



**HAL**  
open science

## Economic Experiments in Honor of Thorstein Veblen

Robin Goldstein

► **To cite this version:**

Robin Goldstein. Economic Experiments in Honor of Thorstein Veblen. Economics and Finance. Université de Bordeaux, 2019. English. NNT : 2019BORD0075 . tel-02419481

**HAL Id: tel-02419481**

**<https://theses.hal.science/tel-02419481>**

Submitted on 19 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **THE BULLSHIT HORIZON**

Essays on consumer demand for food and drink

By Robin S. Goldstein

université  
de **BORDEAUX**

**larefi**  
Laboratoire d'analyse et de recherche  
en économie et finance internationales

Université de Bordeaux  
École doctorale Entreprise, économie et société  
Ph.D. thèse, sciences économiques

Soutenue le 03/06/2019  
Larefi, Laboratoire d'analyse et de recherche en économie et finance internationales  
bâtiment Recherche Économie - 1er étage  
avenue Léon Duguit, 33 608 Pessac

Membres du jury:

Jean-Marie Cardebat, Université de Bordeaux (Directeur de thèse)  
Julian Alston, University of California, Davis (President)  
Johan Swinnen, Katholieke Universiteit Leuven (Rapporteur)  
Steven Ziliak, Roosevelt University (Rapporteur)





For Granddad  
*who spoke less bullshit than anyone else*

## Table of contents

Summary of the dissertation in English and French.....	2
Preface and acknowledgments .....	10
Introduction	
<i>The case for an economics of bullshit</i> .....	12
Chapter 1. Do more expensive wines taste better?	
<i>A large-scale blind tasting experiment on the role of bullshit in retail wine prices</i> .....	18
Chapter 2. Can people distinguish pâté from dog food?	
<i>An experiment to test the sensory impact of food labeling</i> .....	27
Chapter 3. Hide the label, hide the difference?	
<i>Two experiments on consumer preferences for bullshit in lager beer</i> .....	34
Chapter 4. An impostor at the Great American Beer Festival	
<i>A lager drinker shows up to sign books in craft beer heaven.</i> .....	80
Chapter 5. Craft beer is wine	
<i>A question, a poem, and a song</i> .....	92
Chapter 6. Many laws of many prices?	
<i>Premiumization and multimodal distributions in stated tendency to pay for beer and wine</i> .....	98
Chapter 7. The stench of cheap perfume	
<i>A half-blind tasting method for comparing placebo and nocebo responses to wine bullshit</i> .....	136
Chapter 8. Capuchin monkeys do not show human-like pricing effects	
<i>Four lab experiments to test whether demand responses to price bullshit are uniquely human</i> .....	146
Chapter 9. The costs and benefits of excellence	
<i>A bullshit-on-bullshit investigation of wine awards</i> .....	171
Chapter 10. Opportunities in immortality	
<i>On strategic planning in the Omega-3 fish oil industry</i> .....	218
Afterword .....	245
References.....	248

## Summary of the dissertation in English

This dissertation is about the economics of bullshit. My particular focus is on the bullshit attributes of consumer products. I propose some methods of identifying and measuring bullshit in retail markets, I test these methods using data from the U.S. food and drink markets, I report some results, and I suggest some future directions for the research program. The dissertation is organized as follows.

First, in the introduction, “The case for an economics of bullshit,” I review a few of the leading treatments of bullshit in the academic literature, clarify the difference between bullshit and information, and delineate a category of characteristics of consumer goods that I call the “bullshit attributes.”

In Chapter 1, “Do more expensive wines taste better?” (co-authored with Johan Almenberg et al. and published in the *Journal of Wine Economics*), I report results from a nationwide blind tasting experiment with more than 6,000 wine ratings, suggesting that individuals who are unaware of the price do not derive more enjoyment from more expensive wine. We find that the correlation between price and overall rating is small and negative, suggesting that individuals on average enjoy more expensive wines slightly *less*. For individuals with wine training, however, we find indications of a non-negative relationship between price and enjoyment, with marginal statistical significance. Our results suggest that for non-experts, variation in wine prices may be better explained by Veblen attributes (decorative and invisible attributes) than by useful attributes.

