# Supplementary Appendix

*[FOR ONLINE PUBLICATION ONLY]*

This appendix provides additional details to supplement the main text of the paper “Willingness to Pay for Wine Bullshit: Some New Estimates.” 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: <<https://osf.io/ycgbh/>>.

# I. More on the Introductory Remarks

A few remarks related to those in the introduction are as follows.

● When Quandt was compiling his list of bullshit terms from the 24 wine descriptions he collected, he excluded what he considered to be “bullshit-free” terms that are “broadly understood by all to have a common meaning” (Quandt, 2007, p. 131). Note that if the full-text versions of the 24 descriptions were available, then it would be possible to reconstruct terms he considered bullshit free. Marginal willingness to pay (MWTP) for wines described by those terms could be explored, too, perhaps. Unfortunately, except for the three descriptions reproduced in his article, the full-text versions of those descriptions have been lost over time (personal correspondence with Richard Quandt on June 14, 2018).

● Lists of wine descriptors even more extensive than Quandt’s (2007) have been compiled. Klem (2009), in particular, compiles almost 40,000 wine descriptors culled from real reviews by wine critics. Klem does not describe his descriptors as “bullshit,” but he admits “I don’t understand a lot of them” (ibid., p. 1), and the list is meant to be somewhat humorous.

● A notable list of wine descriptors somewhat similar to Quandt’s (2007) is from Amerine and Roessler (1976). As a preamble to their list, they write: “Unfortunately, existing wine terminology abounds in words and phrases that have little or no clearly definable meanings with respect to the sensory evaluation of wines. […Q]uite a few [terms] are simply ridiculous” (ibid., p. 203). They go on to write: “It is not our intent to condemn the following terms (although some of them deserve it) for your wine vocabulary, but merely to warn you to use them with caution, if at all” (ibid., p. 204). They then list 113 terms ranging from “austere” to “zestful.” Terms that overlap with Quandt’s list include the aforementioned “austere,” as well as “chewy,” “dumb,” “meaty,” “oily,” “silky,” and “smoky” (Amerine and Roessler, 1976, p. 204; Quandt, 2007, p. 133).

● In work on wine prices and descriptions that was reported in a conference presentation (Krumme, 2009) and popular piece (Krumme, 2011), Krumme used a Naive Bayesian classifier approach to identify descriptors that strongly predicted whether wines were high or low priced.

Some details on how that was done are unclear, including where the data came from (beyond “an online aggregator of reviews” mentioned in Krumme 2011) and what price level demarcated high-priced wines from low-priced ones (a median price or something else, perhaps), but similar approaches have been applied to wine descriptions since her groundbreaking work. See, for example, Chen et al. (2018) and Croijmans et al. (2019) for more recent applications of similar techniques to wine descriptions.

Descriptors identified by Krumme (2011) as strongly predicting that a wine would be high- rather than low-priced—or “expensive words,” as she called them—included “velvety,” “chocolate,” “elegant,” and others. Like any other set of descriptors, a hedonic regression or matching approach could perhaps be used to try to estimate price premiums for those words. Of course, the fact that prices were used to identify the “expensive words” already suggests the conclusion that should be found. The use of prices to try to select variables to try to explain prices would also seem to jeopardize the validity of standard approaches to statistical inference.

● For their hedonic regression, Chen and McCluskey (2018) include indicators for six descriptors as explanatory variables (as mentioned in the main text of my paper). The six descriptors were “berry,” “cherry,” “currant,” “finish,” “spice,” and “tannin.” Those descriptors were selected because they occurred frequently throughout their dataset. The descriptors were also seen as “sensory characteristics” (ibid., p. 392). Chen and McCluskey (2018) do not control for any other descriptors in a wine’s description besides those six.

None of those six descriptors are *identical* to any of the ones that Quandt deemed to be bullshit (as stated in the main text), but each except “finish” does appear as *part* of some of the Quandt descriptors. “Tannin” appears in “silky tannins” and “velvety tannins,” for example. Some of the Quandt descriptors would therefore count as incidences of the descriptors considered by Chen and McCluskey (2018), although many other incidences would count, too.

● When applying their hedonic approach, Chen and McCluskey (2018) find effects on the order of at most $1 for the six descriptors they considered (as stated in the main text). To expand on that: They consider all the wines in their sample pooled together and, also, split those wines into four subsamples based on their prices. For the pooled sample, the largest effect they find in dollar terms is a $1.53 change in price if one of the descriptors (specifically, “tannins”) appears in the description (ibid., p. 397). Of note, relative to the average price of the wines in their pooled sample (which was $30.17; ibid., pp. 391, 393), that change would be about 5% (specifically, $1.53/$30.17). The marginal effects of the other five descriptors were all estimated to be less than $1.

When they split their data into four subsamples, the largest effect they found in dollar terms was that, among “ultra-premium” wines that cost at least $40 per bottle, there is a $2.34 change in price if one of the descriptors (which was “tannins” again) appears in the description of a wine (ibid., p. 397). Of note, the largest that effect could be in percentage terms would be relative to the minimum price of an ultra-premium wine, so it could be as large as about 6% (specifically, $2.34/$40).

Instead of subdividing a dataset, quantile regression (Koenker and Hallock, 2001) on the entire dataset would be another way to explore whether an explanatory variable’s effect (or association) varies with price. Quantile regression may be the only way to explore that for rare descriptors; the more finely a dataset is subdivided and the rarer the descriptor, the less likely the descriptor is to appear in any given subsample. See below for how my ordinary least squares (OLS) hedonic regression results reported in the main text compare to some quantile regression results.

# II. More on the Hedonic Regression

## A. More on the Data and Methods for the Hedonic Regression

### 1. More on selecting the data source

● Recall I used information provided by the website of one of the three retailers considered by Quandt (2007 to create a dataset of wine prices and descriptions. I used the website for K&L Wine Merchants (hereafter, K&L). Their website is <<https://www.klwines.com/>>. The website for the Chicago Wine Company is <<https://www.tcwc.com/>>. The website for the third retailer (now Central Wine Merchants) is <<https://www.centralwinemerchants.com/>>. K&L’s site offered substantially more information than either of the other two retailers’ sites, at least as of writing. K&L has roughly 10,000 wines for purchase (either in stock or pre-arrival) as of writing, which is more than the roughly 5,000 available from the Chicago Wine Company, and an order of magnitude more than the few hundred available from the other retailer. K&L’s online store also has about 200,000 out-of-stock wines as of writing.

Other online retailers more comparable to K&L in terms of their wine selection and aggregation of ratings and descriptions include Wine.com, which has about 15,000 for-sale wines and almost 400,000 out-of-stock wines as of writing. It is difficult to ascertain how many wines have expert descriptions on Wines.com without scraping their website, but according to the site’s built-in search function, Wine.com appears to have about 70,000 wines with a rating by one of the critics they aggregate. K&L appears to have about 100,000 wines with a rating, by comparison (including out-of-stock wines for both sites). K&L’s site therefore seems to offer more information than—or at least a comparable amount of information to—Wine.com’s site.

There may be other online retailers that aggregate wine descriptions, but to the extent there are, K&L’s selection is hopefully comparable. Krumme (2009, 2011, and 2013) used an unnamed online aggregator of expert wine reviews. The author of those works did not respond to a request to clarify the data source (personal correspondence on July 19, 2018), although it was presumably one of the ones cited above or, again, hopefully comparable to K&L’s.

● Instead of using data from only *one* online wine retailer, information from *more than one* could be used. Doing so could provide information on more wines (to the extent that different retailers offer different wines), as well as more information on the same wines (to the extent different retailers offer the same wines and different information about them). There would however be additional complications to trying to do so. Determining which wines are the same or different across retailers could be difficult if the same wines can be listed under slightly different names. Exactly how information on overlapping wines should be combined would also need to be considered. See below for how I dealt with duplicate wines on K&L’s site.

● As a side note on other possible data sources: Several studies have independently used *Wine Spectator*’s or *Wine Enthusiast*’s sites to construct their own datasets. There are issues of size and/or availability for each of those datasets.

Chen and McCluskey (2018) and Ramirez (2010) constructed datasets from *Wine Spectator* reviews. Their datasets are relatively small (6,085 and 2,700 wines, respectively), so they would not necessarily be the best to use even if they were available for use. The author of Ramirez (2010) did not respond to a request for his dataset (personal correspondence on July 30, 2018). The corresponding author of Chen and McCluskey (2018) understandably preferred to have the author who constructed the dataset be a co-author on future work derived from that dataset (personal correspondence on July 18, 2018).

Chen et al. (2018) constructed a large dataset of over 100,000 wines (specifically, 105,085 wines) from *Wine Spectator*, but at least for the version of their dataset they provided upon request, the full text of wine reviews were not included. Only indicator variables for less than 1,000 descriptors were included (ibid., p. 3). Importantly for my study, many of the descriptors that Quandt (2007) deemed to be bullshit were not included.

Croijmans et al. (2019) is another example of a study that constructed a large dataset of wines from *Wine Enthusiast* reviews (specifically, 76,410 reviews; ibid., p. 4). They declined to share their dataset due to possible legal concerns over doing so (personal correspondence with corresponding author on June 2, 2019).

Looking beyond any issues with the size or availability of those datasets constructed by previous studies, I could have independently used the websites of *Wine Spectator*, *Wine Enthusiast*, or another wine reviewer like *Wine Advocate* to construct a dataset. There are however at least two advantages to using K&L’s site rather than any one of those sites.

The first advantage relates to the fact that K&L’s site aggregates reviews from several reviewers. Although not all of the wines in the dataset I constructed from K&L’s site have multiple reviews, many do. For sites like *Wine Spectator*, in contrast, they have (or generally have) only one review per wine. Any one review is unlikely to include synonymous descriptors, assuming reviewers wish to avoid repeating exactly or nearly equivalent descriptors. Indeed, the closer descriptors are to being synonymous, the less likely it should be that they co-occur in a given critic’s review of a given wine. If *multiple* reviewers are reviewing the same wine, then each reviewer should be as likely to use one synonym as any other (or, at least, that should be the case if there is some aspect of a wine that wine reviewers would identify and they have a shared language of synonyms to describe that aspect).

The second advantage to using K&L’s site relates to the fact that it is an online wine retailer. *Wine Spectator* and other sites have their own business models (subscription paywalls, affiliated linking, advertising space, or other practices), but with any site that is not an online retailer, it becomes much less obvious which if any wine reviews may be informing consumers’ purchasing decisions. With K&L’s site, a wine’s price and other information about it are all presented at the same time in the same place. There is a direct, obvious way in which reviews could be informing purchasing decisions.

● Instead of using data from only *one* wine reviewer’s site, information from *more than one* could be used. Yet that is what one is doing with greater ease by using data from K&L’s site or another similar site that aggregates wine reviews.

● As another side note on another possible data source: To apply his hedonic approach, McCannon (2020) constructed two datasets of 124 and 238 relatively homogenous wines. Even if those two datasets were combined, they would still be quite small. The data source he used to construct his datasets was apparently a website called “winereviewsonly.com” (ibid., p. 7). That website is not currently available, and none of it was ever archived by Internet Archive (<[https://web.archive.org/web/\*/winereviewsonly.com/\*](https://web.archive.org/web/*/winereviewsonly.com/*)>), so that was not an option for my study.

● It should be recognized there may be ethical and legal questions involved with using information provided by any website (Gold and Latonero, 2018) and additional questions when using information provided by a website that aggregates the work of Robert Parker and others. The approach of my study and similar previous ones (Chen and McCluskey, 2018; Chen et al., 2018; Croijmans et al., 2019; McCannon, 2020; Krumme, 2009, 2011, and 2013; Ramirez, 2010) may not be models to follow. Quandt’s (2007) use of information contained in promotional emails from wine retailers is similar in substance, although not scale, to my approach.

### 2. More on constructing the dataset

● Recall that, when creating that dataset, I chose to include for-sale and out-of-stock wines that were being sold or had been sold as a single 750ml bottle, had at least some description, had at least one numerical rating from any of the wine critics, and were being or had been sold for between $5 and $200. Those choices can be explained as follows.

