Journal of Wine Economics, Volume 14, Number 3, 2019, Pages 321–331
doi:10.1017/jwe.2019.40

Half-Blind Tasting: A Deception-Free Method for Sizing Placebo and Nocebo Responses to Price and Packaging Attributes

Robin S. Goldstein [a, b]

Abstract

Information conveyed on the price tag or label of a consumable packaged good is widely thought to change the consumer's sensory experience of consuming the good. Can the positive "placebo" effects of high prices and negative "nocebo" effects of low prices on consumer experience be isolated and observed in a controlled experiment without using deception? In a pilot wine experiment using a method I call "half-blind tasting," I observe that the nocebo response to a $5 price tag is stronger than the placebo response to a $50 price tag. To interpret these preliminary results, I borrow some insights from prospect theory. (JEL Classifications: C91, D81, L66, M31, Q11)

Keywords: blind tasting, nocebo, placebo, prospect theory, wine prices.

I. Introduction

In this article, I present the results of a pilot study using an experimental method that I call a "half-blind tasting." The purpose of the method is to address three questions: (1) Can consumers' sensory placebo (positive) and nocebo (negative) responses to price and packaging attributes (e.g., label information) be measured, in a laboratory setting, controlling for sensory wine characteristics, without using deception? (2) If consumers' placebo and nocebo responses to prices and labels are measurable,

I thank an anonymous referee and Karl Storchmann for their helpful comments. I also thank David Card for hosting me as a visiting scholar at UC Berkeley, where I conducted this experiment with his enthusiastic assistance; Matthew Rabin at Berkeley for his comments on my half-blind methodology; and Hilke Plassmann, Johan Almenberg, Anna Dreber, Shane Frederick, Laurie Santos, Uri Gneezy, and Alexis Herschkowitsch for conversations that led to the development of the ideas and methods presented here.
[a] University of California Agricultural Issues Center, 1 Shields Avenue, Davis, CA 95616-8514, USA
[b] Université de Bordeaux, Laboratoire d'Analyse et de Recherche en Économie et Finance Internationales, Pessac, France; e-mail: robin.s.goldstein@gmail.com.

how do their magnitudes compare? (3) Do placebo and nocebo responses to prices and labels behave like gains and losses in prospect theory?

My method is intended to control for the sensory characteristics of wine without using deception. The procedure is simple. Subjects are presented with two bottles of wine. One of the bottles has its supermarket price tag and label exposed (the "non-blind bottle"). The other bottle has its price tag and label concealed by brown paper (the "blind bottle"). Subjects taste the two wines side by side and write which wine they prefer and how much they would pay for each wine. The trick, unbeknownst to subjects, is that the two bottles are identical, so the only experimental manipulation is the exposure of the price and label. I, thus, take the positive or negative influence of the price and label on the subject's stated preference as a pure placebo or nocebo effect of price and label.

In my pilot study, conducted at the University of California, Berkeley, 53 subjects receive either a "placebo" treatment where the identical wines are each worth $50 (generating a positive bias and swaying tasters toward the exposed bottle) or a "nocebo" treatment where the identical wines are each worth $5 (generating a negative bias and swaying tasters toward the mystery bottle). Deception is not used with either treatment; the (non-blind) wines' actual prices are shown on the non-blind bottle. In spite of the small sample size, I observe statistically significant placebo and nocebo responses. In particular, I find that the "nocebo" response to a $5 price signal (with about 75% of subjects choosing the mystery bottle) is stronger than the "placebo" response to a $50 price signal (with about 60% choosing the exposed bottle). I draw from prospect theory to explain and illustrate these results, and I suggest a future research program.

The canonical 21st-century work on placebo and nocebo responses to consumer prices is Plassmann et al. (2008)'s brain-scanning study, where subjects tasted and rated wines while in an fMRI machine while being given true or false information about the wines' prices. For instance, subjects would taste the same $45 wine several times, sometimes being told (truthfully) that it was a $45 wine, and other times being told (falsely) that it was a $5 wine. Plassmann et al. hypothesis was that subjects' experiences would be negatively impacted by the low price ($5) and positively impacted by the high price ($45), even though the sensory input was otherwise identical in both cases. Plassmann et al. observed significant placebo and nocebo effects not only in differences between subjects' self-reported wine ratings, but also in the activation of frontal areas of the cerebral cortex that are associated with pleasure and reward. The implication is that tasters are not just "pretending" when they claim to derive more sensory pleasure from wines they believe to be more expensive: it is their *primary sensory experiences* that are influenced by the non-sensory information.[1]