In Chapter 2, “Can people distinguish pâté from dog food?” (co-authored with John Bohannon and Alexis Herschkowitsch and published in *Chance*), I report results from a blind tasting experiment in which subjects are presented with five visually identical processed meat samples (pâté, liverwurst, etc.). One of the processed meat samples is Newman’s Own dog food. Subjects perform no better than random chance at correctly identifying which sample is dog food. To equalize the visual attributes of the five samples, we blend them each in food processor, garnish them with parsley, and serve them on Carr’s Water Crackers. We ask 21 subjects to identify which of the five samples is dog food. We find that subjects perform no better than random chance at correctly identifying which sample was dog food.

In Chapter 3, “Hide the label, hide the difference?”, I first present the results of a “triangle test” to determine whether 138 subjects at a beer bar in Portland, Oregon can detect any sensory differences between sensory attributes of competing European beer brands (Heineken, Stella Artois, and Budvar), served on draft. Subjects’ performance on the triangle test is statistically indistinguishable from chance. I conclude that the price differences between Heineken, Stella, and Budvar cannot be explained by the sensory

characteristics of the beer, and must instead be explained by packaging and bullshit attributes. Next, in a multiple-choice survey experiment with about 3,000 subjects, I verify that consumers are highly heterogeneous in their preferences for Heineken, Stella, Budvar, and Bud Light. I account for the results in terms of differences between brands in their bullshit attributes. I compare preferences for “buying” vs. preferences for “drinking,” and I discuss some differences in the brands’ advertising strategies.

In Chapter 4, “An Impostor at the Great American Beer Festival,” I take a detour from the academic writing style and make some qualitative observations about bullshit in the rapidly expanding U.S. craft beer segment. I observe bullshit through the lens of the world’s biggest craft beer convention, in my capacity as the co-author of a book called *The Beer Trials*.

In Chapter 5, “Craft beer is wine,” I make some qualitative observations about the amount of bullshit in contemporary craft beer, first in prose, then in a rhyming poem about craft beer, and finally in a song about the decline of normal beer in California.

In Chapter 6, “Many Laws of Many Prices?,” I first analyze results from an online survey of 3,300 U.S. consumers nationwide, conducted using Google Surveys, that elicits what I call “stated tendency to pay” (“STTP”) for retail beer and wine. I discuss some differences between beer and wine with respect to their levels of “premiumization.” I define premiumization in terms of the multi-modality of price frequency distributions, and I formalize my definitions. I propose that multi-modality may be one marker of Veblen attributes in the markets for retail consumer goods. I identify multi-modality in beer and wine STTP frequency distributions, and I report some gender differences and other demographic effects. Finally, I discuss some effects in neuroscience that may also help to explain multi-modal price frequencies and natural number frequencies.

In Chapter 7, “The stench of cheap perfume,” I ask: can the biasing effect of positive or negative price information on consumer preferences be isolated and observed directly? In this chapter, I describe an experimental design that I call a “half-blind tasting,” meant to measure and compare the positive (“placebo”) consumer responses to high wine prices and negative (“nocebo”) consumer responses to low wine prices. In a simple pairwise-choice task, subjects taste and choose which of two wines they prefer: one bottle with the price concealed, and an identical bottle with the price exposed. I expect subjects to be drawn toward bottles with high price tags (and away from their identical twins with unknown prices) and repelled by bottles with low price tags (and toward their identical twins with unknown prices). In a 53-subject pilot experiment, I try out the half-blind tasting method and observe effects consistent with expectations. In particular, I observe that the “nocebo” response to a \$5 price signal (with about 75% of subjects choosing the identical price-concealed bottle) is stronger than the “placebo” response to a \$50 price signal (with about 60% choosing the identical price-exposed bottle). I suggest that larger

half-blind tasting experiments to compare effects at many price points could yield fruitful results.