In terms of including out-of-stock wines: It is admittedly not ideal to do so. The prices at which out-of-stock wines are listed are not necessarily the prices at which they would be bought and sold if they were for sale. K&L’s site also does not make it clear when a wine went out of stock, so it is unclear how dated or up-to-date a listed price might be. Yet ignoring out-of-stock wines would dramatically reduce the size of the dataset. Similar concerns about potentially misleading prices also exist for other possible data sources. Chen and McCluskey (2018) use wine descriptions written over a 16-year period with prices that were only current at the time when a description was written (ibid., p. 391). Those prices are not necessarily the prices at which the wines would be bought and sold today, either, and not necessarily comparable to each other.[[1]](#footnote-1)

In terms of only considering wines sold as a single 750ml bottle: Although selling the same wine in a different-sized bottle or in several bottles may affect the wine’s per-unit price, it is not the purpose of this study to try to examine any bottle-size effects (as in Drichoutis, Klonaris, and Papoutsi, 2017) or quantity discounts. The wine descriptions aggregated on K&L’s site apply (or presumably almost always apply) to the wines themselves rather than the size of their bottles or the number of bottles packaged together.

In terms of only considering wines that have at least some description: The alternative approach would be to include all wines, even those without any description. In that case, a descriptor could be absent because: no wine reviewers wrote a description of the wine; a reviewer wrote a description that may or may not have contained the descriptor, but K&L did not include that description on its site; or a reviewer wrote a description, it is included on K&L’s site, and it does not contain the descriptor. The question of whether a descriptor has been ascribed to a wine would therefore be wrapped up with the question of whether any reviewers have described the wine and which if any reviews have been included on K&L’s site. I choose to focus on the simpler question of whether, for wines that have descriptions on K&L’s site, a descriptor is included in the description or not.

In terms of only considering wines that have at least one rating by a reviewer: That ensures at least one of the wine critics whose ratings and descriptions are aggregated by K&L’s site will have reviewed the wine if only by giving it a rating and presumably by also writing a description. It also ensures that each wine will have at least one rating that can be used as a control in a hedonic regression.

In terms of only considering bottles between $5 and $200: That is the same price range used by Krumme’s (2011) study of wine prices and descriptions. That price range is somewhat arbitrary, but there are reasons to only consider bottles in such a range. A practical reason to exclude bottles below $5 is that some out-of-stock wines on K&L’s site are listed at incredibly low prices, including as low as $0. A floor of $5 appears to be high enough to exclude such misleading prices. A more principled reason is that anything too cheap may be part of a different market segment. Similarly, anything too expensive may be too different.[[2]](#footnote-2) More practical reasons to exclude wines that are too expensive include that, for a hedonic regression or matching approach, results might be highly sensitive to a few high-priced wines and, for a stated-preference survey, few respondents are likely to spend so lavishly on wine. Bottles studied by Chen and McCluskey (2018) cover a similar price range, specifically, $5 to $185 (p. 391).

● To clarify how I collected the data: I used Scraping Hub’s Portia 2.0 to create a spider to crawl K&L’s website. Scraping Hub was also used to run that spider. To crawl K&L’s site efficiently, I used that site’s built-in search function to generate several hundred pages of search results. I then crawled each page of those results, going one level deeper to each wine in the results.

The options I selected for K&L’s built-in search were, first of all, that I was searching for wines (not spirits). Among wines, I was searching for ones that were in 750ml bottles, that had a wine review score anywhere from 1 to 100 (thereby ensuring at least one numerical rating and generally ensuring at least some description), and that were priced between $5 and $200. Finally, I opted to include out-of-stock items in the search results.

As a technical note, there is an upper limit of 100,000 for the number of items that can be returned by the site’s built-in search function. Pages of search results beyond the 100,000th item return a “Sorry, an error occurred while processing your request” message. That upper limit is the main reason why I restricted the search results rather than trying to return all wines and then narrow them down. Another reason is that I can defer to K&L’s own structuring of their data (with respect to the sizes of bottles, expert ratings assigned to wines, and the prices of the wines) rather than trying to recreate such structure.

After obtaining all that information, I processed it further using a Python 2.7 program that can be found through the link at the start of this appendix.

● Another choice made when constructing the dataset was as follows. Along with its normal retail operations, K&L’s online store has some wines that are sold through auction. When constructing my dataset, I ignored wines that were being sold—or, for out-of-stock wines, had been sold—through auction. The prices for auctions that are still open could be misleading. The prices for auctions that have closed could be considered, perhaps, although it seems questionable whether retail prices set by K&L should be conflated with prices at which auctions close. There are relatively few auctions for wines or lots of wines that have closed on K&L’s site (less than 6,000, as of writing), so ignoring auctions does not dramatically affect the size of the dataset.

● The final choices made when constructing the dataset were as follows.

Some wines listed under different stock-keeping units (SKUs) have almost the same names, except for a note added to their name. The note is typically in parentheses. For some of those wines, the note is about damage (such as a soiled label) or other distinctions (such as very high shoulders). I ignored those wines because there is an otherwise equivalent wine without that noted distinction in my dataset, and I am not trying to estimate the price effects of a soiled label, shoulder levels, etc.

Another note added (again, to wines listed under different SKUs but having almost the same names except for an added note) is about pre-arrivals. I ignored those wines because, given that there was another wine with the same name except for a pre-arrival note, the wine apparently arrived and is in my dataset.

Other notes include that the wine is a duplicate or the viewer should see a different SKU. I ignore those wines because they do indeed appear to be duplicates, given that there is another otherwise equivalent wine in the dataset.

Some wines listed under different SKUs have the exact same names. I treated those wines as follows. If everything about the wines was the same (including their listed prices, ratings, descriptions, etc.), I dropped all but one of the observations; they are just duplicate listings.

If everything about the wines was the same except their listed prices, I took a simple average of their prices, assigned that average price to one of the wines, and dropped the others from my dataset; using that average price seems more reasonable than systematically or randomly dropping higher or lower priced duplicates.

Some of the wines that had the exact same names but were listed under different SKUs also differed in more dimensions than just their listed price; some differed in their descriptions or ratings, in particular. I chose to leave all of those observations in my dataset. I did so partly because there are a relatively small number of wines for which that is the case; there are 321 observations for which that is the case, all but a few have the same name as only one other wine, and no wines have the same name as more than two other wines. Another, more principled reason to do so is that, if descriptions or ratings of a wine have an affect on or association with price, then I should not necessarily take it upon myself to aggregate descriptions even further than K&L did, aggregate ratings of different critics even further than K&L did, or average ratings of the same critic; I should allow any effects or associations to reveal themselves, even if—and perhaps especially if—different prices, descriptions, and ratings are association with the exact same wine.

After doing all of that, my dataset has 51,043 observations in total (as stated in the main text).

### 3. More on the dataset

● When previous studies have applied the hedonic pricing approach to wine, they have usually applied that approach to relatively homogenous samples in terms of grape varietal, geographic origin, or other dimensions (see, e.g., Chen and McCluskey, 2018; McCannon, 2020; Ramirez, 2010). A reason to do that (beyond the practical reason that it is difficult to collect data on the entire universe of wines) is that estimates based on a homogenous sample may be different from estimates based on a pooled sample of more heterogeneous wines.

My dataset is a heterogeneous sample of wines. For a hedonic regression, it might be attractive to only consider a more homogeneous sample (or, better yet, to test for any differences between a pooled effect estimated for a heterogeneous sample and unpooled effects estimated for more homogenous samples), but the fact that most words occur relatively infrequently makes that difficult or impossible to do so. A descriptor needs to occur at least once in order to generate any estimate of its effect and, ideally, the descriptor would appear many times in order to estimate its effect with any confidence.

A matching approach becomes attractive if there is concern about the comparability of wines described by a given descriptor, on the one hand, and some wines *not* described by that descriptor, on the other hand. It might be unimaginable for experts to describe some wines by some descriptors, for example.

● In addition to aggregating reviews and ratings by well-known wine critics, K&L sometimes provides its own description denoted by “K&L Notes” (as mentioned in the main text). Those descriptions can be an original review by an in-house critic, a reproduction of a review by an outside critic (e.g., “*Wine Spectator* says, …”), or a reproduction of a producer’s description (e.g., “Winemaker’s Notes: …”).

There is a similar issue with at least some of the other sources that K&L aggregates. A given source like *Wine Enthusiast* has more than one person reviewing wines.

I do not attempt to distinguish such variation within a source, largely because of the difficulty of trying to do so. I do not explicitly control for it in my hedonic regression, either. To the extent it is important to control for variation within a source (and not just between sources), the only hope would be that my other control variables in the hedonic regression are somehow able to control for it. Certain reviewers may only review wines from certain regions, for example.

● Producer names could perhaps be extracted from the wine descriptions, but doing so would require at least a semi-supervised approach, so I did not attempt to do that. A producer’s name is one of the many possible aspects of a wine description that my text-matching estimator should at least partly account for.

### 4. More on the allowed variations of the Quandt descriptors

● I allowed slight variations on the Quandt descriptors to count as incidences (as stated in the main text). That choice can be defended as follows.

When counting instances of a given descriptor from Quandt’s, I could have only counted instances where the exact descriptor appears as a whole word. The concern with that approach is that it ignores even the slightest variation on a term; so, for example, even though Quandt discusses “zesty minerals” (plural) as a bullshit term in the text of his article, that term would not count because only “zesty mineral” (singular) appears in his list of 123 descriptors.

The concern with allowing any variation on a term is that seemingly slight variations may be quite different. A “dry” (as in no residual sugar) wine is different from a “drying” (as in tannic) wine, to borrow an example from James (2020, p. 154).

To try to balance both of those concerns, I allowed variation by using a supervised approach to identify what I considered slight variations on the 123 descriptors in Quandt’s list. As stated in the main text, for a term such as “jam” (singular) from Quandt’s list, I allowed “jams” (plural), “jammy,” “jamlike,” “jamminess,” and “jamish” to count as equivalent instances. See the above-mentioned Python program for the exact variations allowed for each descriptor.

● When considering the variations to allow, I tried to err on the side of staying true to the Quandt descriptors. So, for example, for his “wild berry” (singular) descriptor, I allowed “wild berries” (plural), but I did *not* allow a bunch of other possibilities such as “wild blackberry,” “wild blueberry,” “wild boysenberry,” “wild cranberry,” “wild elderberry,” “wild huckleberry,” “wild mulberry,” “wild raspberry,” and “wild strawberry,” which all appear in wine descriptions in my dataset. I did not allow those other possibilities because a type of berry is at least somewhat more specific than just “wild berry.”

● Although it should be recognized that the variations I allow might have somewhat different meanings, it should also be recognized that, depending on the context in which a whole-word descriptor is used, that descriptor could have a meaning that varies at least as much. As part of his own compilation of wine descriptors, Klem (2009, pp. 318–322) lists hundreds of partitive expressions ranging from “abundance of” to “zip of” that could precede and thereby modify the meaning of any given descriptor. The difference between an “abundance of jam” and a “zip of jam” might be at least as great as the difference between “jammy” and “jammish.” The myriad of contexts in which a descriptor could be used is still a concern, but allowing more variations may help address that concern; doing so should tend to minimize the influence of any particular context in which a descriptor is used and hopefully yield more robust hedonic estimates. Ultimately, however, a text-matching estimator like the one I apply—or one that uses even richer representations than my bag-of-words representation—may be the only way of trying to control for the myriad of contexts in which a descriptor is used.

### 5. More on the control variables in the hedonic regression

● I use a large number of control variables in my hedonic regression. The controls hopefully ensure that any estimate of the effect of a descriptor on price is due to the descriptor itself or something non-obvious rather than due to a correlation between the descriptor and something more obvious about the wine (as stated in the main text). It is possible that certain descriptors may be more likely to appear in the descriptions for wines of a certain varietal, geographic origin, or vintage. Certain descriptors may also be more likely to appear in the reviews written by certain reviewers. Certain reviewers may also be more likely to use certain descriptors when they assign certain ratings. Longer reviews may also be more likely to contain any given descriptor. If those correlations exist but are not controlled for, then the estimated effect of a descriptor would be confounded with such things. I therefore include controls for a wine’s varietal, geographic origin, vintage, reviewers, ratings by reviewers, and description length. The control variables in my regression are detailed in *Appendix Table 1* below.

### 6. Some more descriptive statistics

● Pairwise correlations between the variable to be explained by my hedonic regression (namely, wine prices), the relatively large number of explanatory variables of interest (namely, the 106 indicator variables for the Quandt descriptors that appear in the descriptions of at least 50 wines in my dataset), and the even larger number of control variables (all of which except the length of a wine description are dichotomous variables) can be explored. It is cumbersome to present too many correlations, but *Appendix Table 2* presents the correlations between select variables in my dataset.

Only one pair of variables in that table has a correlation coefficient of at least 0.10 and thereby reaches the threshold for at least a “small” correlation according to Cohen’s (somewhat arbitrary but standard) rules of thumb for the effect size of a correlation coefficient. The one pair is wine prices and description lengths, which have a correlation coefficient of 0.23. A positive correlation between prices and descriptions is consistent with the results of Ramirez (2010) who finds that longer wine descriptions cause (or are associated with) higher wine prices.

