**SUPPLEMENTARY APPENDIX**

*[FOR ONLINE PUBLICATION ONLY]*

This appendix provides additional details to supplement the main text. Details relevant to each section of the main text are included in the corresponding section of this appendix. The data and code for the paper can be found through the following link, which has been anonymized for the purposes of blind peer review: <<https://osf.io/8zef3/?view_only=c3fa0f7346bf4441943d1337f17865ab>>.

**I. Additional remarks related to those in Section I of the main text**

As stated in the main text, a number of studies have quantitatively analyzed expert wine descriptions. In addition to the examples already given in the main text, other studies that quantitatively analyze wine descriptions include Hendrickx et al. (2016), which is from the same research team behind the Croijmans et al. (2019) paper cited in the main text, as well as Chen et al. (2018), which is from the same research team behind the Dong et a. (2020) paper cited in the main text. See the end of this section for references to papers not already cited in the main text.

**II. Additional remarks related to those in Section II of the main text**

● The data source used by Krumme (2009, 2011) is unnamed, as already stated in the main text. The only information explicitly provided about the data source in Krumme (2009) or Krumme (2011) is that it was “an online aggregator of reviews” (Krumme, 2011). The Krumme (2011) piece, which was published online, includes hyperlinks to two online wine retailers, CellarBid.com and TheWineBuyer.com. At least as of writing, the former site lists a couple hundred wines, the latter lists about 1,000 wines, and some but not all of the wines listed have expert descriptions from well-known wine critics such as Robert Parker. Either of those sites could be considered an online aggregator of reviews and therefore might have been Krumme’s data source. The author of Krumme (2009, 2011) unfortunately did not respond to a request to clarify the data source (personal correspondence on July 19, 2018). Note that the data source we use in the main text would not be considered an aggregator of reviews (unless *Wine Enthusiast* is seen as aggregating the reviews of the different experts who write for it), so it seems safe to presume that the dataset we have used is different from the one used by Krumme (2009, 2011). At the very least, wines and reviews have presumably been added or removed from the *Wine Enthusiast* website since Krumme was conducting her study about a decade ago.

● It should be admitted that the relative probability *P(wi | cH)* / *P(wi | cL)* defined and used in the main text is not necessarily the statistic that Krumme used to identify “expensive” and “cheap” wine words. Yet that statistic is consistent with the quotes from Krumme (2011) cited in the main text of this paper about trying to identify words that are preferentially used or more likely to be used to describe high-priced wines. That statistic is also consistent with the way in which Krumme (2009) reports at least one of her results, namely, that a particular word (“vintage”) is found to be “six times more likely to describe an expensive wine.” Other statistics should also produce similar results if the results are robust.

● The Dong et al. (2020) study cited in the main text seems to use a Boolean NB (rather than Multinomial NB) classifier. If we used that specification, then *P(wi | cj)* would be estimated as the proportion of any wines in price category *cj* that were described at least once by word *wi* (see, e.g., sec. 2.1 of the Metsis et al., 2006, paper cited in the main text).

● When discussing the groups of expensive and cheap words she identified, Krumme (2009) and/or Krumme (2011) used several labels for those groups.

For what we called the dark vs. light grouping, she used the labels “dark,” “dark words,” “darker words,” and “fullness of flavor” vs. “lightness.”

For the specific vs. general grouping, she used the labels “specific,” “more specific descriptors,” “specificity,” “single flavors,” and “single fruits and flavors” vs. “more generic terms” and “generality.”

For the elite vs. accessible grouping, she used the labels “elite,” “exclusive-sounding words,” and “authenticity or exclusivity” vs. “accessibility” and “simple descriptors.”

● Besides the examples reproduced in the main text of this paper, the only other term mentioned in Krumme (2009) or Krumme (2011) in a way that suggests she may have empirically identified it as an expensive word is “focused cassis” (Krumme, 2011). However, it is not entirely clear that term was empirically identified as such, and unlike all the other cheap and expensive words mentioned in Krumme (2009) or Krumme (2011), it is a bigram phrase rather than a unigram word. We will presume that Krumme only considered unigrams in her empirical work and, in any event, we have only considered unigrams.

**III. Additional remarks related to those in Section III of the main text**

*A. On defining high- and low-priced wines*

Our choice to define wines over 50 dollars as “high priced” and wines of 15 dollars or less as “low priced” was somewhat arbitrary and largely informed by the distribution of prices for the dataset we studied in the main text. It can however be noted that, in a study that is arguably similar to Krumme’s (yet fails to acknowledge her earlier work) and that also uses data scraped from the *Wine Enthusiast* website, Hendrickx et al. (2016) define “expensive” wines as ones that are 50 dollars or more and “cheap” wines as those that are less than 15 dollars. We could therefore point to their study as a justification for choosing similar dollar values to define high- and low-priced wines.