In Chapter 8, “Capuchin monkeys do not show human-like pricing effects,” co-authored with Laurie Santos et al. and published in *Frontiers in Decision Neuroscience*), we take a comparative approach to distinguishing between different accounts of price signal effects and biases. We test whether brown capuchin monkeys (*Cebus apella*) are also susceptible to pricing effects within the context of an experimentally trained token economy. Using a capuchin population previously trained in a token market, we explore whether monkeys attached higher value to higher-priced goods, all other things being equal, across four experiments. Although monkeys demonstrated an understanding of which goods had which prices (consistently shifting preferences to cheaper goods when prices were increased), we observe no evidence that such price information affected their choices between goods when the budget constraint was removed. These results suggest that the pricing effects observed in human beings may involve uniquely human cognitive capacities, such as an understanding of market forces and social signals.

In Chapter 9, “The costs and benefits of excellence,” I present joint work with Orley Ashenfelter and Craig Riddell. We begin by investigating the standards for the *Wine Spectator* “Awards of Excellence” given by the magazine’s expert judges to about 4,000 of the world’s top wine restaurants annually. To be considered for an Award of Excellence, a restaurant must submit a \$250 entry fee. We employ a bullshit-on-bullshit technique that was pioneered by the physicist Alan Sokol (1996) and test the standards for awards through an undercover field experiment. From this experiment, we conclude that awards are available to any restaurant that submits the \$250 entry fee and completes the paperwork correctly. Next, we explore the effect of *Wine Spectator* awards on the prices of New York City restaurants. We use a hedonic price model and a data set of 1,709 restaurant ratings and prices from the *Zagat Survey* to measure the returns to restaurants on their investments in bullshit. Finally, we find that for the *Zagat* data set as a whole, higher meal prices do not predict higher food ratings, and that for restaurants with the lowest third of meal prices (\$32 and under), higher meal prices actually predict *lower* food ratings, controlling for service and décor ratings, location, and cuisine-type fixed effects.

In Chapter 10, “Opportunities in immortality,” I report the results of a firsthand investigation into the Omega-3 fish oil industry. Finally, in the afterword, I consider some of the broadest and most concrete brushstrokes of bullshit in the present-day consumer world, and I illustrate my points with some recent trends of Bordeaux wine prices.

**Keywords:** Prices, Wine, Beer, Veblen Effect, Consumer Behavior, Information

**JEL Classifications:** B15, C9, D12, D83

## Summary of the dissertation in French

Cette thèse est motivée par la question suivante: comment appliquer plus fidèlement la théorie de «bullshit» aux marchés actuels? Dans les chapitres empiriques de cette thèse (chapitres 1, 2, 3, 6, 7, 8, et 9), j'essaie d'observer certains «bullshit attributes» et leurs prix sur les marchés de consommation actuels.

J'utilise diverses techniques expérimentales, notamment des enquêtes auprès des consommateurs, des expériences de dégustation à l'aveugle, des expériences de laboratoire et des expériences d'infiltration. Le travail empirique que je présente dans cette thèse a été mené au cours des onze dernières années, remontant à 2008, lorsque mon intérêt initial pour le sujet découlait d'une dégustation à grande échelle de vins à l'aveugle dans laquelle mes collègues et moi avons servi 6 000 verres de dégustation à environ 500 personnes, avec des bouteilles coûtant de 1,65 \$ à 150 \$. Nous avons trouvé une corrélation inverse faible mais statistiquement significative entre les prix et les préférences pour le groupe de sujets dans son ensemble.

Je présente des travaux empiriques sur les préférences des consommateurs et les prix de détail américains de la bière, du vin, et des aliments. Mes techniques comprennent des expériences de choix de consommateurs en ligne, des expériences de dégustation à l'aveugle de vins et de bières, des expériences d'investigation et des régressions hédoniques des prix.

Au chapitre 1, je présente une réimpression légèrement révisée de mon article de 2008 du *Journal of Wine Economics*, intitulé «Des vins plus chers ont-ils meilleur goût? Preuves provenant d'un large échantillon de dégustations à l'aveugle», nous constatons que les personnes qui ignorent le prix n'ont pas préférence marquée pour le vin plus cher. Sur un échantillon de plus de 6 000 notations de vins à l'aveugle de consommateurs américains, nous constatons que la corrélation entre le prix et la note globale est faible et négative, ce qui suggère que les individus apprécient en moyenne un peu moins les vins plus chers. Pour les personnes ayant une formation en dégustation de vin, cependant, nous trouvons des indications d'une relation non-négative entre le prix et le plaisir, avec une significativité statistique marginale. Mes résultats suggèrent que pour les non-experts, la variation des prix du vin pourrait être mieux expliquée par les attributs de Veblen (attributs décoratifs et invisibles) que par les attributs utiles.