*Appendix Table 2* also shows that there is at least a slightly positive correlation between whether a given descriptor appears in a wine’s description, on the one hand, and the length of the wine’s description, on the other hand. Such a positive correlation would be expected. Indeed, if it is possible for a given descriptor to be used to describe a given wine, then that descriptor should eventually appear if the description is long enough.

For the select Quandt descriptors presented in *Appendix Table 2*, their correlations with each other are all essentially zero. “Silky tannins” and “velvety tannins” have a correlation of 0.01, for example. Some descriptors have stronger correlation with each other than that, as discussed below.

● The lack of any strong correlations among the select descriptors in *Appendix Table 2* is perhaps not surprising. With the exception of “silky tannins,” the descriptors considered in that table are the ones that have statistically significant effects according to my hedonic regression results. Highly multicollinear variables would tend to be statistically insignificant in an OLS regression.

● Some of the Quandt descriptors do have significant positive or negative correlations with each other. Nested descriptors such as “silky” and “silky tannins” are strongly positively correlated, of course, but even some of the non-nested descriptors are correlated, too, as discussed below.

● *Appendix Table 3* does not report the strong correlations between nested descriptors such as “silky” and “silky tannins,” but it shows the non-nested Quandt descriptors that have the strongest positive or negative pairwise correlations with each other. All of those correlations are statistically significant at the 10% level and exceeding Cohen’s 0.10 threshold for at least a “small” correlation coefficient.

Examples of terms with strong *positive* pairwise correlations include “blackberry” and “cassis,” “meaty” and “smoky,” “leather” and “tobacco,” and “black cherry” and “plum.” A wine expert might have been able to guess that those terms go together or go with a certain type of wine.

Terms that have strong *negative* pairwise correlations mostly include “honey” and some of the fruiter descriptors that Quandt deemed to be bullshit like “black cherry,” “plum,” and “raspberry,” but not “quince,” which has a strong positive correlation with “honey.” Trying to rationalize those associations helps illustrate Quandt’s argument that many of these descriptors are at best ambiguous. (Black cherries, plums, and raspberries are perhaps too sour to be associated with honey, … but quince is at least as sour as those other fruits, … unless it is mixed with sugar to be as sweet as honey, … but the other fruits are also often mixed with sugar, too, … so maybe the key difference between those other fruits and quince is that the former fruits have the hue of red wine while quiches and honey have more of the hue of a white wine.)

● To further explore the relationship between the descriptors and the other variables in my hedonic regression, *Appendix Table 4* shows statistics for wines described by select descriptors. So for example, that table shows there are 1,340 wines in my dataset described as having “velvety tannins” and those are about 75 dollars on average, whereas there are 1,940 wines in my dataset described as having “silky tannins” and those are about 71 dollars on average. Note that a comparison of the average price of wines described by each of those descriptors is a naive matching estimate of the descriptors’ relative effect on (or association with) price.

As shown in that same table, wines described as having velvety tannins are similar to wines described as having silky tannins in terms of the length of their descriptions. The table also shows that: a plurality of both wines are also classified as “Cabernet Sauvignon and Blends” by K&L; a plurality have a vintage of 2010; and Parker was the most frequent reviewer of them. Somewhat interestingly, a plurality of wines described as having “velvety tannins” come from France whereas a plurality of wines described as having “silky tannins” come from the US.

● The largest number of Quandt’s 106 descriptors (which again is the number of his descriptors that appear in at least 50 wines in my dataset) to appear in any wine’s description is 26; one is an out-of-stock “2016 Pape Clément Pessac-Léognan” described by nine reviewers on K&L’s site and listed at $99.99. The other is an out-of-stock “2015 Valandraud St-Émilion” described by eight reviewers and listed at $169.99. More information on the former can be found at: <<https://web.archive.org/web/20200716025658/https://www.klwines.com/p/i?i=1299400>>. More information on the latter can be found at: <<https://web.archive.org/web/20200621174040/https://www.klwines.com/p/i?i=1201342>>. According to Wine-Searcher.com, the average price before taxes for a 750ml bottle of the former and latter wines is about $120 and $190 as of writing. Prices listed for out-of-stock wines on K&L’s site do not necessarily reflect what they would be sold for today, as discussed above.

### 7. More on the log-linear form of the hedonic regression

● I use log prices as the variable to be explained in my hedonic regression partly because the unlogged prices are heavy tailed. The median price of the wines in my dataset is $49.99 while the mean is about $65, as stated in the main text. By comparison, the mean of the log prices is about 3.93, and the median of the log prices is similar to that (specifically, about 3.91).

● There are other reasons to use log prices, too. Doing so implies the effects of the explanatory variables on price (or, at least, the associations between the explanatory variables and price if causality cannot be assumed) are in terms of *percentage* changes in price rather than *dollar-value* changes. That seems important, given that half the bottles in my dataset are about $50 or more. Most wine consumers do not typically pay that much for a bottle, according to my survey. To the extent that the same effects apply to both relatively high- and low-priced wines, the effects would presumably need to be in percentage terms; otherwise, even small dollar-value changes in high-priced wine might be incredibly large for low-priced wines and even large dollar-value changes in low-priced wines might be trivially small for high-priced wines.

### 8. More on other methods for incorporating wine descriptions into hedonic regressions

● The main text of my paper mentions that McCannon (2020) uses simplified representations of wine descriptions in his hedonic regressions. To be more specific: He starts by estimating a standard “topic model” for his wine descriptions. The topic model assumes each wine description is a probabilistic mixture of a given number of latent “topics” and that each topic is a probabilistic mixture of terms. Or, to put it in more intuitive terms: The model assumes that, if certain terms are used and used together across different wine descriptions, then wine descriptions with those terms are more likely to be about the same “topic.” The main parameterization of the model that a research must decide upon is the number of latent topics to presume. McCannon (2020) considers as few as three for a “coarse division” of topics (p. 6) and as many as 10 for a “finer categorization” (p. 20).

When preparing the text of wine descriptions for his topic model, McCannon (2020) used a “light touch” by removing basic punctuation but not manipulating the descriptions in any other way (such as spell checking, removing commonly occurring terms, or stemming or lemmatizing; p. 8).

After estimating the topic model, point estimates for the probability that a given wine description is about a given topic can be obtained from that model. McCannon (2020) uses those estimated “topic probabilities” as explanatory variables in otherwise standard hedonic regressions.

● McCannon (2020) treats topic probabilities as the main explanatory variables of interest in his otherwise standard hedonic regressions, but using those topic probabilities in that way raises issues. The standard errors for the hedonic estimates should be adjusted to account for the fact that the topic probabilities are themselves estimated. In addition, although the hedonic estimates can be used to say that an increase in a given topic probability is associated with some change in price, as McCannon (2020) says, what is left unsaid is how the probability of that latent topic could increase in the first place. The only way it could increase is if a term is added to (or removed from) a description. But the presence (or absence) of a term could and generally would affect *all* topic probabilities because a topic model makes those probabilities a function of all terms (Gentzkow, Kelly, and Taddy, 2019, sec. 3.2.1).

● Other possible approaches to generating a lower-dimensional representation of the wine descriptions include an approach by Ramirez (2010). In addition to using the length of a wine description, he uses the amount of “analytical” and “non-analytical” terms in the description. As he defines them, an “analytical” term is any term defined in a wine glossary from the same source as the wine descriptions he studied, namely, *Wine Spectator* (ibid., p. 145). “Non-analytical” terms are everything else. I am using reviews written by several reviewers, so his approach would need to be adapted somewhat, perhaps by using several wine glossaries. Some analytical terms would also need to be ignored to avoid confounding because some of the Quandt descriptors are defined in wine glossaries (Quandt, 2007, p. 133). Yet even if Ramirez’s approach could be directly adopted, he recognized that different terms may have different effects on (or associations with) price (Ramirez, 2010, p. 159), so a richer representation of the text of the wine descriptions would be better.

● Given that Ramirez (2010) defines wine-glossary terms as “analytical,” and given that some of the Quandt descriptors are defined in wine glossaries, Ramirez (2010) implicitly assumes some of the Quandt descriptors are analytical. A similar issue exists with Klimmek’s (2013) approach to measuring the “information content” of wine descriptions. In his approach, descriptors compiled by an association that certifies people in wine tasting count towards a description’s information content. Yet some of those same descriptors are some of the descriptors Quandt deemed to be bullshit.

## B. More on the Results and Discussion for the Hedonic Regression

### 1. More on results reported in the main text

A few remarks on the hedonic results reported in the main text are as follows.

● Given the log-linear form of my hedonic regression, I use Kennedy’s (1981) and Van Garderen, Jan, and Shah’s (2002) approach to estimate the effect of an indicator variable, as mentioned in the main text. The main text also mentions that I accounted for the fact that some descriptors such as “silky tannins” have other descriptors such as “silky” nested in them. Accounting for that is necessary because any instance of the “silky tannins” descriptor in a wine’s description would also always count as an instance of the “silky” descriptor, too. To account for that, I used the sum of the relevant coefficients and the standard error of that sum when applying Kennedy’s (1981) and Van Garderen, Jan, and Shah’s (2002) approach. I used the same approach when comparing the effect of one descriptor relative to another, except I use the difference between (rather than the sum of) the relevant coefficients.

● For any log-linear regression of the form *log{y} = b x + u*, if *dx* denotes a discrete change in *x*, then the percentage effect on *y* of a discrete change in *x* (relative to the level of *y* before the change in *x*) is *exp{b dx} - 1*. That percentage effect could be estimated by simply substituting a point estimate for *b* into the expression given above, but an adjustment for small-sample bias was popularized by Kennedy (1981). He suggested that the percentage effect be estimated as *exp{β dx – 0.5 (σ dx)2} – 1*,where *β* and σ denote a point estimate of *b* and its standard error. Kennedy’s suggestion was based on the idea that, if the point estimate is normally distributed with mean *β* and standard error *σ*, then an exponential transformation of it will be log-normally distributed with mean *exp{β + 0.5 σ2}*. Van Garderen, Jan, and Shah (2002) suggest that the variance of Kennedy’s (1981) point estimate of the percentage effect can be estimated as *exp{2 β dx} (exp{–(σ dx)2} – exp{–2 (σ dx)2}).*

### 2. Some other results from the same hedonic regression

A few remarks on some other results related to my hedonic regression that were *not* reported in the main text are as follows.

● As discussed in the main text, my hedonic results imply that the difference with respect to a wine’s price of having “velvety tannins” rather than “silky tannins” appear in the wine’s description is only about 1.4% (se=1.4%) and not statistically different from zero (p-value=0.32). As also already discussed in the text, that finding would be supported by the fact that velvety and silky are synonyms, according to Parker’s wine glossary. The glossary says that “*velvety*” is “a textural description and synonym for *lush* and *silky*” (*Wine Advocate*, 2020; emphasis added). I estimate that “lush tannins”—which is another of the Quandt descriptors for at least 50 wines in my dataset—has an effect that is not statistically different from zero, the effect of velvety tannins, or the effect of silky tannins (with p-values of 0.60, 0.38, and 0.52, respectively). Velvety, silky, and lush may all be synonymous, therefore.

● Ramirez (2010) finds that wines with longer descriptions (as measured by the log of their character count, which is the same measure I use) tend to have higher prices. Consistent with that, my hedonic regression suggests that a one-standard deviation increase in the length of a wine’s description (i.e., a 0.65 increase in the log of the character count) increases the wine’s price by about 2% (se=0.4%, p-value<0.01).

### 3. The (non-) effect of adding estimated topic probabilities as controls

● As discussed above, McCannon (2020) uses estimated topic probabilities from a topic model as explanatory variables in an otherwise standard hedonic regression. I explored adding estimated topic probabilities to my hedonic regression as control variables. My approach and results were as follows.

When preparing the text of the wine descriptions in my dataset for topic modeling, I used the same “light touch” approach that McCannon (2020) used. The only exception is that I removed the 106 Quandt descriptors that appear in the descriptions of at least 50 wines in my dataset. If I did not do that, then the estimated topic probabilities that were to be entered into the hedonic regression could perhaps end up being conflated with my indicator variables for those descriptors.

I also used the same topic model he used (namely, a Latent Dirichlet Allocation topic model). For the number of topics to presume, McCannon (2020) assumed as many as 10 (as noted above), so I assume that many, too. For the other parameters of the topic model (i.e., the α and β hyperparameters in the notation of his paper; McCannon, 2020, p. 76), I used the same parameterization he used.

Fitting the model even once is computationally intensive, but once fit, the estimated topic probabilities can be used as additional control variables in my hedonic regression.

