additive low rank isotropic smoothers of observed variables (Wood, 2017), which are uni-dimensional for each biophysical variables x_u with $u = 1, \dots, \overline{u}$, and bi-dimensional for the geographic coordinates (v_1, v_2). The maximum allowed transformation basis for each function $b_u(\cdot)$ is shrinked endogenously in the penalized estimation procedure. We set:

$$B(\mathbf{x}_i, \mathbf{v}_i)^\top \psi \equiv \Sigma_{u=1}^{\overline{u}}\, b_u(x_{ui})^\top \psi_u + b_0(v_{1i}, v_{2i})^\top \psi_0.$$

Even with this additive structure, the spline transformations of each variable imply a high number of coefficients to estimate and lead to the curse of dimensionality as in all high-dimensional method. Hence, the variance of errors 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 smoothness of a given variable or group of variable is assessed by the effective degrees of freedom that account for the endogenous penalization of any given dimension reduction (Wood, 2017, p.273). To take into account the uncertainty related to the unknown smoothing parameter, Wood et al. (2016) provide the corrections for inference and goodness-of-fit measures.

Secondly, consider the average counter-factual price that vineyards designated as $d = l, c = m$ would have if they were designated as $d = j$ and $c = k$. We have (for $w_{lm} > 0$):

$$\mathbb{E}[h(X^*_; j, k) \mid d = l, c = m] = \; w_{lm}^{-1}\, \mathbb{E}[d_l \cdot c_m \cdot h(X^*_; j, k)]$$

We can then develop further, following closely Słoczyński and Wooldridge (2018):

$$\begin{aligned}\mathbb{E}[d_l \cdot c_m \cdot h(X^*_; j, k)] &= \mathbb{E}\Big\{d_l \cdot c_m \cdot \mathbb{E}[h(X^*_; j, k) \mid X, V, d, c]\Big\}\\ &= \mathbb{E}\Big\{d_l \cdot c_m \cdot \mathbb{E}[h(X^*_; j, k) \mid X, V]\Big\}\\ &= \mathbb{E}\Big\{\mathbb{P}(d = l, c = m \mid X, V) \cdot \mathbb{E}[h(X^*_; j, k) \mid X, V]\Big\}.\end{aligned}$$

The first equality comes from LIE, the second comes from the assumption of spatial continuity SRD.1, and the third is again the LIE. We have also:

$$\begin{aligned}\mathbb{E}[d_j \cdot c_k \cdot y \mid X, V] &= \mathbb{E}\Big\{d_j \cdot c_k \cdot [\Sigma_{j'=0}^{J}\Sigma_{k'=0}^{K}\, d_{j'} \cdot c_{k'} \cdot h(X^*_; j', k') + \varepsilon] \mid X, V\Big\}\\ &= \mathbb{E}\Big\{d_j \cdot c_k \cdot \mathbb{E}[h(X^*_; j, k) \mid X, V, d, c] \mid X, V\Big\}\\ &= \mathbb{P}(d = j, c = k \mid X, V) \cdot \mathbb{E}[h(X^*_; j, k) \mid X, V].\end{aligned}$$

The first equality comes from the structural model, the second from spatial ignorability SRD.2, and the third from spatial continuity SRD.1 with LIE. By expressing $\mathbb{E}[h(X^*_; j, k) \mid X, V]$, plugging it in the equation above, and using LIE a last time, we obtain Equation 15 of the main text:

$$\mathbb{E}[h(X^*_; j, k) \mid d = l, c = m] = w_{lm}^{-1}\, \mathbb{E}\left[\frac{\mathbb{P}(d = l, c = m \mid X, V)}{\mathbb{P}(d = j, c = k \mid X, V)}\, d_j \cdot c_k \cdot y\right]$$

The regression adjustment part of the DR proof follows from Assumption DR.2 of the main text.

Thirdly, we keep the same structure of the proof to derive the pooled DR estimations:

$$\begin{aligned}\mathbb{E}[h(X^*_;j,c) \mid d=l] &= w_{l\cdot}^{-1}\ \mathbb{E}\left[d_l \cdot \Sigma_{k=0}^K\ c_k \cdot h(X^*_;j,k)\right] \\ &= w_{l\cdot}^{-1}\ \mathbb{E}\left\{d_l \cdot \mathbb{E}\left[h(X^*_;j,m) \mid X,V,d,c=m\right]\right\} \quad (\forall m) \\ &= w_{l\cdot}^{-1}\ \mathbb{E}\left\{\mathbb{E}(d_l \mid X,V,c=m) \cdot \mathbb{E}\left[h(X^*_;j,m) \mid X,V\right]\right\}.\end{aligned}$$

The first equality is definitional, the second comes from LIE, and the third from spatial continuity SRD.1 and LIE. It is crucial to note that the second equality is verified for all $m = 0,\ldots,K$, according to which the choice of m is arbitrary. We have also (still for any m):

$$\begin{aligned}\mathbb{E}[d_j \cdot y \mid X,V] &= \mathbb{E}\left\{d_j \cdot \mathbb{E}\left[h(X^*_;j,m) \mid X,V,d,c=m\right] \mid X,V\right\} \quad (\forall m) \\ &= \mathbb{E}\left\{\mathbb{E}(d_j \mid X,V,c=m) \cdot \mathbb{E}\left[h(X^*_;j,m) \mid X,V\right] \mid X,V\right\} \\ &= \mathbb{E}[\mathbb{E}(d_j \mid X,V,c=m) \mid X,V] \cdot \mathbb{E}\left[h(X^*_;j,m) \mid X,V\right].\end{aligned}$$

Combining the two previous results by substituting $\mathbb{E}\left[h(X^*_;j,m) \mid X,V\right]$ gives:

$$\mathbb{E}[h(X^*_;j,c) \mid d=l] = w_{l\cdot}^{-1}\ \mathbb{E}\left[\frac{\mathbb{P}(d=l \mid X,V,c=m)}{\mathbb{P}(d=j \mid X,V,c=m)}\ d_j \cdot y\right]$$

where the choice of m is again arbitrary. From vineyard data at hand, we find that averaging the probability ratio among *communes* m is the most efficient method (while the differences between the m are not significant). The DR weights become (as for Equation 19 in the main text):

$$\tilde{\omega}_{j|l}^{DR}(X,V) = (K+1)^{-1}\sum_{m=0}^{K}\frac{\Phi[\alpha_l+\mu_m-L(X,V)^\top\psi]-\Phi[\alpha_{l-1}+\mu_m-L(X,V)^\top\psi]}{\Phi[\alpha_j+\mu_m-L(X,V)^\top\psi]-\Phi[\alpha_{j-1}+\mu_m-L(X,V)^\top\psi]}.$$

For the regression adjustment part of the DR estimation, we have (under DR.2):

$$\begin{aligned}\mathbb{E}[h(X^*_;j,c) \mid d=l] &= \mathbb{E}\left\{\Sigma_{k=0}^K\ c_k \cdot \mathbb{E}[h(X^*_;j,k) \mid X,V] \mid d=l\right\} \\ &= \Sigma_{k=0}^K\ (w_{lk}/w_{l\cdot})\left\{\kappa_{jk}+\rho_{jk}\cdot\mathbb{E}[B(X,V)^\top\psi \mid d=l,c=k]\right\},\end{aligned}$$

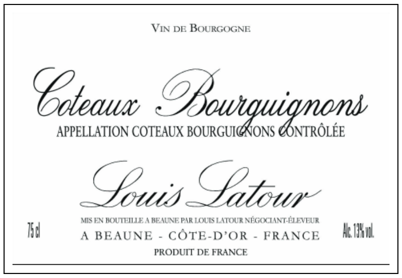
from which we recover:

$$AIVP_j = \Sigma_{k=0}^K\ (w_{jk}/w_{j\cdot})\left\{(\kappa_{jk}-\kappa_{(j-1)k})+(\rho_{jk}-\rho_{(j-1)k})\cdot\mathbb{E}[B(X,V)^\top\psi \mid d=j,c=k]\right\}.$$

Figure A.1: **Exemples of Wine Labels from the *Côte d'Or* Region under Study.**

Notes: The labels below are from the same producer-trader (*négociant-éleveur*) Louis Latour based on the *commune* of *Beaune* (https://www.louislatour.com/en/wines/). Its portfolio of about 100 names from the *Côte d'Or* is among the more diverse, other wine producers have on average about 10 different wine names. The two first panels (a) and (b) present labels from the two lowest vertical levels *Coteaux* and *Région* without the mention of the horizontal names. The following six panels present two horizontal levels for each of the three remaining highest vertical levels.

(a) *Coteaux* vertical level

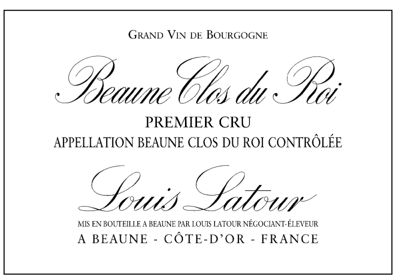


(b) *Région* vertical level

(c) *Village* level from *Gevrey-Chambertin*

(d) *Village* level from *Beaune*

(e) *Premier cru* level from *Beaune*



(f) *Premier cru* level from *Chambolle-Musigny*

(g) *Grand cru* level from *Gevrey-Chambertin*

(h) *Grand cru* level from *Aloxe-Corton*

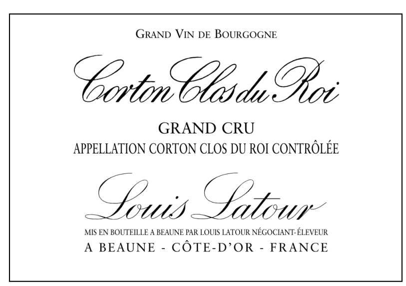


Figure A.2: **Topography and Wine Names for the Vineyards of the *Côte d'Or*.**

Notes: The elevation on the left side map is discretized in eight classes of 50 m intervals. From the East to the West, the elevation is first convex then concave, so the highest slopes are observed for average elevations. On the right side map, highest vertical wine names are located on these highest slopes. The precision of the vertical delineations is such that best vineyards, classified as *Grand cru*, are not visually well-separated from just below *Premier cru*. This map also shows the names of the 31 *communes* of the area, which correpond to the horizontal dimension of wine names.

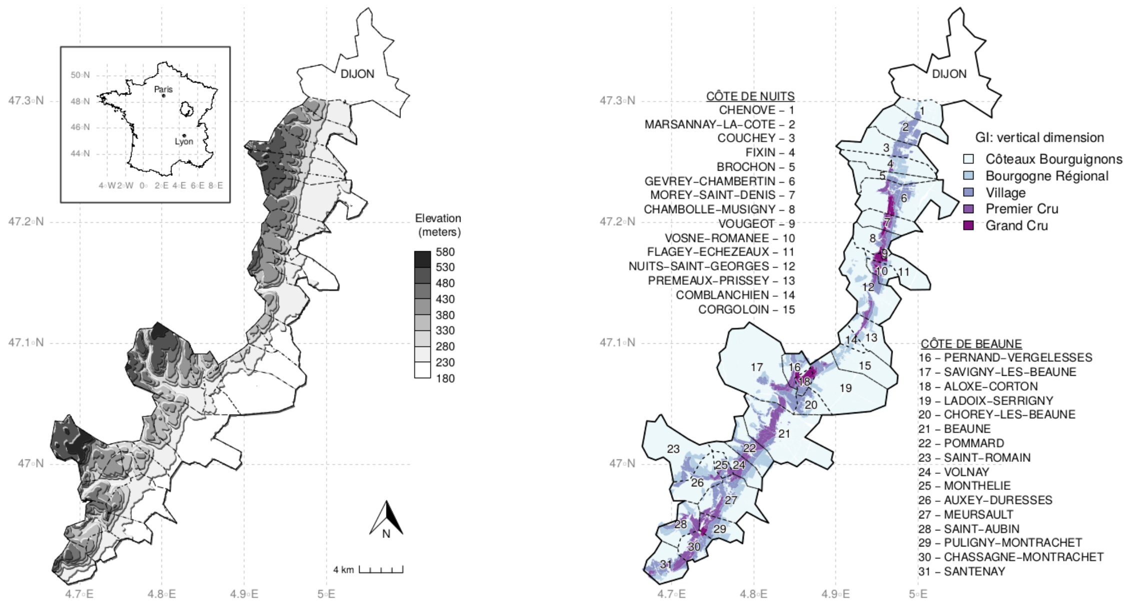


Table A.1: **Summary Statistics for Main Variables of Plot and Sale Data.**

	Plot Data (N = 59 967)						*Sale Data* (N = 8 987)					
Variables	Min	Q1	Q2	Mean	Q3	Max	Min	Q1	Q2	Mean	Q3	Max
Price (M euro/ ha)							0	0.04	0.14	0.35	0.39	23.39
Surface (ha)	0	0.05	0.11	0.19	0.22	9.61	0	0.06	0.13	0.2	0.23	8.2
Longitude (WGS84)	4.66	4.74	4.81	4.84	4.96	5.02	4.67	4.76	4.81	4.84	4.95	5
Latitude (WGS84/ 10)	4.69	4.7	4.7	4.71	4.72	4.73	4.69	4.7	4.7	4.71	4.72	4.73
Elevation (100 m)	2.1	2.41	2.71	2.86	3.18	5.05	2.1	2.36	2.61	2.77	3.07	4.5
Slope (degree)	0	1.54	3.43	5.73	8.69	36.97	0.06	1.42	2.72	4.95	7.67	19.99
Solar Radiation (M Joule)	0.58	1.05	1.06	1.06	1.08	1.23	0.83	1.05	1.06	1.06	1.07	1.21
Exposition (degree)	-0.01	0.93	1.22	1.34	1.64	3.58	0.02	0.92	1.22	1.31	1.62	3.55

Horizontal / Vertical names	*Coteau*	*Région*	*Village*	*P.cru*	*G.cru*	*Coteau*	*Région*	*Village*	*P.cru*	*G.cru*
Marsannay	362	1523	2216	0	0	26	158	387	0	0
Fixin	117	627	1128	90	0	23	93	157	1	0
Gevrey-Chambertin	176	389	2734	567	391	12	62	434	99	24
Morey-Saint-Denis	657	177	688	314	146	146	17	87	26	3
Chambolle-Musigny	27	199	957	507	128	8	51	161	66	8
Vosne-Romanee-Vougeot	333	309	864	360	482	52	36	69	35	67
Nuits-Saint-Georges	326	575	1349	440	0	27	60	195	54	0
Cote-de-Nuits-Village	359	712	618	0	0	33	90	118	0	0
Aloxe-Corton-Ladoix	301	1259	1968	517	586	12	153	399	69	90
Savigny-Chorey-les-Beaune	325	664	1366	301	0	18	110	362	40	0
Beaune-Cote-de-Beaune	261	411	806	877	0	6	29	110	182	0
Pommard	435	1566	1147	433	0	58	439	314	69	0
Monthelie-Volnay	391	586	1575	708	0	84	135	453	118	0
Auxey-Duresses-Saint-Romain	3050	1755	1797	114	0	186	58	208	15	0
Meursault	1176	988	1373	412	0	163	204	208	60	0
Puligny-Montrachet	226	985	598	556	92	89	282	138	110	21
Saint-Aubin	816	1799	914	1119	0	63	203	102	164	0
Chassagne-Montrachet	201	290	1163	913	81	14	49	193	102	13
Santenay	231	293	1071	490	0	15	38	109	45	0

Notes: The top panel displays summary statistics for continuous variables both for the population of plots (left part) and for the sample of sales (right part). The bottom panel reports the interacted frequencies between the horizontal names (in rows) and the vertical levels (in columns) for each data set. The horizontal names reported in the Table are the 19 groups of *communes* aggregated from the 31 administrative *communes* on the basis of pairwise homogeneity tests from hedonic models. *Marsannay* counts the *communes* of *Chenove*, *Marsannay-la-Côte*, and *Couchey*. *Fixin* counts the *communes* of *Fixin* and part of *Brochon*. *Vosne-Romanée-Vougeot* counts *Vosne-Romanée*, *Vougeot*, and *Flagey-Echezeaux*. *Cote-de-Nuits-Village* counts *Premaux-Prissey*, *Comblanchien*, and *Corgoloin*. *Aloxe-Corton-Ladoix* counts *Pernand-Vergeless*, *Aloxe-Corton*, and *Ladoix-Serrigny*. *Savigny-Chorey-les-Beaune* counts *Savigny-les-Beaune* and *Chorey-les-Beaune*. *Monthelie-Volnay* counts *Monthelie* and *Volnay*. *Auxey-Duresse-Saint-Romain* counts *Auxey-Duresse* and *Saint-Romain*.