[1] Although the price of a wine could theoretically be construed as "sensory information" insofar as it is perceived using the sense of sight, the stimulus is representational and cannot be understood by the perceiver in a way that would trigger the effect without higher-level cognitive processing and interpretation of

Packaging attributes can take many different material and informational forms beyond price. For instance, Bekkerman and Brester (2019) find effects of cork type on consumer wine preferences, and Drichoutis, Klonaris, and Papoutsi (2017), in an auction experiment, find that bottle size impacts revealed willingness to pay per unit volume for grape wine and pomegranate wine. Information, packaging, and sales context effects have also been widely observed in beer, as in Lee, Frederick, and Ariely (2006), who find effects of information about whether beer has been spiked with balsamic vinegar. Bohannon, Goldstein, and Herschkowitsch (2010) find that consumers cannot distinguish between dog food and pâte in a blind tasting. Hart (2018), in another blind tasting experiment, finds sensory effects of origin and firm size (i.e., whether a beer is "local" or "craft") on stated hedonic liking and stated willingness to pay for beer, controlling for sensory characteristics. Malone and Lusk (2019) find effects of "choice overload," that is, negative effects of too many choices on willingness to pay, that are mitigated by market interactions on the part of the seller.

Consumers do not prefer more expensive wines to cheaper wines in blind tastings (Goldstein et al., 2008; Ashton, 2014). Experts do not agree with each other on sensory "quality" (Hodgson, 2009; Ashton, 2017; Luxen, 2018), packaging and price attributes may play a bigger role in shaping consumer preferences for competing wine brands (or beer brands) than sensory wine (or beer) attributes.

The pilot experiment presented in the remainder of this article aims to measure the influence of price on wine preferences by eliciting pairwise choice and stated willingness to pay in a non-deceptive lab experiment. In Section II, I will describe the design of my "half-blind pairwise choice" method. In Section III, I will report the results of a modest-sized pilot study to test the method. In Section IV, I will discuss some drawbacks and possible improvements to the method. Finally, in Section V, I will interpret my results in the context of prospect theory and suggest some directions for future research.

II. Method

The important result of Plassmann et al. (2008), which arguably helped to jump-start a whole new research program within economics, was obtained using the technique of deception (i.e., together deceiving subjects with false price information) to generate a strong enough signal to overcome some of the challenges of fMRI methods (small number of subjects, noisy signals, etc.). The prestigious *Proceedings of the National Academy of Sciences,* which published Plassmann et al. (2008), like most peer-reviewed scientific journals outside of economics, accepted the use of deception as an experimental methodology.

the pattern of visual inputs as a representation of price, which in turn requires an understanding of the money system, number line, and other abstract elements of arithmetic and markets.

Unlike neuroscientists or psychologists, economists typically frown on the use of deception in experiments. "This rule exists," according to Jamison, Karlan, and Schechter (2008, p. 477), "in order to protect a public good: the ability of other researchers to conduct experiments and to have participants trust their instructions to be an accurate representation of the game being played…Two of the original experimental economics textbooks, Davis and Holt (1993) and Friedman and Sunder (1994), among others, proscribe the use of deception in experiments" (see also Levitt and List, 2007). Although there has been recent disagreement from Cooper (2014), Lusk (2019), and others on the value of the deception ban in economics, including an interesting focus debate in the February 2019 issue of *Food Policy,* the anti-deception principle remains a norm within economics. In the context of that norm, new methods that are able to elicit measurable cognitive bias effects continue to be useful for experimental economics. In this article, I propose a new non-deceptive method for eliciting placebo and nocebo responses to non-sensory attributes such as price. My method is meant to offer an alternative to deceptive methods such as the ones in Plassmann et al. (2008).

I call my experimental design a "half-blind tasting," built around a classic pairwise-choice comparison. Two identical bottles of wine are served and presented to the subject, with the only manipulation being that one of the two bottles is enclosed in a brown paper bag, concealing all of its packaging and price attributes, whereas the other bottle has its bottle and label exposed, with the price tag on the bottle.