Adding those topic probabilities added almost nothing in terms of explanatory power. The unadjusted R-squared of the hedonic regression increased from about 0.67 to 0.68.

The estimated effects of the Quandt descriptors were also largely unchanged. Thirty eight (instead of 43) of the 106 descriptors were statistically significant at the 10% level. *Appendix Table 5* shows the descriptors with the most significant price effects. With a few exceptions, the most significant descriptors are still the same ones shown in Table 1 of the main text. “Smoked game” and “precocious” are still estimated to have some of the most significant positive effects. “Unctuous” and “brawny” still have some of the most significant negative effects. And although it is not shown in that table, “velvety tannins” is still estimated to have an effect that is statistically significantly different from zero, but small, and not statistically significantly different from the effect of “silky tannins.”

Thus, overall, the hedonic results presented in the main text are largely unchanged if topic probabilities are added to the hedonic regression as additional controls. Whether that non-effect is because the wine descriptions are not associated with price or simply because a small number of topic probabilities do not capture the richness of the wine descriptions is unclear without an approach that more fully captures the wine descriptions. Mozer et al.’s (2020) comparison of different text-matching approaches suggests that matching based on topic probabilities from a topic model is inferior to the alternative approach they advocate and I largely adopt. See the main text of my paper, as well as below in this supplementary appendix, for more on my matching approach.

### 4. Some quantile regression results

● As mentioned above, quantile regressions can be run to explore whether an explanatory variable’s effect (or, at least, association) varies with price. The main question of interest for my study is whether the effects that apply at the (conditional) mean of log prices, according to the OLS results reported in the main text, apply similarly at lower price levels, too. Recall from the main text that the mean price of the wines in my dataset is about $65 while the average price that my survey respondents said they typically paid for a bottle of wine is only about $14.

For the dataset used for my hedonic regression, the (unconditional) distribution of log prices reaches around $14 at its 5% quantile, so I ran a quantile regression for that 5% quantile. At that quantile, the estimated effect of having velvety rather than silky tannins are similar to the OLS results reported in the main text. Specifically, the effect on price is estimated to be about 5% (se=2%, p-value=0.02), which unlike before is statistically significant at conventional levels, but it is only slightly larger than the 1.4% effect suggested by the OLS results. Thus, at least for that main result of interest, quantile regression results at a relevant low price are similar to the OLS regression results reported in the main text.

# III. More on the Matching Estimates

## A. More on the Data and Methods for the Matching Estimates

● Quandt (2007) mentions that the 24 descriptions from which he identified bullshit descriptors were for “Rhône wines, Sauternes, red Bordeaux wines, Burgundies, American reds, and Spanish wines” (p. 131). My dataset and thereby my hedonic regression covered those types of wines and more.

It could perhaps be argued that it is impossible for some of the Quandt descriptors to be used to describe some types of wine.

That argument is not entirely convincing. The Quandt descriptors seem ambiguous or nonsensical enough that it seems possible that they could be used to describe any wine. Furthermore, if a descriptor fails to appear in any of the descriptions for a certain type of wine in my dataset, then that may be due to the infrequency rather than the impossibility of its use to describe that type of wine.

Yet if there is concern that a given descriptor could not appear in the descriptions of some wines—or other concerns about the comparability of wines for which a given descriptor appears and some of the wines for which a descriptor does *not* appear—then a matching estimator would be attractive.

● As stated in the main text, I largely adopt the text-matching approach that Mozer et al. (2020) recommend. They recommend representing the rich texts associated with treated and control subjects in terms of bounded term-document matrices and then performing one-to-one matching on the cosine similarities of those representations.

To clarify: A *term-document matrix* (also called a document-term matrix) is a matrix where unique documents (or, for me, wine descriptions) are indexed along one dimension of the matrix, unique terms that appear in those documents are indexed along the other dimension, and each cell in the matrix is a measure of whether or how frequently a given term appears in a given document.

Terms that a researcher might keep track of include ngrams or skipgrams. For a text that reads “new saddle leather,” the unigrams would be “new,” “saddle,” and “leather.” The bigrams would be “new saddle” and “saddle leather.” The one-skip-one-gram would be “new leather.” Note that an ngram longer than a unigram or a skipgram can capture some aspects of the order of words in a text, but beyond that, a term-document matrix ignores the order of terms in a document.

Measures of frequency that a researcher might use include how frequently a given term appears in a given document (“term frequency”) or how frequently a given term appears in a given document relative to how frequently it appears across all documents (“term frequency-inverse document frequency”). Mozer et al. (2020) consider both.

A *bounded* term-document matrix is one for which terms that occur most frequently or infrequently across the documents are ignored. A researcher determines the bounds. Boundings that Mozer et al. (2020) consider in their study include ignoring terms that occur in less than a handful of documents or more than 1,000 of the roughly 3,000 documents in their dataset.

Cosine similarity is a distance metric that can be calculated for any two matrices, including any (bounded or unbounded) term-document matrices. Mozer et al. (2020) consider other distance metrics such as Euclidean and Mahalanobis, but recommend using cosine similarity.

For the one-to-one matching, Mozer et al. (2020) use so-called “optimal” matching as opposed to nearest neighbor matching.

One-to-one matching based on the cosine similarities of term-document matrices is a relatively simple approach compared to some of the other approaches considered by Mozer et al. (2020). Yet at least for the application they explore, they find that it generates better matches between treated and control subjects than other possible approaches such as topic modeling. They suggest that the strength of their recommended approach lies in the high-dimensional nature of term-document matrices; lower dimensional representations will have difficulty capturing the richness of rich texts.

● Closely but not completely following the approach that Mozer et al. (2020) recommend, my text-matching approach was as follows. For each treated wine that had a given descriptor appear in its description and for each possible control wine that did *not* have the given descriptor appear, I took the text of the wine’s description as aggregated by K&L. To that, I added the name of the wine (which generally included a vintage year and the producer’s name), its varietal as classified by K&L, and its geographic origin as classified as K&L.

I followed McCannon (2020) by using a “light touch” with those descriptions; I did not remove commonly occurring words, stem or lemmatize the words, or otherwise simplify the text.

Next, I created bag-of-words representations of the descriptions (as stated in the main text). To be more specific: I created unbounded term-document matrices for each of the treated and possible control wines. When doing that, I only considered unigrams and whether they appeared or not. I did not consider anything besides unigrams, and I did not consider the frequency with which they appeared within or across descriptions. I did not ignore any of the most frequently or infrequently occurring terms; the matrices were unbounded in that sense. Using unbounded matrices involves a “lighter touch” than trying to specify bounds for terms that appear too frequently across descriptions or not frequently enough.

I then matched each treated wine to a control wine based on the cosine similarity of the bag-of-words representations of their text. To do that, I used one-to-one nearest neighbor matching without replacement.

After matching, the effect of the given descriptor on a wine’s price could then be estimated by using Eq. 2 in the main text.

The extent to which matched treatment and control wines are similar in terms of their observed covariates can and should be examined, too (as mentioned in the main text of my paper).

● When applying one-to-one matching, I used nearest neighbor matching without replacement. That is to say, treated subjects were considered in a specific order (see below on that), a given treated subject was matched to the control subject that was the closest (in terms of the cosine similarity of bag-of-words representations of their text), and that matched control subject was no longer available to be matched to other treated subjects. I used that approach rather than the so-called “optimal” matching used by Mozer et al. (2020). Optimal matching involves searching for a set of matches between treated and control subjects that minimizes the average distance between all matched subjects. I did not use optimal matching partly because it is more computationally intensive than nearest neighbor matching. Previous research on propensity score matching by Austin (2014, sec. 3) also suggests optimal matching yields at best marginally better matches than one-to-one nearest neighbor matching without replacement.

● When applying nearest neighbor matching, treated subjects are considered in a specific order. One approach could have been to consider them in a random order. Austin’s (2014) Monte Carlo simulation study of different propensity score matching approaches suggests one-to-one random nearest neighbor matching without replacement is nearly equivalent to optimal matching. I however used an approach that is deterministic rather than random (except for any randomly broken ties if more than one control subject happens to be equally close to a treated subject) by matching wines starting with the closest matches first. According to Austin (2014), such a “closest distance” approach (albeit in terms of propensity scores for his study) is nearly equivalent to optimal matching, too.

## B. More on the Results and Discussion for the Matching Estimates

● For the “smoked game” descriptor as the treatment and the lack of that descriptor as the control, the mean of the cosine similarities between the matched wines was 0.49 (se=0.08), as stated in the main text. To try to illustrate what that cosine similarity (again, in terms of my bag-of-words representation the text of their wine descriptions) corresponds to, the following two wines were the matched wines with a cosine similarity closest to 0.49 and indeed almost exactly equal to 0.49.

The treated wine was a 2001 California Zinfandel that Parker awarded 91 points and described in part as having “notes of *plum*, figs, *smoked game*, sweet and *sour cherries*, *black currants*, and *earth*,” where I have emphasized the descriptors that Quandt deemed to be bullshit. The wine was classified by K&L as being from the “Alexander Valley/Russian River” appellation. The matched control wine was a 2004 Napa Valley Cabernet Sauvignon that Parker awarded 90 points and described in part as having “notes of sweet *black cherry*, currant, licorice, *cedar*, and *earth*” (where I have again emphasized the Quandt descriptors). The former wine was listed at $69.99 on K&L’s site while the latter was listed at $59.99. Both were among the out-of-stock wines.

The extent to which those two wines are similar or dissimilar is debatable, of course, but again they have a cosine similarity of 0.49, which is essentially the same level as the mean of the pairwise cosine similarities between the matched treated and control wines.

● As mentioned above and in the main text, assessing the quality of matches is important. The match quality for the “smoked game” treatment is discussed in the main text. As also mentioned in there, the match quality was roughly similar for the other treatments. For example: When the appearance of the “brawny” descriptor was taken as the treatment and the absence of that descriptor was taken as the control, the highest cosine similarity for any match was 0.75. That match was for two different wines (named “Numanthia” and “Termanthia”) by the same producer (Bodegas Numanthia of the Toro region of Spain) with the same vintage year (2007). The mean of the cosine similarities between the matched wines was 0.46 (se=0.08). The vast majority of the matched wines (ranging from 80% to 100%) were the same in terms of their varietal, geographic origin, and whether a given reviewer reviewed the wine or not. A treated wine’s description length was about 88% of its matched control wine’s description length on average (se=28%). The matches were again less similar in terms of their vintage (with only 42% of the matched wines sharing the same vintage) and numerical ratings by given reviewers. Those statistics are all roughly similar to the ones reported in the main text for the “smoked game” descriptor as the treatment.

# 

# IV. More on the Stated Preference Survey

## A. More on the Questions and Sample for the State Preference Survey

### 1. More on the survey design

Here I will reproduce the full text of my survey while providing some additional remarks on its design. The online survey opened with a page that contained a standard consent form. The page looked essentially as follows, except I have redacted identifying information for the purposes of blind peer review.

Please complete this survey if you consent to do so. The consent form is given below.

CONSENT FORM

You are invited to participate in a study conducted by [PRINCIPAL INVESTIGATOR’S (P.I.’S) NAME AND INSTITUTIONAL AFFILIATION]. By conducting this study, [P.I.’S NAME] hopes to learn about consumers' willingness to pay for wines. You were selected to participate in this study by the survey company contracted to obtain participants. If you decide to participate, please complete the survey provided to you. There are no foreseeable risks to completing the survey. [P.I.’S NAME] cannot guarantee you will receive any benefits from completing the survey, although you will hopefully find the experience interesting and informative. Any benefits offered by the survey company contracted to obtain participants are determined by and the responsibility of that company. Any information obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. Your decision whether or not to participate will not prejudice your future relations with [P.I.’S AFFILIATED INSTITUTION]. If you decide to participate, you are free to withdraw your consent and to discontinue participation at any time. The Committee on the Protection of Human Subjects at [P.I.’S AFFILIATED INSTITUTION] has reviewed and approved the present research. If you have questions now or in the future, [P.I.’S NAME AND CONTACT INFORMATION] will answer them. Questions regarding the rights of research subjects may be directed to [CONTACT INFORMATION FOR INSTITUTIONAL REVIEW BOARD AT P.I.’S AFFILIATED INSTITUTION]. If you would like a copy of this form to keep, please make a copy for yourself at this time. You are making a decision whether or not to participate. Completing this survey indicates you have decided to participate, having read the information provided above.

After clicking a “Next” button, respondents moved onto the next page of the survey. For each page, respondents were not able to move to the next page until answering all the questions on the current page Respondents were also not able to back up to a previous page once they moved forward.

The next page of the survey had the first question, which was about a respondent's subjective wine knowledge. The exact wording of the question was as follows.