Table A.2: **Estimated Coefficients from Control Variables in Hedonic Models of Table 1.**

	Dependent variable: logarithm of per-ha vineyard prices					
	(1)	(2)	(3)	(4)	(5)	(6)
Log. of plot size	-0.090^{***}	-0.248^{***}	-0.072^{***}	-0.160^{***}	-0.050^{***}	-0.012
	(0.017)	(0.023)	(0.016)	(0.021)	(0.015)	(0.015)
Occuped plot	0.318^{***}	0.626^{***}	0.252^{***}	0.408^{***}	0.167^{***}	0.159^{***}
	(0.039)	(0.054)	(0.037)	(0.048)	(0.031)	(0.030)
Tenured plot	-0.119^{**}	0.033	-0.107^{**}	0.003	-0.069	-0.102^{**}
	(0.052)	(0.076)	(0.051)	(0.062)	(0.042)	(0.040)
Type of Buyer (ref= Not mentioned)						
Agricultural holding (GFA)	0.390^{***}	0.930^{***}	0.364^{***}	0.808^{***}	0.408^{***}	0.403^{***}
	(0.105)	(0.163)	(0.104)	(0.131)	(0.094)	(0.081)
Agricutural corporation	0.549^{***}	1.119^{***}	0.538^{***}	0.938^{***}	0.537^{***}	0.485^{***}
	(0.101)	(0.161)	(0.100)	(0.126)	(0.091)	(0.078)
Agricultural person	0.111	0.363^{**}	0.111	0.335^{***}	0.149^{*}	0.190^{**}
	(0.097)	(0.155)	(0.096)	(0.121)	(0.088)	(0.075)
Agricultural retired	0.196	0.477	0.144	0.328	0.056	0.036
	(0.219)	(0.360)	(0.215)	(0.344)	(0.227)	(0.221)
Non-agricultural holding	0.278	-0.123	0.241	0.223	0.383^{**}	0.470^{**}
	(0.218)	(0.242)	(0.204)	(0.197)	(0.173)	(0.183)
Non-agricultural corporation	0.705^{***}	1.264^{***}	0.678^{***}	1.025^{***}	0.584^{***}	0.512^{***}
	(0.135)	(0.206)	(0.133)	(0.166)	(0.118)	(0.101)
Non-agricultural person	0.176^{*}	0.424^{***}	0.208^{**}	0.443^{***}	0.309^{***}	0.285^{***}
	(0.106)	(0.164)	(0.104)	(0.129)	(0.096)	(0.082)
Non-agricultural retired	0.067	0.329^{*}	0.158	0.346^{**}	0.318^{***}	0.295^{***}
	(0.134)	(0.183)	(0.127)	(0.153)	(0.121)	(0.101)
Other types	0.267^{**}	0.735^{***}	0.274^{**}	0.649^{***}	0.380^{***}	0.330^{***}
	(0.114)	(0.178)	(0.112)	(0.143)	(0.102)	(0.087)
Type of Seller (ref= Not mentioned)						
Agricultural holding (GFA)	-0.148	0.071	-0.127	-0.089	-0.054	-0.090
	(0.166)	(0.252)	(0.166)	(0.220)	(0.162)	(0.138)
Agricutural corporation	-0.033	0.223	0.032	0.020	0.032	0.011
	(0.159)	(0.230)	(0.157)	(0.204)	(0.154)	(0.128)
Agricultural person	-0.069	0.193	0.010	-0.069	-0.004	-0.023
	(0.148)	(0.214)	(0.148)	(0.191)	(0.147)	(0.122)
Agricultural retired	-0.076	0.082	-0.041	-0.141	-0.037	-0.128
	(0.239)	(0.329)	(0.236)	(0.295)	(0.212)	(0.181)
Non-agricultural holding	-0.598^{**}	-0.690^{*}	-0.428	-0.700^{**}	-0.275	-0.145
	(0.301)	(0.360)	(0.284)	(0.297)	(0.287)	(0.273)
Non-agricultural corporation	-0.219	-0.044	-0.148	-0.297	-0.133	-0.197
	(0.206)	(0.304)	(0.203)	(0.269)	(0.184)	(0.156)
Non-agricultural person	-0.266^{*}	-0.045	-0.181	-0.219	-0.179	-0.256^{**}
	(0.151)	(0.217)	(0.151)	(0.194)	(0.150)	(0.126)
Non-agricultural retired	-0.359^{**}	-0.213	-0.283^{*}	-0.338^{*}	-0.184	-0.258^{**}
	(0.149)	(0.214)	(0.149)	(0.192)	(0.148)	(0.123)
Retrocession SAFER	-0.133	0.457^{**}	-0.019	0.240	0.107	0.040
	(0.149)	(0.214)	(0.149)	(0.192)	(0.147)	(0.123)
Other types	-0.327^{**}	-0.101	-0.254^{*}	-0.310	-0.171	-0.232^{*}
	(0.149)	(0.214)	(0.149)	(0.191)	(0.148)	(0.123)
Controls	X	X	X	X	X	X
Biophysical variables			X	X	X	X
Vertical names	X		X		X	X
Horizontal names		X		X	X	X
Interactions btw names						X
Observations	8,987	8,987	8,987	8,987	8,987	8,987
R^2	0.712	0.372	0.730	0.563	0.785	0.811

Standard errors (reported in parenthesis) are clustered by vineyard sales: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
Notes: Control variables include year of sale, see Figure A.3a. Biophysical variables include elevation, slope, solar radiation, and exposition, in third order polynomials, see Figure A.4 for their marginal effects.

Table A.3: **Estimated Coefficients from Wine Names in Hedonic Models of Table 1.**

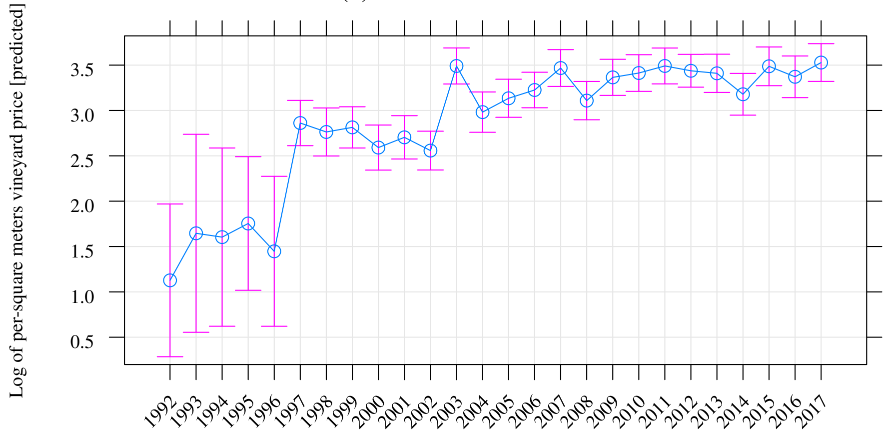
	Dependent variable: logarithm of per-ha vineyard prices					
	(1)	(2)	(3)	(4)	(5)	(6)
Vertical Levels (ref= *Coteaux*)						
Région	0.892***		0.913***		0.922***	0.960***
	(0.062)		(0.062)		(0.067)	(0.077)
Village	2.569***		2.634***		2.639***	2.572***
	(0.059)		(0.063)		(0.068)	(0.091)
Premier cru	3.409***		3.469***		3.312***	3.538***
	(0.065)		(0.072)		(0.076)	(0.145)
Grand cru	4.562***		4.606***		4.482***	4.802***
	(0.114)		(0.119)		(0.120)	(0.191)
Horizontal Names (ref= *Marsannay*)						
Fixin		0.243**		0.235**	0.417***	0.299***
		(0.105)		(0.102)	(0.099)	(0.080)
Gevrey-Chambertin		1.479***		1.752***	1.009***	1.020***
		(0.086)		(0.079)	(0.074)	(0.102)
Morey-Saint-Denis		0.113		0.766***	0.934***	−0.020
		(0.174)		(0.158)	(0.115)	(0.202)
Chambolle-Musigny		1.565***		1.839***	1.298***	1.118***
		(0.140)		(0.125)	(0.092)	(0.136)
Vosne-Romanee-Vougeot		1.423***		2.257***	1.454***	0.374*
		(0.165)		(0.143)	(0.103)	(0.213)
Nuits-Saint-Georges		1.035***		1.808***	1.176***	0.973***
		(0.117)		(0.110)	(0.086)	(0.117)
Cote-de-Nuits-Village		0.017		0.901***	0.506***	0.062
		(0.113)		(0.112)	(0.109)	(0.098)
Aloxe-Corton-Ladoix		0.528***		1.351***	0.293***	−0.071
		(0.099)		(0.103)	(0.090)	(0.117)
Savigny-Chorey-les-Beaune		0.474***		1.540***	0.394***	0.305***
		(0.087)		(0.094)	(0.082)	(0.082)
Beaune-Cote-de-Beaune		1.039***		2.015***	0.512***	−0.191
		(0.109)		(0.127)	(0.103)	(0.284)
Pommard		0.813***		1.746***	1.228***	0.642***
		(0.096)		(0.094)	(0.083)	(0.093)
Monthelie-Volnay		0.776***		1.562***	0.891***	0.609***
		(0.102)		(0.104)	(0.087)	(0.117)
Auxey-Duresses-Saint-Romain		−0.601***		0.677***	0.274**	−0.319*
		(0.127)		(0.126)	(0.111)	(0.163)
Meursault		0.511***		1.977***	1.365***	0.590***
		(0.114)		(0.107)	(0.090)	(0.131)
Puligny-Montrachet		0.735***		2.368***	1.389***	0.401***
		(0.127)		(0.126)	(0.098)	(0.150)
Saint-Aubin		0.385***		1.148***	0.502***	−0.383*
		(0.127)		(0.130)	(0.104)	(0.199)
Chassagne-Montrachet		1.973***		3.225***	1.730***	1.291***
		(0.108)		(0.110)	(0.092)	(0.167)
Santenay		0.645***		1.610***	0.625***	0.217
		(0.118)		(0.138)	(0.104)	(0.162)
Controls	X	X	X	X	X	X
Biophysical variables			X	X	X	X
Vertical names	X		X		X	X
Horizontal names		X		X	X	X
Interactions btw names						X
Observations	8,987	8,987	8,987	8,987	8,987	8,987
R^2	0.712	0.372	0.730	0.563	0.785	0.811

Standard errors (reported in parenthesis) are clustered by vineyard sales: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
Notes: Control variables include year of sale, plot size, type of seller, type of buyer, tenure status, and current occupation of vineyards. They are included in all models, see Table A.2 and Figure A.3a for the details of estimation. Biophysical variables include elevation, slope, solar radiation, and exposition in third order polynomials, see Figure A.4 for their marginal effects. The interaction effects are centered, $\tilde{a}_{jki} \equiv (d_{ji} - n_{jk}/n_j)(c_{ki} - n_{jk}/n_k)$ for $j = 1, 2, 3, 4$ and $k = 1, 2, \ldots, 19$, in order to maintain the interpretation of the main coefficients reported in the Table for model with interaction.

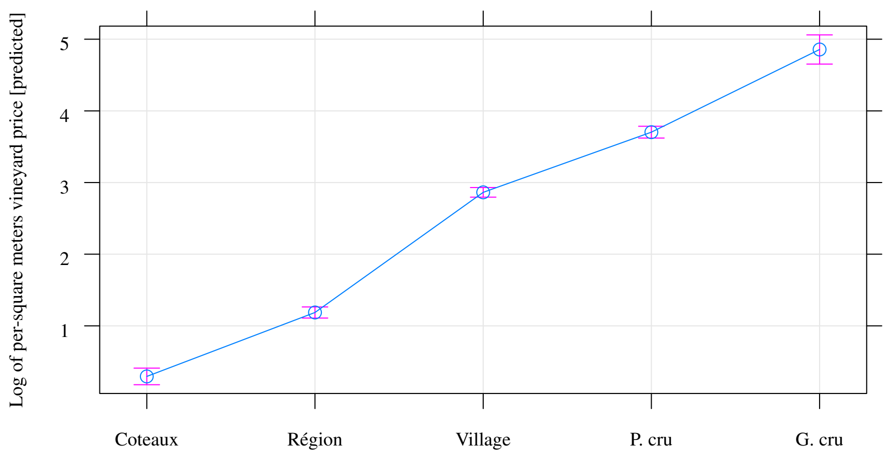
Figure A.3: **Time and Names Coefficients from Hedonic Models (1) and (2) of Table 1.**

Notes: The trend break in 1996 comes from a change in data gathering, controled in the analysis by year fixed effects. Horizontal names of the Panel (c) are ordered from the northernmost (at the left) to the southermost (at the right).