I call the methodology "half-blind" because one of two wines is being tasted blind (with the price and brand concealed by a paper bag), while the other is being tasted non-blind (with the price and/or brand exposed). Other types of "half-blind" tastings not involving pairwise choice could also be imagined.

My procedure in the pilot half-blind pairwise choice experiment is relatively simple: two glasses of wine are poured by the experimenter (in this case, me) as the subject watches. One glass is poured from each of the two bottles. The subject is presented with the price-and-label-exposed (unbagged) bottle to inspect. (There is nothing to inspect on the bagged bottle: it is just a sealed brown paper bag.) The subject is asked to state which of the two wines he or she likes better; and how much he or she would be willing to pay for each of the two bottles. Subjects record gender, age group, and frequency of drinking wine. As a reference point, I also elicit how much subjects typically spend on wine.

The trick is that the two bottles are identical. Therefore, the experiment does not require straightforward deception. It may be argued that subjects are implicitly deceived into assuming that the two wines are different. On the other hand, providing subjects with the information that two products have identical sensory properties could render futile *any* choice task that is aimed at eliciting non-sensory preferences. The form I used to elicit preferences in the pilot experiment is shown in Figure 1.

Figure 1

Berkeley Half-Blind Pairwise Choice Experiment Form

GROUP 2 (3:30-5:30)

A FEW QUESTIONS

1. Contact (optional): Write your email here if you'd like to be debriefed later:

____

2. Sex: (circle one) Male Female **3. Age:** ________________

4. How often do you drink wine? (circle one)

Several times per week Once or twice per week Occasionally Never

5. How much do you usually spend on a bottle of wine in a store? $_______

NOW TASTE THE TWO WINES…Be sure to eat a cracker and have a drink of water in between the two.

YOUR PREFERENCES

Which wine did you like better? (circle one)

Wine A1 Wine A2

How much would you be willing to pay for each at a store?

Wine A1: $_______ Wine A2: $_______

When the wine's price tag is low, the predicted result is that consumers will favor the unlabeled wine. I call this the "nocebo response," or the choice-likelihood (or willingness to pay) value of the non-sensory attribute. When the price tag is high, on the other hand, the predicted result is that consumers will favor the labeled wine. I call this the "placebo response," or the choice-likelihood value of the non-sensory attribute.

My experiment included 53 subjects, who were randomly divided into two groups based on what time in the afternoon they took the test.

Eighteen subjects were placed into the "placebo" (high-wine-price) group, in which subjects compared an expensive ($50) exposed bottle of white wine against the same bottle in a brown paper bag. In this group, consumers were expected to exhibit the placebo response by disproportionately choosing the exposed bottle over the bottle in the brown paper bag.

The remaining 35 subjects were placed into the "nocebo" (low-wine-price) group, in which a subjects compared a cheap ($5) exposed bottle of white wine against the same bottle in a brown paper bag. In this group, consumers were expected to exhibit nocebo responses by disproportionately choosing the brown paper bag over the exposed bottle.

III. Results

Main results for the pilot study are shown in Table 1. The direction of both the nocebo ($5 price tag) and placebo ($50 price tag) effects come out as predicted: the non-sensory impact of price and packaging information is negative for the $5 wine and positive for the $50 wine.

Overall, as shown in the first row ("All wine drinkers") of each condition (Low-Price "Nocebo" and High-Price "Placebo"), 59% of subjects state a preference for a wine exposed as being worth $50 (the "placebo wine") to an identical wine whose bottle is concealed, whereas 26% of subjects state a preference for a wine exposed as being worth $5 (the "nocebo wine") to an identical wine whose bottle is concealed. The nocebo response for this pairwise-choice component of the experiment is statistically significant in a t-test against the null hypothesis that subjects are choosing between the two bottles randomly. The placebo response, however, does not reach statistical significance, although almost 60% of subjects choose the placebo, because the sample size for this condition, due to budgetary limitations for buying $50 wines, was limited to 18 subjects.