[Question #1.] How would you rate your knowledge of wine?

◯ Far above average

◯ Somewhat above average

◯ Average

◯ Somewhat below average

◯ Far below average

I then presented them with a standard “cheap talk” script and an associated attention-check question. The script and question were as follows.

As part of this survey, you are going to be asked about your willingness to pay for wine in a hypothetical scenario. *Please treat the hypothetical scenario as if it was real.* Previous experience suggests people often say they would do one thing in a hypothetical situation but they would actually do something else if faced with the same situation in the real world. In particular, the amount that survey respondents say they would pay for a wine is often higher than the amount they would actually be willing to pay. So please make your choices as if you were making them in the real world without overstating how much you’d be willing to pay.

[Question #2.] According to the information given above, the amount that survey respondents say they would pay for a wine is often higher, lower, or exactly equal to the amount they would actually be willing to pay?

◯ higher

◯ lower

◯ exactly equal to

Note that the “cheap talk” script is about how, when faced with a hypothetical situation, people often say they would pay an amount higher than what they would actually be willing to pay. Such scripts have been shown to be somewhat effective at mitigating hypothetical bias (Penn and Hu, 2019).

For a cheap talk script to have any hope of mitigating hypothetical bias, respondents obviously need to at least read the script. To try to verify they read it, I followed Penn and Hu (2016) by including the above-given attention-check question.

If a respondent answered that attention-check question incorrectly, then they were temporarily diverted to a page that gave them an opportunity to reread the script and reanswer the question. The page looked essentially as follows.

Your answer to the previous question was incorrect. Please try again after more carefully reading the information given below.

As part of this survey, you are going to be asked about your willingness to pay for wine in a hypothetical scenario. *Please treat the hypothetical scenario as if it was real.* Previous experience suggests people often say they would do one thing in a hypothetical situation but they would actually do something else if faced with the same situation in the real world. In particular, the amount that survey respondents say they would pay for a wine is often higher than the amount they would actually be willing to pay. So please make your choices as if you were making them in the real world without overstating how much you'd be willing to pay.

[Question #2, repeated.] According to the information given above, the amount that survey respondents say they would pay for a wine is often higher, lower, or exactly equal to the amount they would actually be willing to pay?

◯ higher

◯ lower

◯ exactly equal to

Malone and Lusk (2019b) advocate for such a “trap and release” approach in which respondents who answer an attention-check question incorrectly are given an opportunity to reanswer. For all of my analysis, I dropped a respondent if they did not answer the question correctly on at least their second try.

Respondents who answered the attention-check question correctly on their first try, as well as the other respondents rejoining them along the main path of the surveyed, were then asked the following question about the price they typically pay for a bottle of wine.

[Question #3.] Imagine you’re shopping for wine. When doing so, about how much would you typically pay for a bottle of wine?

◯ $5

◯ $10

◯ $20

◯ $30

◯ $40

◯ $50

◯ Over $50 (please specify): \_\_\_

There are potential issues with that question. Most notably, it is completely vague about where the person might be shopping, why they might be shopping for wine, and whether they are buying a standard-sized bottle or not, all of which presumably matter to what a person would be willing to pay. Yet exploring the effects (if any) of such things would be a research project unto itself.

The questions that asked respondents to choose between a wine with silky or velvety tannins and to state their MWTP for their preferred wine were then asked.

[Question #4.] Suppose that, while shopping for wine, you find two wines in your typical price range. The wines are almost identical, except one is described as having *silky* tannins and the other is described as having *velvety* tannins. Which wine would you buy?

◯ The wine with silky tannins

◯ The wine with velvety tannins

◯ Either of those wines

◯ Neither of those wines

I chose to list the “velvety tannins” descriptor after the “silky tannins” descriptor based on their alphabetical order. There was no randomization to the wording of the stem or the order of the answer choices.

Note that if a respondent says they would buy “Neither of those wines,” they should *not* be saying that because the wines are beyond their budgets. The wines are in their typical price range, respondents were asked to assume. A respondent should say they would buy neither wine only if they would indeed not buy either wine at the price they typically pay for a wine. So, for example, if a respondent typically pays a relatively high price for a wine, they should be unwilling to buy a wine in their high price range if it sounds like it should be low priced. (If a respondent typically pays a relatively low price for wine, then they should be more than happy to buy a wine in their low price range if it sounds like it should be high priced.)

Asking respondents to assume the wines were in their “typical price range” can be defended as follows. Any relative prices assigned to the wines could affect respondents’ perceptions of and preferences over the wines. Expensive might be taken to mean better. Malone and Lusk (2018, pp. 3, 11) express that concern even when consumers are being surveyed about brand-name beers. That concern seems especially worrisome when there is less non-price information; for my survey, the only other information is descriptors that may be ambiguous or nonsensical. Examining how perceptions and preferences are affected by price is an important research area (Ashton, 2014; Goldstein, 2019; Mastrobuoni et al., 2014), but here I am trying to examine the effect of the descriptors.

Admittedly, respondents’ perceptions and preferences could still be affected by asking them to assume both wines are in their “typical price range.” That supposition suggests the wines are similarly priced and, as such, any anchoring effect may tend to draw respondents’ answers towards a smaller MWTP. Yet it seems less problematic to suggest to respondents that the wines are similarly priced than to suggest that one wine is better, worse, or identical to another by assigning a higher, lower, or equal price.

Another reason to ask respondents to assume the wines were in their “typical price rage” is that it helps maintain as large and as credible a sample as possible. If I were to assign dollar values, then the higher the prices of the wines, the more likely it would be that more respondents would say they would buy *neither* wine because it would be beyond their budgets. Furthermore, the higher the prices of the wines, the more likely it might be that respondents who say they would buy a wine are *overstating* their total willingness to pay. Inferences about MWTP would therefore likely be restricted to a smaller and perhaps more questionable sample.

If a respondent selected either “Either of those wines” or “Neither of those wines” for the above-given question, then a forced-choice follow-up question was asked, as discussed in the main text of my paper. (Again, respondents who selected “Neither of those wines” were asked that forced-choice question so that they followed the same path as other respondents, but they were eventually dropped from the analysis.) The exact wording of the question was as follows:

[Forced-choice version of Question #4.] If you had to choose: Would you buy a wine described as having silky tannins or an almost identical wine described as having velvety tannins, if you were shopping for wine and if both wines were in your typical price range?

◯ The wine with silky tannins

◯ The wine with velvety tannins

By answering the initial or forced-choice question, each respondent selected either the wine described as having silky tannins or the wine described as having velvety tannins. Respondents who selected the wine with *silky* tannins were asked the following question.

[Version of Question #5 for respondents who selected the wine with silky tannins.] You said you'd buy the wine with *silky* tannins. What's the most you'd be willing to pay for that wine, compared to the almost identical wine with velvety tannins?

◯ At most zero dollars more

◯ At most one dollar more

◯ At most 5 dollars more

◯ At most 10 dollars more

◯ Over 10 dollars more (please specify): \_\_\_

Respondents who selected the wine with *velvety* tannins were asked the following analogous question.

[Version of Question #5 for respondents who selected the wine with velvety tannins.] You said you'd buy the wine with *velvety* tannins. What's the most you'd be willing to pay for that wine, compared to the almost identical wine with silky tannins?

◯ At most zero dollars more

◯ At most one dollar more

◯ At most 5 dollars more

◯ At most 10 dollars more

◯ Over 10 dollars more (please specify): \_\_\_

Range or centering effects are a concern when presenting interval choices, but examining any such effects would be a research project unto itself, so I did not attempt to examine them.

The above-given questions were all the questions related to a respondent’s relative preference for a wine described as having “silky tannins” or “velvety tannins.”

Next, a similar series of questions was asked to assess a respondent’s relative preference for a wine described as being “precocious” or “unctuous.” A discussion of these questions was omitted from the main text of my paper because that discussion would have been extended and would not have altered the overall conclusions of my study, but for completeness and transparency, I discuss them here in this appendix. I started by asking a question analogous to Question #4:

[Question #6.] Now suppose that, while shopping for wine, you find two wines in your typical price range. The wines are almost identical, except one is described as being *unctuous* and the other is described as being *precocious*. Which wine would you buy?

◯ The unctuous wine

◯ The precocious wine

◯ Either of those wines

◯ Neither of those wines

Aside from the silky and velvety tannin descriptors highlighted by Quandt in his article, there are many pairs of the Quandt descriptors I could have asked about. (Indeed, with 123 descriptors in his list, there are about 7,500 possible pairs.) I asked respondents to compare “precocious” and “unctuous” because, firstly, that was one of a few pairs with a significant price difference according to the hedonic estimates. It is of interest whether my survey yields a MWTP estimate near the large hedonic one. Secondly, “precocious” and “unctuous” are both defined in wine glossaries like Parker’s (unlike some other possible pairs). Whether such wine-glossary terms are disregarded by consumers is of interest. Third and finally, “precocious” and “unctuous” share the same number of syllables (unlike “angular” and “brawny,” which would be another possible pair).

When asking the above-given question, I chose to list the “precocious” descriptor after the “unctuous” descriptor so that, in both this question and the earlier one, the descriptor that was supposed to have a premium according to the hedonic estimates was listed second. To the extent there is any response-order bias, the same bias should apply to this question and the earlier one.

Just like before, a forced-choice follow-up question was asked of respondents who selected “Either of those wines” or “Neither of those wines.”

[Forced-choice version of Question #6.] If you had to choose: Would you buy a wine described as being *unctuous* or an almost identical wine described as being precocious, if you were shopping for wine and if both wines were in your typical price range?

◯ The unctuous wine

◯ The precocious wine

Then based on their initial or forced-choice answer, respondents were asked about their MWTP for their preferred wine. Respondents who selected the *unctuous* wine were asked the following.

[Version of Question #7 for respondents who selected the unctuous wine.] You said you'd buy the *unctuous* wine. What's the most you'd be willing to pay for that wine, compared to the almost-identical precocious wine?

◯ At most zero dollars more

◯ At most one dollar more

◯ At most 5 dollars more

◯ At most 10 dollars more

◯ Over 10 dollars more (please specify): \_\_\_

An analogous question was asked of respondents who selected the *precocious* wine.

The last page of the survey had the last three questions. The first was one of two objective wine knowledge questions taken from Aqueveque’s (2018) test.

[Question #8.] Based on your knowledge of wine, which of the following aspects of a wine is most affected by its tannins?

◯ sweetness

◯ astringency

◯ acidity

◯ Don't know

The order of each of those possible answers, except the last “Don’t know” answer, was randomized for each respondent. The possible answers are taken verbatim from the test; the correct answer is “astringency,” according to Aqueveque (2018, p. 182). I did slightly rewrite the wording of the question’s stem. According to Aqueveque (2018, p. 182), the stem was a complete-the-sentence stem worded as “Tannins are responsible of [*sic*] wine’s...”, although that might be an English translation of a Spanish version of the test; Aqueveque administered his survey to Chilean university students. I rewrote that stem in the manner given above.

The next question was another objective wine knowledge question taken from the same test.

[Question #9.] Based on your knowledge of wine, which of the following groups of descriptors is most commonly associated with Chardonnay wine?

◯ pineapple, melon, peach, lemon

◯ tomato, peppers, cherry

◯ strawberry, vanilla, lime, cinnamon

◯ Don’t know

The order of each of those possible answers, except the last “Don’t know” answer, was again randomized for each respondent. As with the other question, I rewrote the stem slightly. According to Aqueveque (2018, p. 182), the stem was originally worded as “Which of the following groups of characteristics is commonly associated to [*sic*] Chardonnay wine?” The possible answers are taken almost verbatim from Aqueveque (2018, p. 182), except I dropped one descriptor to try to make the correct answer less ambiguous; “apple” appeared in both a correct and incorrect answer. “Pineapple, melon, peach, lemon” (and apple) is the correct answer, according to Aqueveque (2018, p. 182).

I asked those two questions because I did not want to ask too many wine-knowledge questions and those two seemed like the most relevant ones from Aqueveque’s (2018) seven-question test. The other five questions from that test were about which grape varietals are red, “full bodied,” or Chilean, and what it means to aerate or rack wine. Other tests of wine knowledge include Frøst and Noble’s (2002, p. 278) 11-question “wine trivia quiz,” which has questions about alcohol content, varietals, and appellations.

The last question in my survey was another attention-check question.

[Question #10.] Please select “Portugal” to demonstrate you are paying attention.

◯ France

◯ Italy

◯ Spain

◯ Germany

◯ Portugal

Those possible answers are all non-US, wine-producing countries. They were presented in a randomized order for each respondent. Note that, if a respondent was randomly answering questions, they should have only a 20% chance of answering that question correctly. In my analysis reported in the main text, I dropped a respondent if they did not select “Portugal.”