(a) Year Fixed Effects



(b) Vertical Names Fixed Effects



(c) Horizontal Names Fixed Effects

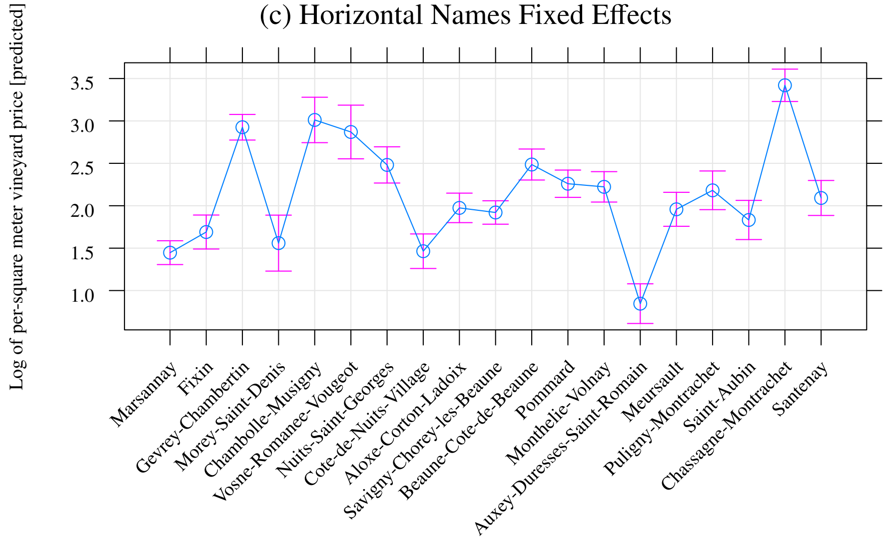
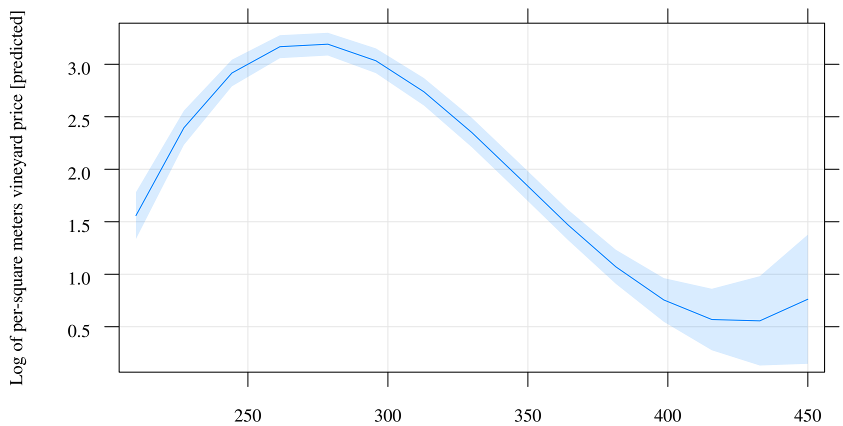


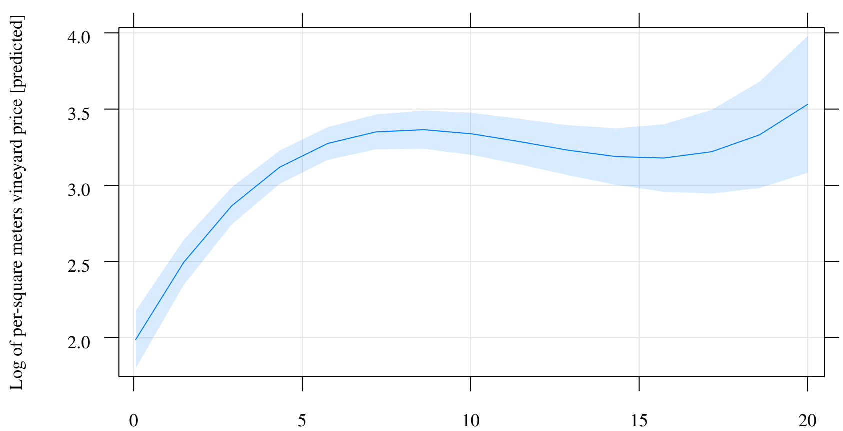
Figure A.4: **Marginal Effects of Biophysical Variables from Model (0) of Table 1.**

Notes: The marginal effects displayed come from third-order polynomials specifications of biophysical variables, with other explanatory variables fixed at their sample means. Pointwise confidence intervals are computed from the variance-covariance matrix of coefficients (Friendly et al., 2013), clustered between vineyard sales.

(a) Elevation Effects [meters]



(b) Slope Effects [degrees]



(c) Solar Radiation [scaled]

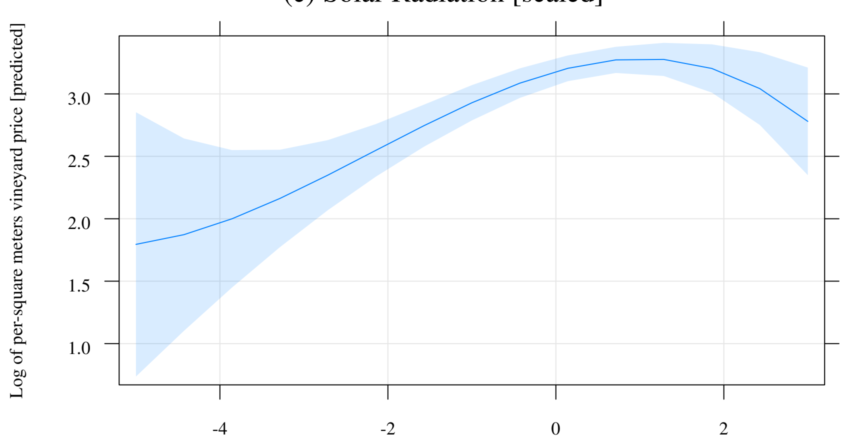


Figure A.5: **Signaling Values from the Hedonic Model with Parsimonious Interactions.**

Notes: The Figure reports on a log y-scale the predicted vineyard price for each existing combination of horizontal (x-axis, from the North to the South) and vertical names (lines and points), with other explanatory variables fixed at their sample means. As explained in the text, the interactions are limited to the intermediate vertical levels (*Village* and *Premier cru*) because the two lowest levels (*Coteaux* and *Régional*) are precluded to put the horizontal names on labels. The *Premier cru* line is discontinous at the two *commune* names that do not have this vertical level. Predicted price for the highest level *Grand cru* is estimated from the seven *communes* that are marked at the top of the Figure.

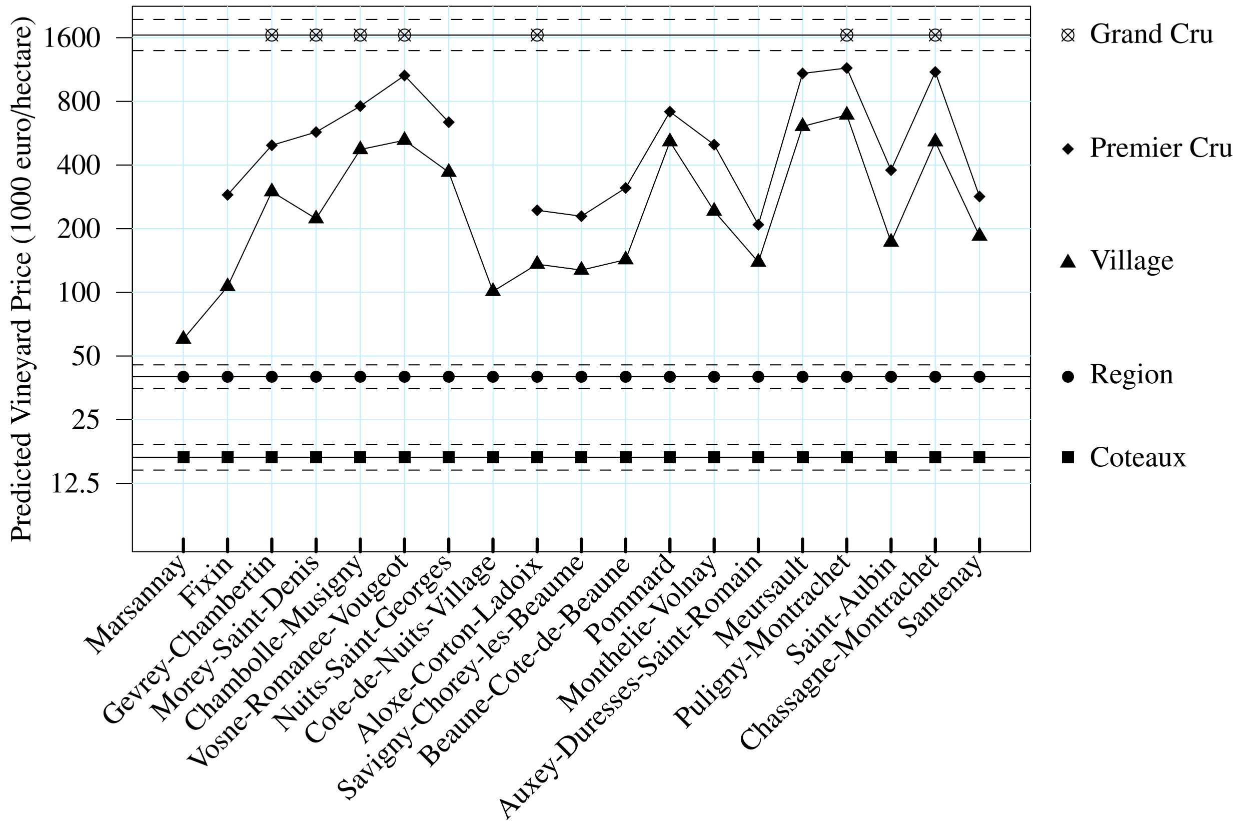
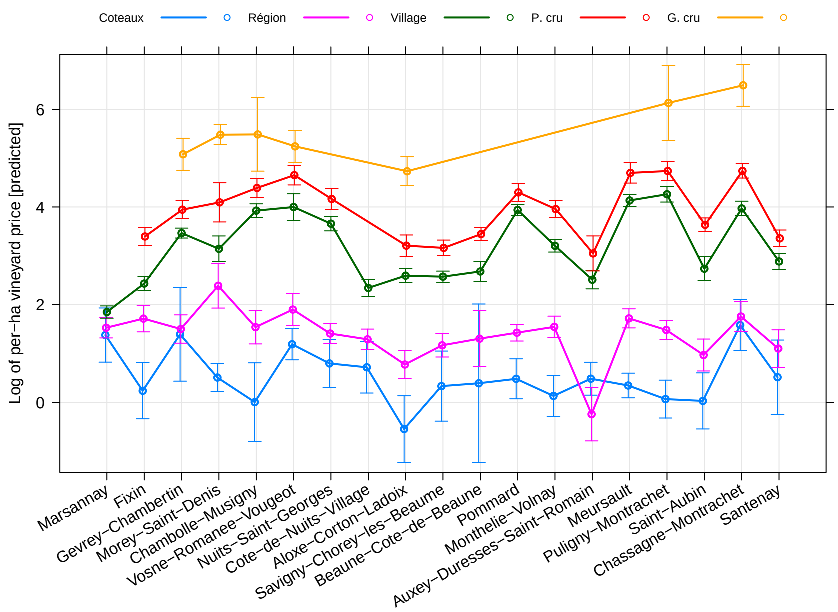


Figure A.6: **Name Premiums from Hedonic Model with Full Interactions.**

Notes: The same names premiums are presented in two different ways. Panel (a) displays horizontal names on the x-axis with a different line for each vertical name, and panel (b) do the opposite. On panel (a), the only (non-significant) gap from vertical monotonicity is for the *Région* of *Auxey-Duresses-Saint-Romain* with a line crossing. On panel (b), horizontal monotonicity can be assessed from plot alignement for each vertical level (on the x-axis).

(a) Monotonicity Assessment for Vertical Names



(b) Monotonicity Assessment for Horizontal Names

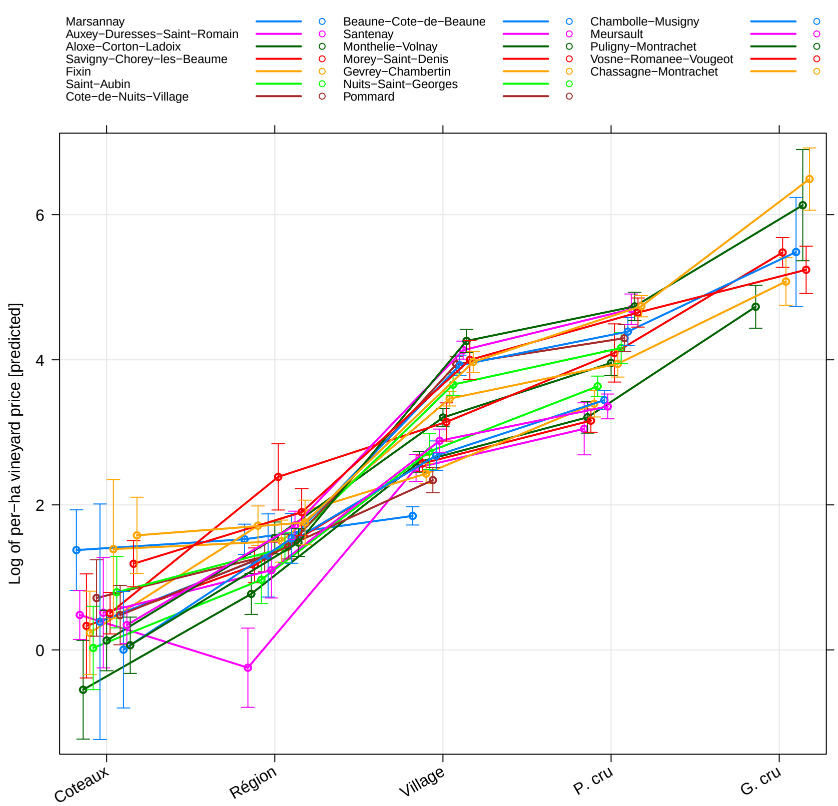


Figure A.7: **Hedonic Evidences about Complementary Wine names.**

Notes: Reported name premiums are averaged as described in Section A1.1 of OA. For each horizontal name, vertical premiums on the x-axis are acreage-weighted average of vertical premiums according to their within distribution. Premiums are expressed in % of increase relatively to the less expensive vertical and horizontal levels: the *Coteaux* of *Marsannay* located at the North of the area (Figure A.2). The slope coefficient is significant at 99% ($t = 3.98$).