In terms of stated willingness to pay (indicated as "Mean WTP" in Table 1), in both the placebo and nocebo conditions, mean differences between the exposed and concealed bottles (in dollars) are statistically significant in one-tailed t-tests against a null hypothesis of zero or inverse dollar differences. Overall, as shown in the first row ("All wine drinkers") of each condition (Low-Price "Nocebo" and High-Price "Placebo"), subjects are willing to pay an average of $4.78 more for the exposed $50 wines than for the concealed $50 wines, and are willing to pay an average of $2.19 less for the exposed $5 wines than for the concealed $5 wines.[2]

Table 1 also reports some means comparisons between groups based on gender, self-reported frequency of wine drinking, and stated willingness to pay, but I choose not to test their statistical significance of between-means comparisons for gender, wine-expertise, or wine-buying effects in reporting these pilot results. Given the small number of total subjects (53), the particularly small number of

[2] Note that these two differences are not directly comparable as magnitude estimates, given the expected "anchoring" effect leading consumers to expect that both wines are worth more when one of their prices is exposed as high, or less when one is exposed as low. The anchoring effect is illustrated by the fact that mean willingness to pay for the concealed wine is significantly higher in the placebo condition ($11.44) than in the nocebo condition ($8.89).

Table 1

Results of Berkeley Half-Blind Pairwise Choice Experiment

Cheap Wine (Nocebo Response)

Wine with $5 Price Label vs. Identical Wine with Unknown Price

		Known Price = $5		*Unknown Price*		
	n	*% Prefer in Pairwise Choice*	*Avg. Stated WTP*	*% Prefer in Pairwise Choice*	*Avg. Stated WTP*	*WTP Difference*
All wine drinkers	35	0.257***	$6.75	0.742***	$8.89	–$2.19**
Frequent wine drinkers	8	0.250*	$7.06	0.750	$6.68	$0.38
Infrequent wine drinkers	23	0.261***	$6.57	0.739***	$8.93	–$2.37**
Males	21	0.333*	$7.93	0.667*	$9.14	–$1.21
Females	14	0.140**	$4.85	0.860**	$8.50	–$3.64**
Usually spend <= $10 on wine	23	0.261***	$5.50	0.739***	$7.96	–$2.46**
Usually spend >$10	12	0.250**	$9.00	0.750**	$10.67	–$1.67

Expensive Wine (Placebo Response)

Wine with $50 Price Label vs. Identical Wine with Unknown Price

		Known Price = $50		*Unknown Price*		
	n	*% Prefer in Pairwise Choice*	*Avg. Stated WTP*	*% Prefer in Pairwise Choice*	*Avg. Stated WTP*	*WTP Difference*
All wine drinkers	18	0.592	$15.89	0.421	$11.44	$4.78***
Frequent wine drinkers	7	0.571	$12.00	0.429	$8.43	$3.57**
Infrequent wine drinkers	9	0.667	$19.56	0.333	$12.44	$7.11***
Males	13	0.571	$17.57	0.429	$12.08	$6.08**
Females	5	0.600	$11.20	0.400	$9.80	$1.40
Usually spend <=$10 on wine	10	0.636	$11.55	0.364	$9.10	$2.60*
Usually spend >$10	8	0.500	$21.88	0.500	$14.38	$7.50***

*** Statistically significant in t-test comparing two means at $p < 0.001$ level.

** Statistically significant at $p < 0.01$ level.

* Statistically significant at $p < 0.05$ level.

https://doi.org/10.1017/jwe.2019.40 Published online by Cambridge University Press

female subjects (19), and the considerable selection bias in the subject pool (UC Berkeley economics graduate students and professors), I choose not to report the results of statistical significance tests on any of the between-groups comparisons (other than the basic experimental manipulation). To report statistics on these comparisons would imply a greater degree of external validity than I feel is warranted given the nature of the pilot study and data. I simply infer, from the differences observed (particularly the strong female reaction, in terms of pairwise choice, to the nocebo price cue), that the full-scale version of this experiment should be designed to have the power to detect male-female differences in nocebo and placebo responses.

IV. Drawbacks

My pilot experiment has many drawbacks beyond its low power. First, the nocebo-placebo difference I claim to be observing is highly sensitive to the choice of the two price points and the assumed neutral reference points of subjects. If the reference point is closer to the placebo price point than to the nocebo price point, this could explain the greater magnitude of placebo effects, although this is unlikely given that the stated tendency to pay for wine for all subjects, and both experimental groups, was around \$15, closer to \$5 than to \$50.