Those were all of my questions, but respondents’ basic demographic information is also provided by SurveyMonkey when responses are purchased through its panel. A respondent’s age is reported in one of five ranges (specifically, under 18, 18–29, 30–44, 44–60, or over 60 years of age). Gender is reported as “Male” or “Female.” Annual household income is reported in 10 ranges (the lowest being $0–$9,999 and the highest being $200,000 or more). A respondent’s geographic region is reported as one of nine regions (ranging from “New England” to “Pacific”). Some of that basic demographic information was not provided for some respondents; those respondents could apparently decline to provide such information and still complete a survey. My main interest is in whether experts and novices have any differences in MWTP, but the MWTP of other subgroups could be explored, too, especially with larger sample sizes.

Some final comments on the design of my survey are as follows.

● By answering my survey’s key questions, a respondent can directly say whether they prefer one wine over the other in a pair and, if so, the most they would be willing to pay for their preferred wine relative to the other one. A respondent can also directly say whether they are indifferent and would pay no more for one over the other. Whether a respondent is indifferent is of particular interest. Again, if someone disregards descriptors like silky or velvety tannins as complete bullshit, then they should not favor either wine. A respondent can also directly say whether they disfavor both wines in a pair by saying they would buy neither.

● A similar alternative to my stated-preference survey would be a discrete choice experiment (DCE) survey. For a DCE survey, respondents would be asked to choose between two or more wines with different characteristics and prices. The characteristics and prices would be varied by the researcher so MWTP for a characteristic could be estimated from respondents’ choices. That approach is popular for studying MWTP for characteristics of wine and other goods. It is a possible approach for future research on MWTP for wine bullshit.

I favor my approach over a DCE survey, at least as a first attempt to estimate MWTP for some of the Quandt descriptors, because: (1.) I do not need to specify the relative prices of the wines and thereby largely avoid the issue discussed above of prices potentially affecting perceptions of and preferences over wines; (2.) I do not need to specify the absolute prices of the wines and thereby maintain as large and credible a sample as possible for the reasons discussed above; (3.) I do not need to ask individual respondents to consider a large number of potentially maddening choice situations with the same descriptors and slightly different prices, which reduces the number of questions in the survey and thereby both the fatigue to the subject and the cost to the researcher; (4.) although a DCE could be used to ask about more than two wines at once, asking respondents to consider more than two might lead to the “choice overload” discussed by Malone and Lusk (2019a); (5.) I can obtain a larger amount of information from a smaller number of responses because respondents who are indifferent can directly say that and those with slight or strong preferences can directly express them; and (6.) individual responses can be directly examined in their own right and not simply as a partial contribution to an estimate of MWTP for a characteristic across respondents.

Note that hypothetical and inattention biases would still need to be addressed, regardless of whether the survey was a stated-preference or DCE survey.

● As already suggested above, it would be of interest to explore how sensitive the results of my survey are to the design of it. In particular, it could be interesting to explore whether price information affects perceptions and preferences. Again, however, doing that would be a research project unto itself. Moreover, if perceptions and preferences are sensitive to price information, then that has implications that extend far beyond MWTP for potentially bullshit descriptors. If it can be shown that consumers are willing to pay more for a wine with a descriptor that is complete gibberish when a researcher assigns as a high price to a wine with the descriptor, then that would be interesting, but it would seem to mostly be a finding about a bias in consumer decision-making rather than a finding about willingness to pay for gibberish. The fact that the descriptor is gibberish is what makes it clear there must be a bias.

### 2. More on the sample

The number of survey responses I purchased was 500 (as noted in the main text). Given the number of questions in the survey and the targeting of respondents who purchased wine in the last 30 days, the cost was $2 per completed survey. I purchased those responses with professional development funds awarded to me by my university for purposes like purchasing data.

## B. More on the Results and Discussion for the Stated Preference Survey

A few remarks on the survey results are as follows.

### 1. More on the respondents as wine consumers

● Recall I targeted people who self-reported purchasing wine in the last 30 days. In order to legally buy alcohol in the US, a person must be at least 21 years of age. Consistent with that, I did not have any respondents who reported being in the 18 or under age range.

### 2. More on the expertise of the respondents

● Recall I had 469 respondents who passed my attention checked and I classified 103 of them as “experts” (rather than “non-experts”) if they self-reported at least an average level of knowledge and also correctly answered the questions I asked from Aqueveque’s (2018) test. Although there are other ways to split the sample besides that, I did not require experts to self-report a strictly above-average knowledge partly because that would have cut the number of experts in half (down to 50) and also because Aqueveque (2018) argues experts may understate their expertise. I do not allow any wrong answers for the objective assessment of wine knowledge partly because allowing one wrong answer would have more than doubled the number of experts (up to 256) and also because I wished to ensure experts have expertise.

● In terms of the subjective and objective assessments of wine knowledge: A majority of the respondents (334 or about 71% of the 469 respondents) rated their knowledge as at least average. A majority of those respondents (231 or about 69% of them) got at least one of the test questions wrong.

● For the two questions I asked from his test, I can compare my sample’s performance on the objective assessment of wine knowledge to the performance of Aqueveque’s (2018) sample. He administered his survey to a convenience sample of 204 students recruited to “a study oriented to investigate wine consumption habits” (p. 182). The students were “all enrolled at a large Chilean university in different academic programs for executives” (ibid.). Their ages ranged from 26 to 55 years (ibid.). After dropping 11 respondents with a suspicious response pattern (specifically, they selected the middle option on a 5-item Likert scale for all four of four questions about their subjective wine knowledge), his final sample was made up of 193 respondents (ibid.). Aqueveque’s (2018) published paper does not report how those 193 respondents did on each of the objective wine knowledge questions, but in personal correspondence, he kindly provided that information (personal correspondence on May 4, 2020), which allows me to compare my sample’s performance to his sample’s.

For the question that assessed whether a respondent knew that tannins affect astringency, 37% (se=2%) of my sample answered correctly, which is about 10 percentage points more than the 26% (se=3%) of his sample who correctly answered. For the question that assessed whether a respondent knew common descriptors for Chardonnay, 55% (se=2%) of my sample answered correctly, which is about 15 percentage points more than the 39% (se=4%) of his sample who correctly answered. Both of those differences are statistically significant (based on a two-sided z-test of independent proportions; p-value<1%). The respondents to my sample therefore did somewhat better, although neither sample did particularly well. Of course, a test should be neither too easy nor too hard if the goal is to split test-takers based on their knowledge.

### 3. Additional survey results

The survey results related to the “precocious” and “unctuous” pair of descriptors were as follows.

● When respondents were asked whether they would prefer the precocious or unctuous wine, about 28% of the 469 respondents said they would prefer *neither* wine (se=2%). Similar percentages of experts and non-experts said the same; about 32% of experts (se=5%) and 27% of non-experts (se=2%) said they would prefer neither wine, which is not a statistically significant difference (p-value=0.29).

About 21% said they would prefer *either* wine (se=2%). When forced to choose, about 74% (se=4%) of those respondents chose the precocious rather than unctuous wine, which is more than half of them (p-value<0.01). The fact that respondents who initially expressed indifference between the two wines did not break evenly between them when forced to do so is curious. At least some of those respondents apparently did have a preference.

The remaining 51% of respondents initially said they preferred one wine over the other without being forced to choose. Among those respondents, about 72% (se=3%) said they preferred the precocious rather than unctuous wine, which is more than half (p-value<0.01). The fact that those who were forced to choose broke in a very similar way to those who freely chose suggests both groups have similar preferences, even if they gave different answers to the initial question.

● Similar to Table 3 of the main text, *Appendix Table 6* below summarizes respondents’ preferences for a precocious or unctuous wine, including the forced choices of respondents who initially said they preferred either wine and excluding respondents who said they preferred neither wine. As shown in that table, respondents tended to prefer the precocious rather than unctuous wine. About 72% of all respondents, 64% of experts, and 75% of non-experts preferred that (with standard errors of 2%, 6%, and 3%), where all of those percentages are statistically different from one half (with p-values of at most about 1%). The percentage of experts who preferred the precocious wine was somewhat lower than the percentage of non-experts (p-value=0.09), so experts were not quite as enamored with precociousness or adverse to unctuousness as the non-experts. Note that both experts and non-experts are leaning towards the descriptor that has a premium according to my hedonic estimates.

Although respondents were more decisive than before, their MWTP for their preferred wine relative to the other was still typically zero. As seen in *Appendix Table 6*, for all respondents, experts, and non-experts who preferred the precocious wine, the mode of their MWTP for their preferred wine relative to the other was $0. Similarly for those who preferred the unctuous wine.

● Like Figure 1 of the main text, *Appendix Figure 1* shows the distribution of MWTP for a precocious rather than unctuous wine among all the respondents, where I again follow Hanley et al. (2009) by assuming the *negative* of a respondent’s MWTP for an *unctuous* rather than precocious wine is their *negative MWTP* for *a precocious* rather than unctuous wine.

There is mostly indifference. A near-majority of all the respondents, the experts, and the non-experts reported a MWTP of exactly $0 (specifically, 45%, 43%, and 46% with standard errors of 3%, 6%, and 3%). And about 70% had a MWTP within $1 of $0 (specifically, 70%, 69%, and 71% with standard errors of 3%, 6%, and 3%).

Yet not all respondents had a near-zero MWTP. Four outliers lie beyond the range shown in *Appendix Figure 1*. Two respondents, including one expert and one non-expert, reported willingnesses to pay for an unctuous rather than precocious wine of $15 and $20. Two more respondents, both non-experts, reported willingnesses to pay for precocious rather than unctuous wine of $30 and $50.

Accounting for those and the other responses, the mean MWTP for the precocious rather than unctuous wine is $1.06 for all respondents, 66¢ for experts, and $1.16 for non-experts (with standard errors of 26, 46, and 30 cents). Those means are all statistically different from zero at conventional levels, except the mean MWTP among experts, which is not (p-value=0.16), although that mean MWTP among experts is also not statistically different from the mean MWTP among non-experts (p-value=0.36).

Such dollar values are fairly small, but they are larger than what respondents said they would pay on average for a wine with velvety rather than silky tannins, and those dollar values are not too trivial relative to the average price respondents said they typically pay for a bottle of wine. Recall respondents said they typically pay about $14 on average for a bottle of wine. A $1 change in a $14 bottle of wine is about a 7% price difference.

● To examine respondents’ MWTP as a percentage of the price they report typically paying for a bottle of wine, *Appendix Table 7* recalculates mean MWTP in those terms. Recalculating MWTP in those terms also allows a more direct comparison to my hedonic estimates.

As seen in the table, the mean MWTP in percentage terms for a precocious rather than unctuous wine is 7% (se=2%) for all respondents, which is significantly different from zero (p-value<0.01). That survey estimate is about half the size of the earlier hedonic estimate (which again was about 14%), but it is in the same direction as and not too dissimilar from the hedonic estimate. The mean MWTP among experts was similar to that of non-experts; the latter’s MWTP was slightly larger, but not statistically different from the former’s (p-value=0.48).

● Recall that the mean MWTP among both experts and non-experts for velvety rather than silky tannins was neither statistically nor substantively different from zero in either dollar terms or as a percentage of the price they typically pay for a bottle of wine. Their mean MWTP for a precocious rather than unctuous wine was statistically different from zero and, although arguably small even in percentage terms, it is bigger than the mean MWTP for velvety rather than silky tannins.

# V. More on the Concluding Remarks

As suggested at the conclusion of the main text, experimental methods may help disentangle whether consumers who are willing to pay for wines described by potentially bullshit descriptors are paying for real, objective aspects of a wine (detectable with mass spectrometers or more naked human senses) or imaginary, wholly subjective aspects.

In the main text, I specifically suggest a modified version of Goldstein’s (2019) “half-blind” tasting approach could be used to try to determine whether consumers are paying for a wholly subjective effect and I provided a brief example. I can expand on that suggestion as follows with another more extended example. As originally proposed, a half-blind tasting involves using two bottles of the same wine. Some information is unmasked for only one of the bottles (such as its price and label). The investigator remains silent as to whether the bottles are identical or not. After sampling from each bottle, the subject reports any preferences over the wines (such as which sample they prefer and how much they would be willing to pay for a bottle of it). An advantage of that approach is the investigator is not explicitly deceiving the subject; the investigator does not lie about the price charged for a wine, the label affixed to it, or anything else. The only somewhat misleading aspect is the investigator may be implicitly suggesting to the subject that the two samples are different wines (ibid., p. 324). Another advantage is that, if a subject has any difference of opinion about the two samples, then the only obvious explanation for that is the unmasked information.