Au chapitre 2, «Les gens peuvent-ils distinguer les pâtés des aliments pour chiens?», qui réimprime une version légèrement révisée d'un article, co-écrit avec John Bohannon et Alexis Herschkowitsch, précédemment publié dans *Chance* (2010). Lors d'une expérience de dégustation à l'aveugle, nous présentons aux sujets quatre échantillons de produits à base de viande mélangés destinés à la consommation humaine, tels que le pâté et le pâtée de foie, ainsi qu'un échantillon constituant la nourriture pour chiens

Newman's Own. Pour égaliser les attributs visuels des cinq échantillons, nous les mélangeons chacun dans un robot culinaire, nous les garnissons de persil et nous les servons sur des biscuits croustillants de Carr's. Nous demandons à 21 sujets d'identifier lequel des cinq échantillons est de la nourriture pour chien. Les sujets ne font pas mieux que le hasard pour identifier correctement quel échantillon était de la nourriture pour chien.

Au chapitre 3, «Cachez l'étiquette, cachez la différence», je reviens à la présentation de nouveaux travaux. Je présente les résultats d'une expérience sur le terrain d'une dégustation à l'aveugle d'un groupe de 138 buveurs de bière américains, qui s'est déroulée dans un bar à bière de Portland, Oregon. Mon expérience utilise une conception de «test de triangle», une technique de science sensorielle mise au point par Amerine et al. (1965) pour déterminer si deux échantillons ont des propriétés sensorielles identiques du point de vue des consommateurs. Je vérifie si les sujets peuvent détecter des différences sensorielles entre les «attributs utiles» (attributs sensoriels de la bière) des marques de bière européennes concurrentes (Heineken, Stella Artois et Budvar), servies à la pression. La performance des sujets au test du triangle est statistiquement impossible à distinguer du hasard. Je conclus que les différences de prix entre Stella et Heineken, et entre Stella et Budvar, ne peuvent pas être expliquées par des attributs utiles, mais doivent plutôt être expliquées par des «attributs de Veblen» (attributs décoratifs et invisibles, tels que le nom de marque, l'emballage, l'étiquetage, les notes, etc.).

Partant du principe que l'hétérogénéité des consommateurs dans les préférences entre les marques de Stella, Heineken et Budvar ne peut s'expliquer que par les attributs de Veblen, je pose deux questions: premièrement, les consommateurs américains ont-ils des préférences cohérentes et hétérogènes pour ces trois marques? Deuxièmement, si oui, quels modèles peut-on observer dans de telles préférences? Dans le cadre d'une expérience d'enquête à choix multiples, menée via Google Consumer Surveys, j'ai divisé un groupe d'environ 3 000 consommateurs américains en deux, posant à chaque sous-groupe une question légèrement différente: (V1) «Quelle bière auriez-vous le plus tendance à acheter? (V2) «Quelle bière préférez-vous boire?» Les choix proposés sont les trois marques européennes de bière blonde testées au chapitre 4 (Stella, Heineken et Budvar), ainsi que la bière la plus vendue en Amérique (Bud Light). J'observe diverses différences systématiques entre les préférences «d'achat» et «de consommation», ainsi que des différences démographiques entre les préférences des consommateurs en matière de marque, y compris des effets statistiquement significatifs sur le genre. Je rends compte des résultats en termes de différences entre les marques dans leurs attributs Veblen et je complète la discussion avec des exemples tirés des campagnes publicitaires de chacune des quatre marques de bière.