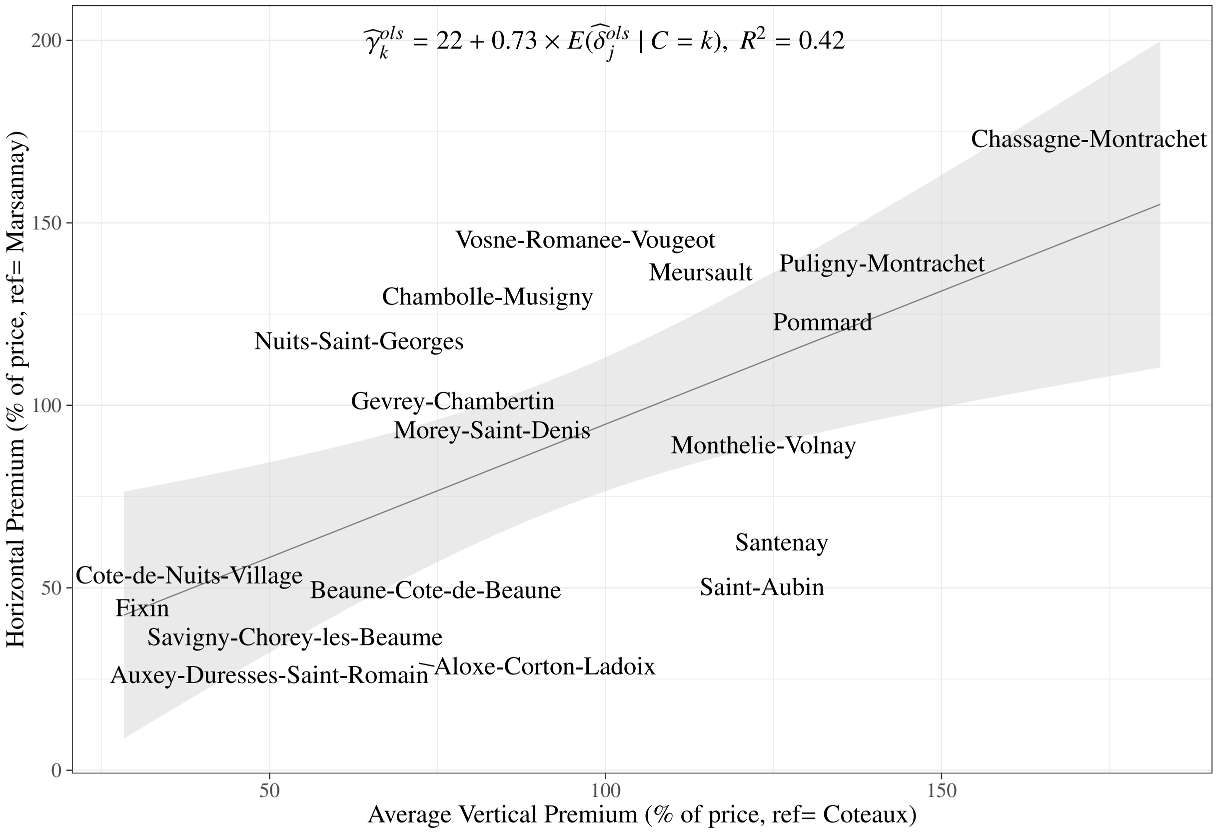


Figure A.8: **Spatial Distribution of Borders between Horizontal and Vertical Names.**

Notes: For the region of interest, the area shadded in grey corresponds to the population vineyards plots with an official wine name. The map below also reports the administrative borders of the *communes* (in black) that correspond to the horizontal names and the incremental borders (in colors) that separate vineyards of different vertical levels.

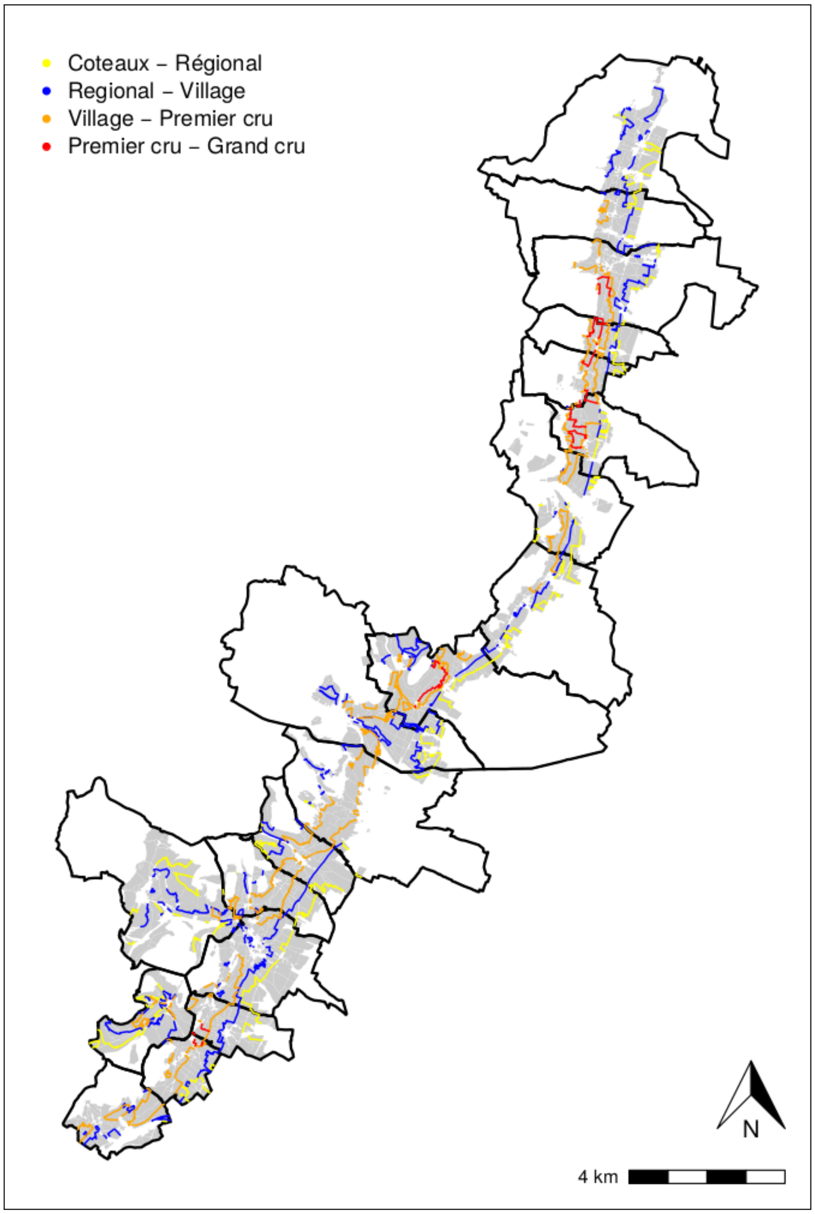


Figure A.9: **Zoom on *Gevrey-Chambertin*, *Morey-Saint-Denis*, and *Chambolle Musigny*.**

Notes: This map displays the horizontal and vertical borders for three northern *communes* of the area (see Figure A.2 of this Online Appendix). For clarity, the full geometries of vineyards plots appear only for the *commune* of *Morey-Saint-Denis* in the middle of the map. The *commune* of *Gevrey-Chambertin* is further north and the *commune* of *Chambolle-Musigny* is further south. Black lines are the administrative *commune* borders, where some delineates both horizontal and vertical names but not all. Centroids of vineyard plots sold on the period are marked by black dots.

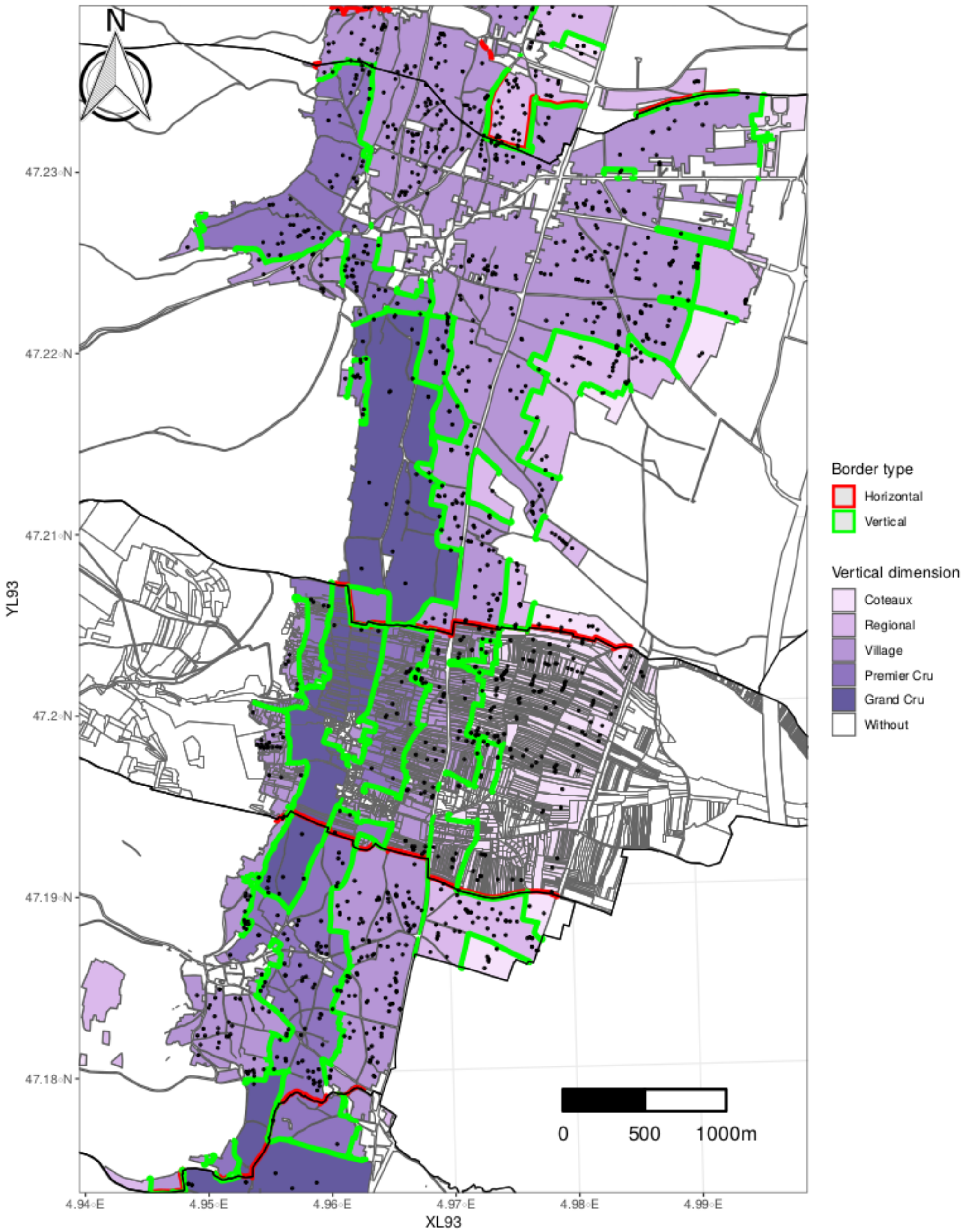


Figure A.10: **Population weights $w_{jk}/w_{j\cdot}$ and implicit weights from pooled SRD ω_{jk}^{SRD}.**

Notes: For each couple (j,k) of wine names, each dot represents the correspondig share of acreages (x-axis) and the corresponding share of incremental borders (y-axis). The obserevd dependance depends directly on the shape of name polygons. If the points are aligned on the first diagonal, the implicit weights from pooled SRD estimation are equal to population weights, which means the only bias come from the local nature of SRD estimates and not their aggregation.

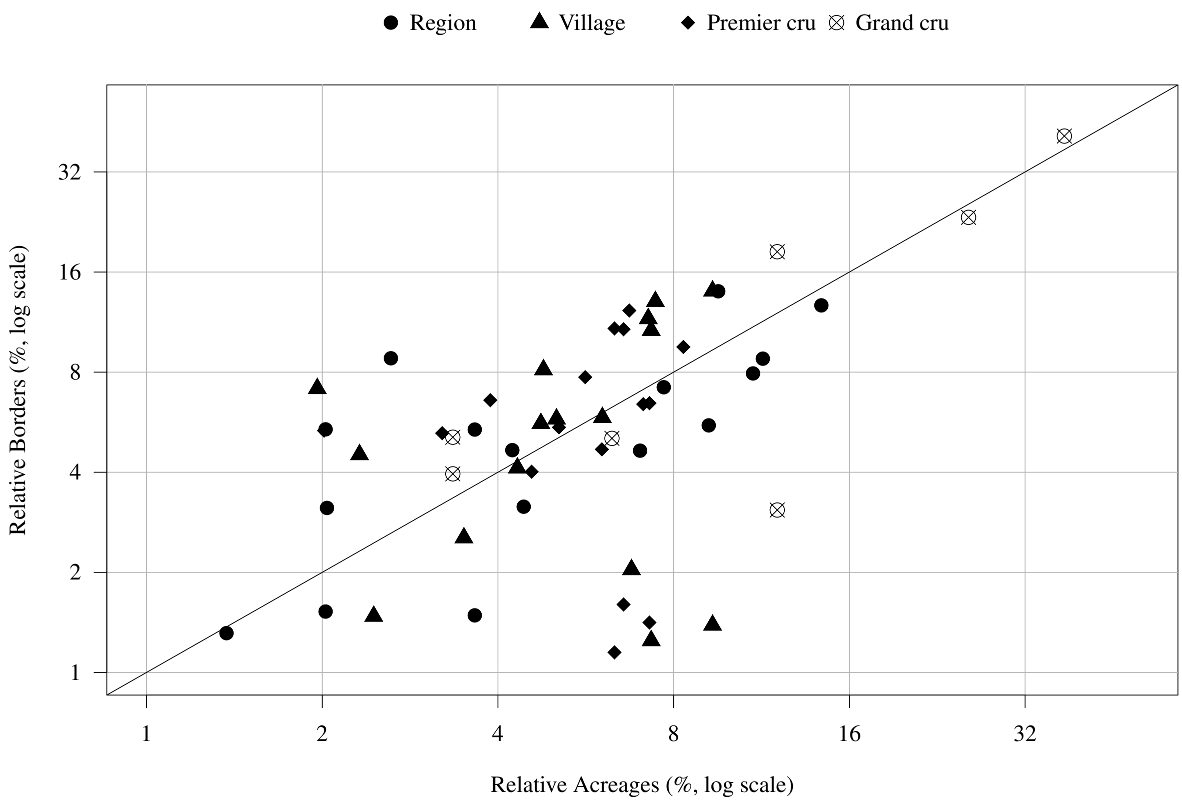
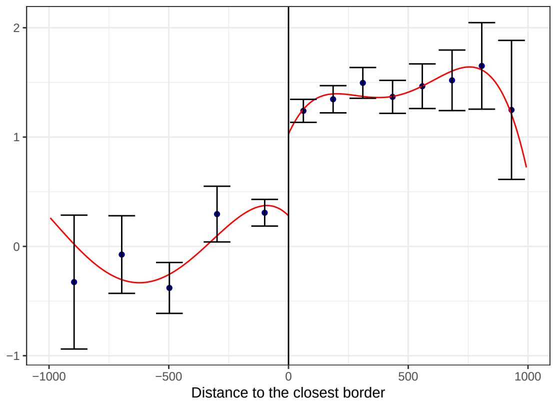


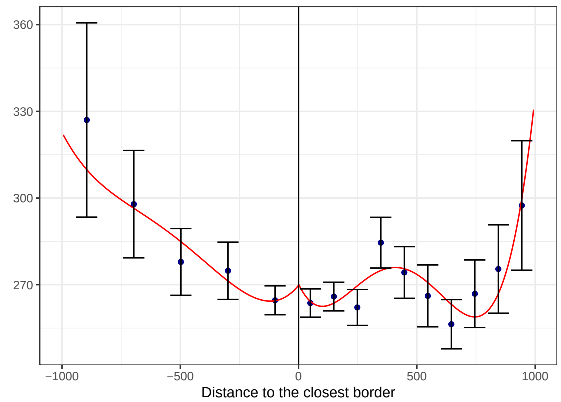
Figure A.11: **SRD Average Incremental Vertical Premium for *Région* wrt *Coteaux*.**

Note: The SRD estimators and their standard errors are reported in Table 3 of the main text.