A variation on that approach would be to use one wine that has been described by a wine critic with two descriptors of interest. It could be difficult to find such a wine for some descriptors, but even for descriptors like “precocious” and “unctuous” that occur relatively infrequently and even less often together, there are wines that have been described with both descriptors; the “1997 Bruno Giacosa Asili Barbaresco” listed at $159.99 on K&L’s site has been described by Parker as both precocious and unctuous, for example.[[3]](#footnote-3) It would be extremely expensive but not explicitly deceitful to present a subject with two samples from a bottle of the 1997 Bruno Giacosa Asili Barbaresco masked of its label and other information, say that Parker had described the wine in one sample as precocious, and say that Parker had described the wine in the other sample as unctuous. The only somewhat misleading aspect is the researcher may be implicitly suggesting to the subject that Parker said those things about different wines. If the subject has any difference of opinion about the two samples, then the only obvious explanation for that is the descriptors ascribed to the samples.

# References for Supplementary Appendix

Amerine, M.A. and Roessler, E.B. (1976). *Wines: Their Sensory Evaluation*. San Francisco, CA: W.H. Freeman and Company.

Aqueveque, C. (2018). Ignorant experts and erudite novices: Exploring the Dunning-Kruger effect in wine consumers. *Food Quality and Preference*, 65(April), 181–184.

Ashton, R. (2014). Wine as an experience good: Price versus enjoyment in blind tastings of expensive and inexpensive wines. *Journal of Wine Economics*, 9(2), 171–182.

Austin, P. (2014). A comparison of 12 algorithms for matching on the propensity score. *Statistics in Medicine*, 33(6), 1057–1069.

Bosker, B. (2017). *Cork Dork.* New York, NY: Penguin Random House.

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

Chen, K.J. and McCluskey, J. (2018). Impacts of expert information on prices for an experience good across product segments: Tasting notes and wine prices. *Journal of Agricultural and Resource Economics*, 43(3), 388–402.

Croijmans, I., Hendrickx, I., Lefever, E., Majid, A., and Van Den Bosch, A. (2019). Uncovering the language of wine experts. *Natural Language Engineering*, 1–20.

Drichoutis, A.C., Klonaris, S., and Papoutsi, G.D. (2017). Do good things come in small packages? Bottle size effects on willingness to pay for pomegranate wine and grape wine. *Journal of Wine Economics*, 12(1), 84–104.

Frøst, M.B., and Noble, A.C. (2002). Preliminary study of the effect of knowledge and sensory expertise on liking for red wines. *American Journal of Enology and Viticulture*, 53(4), 275–284.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3), 535–574.

Goldstein, R.S. (2019). Half-blind tasting: A deception-free method for sizing placebo and nocebo responses to price and packaging attributes. *Journal of Wine Economics*, 14(3), 321–331.

Gold, Z., and Latonero, M. (2018). Robots welcome? Ethical and legal considerations for web crawling and scraping. *Washington Journal of Law, Technology, and the Arts*, 13(3), 275–312.

James, V. (2020). *Wine Girl: The Obstacles, Humiliations, and Triumphs of America's Youngest Sommelier.* Harper Collins: New York, NY.

Klem, B. (2009). *WineSpeak: A vinous thesaurus of (gasp!) 36,975 bizarre, erotic, funny, outrageous, poetic, silly, and ugly wine tasting descriptions. Who knew?* Stamford, CT: WineSpeak Press.

Klimmek, M. (2013). On the information content of wine notes: Some new algorithms? *Journal of Wine Economics,* 8(3), 318–334.

Koenker, R., and Hallock, K. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4), 143–156.

Krumme, C. (2009). A nose by any other name: Descriptors as signals for wine price. Paper presented at the 3rd Annual Conference of the American Association of Wine Economists in Reims, France, June 18–20. Available at <<https://pdfs.semanticscholar.org/9299/b521e4720b92f9028b9b7d03da8a1a54256b.pdf>> (Accessed 1 June 2018).

Krumme, C. (2011). Velvety chocolate with a silky ruby finish, pair with shellfish. *Slate*, February 23. Available at <<http://www.slate.com/articles/life/drink/2011/02/velvety_chocolate_with_a_silky_ruby_finish_pair_with_shellfish.html>> (Accessed 23 February 2011).

Krumme, C. (2013). Volatile elements: the dynamics of wine descriptors from 2001–2011. Paper presented at the 7th Annual Conference of the American Association of Wine Economists in Stellenbosch, South Africa, June 26–29. Available at <<https://www.wine-economics.org/wp-content/uploads/2013/07/coco_krumme.pdf>> (Accessed 1 June 2018).

Malone, T, and Lusk, J.L. (2018). An instrumental variable approach to distinguishing perceptions from preferences for beer brands. *Managerial and Decision Economics*, 39(4), 403–417.

Malone, T. and Lusk, J.L. (2019a). Mitigating choice overload: An experiment in the US beer market. *Journal of Wine Economics*, 14(1), 48–70.

Malone, T., and Lusk, J.L. (2019b). Releasing the trap: A method to reduce inattention bias in survey data with application to US beer taxes. *Economic Inquiry*, 57(1), 584–599.

Mastrobuoni, G., Peracchi, F. and Tetenov, A. (2014). Price as a signal of product quality: Some experimental evidence. *Journal of Wine Economics*, 9(2), 135–152.

McCannon, B.C. (2020). Wine descriptions provide information: A text analysis. *Journal of Wine Economics*, 15(1), 71–94.

Mozer, R, Miratrix, L, Kaufman, A.R., and Anastasopoulos, L.J. (2020). Matching with text data: An experimental evaluation of methods for matching documents and of measuring match quality. *Political Analysis*, 28(2), 445–468.

Oczkowski, E., and Pawsey, N. (2019). Community and expert wine ratings and prices. *Economic Papers*, 38(1), 27–40.

Quandt, R. (2007). On wine bullshit: Some new software? *Journal of Wine Economics*, 2(2), 129–135.

Penn, J., and Hu, W. (2016). Making the most of cheap talk in an online survey. Paper prepared for Agricultural and Applied Economics Association Annual Meeting, Boston, MA, July 31–August 2. Available at <<https://ageconsearch.umn.edu/record/236171>>.

Penn, J., and Hu, W. (2019). Cheap talk efficacy under potential and actual hypothetical bias: A meta-analysis. *Journal of Environmental Economics and Management*, 96(July), 22–35.

Ramirez, C.D. (2010). Do tasting notes add value? Evidence from Napa wines. *Journal of Wine Economics*, 5(1), 143–163.

# TABLES AND FIGURES FOR SUPPLEMENTARY APPENDIX

*Appendix Table 1*

**Details on the control variables for the hedonic regression**

|  |  |
| --- | --- |
| *Variable(s)* | *Description* |
| Varietals | 37 indicators for the varietal of a wine as classified by K&L, including in order of prevalence: Pinot Noir (21%), Cabernet Sauvignon and Blends (20%), Chardonnay (12%), and 34 others that each cover less than 10% of the observations. About 1% have no varietal classification. |
| Countries | 22 indicators for the country-of-origin for the wine as classified by K&L, including: the United States (42%), France (32%), Italy (10%), and 19 others that each cover less than 5% of the observations. Less than 1% have no country classification. |
| Subregions | 51 indicators for the subregion of a wine as classified by K&L, including: California (37%), Burgundy (14%), no subregion classification (8%), Bordeaux (6%), Rhone (6%), and 46 others that each cover less than 5% of the observations. |
| Appellations | 57 indicators for specific appellations as classified by K&L, including: no appellation classification (35%), Napa Valley (15%), Sonoma County (6%), and 54 others that each cover less than 5% of the observations. |
| Vintages | 68 indicators for the vintage of a wine, including: 2010 (6.5%), 2015 (6.4%), 2009 (6.3%), 2005 (6.3%), 2012 (6.0%), and 63 others that each cover at most 6% of the observations. Less than 1% have no vintage. The year in a wine’s name is assumed to be its vintage. |
| Reviewers | 22 indicators for whether a reviewer reviewed a wine, including: Robert Parker’s *Wine Advocate* (60%), *Wine Spectator* (45%), Stephen Tanzer’s *International Wine Cellar* (29%), *Vinous* (25%), K&L Notes (22%), *Wine Enthusiast* (22%), Allen Meadows’s *Burghound* (13%), James Suckling (13%), and 14 others that each appear in less than 10% of the observations. To avoid trying to deaggregate the reviews aggregated by K&L, I assume a reviewer reviewed a wine if K&L provides a rating from that reviewer or their name appears in the description. About 25% of the wines have only one reviewer. The median number of reviewers per wine is two. The mean (standard deviation) is about 2.6 (1.5) reviewers. There are as many as 11 reviewers for a wine; three wines were reviewed that much. None of the indicator variables for the reviewers need to be dropped from my regression because they are not mutually exclusive. |
| Ratings by reviewers | 21 sets of indicators (one for each reviewer, except K&L itself, which does not give numerical ratings) for whether a given reviewer gave a wine a given rating. All ratings, as well as no rating by a reviewer who reviewed a wine, are treated as indicators to allow for non-linearity in numerical ratings (89 vs. 90 may be quite different from 90 vs. 91, e.g.), ranges (such as a rating of 90–92), and any information provided by the lack of a rating. The use of indicators should also minimize the influence of any mistakes in K&L’s site (such as the apparent mistake of saying that Parker gave one wine a rating of 32). |
| Description length | Log of the number of characters in a wine’s description, following Ramirez (2010). The mean and median of the log characters is about 6.86 and 6.94, respectively. |

*Note:* This table describes the control variables included in the hedonic regression, which are all indicator variables, except the continuous variables for the log description length. The percentage in parentheses are the frequencies with which an indicator variable occurs in the dataset. For the indicator variables, the indicator with the highest frequency was dropped to avoid the dummy variable trap, unless otherwise noted.

*Appendix Table 2*

**Pairwise correlations between select variables in my dataset**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Variable* | *Price* | *Descrip- tion  length* | *Smoked game* | *Precocious* | *Velvety tannins* | *Silky  tannins* | *Unctuous* | *Brawny* |
| Price | 1 |  |  |  |  |  |  |  |
| Description length | 0.23\* | 1 |  |  |  |  |  |  |
| Smoked game | 0.01 | 0.02\* | 1 |  |  |  |  |  |
| Precocious | 0.04\* | 0.03\* | −0.00 | 1 |  |  |  |  |
| Velvety tannins | 0.05\* | 0.06\* | 0.00 | 0.00 | 1 |  |  |  |
| Silky tannins | 0.03\* | 0.07\* | 0.01 | 0.00 | 0.01 | 1 |  |  |
| Unctuous | 0.05\* | 0.05\* | 0.01 | 0.00 | 0.00 | −0.01 | 1 |  |
| Brawny | 0.01 | 0.03\* | −0.00 | 0.00 | −0.00 | 0.01 | 0.01 | 1 |

*Note:* This table shows the Spearman correlation coefficients between select variables, including the log of a wine’s price, the log of the length of the wine’s description, and indicator variables for select Quandt descriptors. The Spearman correlation coefficient was used because the variables are a mix of continuous and dichotomous ones. Prices and description lengths were both measured in log terms here, but the results would be similar in unlogged terms. The descriptors shown here were selected based on my hedonic results; these descriptors include some of the ones with the most significant effects (in the case of “smoked game,” “precocious,” “unctuous,” and “brawny”), statistically significant but substantively insignificant effects (in the case of “velvety tannins”), and insignificant effects (in the case of “silky tannins”) on price according to those hedonic results. Some descriptors have stronger correlation with each other than the correlations shown here (see *Appendix Table 3* on that). By Cohen’s rules of thumb for interpreting the effect size of correlation coefficients, the only correlation that is non-trivial in size is the one between price and description length. Each indicator variable has a slight positive relationship with the description length. “Silky tannins” and “velvety tannins” have a correlation of essentially zero (about 0.01).

\* Statistically significantly different from zero with the Bonferroni correction at the 10% level.