Le chapitre 4, «Un imposteur au Great American Beer Festival», traite de mon expérience en tant qu'invité à la plus grande convention de bière artisanale au monde, à Denver, au Colorado, en ma qualité de co-auteur d'un livre intitulé *The Beer Trials*. J'appelle ce chapitre (ainsi que le chapitre 12) une «expérience» dans le sens très large puisqu'il s'agit d'une manipulation, mais il ne s'agit en fait que d'un essai journalistique à la première personne, de nature ni empirique ni économique, qui devraient être considérés comme des lectures facultatives pour mes directeurs de thèse.

Le chapitre 5, «Craft beer is wine», traite des essais et chansons de bière «craft».

Au chapitre 6, «De nombreuses lois à plusieurs prix?», je présente les résultats d'un sondage en ligne mené auprès de 3 300 consommateurs américains, à l'aide de Google Surveys, qui permet d'obtenir ce que j'appelle la «tendance déclarée à payer» (STTP en anglais) pour la bière et le vin dans le commerce de détail. Je discute de certaines différences entre la bière et le vin en ce qui concerne leur niveau de «premiumisation». Je définis la premiumisation en termes de multimodalité de la distribution de la fréquence des prix, et je formalise mes définitions. Je propose que la multimodalité puisse être un marqueur des attributs de Veblen sur les marchés des biens de consommation au détail. J'identifie la multimodalité dans les distributions de fréquence STTP de la bière et du vin et j'identifie certaines différences entre les sexes et d'autres effets démographiques. Enfin, je discute de certains effets en neuroscience qui peuvent également aider à expliquer les fréquences de prix multimodales et les fréquences de nombres naturels.

Au chapitre 7, «The stench of cheap perfume», je propose une nouvelle méthode expérimentale, que je qualifie de «dégustation à demi aveugle», et je rapporte les résultats d'une première expérience pilote menée à l'Université de Californie à Berkeley dans le but de tester cette méthode. 53 sujets choisissent lequel de deux vins identiques ils préfèrent, et précisent combien ils paieraient pour chacun. Une des bouteilles est présentée avec son étiquette et l'étiquette de prix de supermarché exposée. Un duplicata du même vin est emballé dans un sac en papier brun avec tous ses attributs d'emballage cachés. Dans le groupe «placebo», les vins identiques valent 50 dollars, générant un biais positif et incitant les dégustateurs à se tourner vers la bouteille exposée. Dans le groupe «nocebo», les vins identiques valent 5 \$, générant un biais sensoriel négatif et incitant les dégustateurs à se tourner vers la bouteille mystère. La tromperie n'est utilisée avec aucun des deux groupes; les prix réels des vins sont indiqués. En comparant les deux, je trouve que la réponse «nocebo» à un signal de prix de 5 \$ (environ 75% des sujets choisissant la bouteille mystère) est plus forte que la réponse «placebo» à un signal de prix de 50 \$ (avec environ 60% des sujets choisissant la bouteille exposée) et plus forte chez les femmes que chez les hommes. Dans le cas «nocebo», 86% des femmes préfèrent les attributs sensoriels du vin dans la bouteille mystère aux attributs sensoriels du vin identique dans la bouteille où l'étiquette indique un prix de 5 \$.



Le chapitre 8, «Les singes capucins ne montrent pas d'effets de tarification semblables à ceux de l'homme», réédite une version légèrement révisée d'un article, co-écrit avec Laurie Santos, Rhia Catapano, Nicholas Buttrick et Jane Widness, qui a été publié dans *Frontiers in Decision Neuroscience* (2014). Nous présentons les résultats d'une expérience qui utilise une approche comparative pour distinguer différents comptes d'effets de signaux de prix et de biais. Nous testons si les singes capucins bruns (*Cebus apella*) sont également sensibles aux effets de tarification dans le contexte d'une économie de jetons formée de manière expérimentale. En utilisant une population de capucins préalablement formés sur un marché de jetons, nous avons cherché à déterminer si les singes accordaient une valeur plus élevée aux produits plus chers, toutes choses étant égales par ailleurs, au cours de quatre expériences. Bien que les singes sachent quels biens avaient quels prix (changement de préférence constant pour les biens meilleur marché lorsque les prix augmentaient), nous n'observons aucune preuve que cette information sur les prix ait une incidence sur leurs choix entre biens lorsque la contrainte budgétaire a été levée. Ces résultats suggèrent que les effets de prix observés chez les êtres humains pourraient impliquer des capacités cognitives humaines uniques, telles qu'une compréhension des forces du marché et des signaux sociaux.