A second major drawback to my pilot design is that it does not isolate or discriminate between the effects of price and label, or more specifically, between the supermarket price tag and the expensive-looking label to which it is affixed. Future half-blind tasting designs could be more precise in the way they conceal or reveal information to observe the effects of price and label separately.

Among all objections to my design, the most serious problem may be that my pairwise choice is forced. Consumers do not have a null "indifferent" option to indicate no preference between the two wines. In defense of my method, I decided not to offer this option because of the small subject pool: offering a neutral option might have reduced power to untenably low levels. None of my 53 subjects asked whether the two wines were the same (my plan had been simply to decline to answer positively or negatively), and none asked whether they could indicate indifference. Nonetheless, forced choice is a valid criticism of this simplest version of half-blind pairwise choice, and it would be useful in future experiments to compare these results with a version in which a no-preference option is included.

V. Discussion and Directions for Future Research

The main purpose of this article has simply been to propose and try out a method for eliciting placebo and nocebo response sizes generated by price signals, and to report results from one initial beta test of this method. From the relatively clean results from the pilot experiment, with mean differences in the expected directions, I conclude that the method is worth exploring further: subjects as a whole exhibit both placebo and nocebo responses with sufficient magnitude to provisionally validate the method.

At present, the validity and usefulness of the method I propose are empirically supported only by the single, relatively small-scale pilot experiment whose results I report here. This provides some limited initial support of my claim that the "half-blind tasting" design warrants further exploration and refinement as a method of eliciting placebo/nocebo responses that does not use deception. In the future, this method should ideally be tested on a larger subject pool that is sufficiently large enough subject pool to test for gender and expertise effects, and with the ability to test signals at a variety of intermediate price points.

Adjusting the prices and comparing the magnitude of responses at a variety of price points could help to illuminate consumer groups' neutral price reference points (or non-numerical packaging attributes), to evaluate the magnitude of their positive and negative responses at different placebo and nocebo price points, and perhaps eventually to construct a curve that describes these response patterns analytically. A purely hypothetical example of what such a curve might look like is shown in Figure 2.

In Kahneman and Tversky's (1979) "prospect theory," subjects in a lab experiment do not maximize expected value but rather over-protect against losses, nonlinearly. This effect is represented by the familiar idealized prospect theory curve shown in Figure 3.

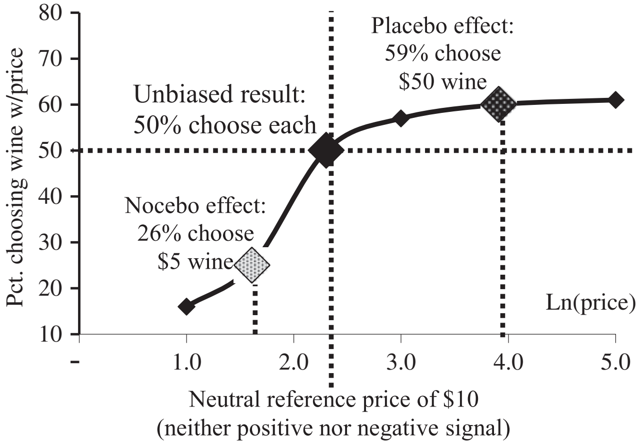
Extending the analogy to prospect theory into the behavioral decision-making realm, one possible explanation for the greater strength of the nocebo response than the placebo response, then, might simply be that consumers process lower-than-usual price signals as losses and higher-than-usual price signals as gains, and that in wine shopping as in psych-lab gambling, the pain of a loss may sting us more than the pleasure of a gain rewards us.

Comparing the upward-sloping curves and their kinks in Figures 2 and 3, a parallel to prospect theory could be drawn if low price signals (nocebo inputs) were interpreted as "losses" (with respect to the neutral reference point, where the curve crosses the "losses-gains" (X) axis at zero value), and high price signals (placebo inputs) were interpreted as "gains." Thus the "losses-gains" axis in Figure 3 would be analogous to the log-wine-price axis in Figure 2. The "value" (Y) axis in Figure 3, meanwhile, could be viewed as analogous to the percent choosing, the Y-axis in Figure 2.