*Appendix Table 3*

**Quandt descriptors with most significant pairwise correlations in my dataset**

|  |  |  |
| --- | --- | --- |
| *One descriptor* | *Another descriptor* | *Correlation* |
| Blackberry | Cassis | 0.26 |
| Meaty | Smoky | 0.24 |
| Blackberry | Chocolate | 0.23 |
| Cassis | Chocolate | 0.21 |
| Cedar | Tobacco | 0.21 |
| Leather | Tobacco | 0.21 |
| Black cherry | Plum | 0.20 |
| Cassis | Graphite | 0.20 |
| Honey | Quince | 0.19 |
| Cedar | Black currant | 0.19 |
| Blackberry | Violets | 0.19 |
| Raspberry | Roses | 0.18 |
| Blackberry | Plum | 0.18 |
| Blackberry | Graphite | 0.18 |
| Black currant | Tobacco | 0.18 |
| Blackberry | Inky | 0.17 |
| … | … | … |
| Black cherry | Honey | −0.11 |
| Honey | Plum | −0.11 |
| Honey | Raspberry | −0.13 |
| Blackberry | Honey | −0.14 |

*Note:* This table shows, for each possible combination of the 106 Quant descriptors that appear in the descriptions of at least 50 wines in my dataset, the 20 pairs of non-nested descriptors with the most significant correlations. The correlations were measured by Spearman’s correlation coefficient as in *Appendix Table 2*. The correlations were the most significant in the sense that they were statistically significant with the Bonferroni correction at the 10% level and, furthermore, the absolute value of their correlation was at least 0.10. There were 138 pairs of descriptors for which that was the case. Only 4 pairs that met those criteria had negative correlations; those 4 pairs with the most significant negative correlations are reported here. Some of the pairs with the most significant positive correlations were for descriptors that were nested in other descriptors. “Silky” is nested in “silky tannins,” for example, and so the correlation between that pair of descriptors is relatively high at 0.48. Such strong positive correlations are to be expected for nested descriptors, so this table only shows correlations between non-nested descriptors. The pairs of descriptors were sorted by the size of their correlation coefficient before rounding.

*Appendix Table 4*

**Statistics for wines in my dataset described by select descriptors**

|  |  |  |
| --- | --- | --- |
|  | *Descriptors* | |
| *Statistic* | *Smoked game* | *Precocious* |
| Number of wines | 55 | 197 |
| Mean (stdev) of price | $83 ($51) | $90 ($49) |
| Mean (stdev) of length | 1,546 (784) characters | 1,578 (885) characters |
| Most frequent varietal | Tie btw. Pinot Noir & Shiraz/Syrah (29%) | Cabernet Sauvignon and Blends (41%) |
| Most frequent country | France (55%) | France (53%) |
| Most frequent subregion | Rhone (31%) | California (28%) |
| Most frequent vintage | 2010 (16%) | 2012 (10%) |
| Most frequent reviewer | Robert Parker’s *Wine Advocate* (85%) | Robert Parker’s *Wine Advocate* (90%) |
|  | *Velvety tannins* | *Silky tannins* |
| Number of wines | 1,340 | 1,940 |
| Mean (stdev) of price | $75 ($44) | $71 ($44) |
| Mean (stdev) of length | 1,433 (786) characters | 1,421 (790) characters |
| Most frequent varietal | Cabernet Sauvignon and Blends (37%) | Cabernet Sauvignon and Blends (31%) |
| Most frequent country | France (32%) | United States (34%) |
| Most frequent subregion | Bordeaux (24%) | California (31%) |
| Most frequent vintage | 2010 (8%) | 2010 (8%) |
| Most frequent reviewer | Robert Parker’s *Wine Advocate* (71%) | Robert Parker’s *Wine Advocate* (71%) |
|  | *Unctuous* | *Brawny* |
| Number of wines | 745 | 202 |
| Mean (stdev) of price | $83 ($51) | $69 ($44) |
| Mean (stdev) of length | 1,433 (745) characters | 1,453 (785) characters |
| Most frequent varietal | Cabernet Sauvignon and Blends (22%) | Cabernet Sauvignon and Blends (31%) |
| Most frequent country | United States (44%) | United States (47%) |
| Most frequent subregion | California (42%) | California (43%) |
| Most frequent vintage | 2015 (10%) | 2012 (9%) |
| Most frequent reviewer | Robert Parker’s *Wine Advocate* (84%) | Robert Parker’s *Wine Advocate* (73%) |

*Note:* This table shows, for six of the Quandt descriptors, statistics on the wines in my dataset that are described by a given descriptor. The statistics shown are: the number of wines described by the descriptor; the mean (and standard deviation) of of the wines’ prices in unlogged terms; the mean (and standard deviation) of the length of the wines’ descriptions in unlogged terms; and the varietals, countries, subregions, vintages, and reviewers for which that descriptor most frequently appears (with the frequency in parentheses). Of note, wines described by the “precocious” descriptor are mostly from France (with 53% from that country), but they come from a number of subregions within that country (including Bordeau, Rhone, and Burgundy), so California is the most frequent “subregion for wines described by that descriptor. Also of note, for all six of the descriptors except “brawny,” the most frequent appellation as classified by K&L is no appellation classification. The most frequent appellation for wines described as “brawny” is Napa Valley; 25% of the brawny wines are from that appellation.

*Appendix Table 5*

**Hedonic results for descriptors with most significant price effects   
if estimated topic probabilities are also included as control variables**

|  |  |  |  |
| --- | --- | --- | --- |
| *Descriptor* | *Marginal effect (%)* | *Standard error (%)* | *p-value* |
| Smoked game † | 12 | 5 | 0.03 |
| Precocious \*† | 10 | 3 | 0.00 |
| Raisin † | 10 | 2 | 0.00 |
| Dried apricot † | 6 | 3 | 0.04 |
| Pain grille † | 6 | 2 | 0.01 |
| Vegetable | 6 | 4 | 0.10 |
| Angular \*† | 6 | 3 | 0.04 |
| Honey † | 4 | 1 | 0.00 |
| Animal | 4 | 2 | 0.05 |
| Espresso † | 4 | 1 | 0.00 |
| ... | ... | ... | ... |
| Pepper † | −3 | 1 | 0.00 |
| Black cherry | −3 | 0 | 0.00 |
| Floral pastille † | −3 | 1 | 0.01 |
| Spice box † | −3 | 1 | 0.01 |
| Black fruits † | −3 | 1 | 0.00 |
| Roasted meat | −3 | 2 | 0.04 |
| Unctuous \*† | −3 | 2 | 0.03 |
| Sour cherry † | −4 | 2 | 0.05 |
| Plum sauce † | −5 | 2 | 0.02 |
| Brawny \*† | −6 | 3 | 0.02 |

*Note:* This table is just like Table 1 of the main text, except these results are for a hedonic regression that also includes (following McCannon, 2020) estimated topic probabilities from a topic model as control variables. Everything else about the hedonic regression was the same as before. Just like Table 1 in the main text, this table shows the Quandt descriptors with estimated effects that are statistically significant at the 10% level and among the 10 most positive or most negative. The descriptors were sorted by the size of their estimated effect before rounding.

\* Defined in Parker’s wine glossary, as noted by Quandt (2007, p. 133)

† One of the most significant descriptors in Table 1 of the main text, too

*Appendix Table 6*

**Survey results for preferences between a precocious or unctuous wine**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Percent of...* | | |  |
|  | *All  (N=338)* | *Experts  (N=70)* | *Non-experts (N=268)* | *Experts same as non-experts?* |
| Prefer precocious | 72 (2) | 64 (6) | 75 (3) | 0.09\* |
| MWTP of  over 10 dollars | 1 (0) | 0 (0) | 1 (1) | 0.47 |
| Ten dollars | 4 (1) | 6 (3) | 4 (1) | 0.46 |
| Five dollars | 17 (2) | 14 (4) | 17 (2) | 0.56 |
| One dollar | 19 (2) | 17 (5) | 19 (2) | 0.67 |
| Zero dollars | 32 (3) | 27 (5) | 34 (3) | 0.31 |
| Prefer unctuous | 28 (2) | 36 (6) | 25 (3) | 0.09\* |
| MWTP of  zero dollars | 13 (2) | 16 (4) | 12 (2) | 0.45 |
| One dollar | 6 (1) | 9 (3) | 6 (1) | 0.36 |
| Five dollars | 7 (1) | 10 (4) | 6 (1) | 0.18 |
| Ten dollars | 1 (1) | 0 (0) | 1 (1) | 0.30 |
| Over 10 dollars | 1 (0) | 1 (1) | 0 (0) | 0.31 |

*Note:* This table shows, for all survey respondents except the 131 who said they preferred neither wine, the percentage who said they preferred a precocious or unctuous wine. The forced choices of the 97 respondents who initially said they would prefer either wine are included in the calculations. The table also shows their MWTP for their preferred wine relative to the other wine. Standard errors are reported in parentheses. Responses are also broken down by whether a respondent was an expert or not. To assess whether there are any differences between experts and non-experts, the p-values for z-tests of independent proportions are reported in the last column.

\* p-value < 10%

*Appendix Table 7*

**More survey results for mean MWTP in percentage terms**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Mean MWTP (%) among...* | | |  |
|  | *All* | *Experts* | *Non-experts* | *Experts same as non-experts?* |
| Precocious  rather than  unctuous | 7 (1)\*\*\* | 5 (3)\* | 7 (2)\*\*\* | 0.48 |

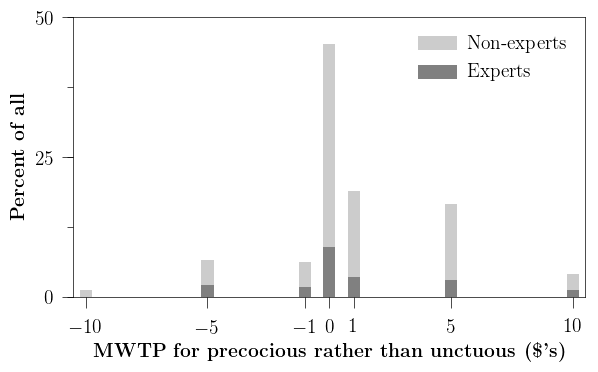
*Note:* Analogous to Table 4 in the main text, this table reports the means (and standard errors) of respondents’ MWTP in percentage terms for a precocious rather than unctuous wine (except the 131 who preferred neither of those). Responses are also broken down by whether a respondent was an expert or not. Along with each mean, the result of a one-sample t-test for whether the mean is different from zero is reported. The p-value for a Welch’s t-test of whether the mean MWTP of experts is different from the mean MWTP of non-experts is also reported in the last column. A respondent’s MWTP in percentage terms was calculated as their MWTP in dollar terms for their preferred descriptor, on the one hand, all over the price in dollars they said they typically pay for a bottle of wine, on the other hand. A negative MWTP was again used to denote a MWTP for silky rather than velvety tannins.

\*\*\* p-value < 1%

\* p-value < 10%

*Appendix Figure 1*

**Distribution of MWTP for a precocious rather than unctuous wine**



*Note:* Analogous to Figure 1 in the main text, this figure shows, for all survey respondents (except the 131 who said they preferred neither wine), a histogram of respondents’ MWTP for a precocious rather than unctuous wine. A negative MWTP denotes a MWTP for an unctuous rather precocious wine. Whether respondents were experts or non-experts is also shown. Four outliers in MWTP (two less than −$10 and two more than $10) lie beyond the range shown above, as discussed in the text of this supplementary appendix.

1. Ramirez (2010) uses descriptions written over a six-year period (ibid., p. 146). Both Chen and McCluskey (2018) and Ramirez (2010) use the price of a wine from *before* the wine’s description was published (Chen and McCluskey, 2018, p. 391; Ramierz, 2010, pp. 145). Ramierz (2010, p. 145) recognizes future descriptions cannot affect past prices, so he also tries to explain the price of a producer’s wine in terms of the description of the *previous* year’s vintage. Future descriptions could be associated with past prices to the extent the descriptions are reflecting past prices or some aspects of the wine that affect price. [↑](#footnote-ref-1)
2. According to Orley Ashenfelter, when you buy a $500 wine, you are not buying wine, you are buying a collectible (quoted in Bosker 2017, p. 173). Although I do not attempt to subdivide the wine market as finely as posited by Chen and McCluskey (2018), see that paper and references therein for more on wine market segmentation. [↑](#footnote-ref-2)
3. That wine is out of stock on K&L’s site. According to Wine-Searcher.com, the average price before taxes for a 750ml bottle of the wine is about $290 as of writing. Part of Parker’s description of the wine reads as follows, where I have emphasized the descriptors that Quandt deemed to be bullshit: “Surprisingly *precocious* and evolved for a wine from this estate [i.e., Bruno Giacosa], it possesses a dark ruby color with an amber edge, and abundant quantities of *tobacco*, *cherry liqueur*, *incense*, *spice box*, and licorice in its flamboyant nose. It is like candy, fleshy, full-bodied, *unctuous*, and *silky*.” For more on the wine, see <<https://web.archive.org/web/20200705230520/https://www.klwines.com/p/i?i=994716>>. [↑](#footnote-ref-3)