*B. Results for the same dataset but a different definition of high- and low-priced wines*

Given that the definitions of high- and low-priced wines used in the main text are somewhat arbitrary, the robustness of our results to different definitions can be explored. As an alternative to the definitions used in the main text, we can define wines greater than 40 dollars as “high priced.” Forty dollars roughly corresponds to the upper quartile of the prices of the wines in our *WE* dataset. It can be noted that the Chen and McCluskey (2018) study cited in the main text suggests that wines over 40 dollars are part of an “ultra-premium” segment of the wine market, so we could point to their study as a justification for choosing that dollar value to demarcate a relatively high price.

A corresponding alternative definition for low-priced wines would be ones under 20 dollars. That roughly corresponds to the lower quartile of the wine prices. The total number of words used to describe any wine under 20 dollars is almost exactly equal to the total number of words used to describe any wine over 40 dollars.

If high- and low-priced wines are redefined in that way, then using the same dataset and other methods as before, our estimates of the relative probabilities for each of the “expensive” and “cheap” words identified by Krumme (2009, 2011) would be the ones reported in *Table A* of this appendix. We can draw the same conclusions as those reported in the main text, with only one exception. The exception is that the cheap word “enjoy” is now statistically significantly different from unity at conventional levels (p=0.01) rather than not. The point estimate is less than unity, as before and as expected.

In terms of the other results, about 56% of the unique words used to describe the over-40 and under-20 dollar wines are non-overlapping (in the same sense as the main text). And there were over twice as many instances of the cheap words as expensive words across the descriptions of the over-40 and under-20 dollar wines (calculating it either of the ways mentioned in the main text).

*C. Results for a different dataset and correspondingly different definition of high- and low-priced wines*

The findings reported in the main text speak to the robustness of Krumme’s results when using methods that are as similar as possible to hers and a different data source, but we can also explore the robustness of our results by using the exact same methods we used in the main text and yet another data source. Another source we can use is a dataset constructed by scraping the website of the wine retailer K&L Wines (hereafter, K&L). The dataset has prices and descriptions for about 50,000 wines priced between 5 and 200 dollars. Compared to the dataset studied in the main text, there are less than half as many wines covered (50,000 vs. the aforementioned 120,000), but the wine descriptions are over three times longer on average (with a median, mean, and standard deviation of 169, 190, and 115 words per description vs. the aforementioned 41, 42, and 11 words per review). The descriptions are longer because K&L is “an online aggregator of reviews” (to use the same phrase that Krumme, 2011, uses to describe her data source). For any given wine, K&L aggregates reviews from various experts such as *Wine Enthusiast*, *Wine Spectator*, *Wine Advocate*, or others. Another notable difference with the dataset studied in the main text is that the wines covered in the K&L dataset are twice as expensive on average (with a median, mean, and standard deviation of 50, 65, and 45 dollars vs. the aforementioned 25, 33, and 25 dollars). Like the dataset studied in the main text, for the K&L dataset, there is a positive correlation between a wine’s price and the length of its description in words (r=0.23, p<0.01).

 Given that the wines in the K&L dataset are more expensive, we will redefine high- and low-priced wines when studying it. The upper quintile of the prices of the wines in the K&L dataset is about 100 dollars, so we will now define wines over 100 dollars as “high priced” ones. We will define wines less than 30 dollars as “low priced” ones because, for the K&L dataset, 30 dollars roughly corresponds to the lower quintile of the wine prices and the total number of words used to describe any wine under 30 dollars is almost exactly equal to the total number of words used to describe any wine over 100 dollars. The dollar amounts we are using to define high- and low-priced wines for the K&L dataset are therefore obviously different in dollar terms from the ones used in the main text when studying the other dataset (30 and 100 vs. the aforementioned 15 and 50 dollars), but they are similar in terms of where they lie in the distribution of prices for each dataset.