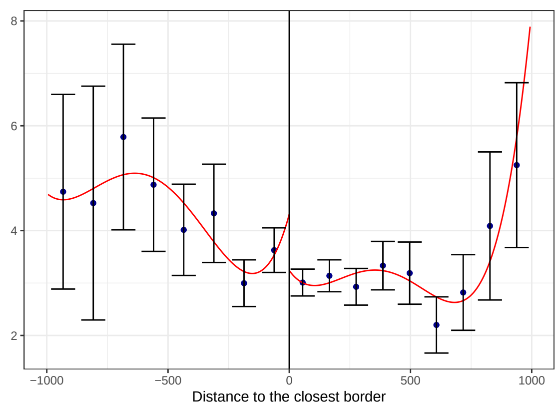
(a) Log Price Discontinuity



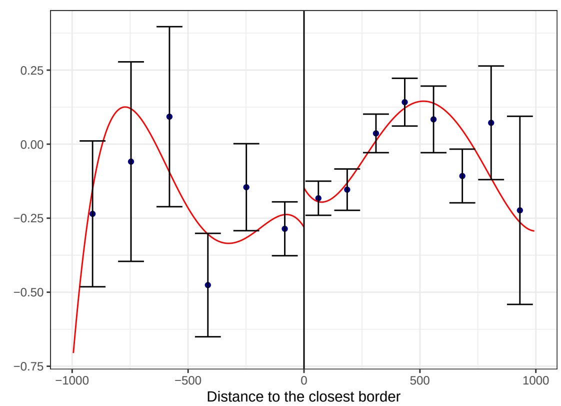
(b) Elevation Discontinuity



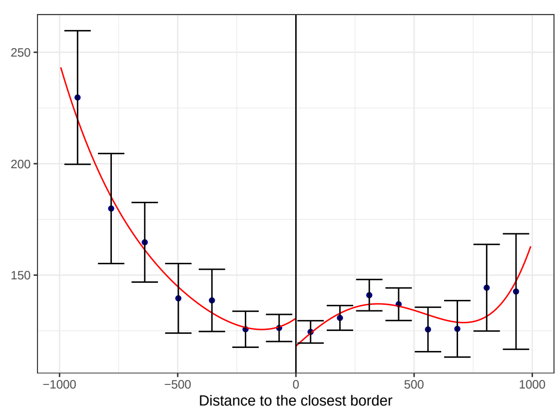
(c) Slope Discontinuity



(d) Solar Discontinuity



(e) Exposition Discontinuity



(f) Plot Size Discontinuity

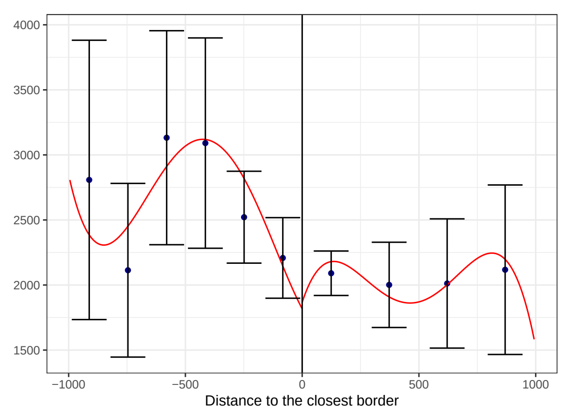
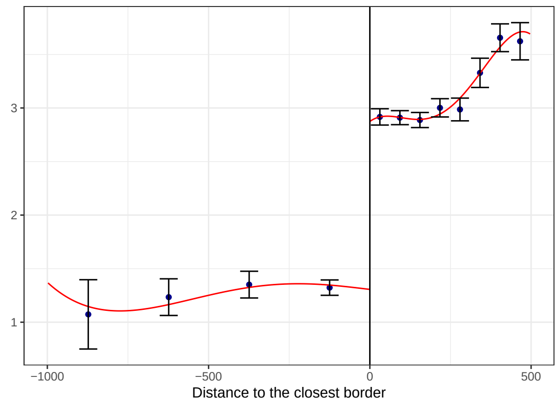


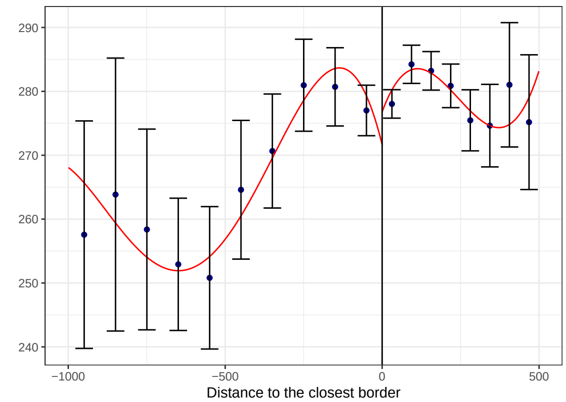
Figure A.12: **SRD Average Incremental Vertical Premium for *Village* wrt *Région*.**

Note: The SRD estimators and their standard errors are reported in Table 3 of the main text.

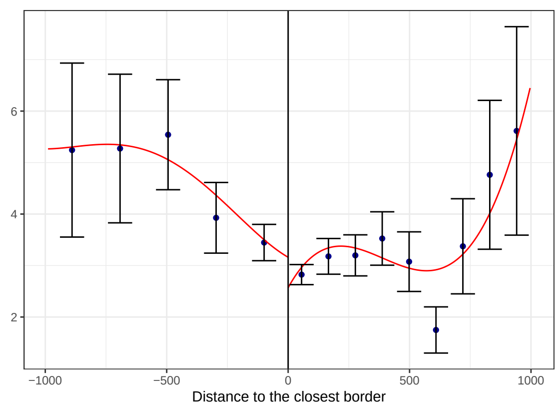
(a) Log Price Discontinuity



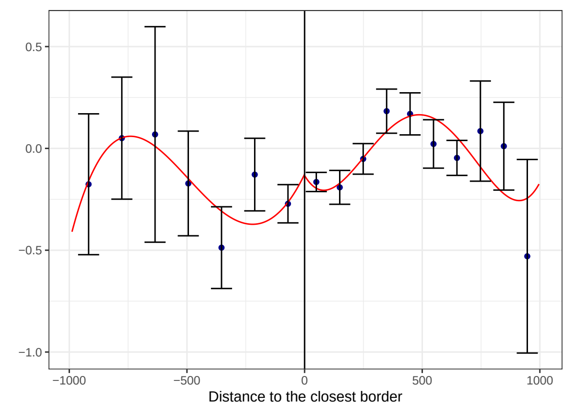
(b) Elevation Discontinuity



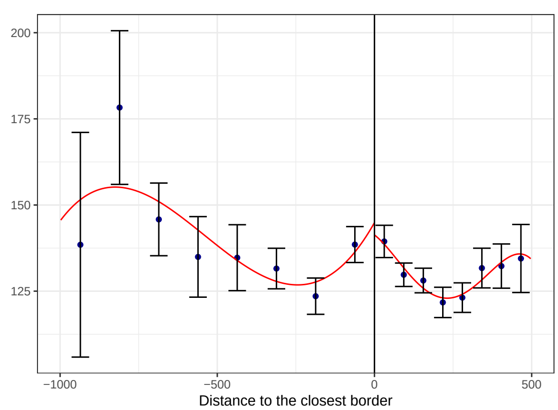
(c) Slope Discontinuity



(d) Solar Discontinuity



(e) Exposition Discontinuity



(f) Plot Size Discontinuity

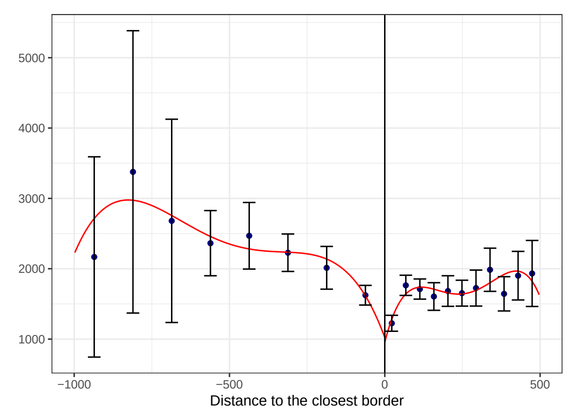
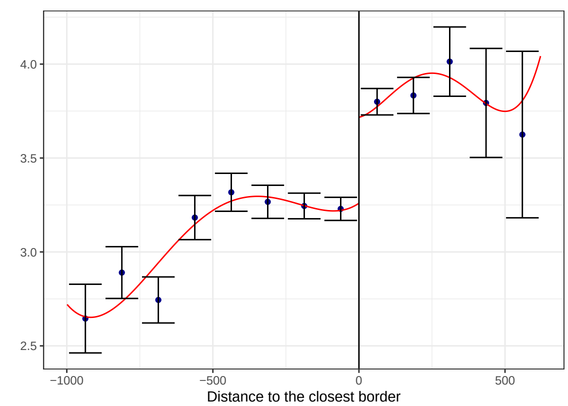


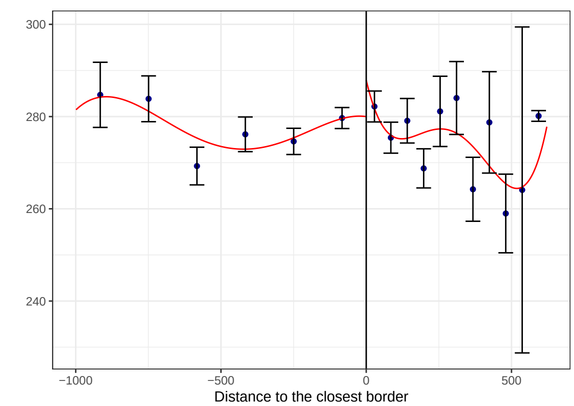
Figure A.13: **SRD Average Incremental Vertical Premium for *Premier cru* wrt *Village*.**

Note: The SRD estimators and their standard errors are reported in Table 3 of the main text.

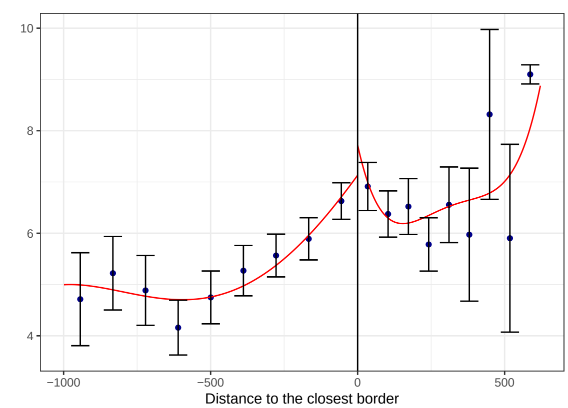
(a) Log Price Discontinuity



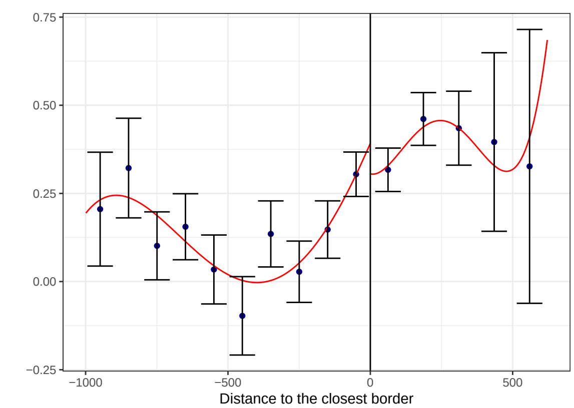
(b) Elevation Discontinuity



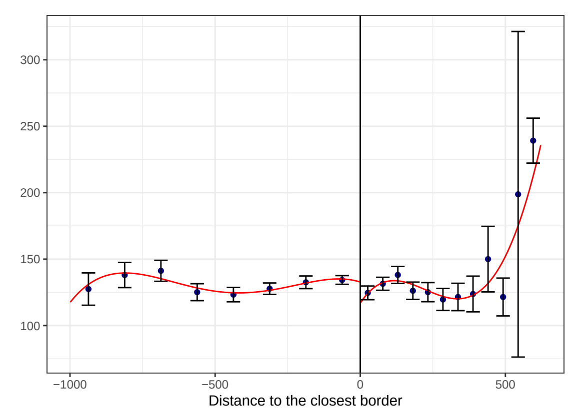
(c) Slope Discontinuity



(d) Solar Discontinuity



(e) Exposition Discontinuity



(f) Plot Size Discontinuity

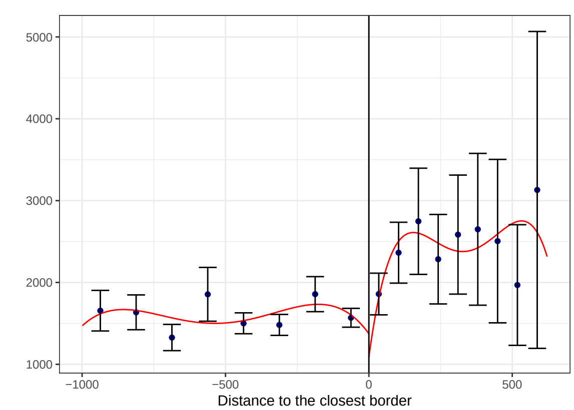
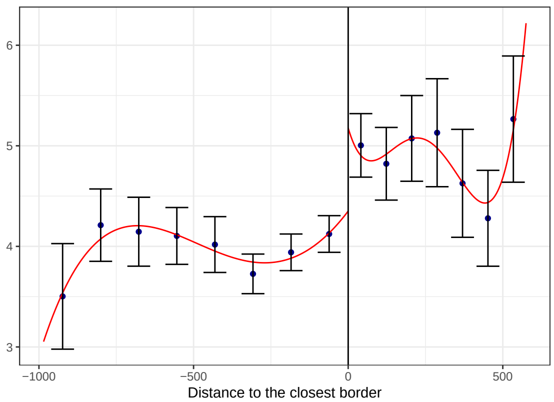


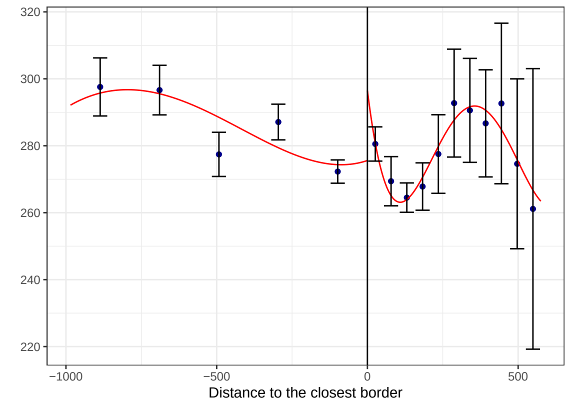
Figure A.14: **SRD Average Incremental Vertical Premium for *Grand cru* wrt *Premier cru*.**

Note: The SRD estimators and their standard errors are reported in Table 3 of the main text.