If the half-blind research program proposed here were extended, the curve could be filled out with more data points and more accurately described by eliciting responses at multiple price points along the X-axis (e.g., testing responses at $5, $10, $15, …). Observing placebo/nocebo curves on the individual and aggregate

Figure 2

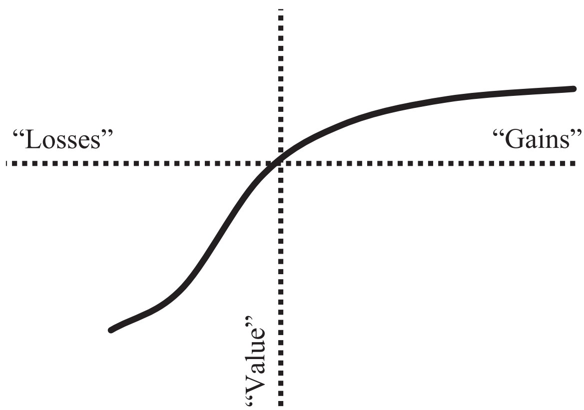
Plotting Half-Blind Tasting Results at Different Price Points on a Placebo/Nocebo Curve



Note: Percent of consumers theoretically preferring a price-exposed wine to the identical price-concealed wine, shown to a price-blind wine, based on half-blind tasting results at several different price points (populated with two data points from experiment results and a purely hypothetical curve going through them).

Figure 3

Prospect Theory Value Curve



Source: Based on Kahneman and Tversky (1979).

levels could help illuminate a predictable relation between price stimuli and experienced pleasure. Other packaging attributes (e.g., information on the label) could also be tested against a neutral reference point and conceptualized as "gains" or "losses" in this way. Such research could eventually help firms make better decisions about pricing and investment in packaging and information attributes.

References

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.

Ashton, R. (2017). Dimensions of expertise in wine evaluation. *Journal of Wine Economics*, 12(1), 59–83.

Bekkerman, A., and Brester, G. W. (2019). Don't judge a wine by its closure: Price premiums for corks in the U.S. wine market. *Journal of Wine Economics*, 14(1), 3–25.

Bohannon, J., Goldstein, R., and Herschkowitsch, A. (2010). Can people distinguish pâté from dog food? *Chance*, 23(2), 43–46.

Cooper, D. J. (2014). A note on deception in economic experiments. *Journal of Wine Economics*, 9(2), 111–114.

Davis, D. D., and Holt, C. A. (1993). *Experimental Economics.* Princeton, NJ: Princeton University Press.

Drichoutis, A. C., Klonaris, S., and Papoutsi, G. S. (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.

Friedman, D., and Sunder, S. (1994). *Experimental Methods: A Primer for Economists.* Cambridge, UK: Cambridge University Press.

Goldstein, R., Almenberg, J., Dreber, A., Emerson, J. W., Herschkowitsch, A., and Katz, J. (2008). Do more expensive wines taste better? Evidence from a large sample of blind tastings. *Journal of Wine Economics*, 3(1), 1–9.

Hart, J. (2018). Drink beer for science: An experiment on consumer preferences for local craft beer. *Journal of Wine Economics*, 13(4), 429–441.

Hodgson, R. T. (2009). How expert are "expert" wine judges? *Journal of Wine Economics*, 4(2), 233–241.

Jamison, J., Karlan, D., and Schechter, L. (2008). To deceive or not to deceive: The effect of deception on behavior in future laboratory experiments. *Journal of Economic Behavior & Organization*, 68(3–4), 477–488.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision making under risk. *Econometrica*, 47(2), 263–292.

Lee, L., Frederick, S., and Ariely, D. (2006). Try it, you'll like it: The influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science*, 17(12), 1054–1058.

Levitt, S., and List, J. A. (2007). What do laboratory experiments measuring social preferences tell us about the real world? *Journal of Economic Perspectives*, 21(2), 153–174.

Lusk, J. (2019). Viewpoint: The costs and benefits of deception in economic experiments. *Food Policy*, 83, 2–4.

Luxen, M. (2018). Consensus between ratings of red Bordeaux wines by prominent critics and correlations with prices 2004–2010 and 2011–2016: Ashton revisited and expanded. *Journal of Wine Economics*, 13(1), 83–91.

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

Plassmann, H., O'Doherty, J., Shiv, B., and Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, 105(3), 1050–1054.