 If we use the K&L dataset and redefine high- and low-priced wines in that way, then our estimates of the relative probabilities for each of the “expensive” and “cheap” words identified by Krumme would be the ones reported in *Table B* of this appendix. We can still draw the same overall conclusion that was stated in the main text, namely, that most of the expensive and cheap words identified by Krumme (2009, 2011) are reconfirmed to be indicative of a wine’s price. That is especially true of the cheap words. All but one of the 17 previously identified cheap words have relative probabilities that are less than unity and statistically significantly so. The one exception is “harvest.” As in *Table 2* of the main text and *Table A* of this appendix, that word has a relative probability that is not statistically significantly different from unity at conventional levels. For the previously identified expensive words, it is still the case that most have relative probabilities that are greater than unity and statistically significantly so, but it is now barely more than half of them. Six of the previously identified expensive words (“smoky,” ‘butterscotch,” “old,” and all three words in the food pairing group) have relative probabilities that are less than unity and statistically significantly so, which is the opposite of what would be expected based on Krumme’s findings. The conclusions when using the K&L dataset are therefore similar to the ones reported in the main text and above to some extent, but they are also more mixed. The at least somewhat different findings for the *WE* and K&L datasets call out for the sort of further study suggested in the conclusion of the main text. Perhaps a truly high priced wine is supposed to be appreciated in its own right and it would be uncouth for an expert to suggest it should be paired with a surf and turf of shellfish and steak.

In terms of the other results, about 40% of the unique words used to describe the over-100 and under-30 dollar wines are non-overlapping (in the same sense as the main text). If we add up every instance of any one of the 17 cheap words from Table 1 of the main text describing either an over-100 or under-30 dollar wine, and if we do the same for every instance of any of the 17 expensive words from Table 1, then those cheap words occur 1.1 times more often than the expensive ones. If we instead ignore the words from *Table B* of this appendix that did not have the relative probabilities we would expect based on Krumme’s findings, then the cheap words occur 1.6 times more often than the expensive ones.

**References not already cited in the main text**

Chen, B., Velchev, V., Palmer, J., and Atkison, T. (2018). Wineinformatics: A quantitative analysis of wine reviewers. Fermentation, 4(4), 82.

Hendrickx I., Lefever E., Croijmans I., Majid A., Van den Bosch, A. (2016). Very quaffable and great fun: Applying Natural Language Processing to wine reviews. Proceedings of the 2015 Meeting of the Association for Computational Linguistics. Available at <https://aclanthology.org/P16-2050.pdf>.

*Table A*

**Relative probabilities for Krumme’s expensive and cheap word examples
using the same *WE* dataset but a different definition of high- and low-priced wines**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | Relative probability of describing over 40 rather than under 20 dollar wine |
| Group | Subgroup | Word | Lower 95% CI | Point estimate | Upper 95% CI |
| Expensive words | Dark | Intense | 1.92 | 2.18\* | 2.48 |
|  |  | Smoky | 1.59 | 1.77\* | 1.98 |
|  |  | Supple | 1.66 | 1.99\* | 2.38 |
|  |  | Velvety | 6.68 | 8.63\* | 11.14 |
|  | Specific | Boysenberry  | 1.25 | 1.54\* | 1.89 |
|  |  | Butterscotch  | 1.01 | 1.34\* | 1.78 |
|  |  | Chocolate  | 2.19 | 2.38\* | 2.59 |
|  |  | Cigars | 0.35 | 1.30 | 4.92 |
|  |  | Tobacco | 3.23 | 3.65\* | 4.13 |
|  |  | Truffle | 11.97 | 25.29\* | 53.43 |
|  | Elite | Cuvee  | 6.34 | 12.33\* | 24.01 |
|  |  | Elegant | 3.04 | 3.46\* | 3.95 |
|  |  | Old | 2.20 | 2.64\* | 3.16 |
|  | Food pairings | Pork | 0.95 | 1.35 | 1.92 |
|  |  | Shellfish  | 0.17 | 0.29\* | 0.50 |
|  |  | Steak | 1.51 | 2.07\* | 2.85 |
|  | Other | Vintage | 3.10 | 3.55\* | 4.08 |
| Cheap words | Light | Bright  | 0.49 | 0.52\* | 0.55 |
|  |  | Fresh | 0.35 | 0.37\* | 0.39 |
|  |  | Light  | 0.30 | 0.32\* | 0.33 |
|  |  | Pink | 0.20 | 0.24\* | 0.29 |
|  |  | Tropical | 0.24 | 0.26\* | 0.30 |
|  | General | Clean | 0.14 | 0.16\* | 0.17 |
|  |  | Fruity | 0.22 | 0.24\* | 0.26 |
|  |  | Good | 0.37 | 0.40\* | 0.42 |
|  |  | Juicy | 0.67 | 0.72\* | 0.77 |
|  |  | Tasty | 0.33 | 0.40\* | 0.49 |
|  | Accessible | Enjoy | 0.92 | 1.07 | 1.26 |
|  |  | Pleasing | 0.61 | 0.78\* | 0.98 |
|  |  | Refreshing | 0.15 | 0.17\* | 0.19 |
|  |  | Value | 0.03 | 0.04\* | 0.05 |
|  | Food pairings | Chicken  | 0.11 | 0.16\* | 0.25 |
|  |  | Pizza | 0.02 | 0.04\* | 0.08 |
|  | Other | Harvest | 0.67 | 1.05 | 1.62 |