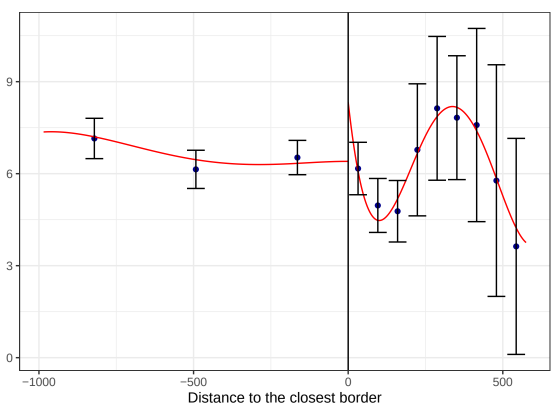
(a) Log Price Discontinuity



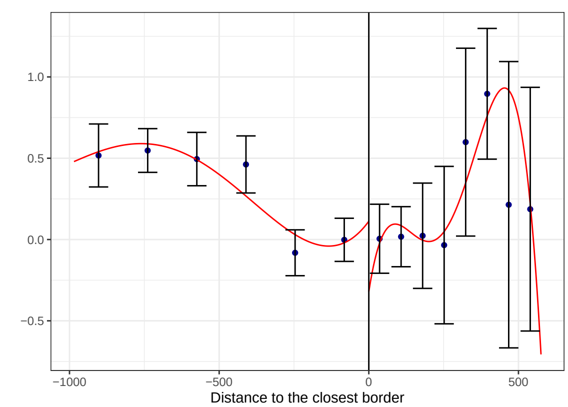
(b) Elevation Discontinuity



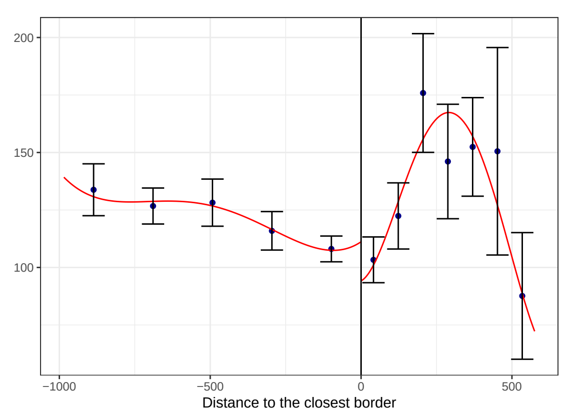
(c) Slope Discontinuity



(d) Solar Discontinuity



(e) Exposition Discontinuity



(f) Plot Size Discontinuity

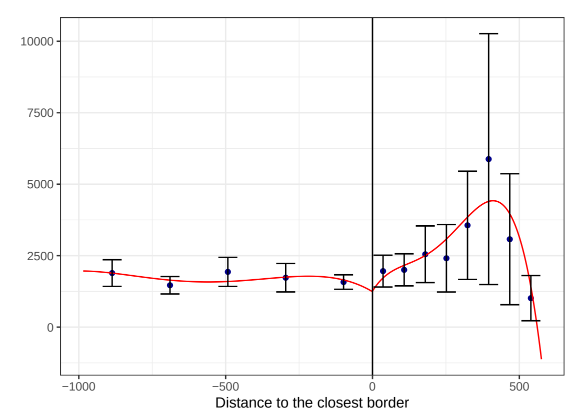


Figure A.15: **Horizontal and Interaction Borders Used in Pooled and Interaction SRD.**

Notes: The left panel (a) presents the delineations of *communes* used to estimate the incremental values of horizontal wine names (the index $k = 0, 1, \ldots, 18$ appear in the boxes of each group of *communes*). The borders between contiguous vineyards with different horizontal names are displayed by blue lines, black lines are the full adminstrative delineations between *communes*. The aggregation of initial *communes* mentionned in the notes below Table A.1 appears though dotted grey lines. The right panel (b) presents the spatial relationships used to estimate the incremental values of vertical and horizontal names. Each straight line represents a contiguity between vineyards used to estimate an incremental name premium (for a total of 51 SRD displayed in the rows of Table A.5 of this Online Appendix).

(a) Horizontal Wine Names (b) Interaction Between Wine Names

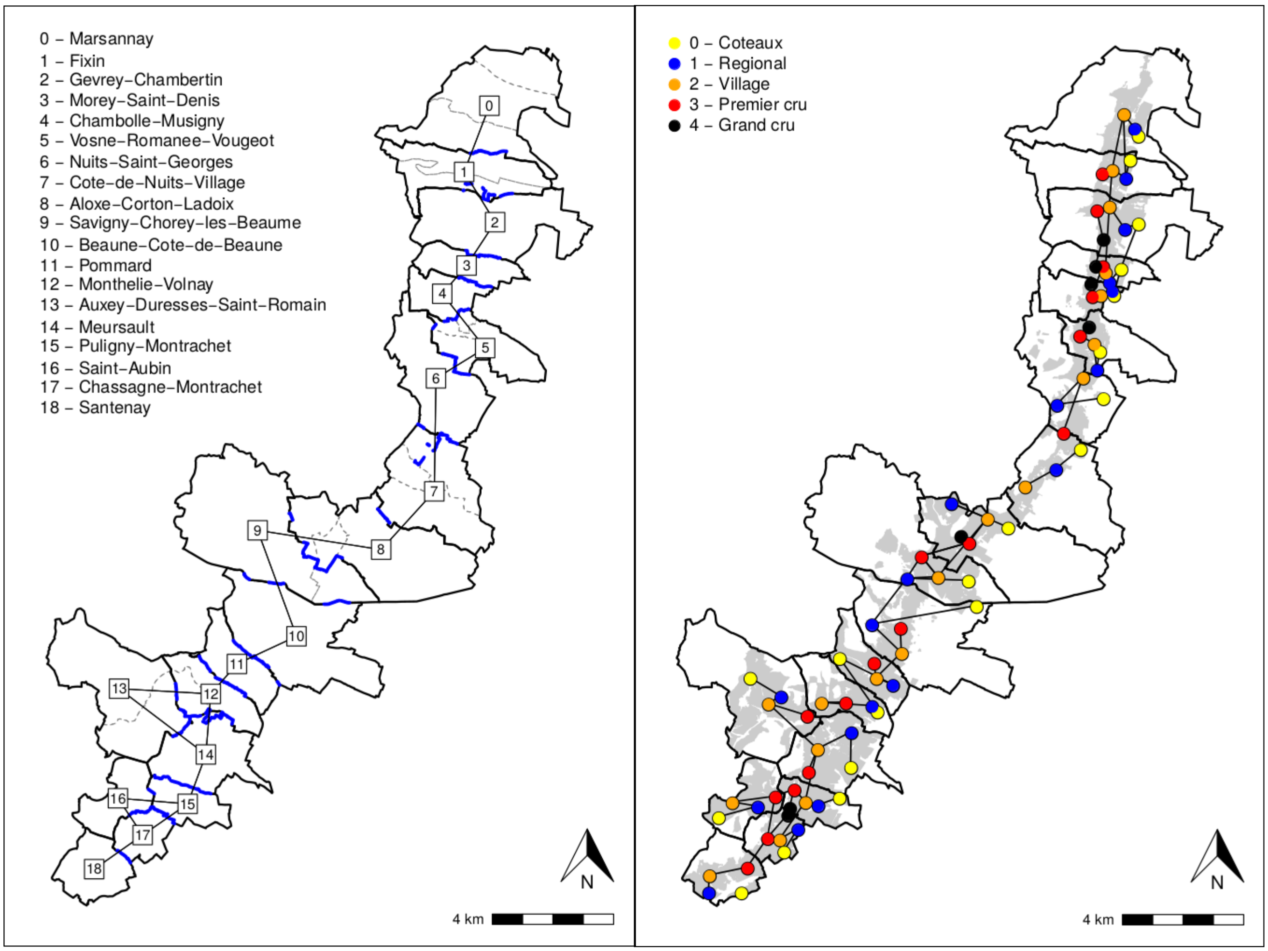


Table A.4: **Average Incremental Vertical Premiums from SRD Estimations.**

	WLS	SRD pool	SRD wfe	SRD wc
Fixin & Marsannay	0.117	0.101	1.809	−0.171
(N = 256)	(0.180)	(2.409)	(2.116)	(2.409)
Gevrey-Chambertin & Fixin	1.045***	2.519***	1.981***	1.076**
(N = 477)	(0.088)	(0.555)	(0.565)	(0.555)
Morey-Saint-Denis & Gevrey-Chambertin	−0.188	1.002	0.888	2.272**
(N = 344)	(0.271)	(1.170)	(0.938)	(1.170)
Chambolle-Musigny & Morey-Saint-Denis	0.130	−1.708	0.957	−3.094**
(N = 349)	(0.177)	(1.255)	(1.157)	(1.255)
Vosne-Romanee & Chambolle-Musigny	−0.215	1.360	0.983	1.422
(N = 299)	(0.240)	(1.248)	(0.986)	(1.248)
Nuits-Saint-Georges & Vosne-Romanee	−0.226	0.054	0.258	0.336
(N = 299)	(0.161)	(0.655)	(0.703)	(0.655)
Cote-de-Nuits & Nuits-Saint-Georges	−0.303**	0.023	1.591	1.056
(N = 213)	(0.169)	(1.645)	(1.167)	(1.645)
Aloxe-Corton-Ladoix & Cote-de-Nuits	−0.761***	−4.892**	−4.977**	−3.755
(N = 138)	(0.274)	(2.621)	(2.269)	(2.621)
Savigny-Chorey & Aloxe-Corton-Ladoix	−0.074	−2.344**	−2.096**	−2.071**
(N = 590)	(0.111)	(1.035)	(0.988)	(1.035)
Beaune-Cote-de-Beaune & Savigny-Chorey	−0.570***	−1.932	−1.793	−1.934
(N = 256)	(0.205)	(1.419)	(1.212)	(1.419)
Pommard & Beaune-Cote-de-Beaune	0.426**	−0.126	−0.789	0.176
(N = 464)	(0.199)	(0.447)	(0.561)	(0.447)
Monthelie-Volnay & Pommard	−0.046	−0.098	0.285	−0.141
(N = 888)	(0.075)	(0.532)	(0.434)	(0.532)
Auxey-Saint-Romain & Monthelie-Volnay	−0.326**	1.211	1.204	1.366**
(N = 503)	(0.149)	(0.799)	(0.879)	(0.799)
Meursault & Auxey-Saint-Romain	0.931***	1.520**	2.114**	1.439**
(N = 248)	(0.195)	(0.739)	(0.833)	(0.739)
Puligny-Montrachet & Meursault	0.328**	0.669**	0.402	0.482**
(N = 525)	(0.089)	(0.342)	(0.273)	(0.175)
Saint-Aubin & Puligny-Montrachet	0.948	−1.776**	−1.240	−2.382**
(N = 1 172)	(0.506)	(1.070)	(0.842)	(1.070)
Chassagne-Montrachet & Saint-Aubin	0.248**	1.776**	1.170	1.968**
(N = 903)	(0.071)	(0.790)	(0.842)	(0.650)
Santenay & Chassagne-Montrachet	−0.376	−1.678	−1.440	−1.873
(N = 578)	(0.201)	(1.234)	(0.827)	(0.99)

Standard errors (reported in parenthesis) are clustered by vineyard sales: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
Notes: For each couple of contiguous horizontal names in rows, the Table reports the spatial regression discontinuity estimators for the first *commune* names relatively to the second (see Figure A.15a for the spatial structure). The first column is a benchmark estimation of the descriptive model of Equation 1 by weighted least squares (see Section A1.1 of the Online Appendix for details). Other columns report the spatial regression discontinuity estimator with the log of per-ha vineyard prices as the outcome variable (see Table 3 in the main paper for meanings of the labels).

Table A.5: **Vertical Incremental Estimations from SRD on Horizontal Subsamples.**

	Région & Coteaux		*Village & Région*		*P. cru & Village*	
Horizontal names	OLS	SRD	OLS	SRD	OLS	SRD
Marsannay	0.406	1.412*	0.343**	1.188**	–	–
	(0.288)	(0.655)	(0.112)	(0.383)		
Fixin	1.374**	−0.438	0.883**	0.827**	–	–
	(0.334)	(0.511)	(0.158)	(0.303)		
Gevrey-Chambertin	0.200	−1.361	2.122**	2.441**	0.504**	0.652*
	(0.516)	(0.971)	(0.150)	(0.193)	(0.093)	(0.346)
Morey-Saint-Denis	1.803**	0.836*	0.972**	−1.724	0.876**	−0.316
	(0.312)	(0.426)	(0.306)	(1.183)	(0.218)	(0.273)
Chambolle-Musigny	–	–	2.437**	2.918**	0.521**	0.326
			(0.173)	(0.564)	(0.123)	(0.218)
Vosne-Romanee-Vougeot	0.470*	1.280*	2.224**	1.710*	0.793**	1.218*
	(0.214)	(0.712)	(0.191)	(1.035)	(0.172)	(0.661)
Nuits-Saint-Georges	0.417	0.080	2.519**	1.960**	0.424**	0.299
	(0.326)	(0.380)	(0.128)	(0.319)	(0.157)	(0.299)
Cote-de-Nuits-Village	0.647*	2.014	1.290**	0.758*	–	–
	(0.300)	(2.865)	(0.145)	(0.405)		
Aloxe-Corton-Ladoix	1.860**	2.211	1.941**	1.855**	0.912**	0.238
	(0.455)	(1.420)	(0.169)	(0.254)	(0.166)	(0.222)
Savigny-Chorey-les-Beaune	0.944*	−0.169	1.458**	1.396**	0.520**	0.594*
	(0.420)	(0.612)	(0.129)	(0.260)	(0.074)	(0.287)
Beaune-Cote-de-Beaune	–	–	1.693**	2.681**	0.832**	0.555**
			(0.413)	(0.690)	(0.113)	(0.174)
Pommard	0.989**	1.254*	2.605**	1.903**	0.555**	0.046
	(0.230)	(0.667)	(0.094)	(0.292)	(0.094)	(0.120)
Monthelie-Volnay	1.439**	0.211	1.635**	1.532**	0.946**	0.485*
	(0.232)	(0.350)	(0.109)	(0.304)	(0.099)	(0.206)
Auxey-Duresses-Saint-Romain	−1.032**	1.211*	3.340**	3.085**	0.158	−0.140
	(0.296)	(0.636)	(0.277)	(0.577)	(0.185)	(0.691)
Meursault	1.316**	1.142**	2.484**	1.807**	0.627**	0.044
	(0.147)	(0.278)	(0.110)	(0.272)	(0.111)	(0.178)
Puligny-Montrachet	1.639**	1.611**	2.889**	2.925**	0.536**	0.788**
	(0.242)	(0.448)	(0.113)	(0.317)	(0.105)	(0.153)
Saint-Aubin	0.958**	1.369*	1.806**	1.927**	1.076**	0.272
	(0.298)	(0.624)	(0.220)	(0.452)	(0.181)	(0.501)
Chassagne-Montrachet	0.135	−1.620**	2.374**	2.079**	0.882**	0.817**
	(0.368)	(0.189)	(0.180)	(0.586)	(0.096)	(0.260)
Santenay	–	–	1.734**	0.348	0.604**	0.016
			(0.195)	(0.332)	(0.120)	(0.288)

Standard errors (reported in parenthesis) are clustered by vineyard sales: ** $p < 0.01$, * $p < 0.05$.
Notes: Each cell of the Table reports a vertical incremental premium (in columns) for each horizontal name (in rows). The OLS estimators are obtained from a vertical dummy on a subsample of sales with $(j-1, j, k)$ names. Dashed cells mark the cases when these borders do not exist, i.e., vertical names are not contiguous (see Figure A.15b above).