Le travail conjoint avec Orley Ashenfelter et Craig Riddell est présenté au chapitre 9, «The costs and benefits of excellence». Qui considère les «Prix d'excellence» de Wine Spectator, attribués par les juges experts du magazine à environ 4 000 des meilleurs cartes de vins de restaurants au monde chaque année. Pour être considéré pour un prix d'excellence, un restaurant doit soumettre des frais d'inscription de 250 \$. Nous commençons par une expérience sur le terrain sous couverture (création d'une fausse carte pour un restaurant fictif), à partir de laquelle nous concluons que les récompenses ne sont pas rationnées, c'est-à-dire qu'elles sont disponibles pour tout restaurant qui soumet le droit d'inscription de 250 \$ et complète correctement le dossier.

Nous examinons ensuite l'effet des récompenses Wine Spectator sur les prix des restaurants de la ville de New York dans un ensemble de régressions multivariées utilisant un modèle de prix hédonique et un ensemble de données de 1709 évaluations de restaurants et de prix provenant de l'enquête Zagat. Nous constatons que, en fonction du lieu, du type de cuisine, des notes des consommateurs sur les restaurants, des décors et des services, un restaurant primé par le Wine Spectator facturera en moyenne plus par repas (selon le niveau de récompense) qu'un restaurant équivalent (ayant les mêmes notes de nourriture, de décor et de service mais aucun prix d'excellence de Wine Spectator). Nous examinons également les effets des attributs du restaurant sur les notes des aliments. Nous constatons que pour l'ensemble des données Zagat, les prix les plus élevés ne prédisent pas les notes les plus élevées, et que pour les restaurants dont le tiers des prix est le plus bas (32 \$ et moins), les prix les plus élevés prédisent en réalité des notes moins élevées.

Le chapitre 10, «Opportunities in immortality», qui (comme le chapitre 6) est une lecture supplémentaire, écrite sous forme d'essai journalistique, ne devrait pas être considéré comme faisant partie de la réalisation de mon doctorat. C'est un rapport d'enquête sur la fête de l'année du secteur des suppléments oméga-3 au Ritz-Carlton Resort, dans les îles Canaries.

## **Preface and acknowledgments**

When I say “the economics of bullshit,” I don’t mean the economics of Freddie Mac’s accountants, Elon Musk’s pronouncements, Bernie Madoff’s dividends, or Boris Johnson’s Brexit plans. Although bullshit (however defined) probably plays a role in all of the above phenomena, I do not aim to expose fraud, deceit, or other violations of laws or market norms. I do not aim to uncover violations of utility theory, the Laws of Supply and Demand, or any other economic theory. I do not aim to uncover sensational new examples of consumers, producers, or markets behaving stupidly or irrationally.

Much of the recent experimental and behavioral literature in economics and management science seems directed toward the same vague conclusion that human beings are irrational fools who naïvely deceive ourselves into holding implausible beliefs or making decisions that are against our own best interests. I find this conclusion unhelpful. My topic is ordinary bullshit, not unusual bullshit. I want to understand the natural taste for bullshit that is embedded into every consumer, including me and you, and see how it might help to explain everyday economic phenomena such as the distribution of retail prices for beer or wine in a supermarket.

There is bullshit everywhere. Turn to your left or right: bullshit is everywhere. Science is not needed to point out that bullshit is everywhere. But for scientists who are trying to understand the prices of consumer goods, and how much consumers are willing to pay for value-added features of premium consumer goods, it may be helpful to consider bullshit. It is in this spirit that I hope that my work may be of some practical use to the field.

I first got interested in economics because I was puzzled by the prices of wines. In 2007, in order to research a consumer wine guidebook that I was writing, I ran a series of 17 blind tastings around the United States, serving more than 6,000 tastes of wine to 500 people. Consumers simply rated wine on a four-point scale: bad, okay, good, great. Wines ranged in price from \$1.50 per bottle to \$150 per bottle. When the label was concealed, my colleagues and I found a small but significant inverse correlation between price and ratings. If people are willing to pay more for packages whose contents they like less, then what are consumers really paying extra for, when they buy expensive wines? Is it the package, the words and numbers on the package, or something else?