*Note:* The upper and lower parts of this table are the same as Tables 2 and 3 of the main text, except these estimates are based on the alternative dataset discussed in the text of this appendix and defining high- and low-priced wines as ones over 100 and under 30 dollars. This table reports, for each word Krumme (2009, 2011) identified as an “expensive” or “cheap” one, our point estimate and a 95% CI for the relative probability of the word being used to describe an over-100 rather than under-30 dollar wine.

\* Statistically significantly different from unity at the 5% level

*Table B*

**Relative probabilities for Krumme’s expensive and cheap word examples
using an alternative dataset and a correspondingly different definition of high- and low-priced wines**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | Relative probability of describing over 100 rather than under 30 dollar wine |
| Group | Subgroup | Word | Lower 95% CI | Point estimate | Upper 95% CI |
| Expensive words | Dark | Intense | 1.44 | 1.53\* | 1.62 |
|  |  | Smoky | 0.75 | 0.81\* | 0.87 |
|  |  | Supple | 0.88 | 0.95 | 1.02 |
|  |  | Velvety | 1.55 | 1.71\* | 1.89 |
|  | Specific | Boysenberry  | 0.80 | 0.93 | 1.08 |
|  |  | Butterscotch  | 0.35 | 0.50\* | 0.71 |
|  |  | Chocolate  | 1.54 | 1.64\* | 1.74 |
|  |  | Cigars | 1.59 | 3.51\* | 7.75 |
|  |  | Tobacco | 1.53 | 1.64\* | 1.76 |
|  |  | Truffle | 2.68 | 3.26\* | 3.98 |
|  | Elite | Cuvee  | 1.16 | 1.28\* | 1.40 |
|  |  | Elegant | 1.05 | 1.11\* | 1.18 |
|  |  | Old | 0.51 | 0.54\* | 0.58 |
|  | Food pairings | Pork | 0.03 | 0.05\* | 0.11 |
|  |  | Shellfish  | 0.05 | 0.10\* | 0.23 |
|  |  | Steak | 0.34 | 0.47\* | 0.66 |
|  | Other | Vintage | 1.46 | 1.54\* | 1.61 |
| Cheap words | Light | Bright  | 0.63 | 0.66\* | 0.69 |
|  |  | Fresh | 0.56 | 0.59\* | 0.62 |
|  |  | Light  | 0.35 | 0.37\* | 0.40 |
|  |  | Pink | 0.12 | 0.15\* | 0.21 |
|  |  | Tropical | 0.20 | 0.24\* | 0.28 |
|  | General | Clean | 0.33 | 0.37\* | 0.41 |
|  |  | Fruity | 0.44 | 0.49\* | 0.55 |
|  |  | Good | 0.81 | 0.84\* | 0.88 |
|  |  | Juicy | 0.41 | 0.44\* | 0.46 |
|  |  | Tasty | 0.14 | 0.19\* | 0.24 |
|  | Accessible | Enjoy | 0.44 | 0.51\* | 0.58 |
|  |  | Pleasing | 0.28 | 0.38\* | 0.51 |
|  |  | Refreshing | 0.15 | 0.17\* | 0.20 |
|  |  | Value | 0.03 | 0.04\* | 0.05 |
|  | Food pairings | Chicken  | 0.07 | 0.12\* | 0.20 |
|  |  | Pizza | 0.00 | 0.03\* | 0.22 |
|  | Other | Harvest | 0.78 | 0.93 | 1.11 |

*Note:* The upper and lower parts of this table are the same as Tables 2 and 3 of the main text, except these estimates are based on the alternative dataset discussed in the text of this appendix and defining high- and low-priced wines as ones over 100 and under 30 dollars. This table reports, for each word Krumme (2009, 2011) identified as an “expensive” or “cheap” one, our point estimate and a 95% CI for the relative probability of the word being used to describe an over-100 rather than under-30 dollar wine.

\* Statistically significantly different from unity at the 5% level