Figure A.16: **Umbrella Branding Intensity from SRD estimators.**

Notes: ERRORS IN VALUES. Reported hedonic premiums are estimated by controlling for sale and vineyards characteristics (Appendix A). Horizontal premiums on the y-axis are raw estimated coefficients from a log-linear model, expected vertical premiums on the x-axis are acreage-weighted coefficients from the same model for each horizontal level. Premium are expressed in % increases relatively to the less famous vertical and horizontal levels (Côteaux of Marsannay) located at the extreme North of the region under study (Figure A1 in Appendix A). The slope coefficient is significant at 99% (t = 2.98).

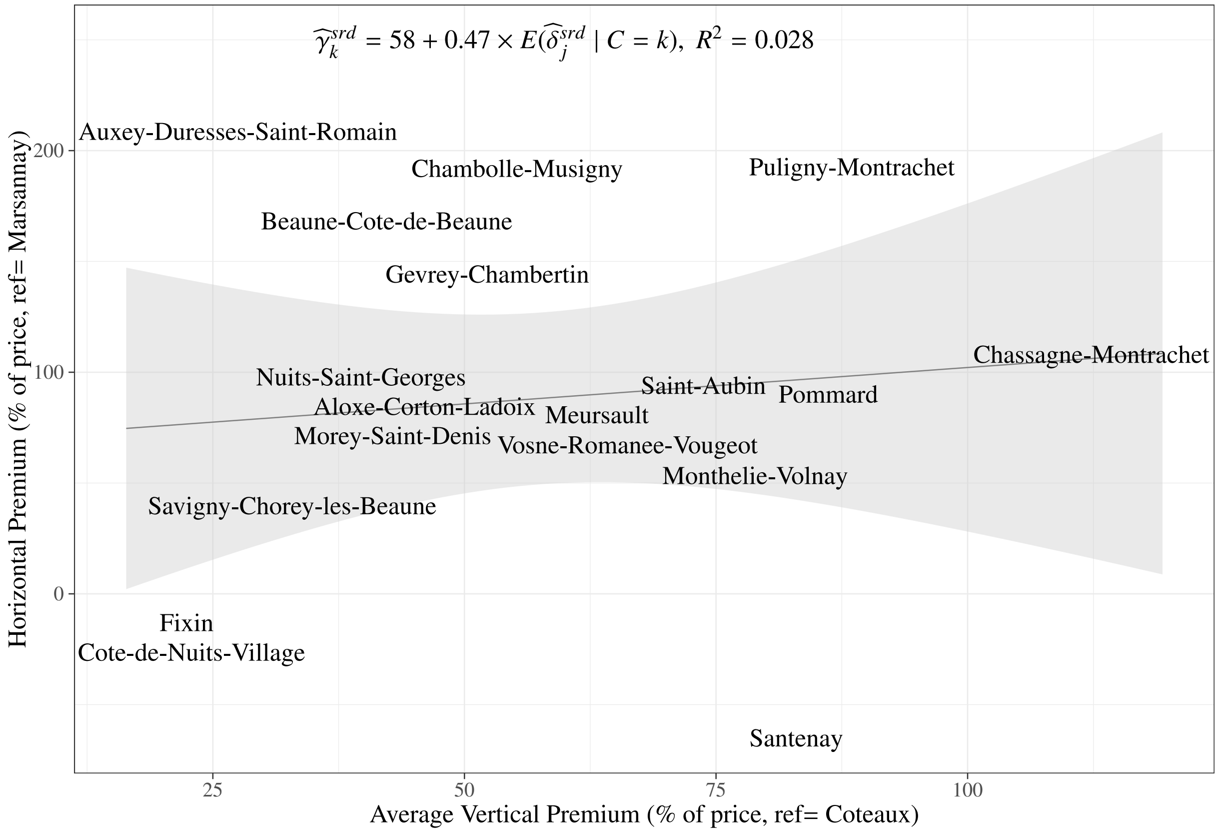


Figure A.17: **Marginal Effects from Designation Model of Table 4 with df= 900.**

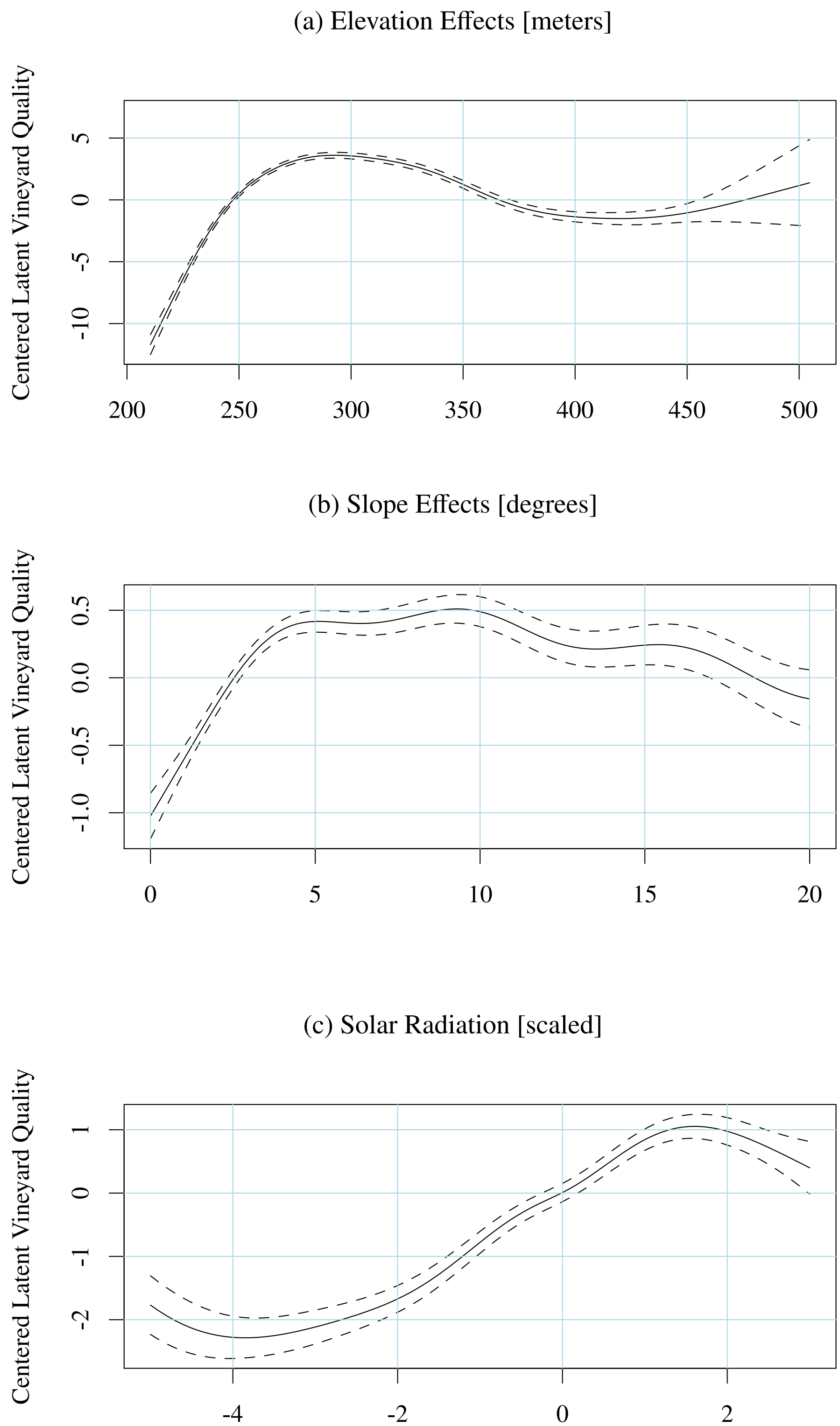
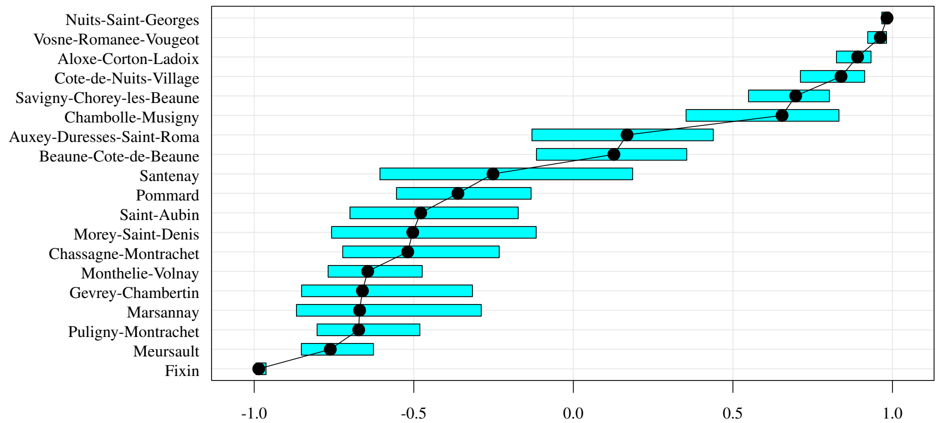


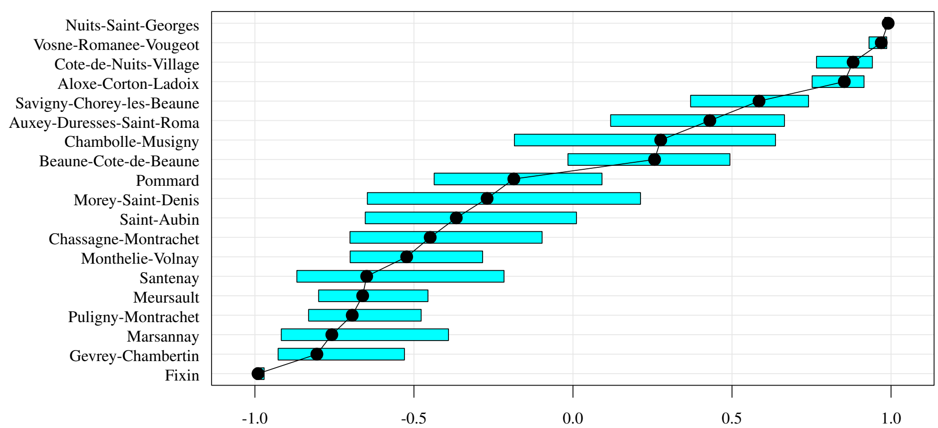
Figure A.18: **Ordinal Superiority Measures from Designation Models of Table 4.**

Notes: For a given *commune* on the y-axis, ordinal superiority measures (Agresti and Kateri, 2017) are the differences between estimated fixed effect μ_c and average fixed effect $\overline{\mu}$ of all *commune* according to: $\Delta_c = 2 \times \Phi[(\mu_c - \overline{\mu})/\sqrt{2}] - 1$.

(a) Ordinal Superiority Measures from Model with df= 500



(b) Ordinal Superiority Measures from Model with df= 700



(c) Ordinal Superiority Measures from Model with df= 900

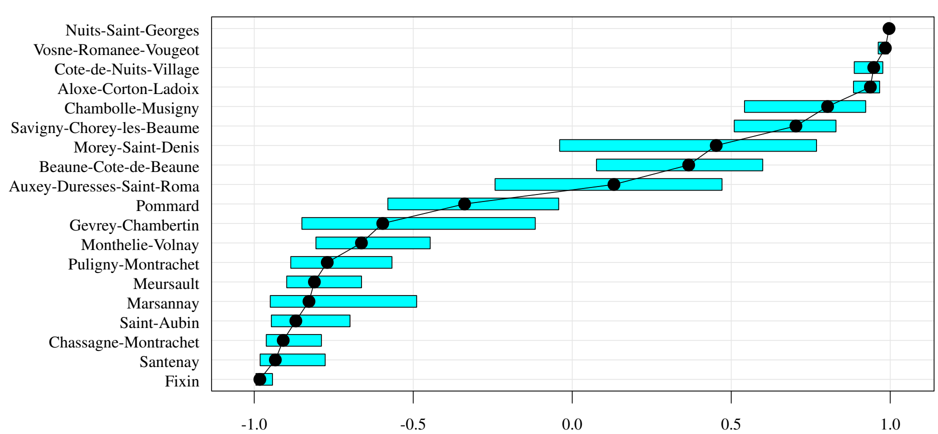


Figure A.19: **Latent Vineyard Quality from the Designation Model with df=900.**

Notes: The northern *Côte de Nuits* is reported in the left panel and the southern *Côte de Beaune* is reported in the right panel to save place. From the notation of the main text, the two maps report the predicted values of the latent vineyard quality $B(\bar{\mathbf{x}}, \mathbf{v})^\top \psi$ for a regular grid $\mathbf{v}$ of 100×100 meters with the biophysical variables set at their sample averages $\bar{\mathbf{x}}$. The latent predictions from an ordered models are unitless, as the vertical bar at the right of the maps.

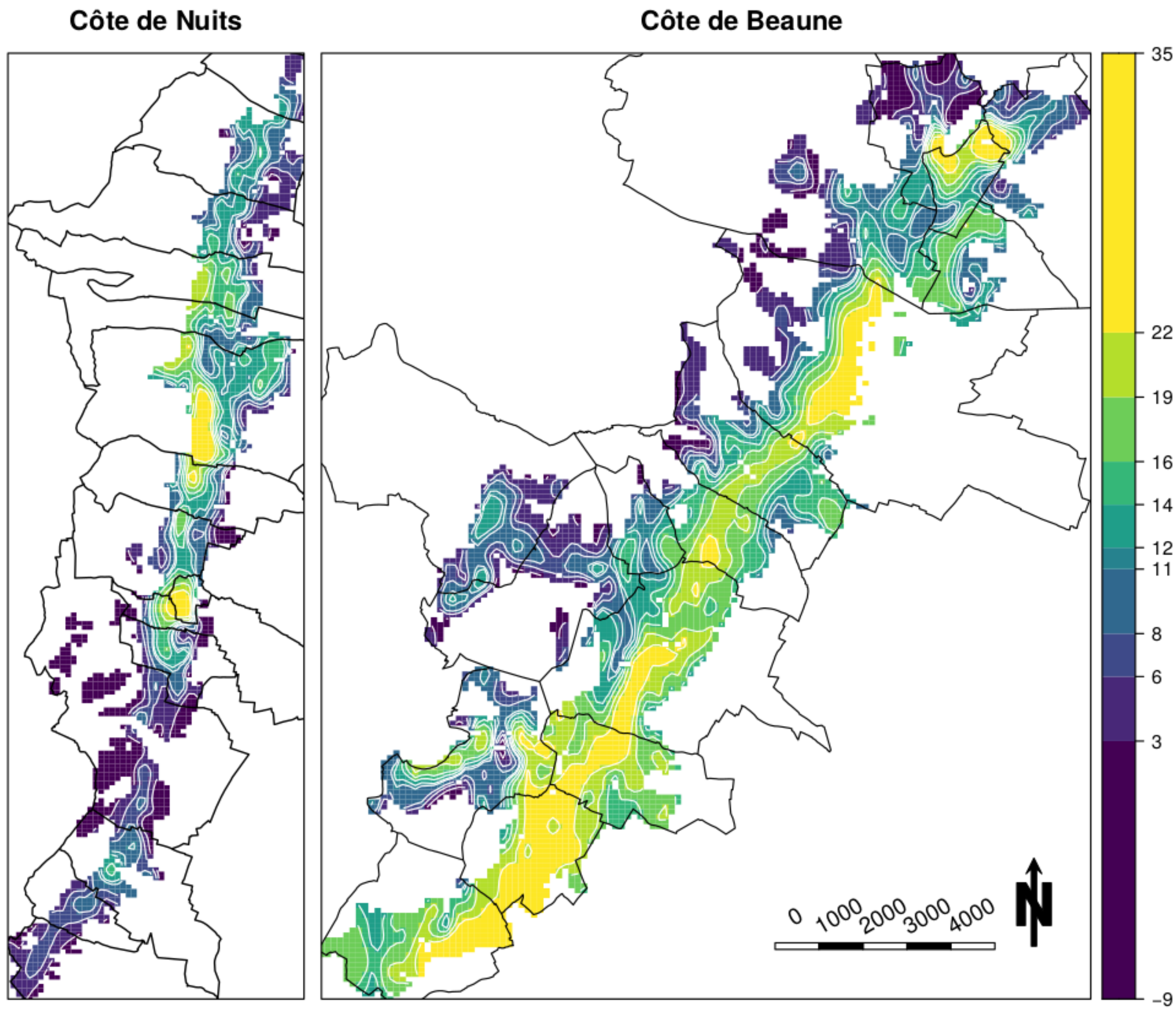
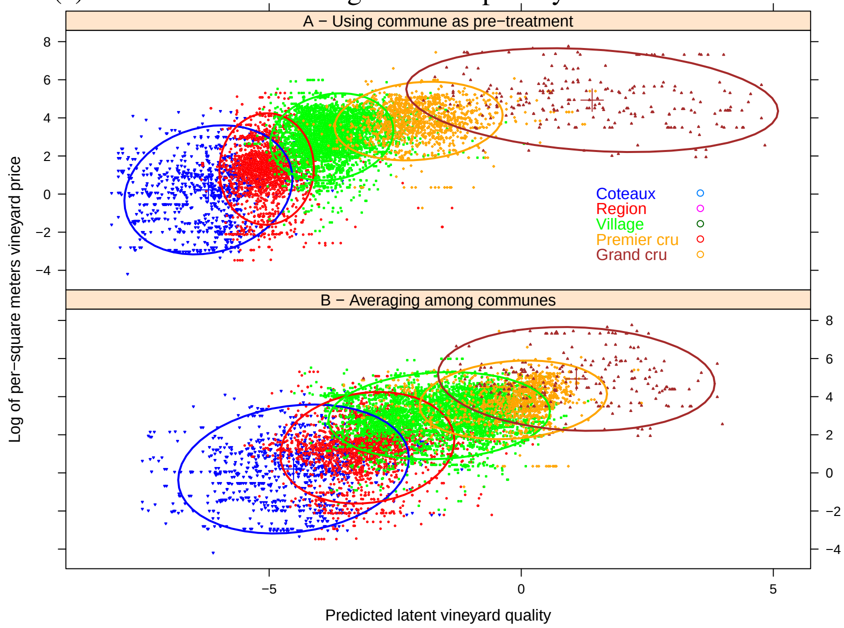


Figure A.20: **Overlap Gains from Averaging Latent Quality Variables Across *Communes*.**

Notes: For two designation models of Table 4 (with df= 500 in (a) and df= 900 in (b)), the Figure reports the relations between observed vineyard price and predicted vineyard quality, respectively with *commune* dummies as pre-treatment variables $L(X,V)^{\top}\psi + \mu_c$ (top panels A of each subfigure) and by an averaging of *communes*' coefficients $L(X,V)^{\top}\psi + \overline{\mu}$ (bottom panels B of each subfigure). We average *commune* coefficients in $\overline{\mu}$ to keep the same scale between predictions and to better compare them. The confidence ellipsoids at 95% are computed from Friendly et al. (2013).

(a) Predictions from Designation Propensity Model with df= 500.



(b) Predictions from Designation Propensity Model with df= 900.

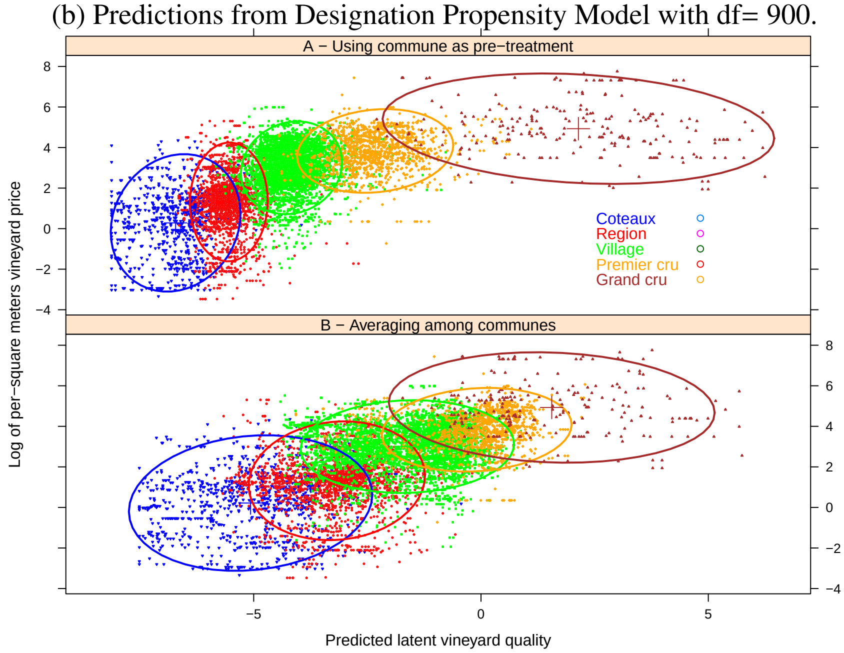


Table A.6: **Additional Diagnostic Statistics for Sub-Sample WLS Regressions.**

		Coteaux	*Régional*	*Village*	*Premier cru*	*Grand cru*
Coteaux	(1)	7.14***	16.96***	7.04***	7.44***	1.29
	(2)	7.24***	5.31***	−0.13	−1.51	0.20
	(3)	0.12	0.36	0.26	0.28	0.25
	(4)	1035	1035	1035	1035	1035
	(5)	100.00	80.68	75.75	63.96	51.21
Régional	(1)	23.00***	20.85***	33.24***	50.47***	94.29***
	(2)	12.27***	9.50***	8.18***	7.99***	5.82***
	(3)	0.17	0.16	0.21	0.29	0.46
	(4)	2267	2267	2267	2267	2267
	(5)	98.76	100.00	95.32	90.60	83.77
Village	(1)	91.59***	103.17***	120.29***	115.08***	108.66***
	(2)	17.95***	21.72***	18.61***	14.34***	10.28***
	(3)	0.28	0.31	0.37	0.39	0.42
	(4)	4204	4204	4204	4204	4204
	(5)	98.76	97.72	100.00	95.98	89.94
Premier cru	(1)	153.28***	119.18***	53.21***	26.79***	38.60***
	(2)	7.51***	−7.27***	−3.37***	9.87***	13.64***
	(3)	0.76	0.74	0.54	0.31	0.43
	(4)	1255	1255	1255	1255	1255
	(5)	93.55	90.52	94.82	100.00	85.26
Grand cru	(1)	43.57***	37.25***	18.20***	7.89***	5.75***
	(2)	−19.66***	−12.73***	−7.55***	−6.85***	1.90*
	(3)	0.80	0.73	0.59	0.47	0.15
	(4)	226	226	226	226	226
	(5)	46.55	88.68	95.36	99.56	100.00

Fisher and Student Statistics in the two first rows are clustered by sales: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.
Notes: For each of the 25 combinations of vertical names, we run a WLS regression for a sub-sample of vineyard sales with a given vertical name (in rows) and a given numerator in the weighting schedule (in columns). The structure of the Table is such that rows correspond to k in Equation 19 and columns correspond to l in the weights . The diagonal reports the OLS regressions as the weights are one when the sub-sample is both the control and the treatment (see Equation 17).
The rows (1) report Fisher statistics about the joint significance of *commune* fixed effects $\boldsymbol{\kappa}_j(l)$.
The rows (2) report Student statistics about the significance of the predicted latent quality index $\rho_j(l)$.
The rows (3) report the full R^2 for the each WLS regression.
The rows (4) report the number of observations in the sub-sample. Sub-sample are not trimmed to reach overlap.
The rows (5) report the underlying overlap area measured in % of treated units (in rows) that have a propensity score within the 95% range of the propensity score of the plots that serve as controls (in columns).

Table A.7: **Horizontal Signaling Values from Sub-Sample Regressions in DR estimates.**

	Coteaux	*Régional*	*Village*	*Premier cru*	*Grand cru*	$AIHP_k$
Marsannay	1.93^{***}	0.85^{***}	-0.55^{***}	-	-	0.21^{*}
	(0.39)	(0.12)	(0.06)			(0.12)
Fixin	0.52	1.11^{***}	0.02	0.56	-	0.46^{***}
	(0.40)	(0.14)	(0.08)	(0.57)		(0.18)
Gevrey-Chambertin	1.96^{***}	1.33^{***}	1.25^{***}	0.69^{***}	0.18	1.06^{***}
	(0.48)	(0.17)	(0.06)	(0.13)	(0.23)	(0.12)
Morey-Saint-Denis	1.17^{***}	1.71^{***}	0.70^{***}	0.68^{***}	0.41	0.90^{***}
	(0.35)	(0.29)	(0.12)	(0.22)	(0.49)	(0.29)
Chambolle-Musigny	1.20^{*}	1.15^{***}	1.44^{***}	1.26^{***}	0.70	1.26^{***}
	(0.57)	(0.19)	(0.09)	(0.16)	(0.48)	(0.19)
Vosne-Romanee-Vougeot	1.64^{***}	1.26^{***}	1.50^{***}	1.34^{***}	-	1.25^{***}
	(0.36)	(0.20)	(0.10)	(0.17)		(0.18)
Nuits-Saint-Georges	1.93^{***}	0.79^{***}	1.17^{***}	1.37^{***}	-	1.18^{***}
	(0.40)	(0.15)	(0.08)	(0.15)		(0.16)
Cote-de-Nuits-Village	1.19^{***}	0.71^{***}	-0.08	-	0.88^{***}	0.25^{***}
	(0.40)	(0.15)	(0.09)		(0.18)	(0.19)
Savigny-Chorey-les-Bea	0.83^{*}	0.19	0.03	0.03	-	0.13
	(0.42)	(0.13)	(0.06)	(0.15)		(0.12)
Beaune-Cote-de-Beaune	0.35	0.54^{***}	-0.11	0.13	-	0.14
	(0.52)	(0.20)	(0.09)	(0.12)		(0.15)
Pommard	0.52	0.46^{***}	1.04^{***}	0.81^{***}	-	0.71^{***}
	(0.37)	(0.11)	(0.06)	(0.13)		(0.12)
Monthelie-Volnay	0.32	0.67^{***}	0.33^{***}	0.41^{***}	-	0.42^{***}
	(0.36)	(0.14)	(0.07)	(0.13)		(0.14)
Auxey-Duresses-Saint-R	0.88^{*}	-0.91^{***}	-0.11	-0.11	-	0.05
	(0.35)	(0.18)	(0.08)	(0.27)		(0.23)
Meursault	0.82^{*}	0.80^{***}	1.19^{***}	1.24^{***}	-	0.98^{***}
	(0.34)	(0.11)	(0.07)	(0.15)		(0.19)
Puligny-Montrachet	1.36^{***}	0.81^{***}	1.20^{***}	0.95^{***}	0.58^{*}	1.00^{***}
	(0.38)	(0.12)	(0.09)	(0.14)	(0.31)	(0.16)
Saint-Aubin	0.04	0.39	0.14	0.33^{***}	-	0.50^{***}
	(0.41)	(0.13)	(0.12)	(0.12)		(0.18)
Chassagne-Montrachet	1.16^{*}	0.97^{***}	0.98^{***}	1.18^{***}	0.92^{***}	1.07^{***}
	(0.48)	(0.17)	(0.08)	(0.13)	(0.33)	(0.16)
Santenay	0.41	0.24	0.15^{*}	-0.21	-	0.09
	(0.44)	(0.18)	(0.09)	(0.15)		(0.15)

Standard errors (reported in parenthesis) are clustered by vineyard sales: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

The Table reports the κ_j from Equation 19 of the main text, estimated from sub-samples of vertical names reported in columns. The reference reference modality is *Aloxe-Corton-Ladoix*, which presents the smallest signaling value on average. The last column reports the aggregation of each individual average incremental horizontal premiums according to the population acreages of each combination of names (see Equation 5 in the main text). These causal statistics are used to assess the monotonicty in the Table 6 of the main text according to the formula provided in OA1.3.

Figure A.21: **Umbrella Branding Intensity from DR estimators.**

Notes: Average vertical premiums on the x-axis are acreage-weighted coefficients of vertical premiums from for each horizontal name. Premiums are expressed in % of increase relatively to the less expensive vertical and horizontal levels: the *Coteaux* of *Marsannay* located at the North of the area (Figure A.2). The slope is significant at 85% ($t = 1.58$).