These questions, turn, naturally drew me to the topic of bullshit. But what is bullshit? Philosopher Harry Frankfurt’s well-known definition emphasizes the difference between lying and bullshit: whereas the liar knows the objective truth and attempts to conceal the truth for his own purposes by stating its opposite, the bullshitter has no interest in whether

his statement is true, false, or neither. The truth or falsity of a statement is simply irrelevant to the bullshitter's purpose or his choice of words.

In the behavioral sciences there are many different partial windows onto bullshit. Psychologists might describe “embodiment effects,” “cognitive dissonance,” or “confirmation bias.” Health scientists might talk about “placebo responses,” or management scientists “mental shortcuts,” or behavioral economists “signaling effects,” or experimental game theorists “warm glow,” which is a variable representing the invisible quantity that compensates a philanthropist for his seemingly irrational and inexplicable generosity. What all of these effects share with Frankfurt's bullshit is that consumer value is generated, or consumer decisions influenced, by statements whose truth or falsity is irrelevant to the value or influence generated.

It is hardly surprising that consumers are willing to pay extra for bullshit—at least not any more surprising than that consumers are willing to enter insurance contracts with negative expected returns in exchange for “peace of mind.” There is also a peace of mind that comes with bullshit. There is a genuine and consistent human craving for bullshit that might be better off accepted and harnessed for good than ridiculed as “irrational behavior.”

In these essays, I use a variety of empirical methods to describe and measure some of the effects of bullshit on consumer demand. My methods for observing prices and eliciting preferences include consumer-choice experiments, blind tastings, hedonic price models, and what I call the “bullshit-on-bullshit” method.

Almost all of my data are on consumer prices and preferences in the U.S. beer, wine, and restaurant markets. I report a variety of empirical results from experiments that aim to cast light on different aspects of present-day consumer demand for bullshit, and I analyze these data in hopes that bullshit might help explain some seemingly “irrational” patterns of demand, such as the “bandwagon,” “snob,” and “Veblen” effects.

I am deeply grateful to my thesis committee, Jean-Marie Cardebat, Julian Alston, Jo Swinnen, and Steve Ziliak, whose thorough and insightful comments, both in writing and at the oral defense in Bordeaux in June 2019, helped me revise a bloated earlier draft into this much more concise one.

I would also like to thank the staff of LAREFI at the University of Bordeaux; Dan Sumner and the University of California Agricultural Issues Center, who supported me and this research throughout the process; the co-authors of the papers in this dissertation, especially Johan Almenberg, Orley Ashenfelter, John Bohannon, Craig Riddell, and Laurie Santos; Karl Storchmann, without whom I would not be an economist; Mom, Dad, Denton, and Tessa, my four most loyal supporters and shoppers in Bordeaux; Christiane, Rosie, Azai, Mahi, Joyce, Nonnie, and the many other colleagues and loved ones that commented on drafts and supported me throughout this process.

## Introduction

### *The case for an economics of bullshit*

The once and future king of bullshit, the work at the center of every current academic debate on bullshit, is a reflective essay by the philosopher Harry G. Frankfurt entitled “On bullshit,” first published in 1986, to little fanfare, in the *Raritan Review*. Almost two decades later, in 2005, Frankfurt released the same essay, repackaged into a mini-hardcover book called *On Bullshit* with line spacing so liberal that a few sentences could take up an entire page. Repurposed in this way, the essay went on to become the biggest-selling book in the history of Princeton University Press, with half a million copies in print. At 67 text pages and 7,800 words, it was the shortest #1 *New York Times* bestseller of all time (Leibowitz, 2017), and possibly also the smallest (4.2 in x 6.2 in., 4.8 oz., according to Amazon.com).

Bullshit, from the perspective of Frankfurt (1986), is a statement made without regard to its truth or falsity. Frankfurt’s bullshit may be true, false, or indeterminate, but the “truth-value” of the statement is unrelated to the bullshitter’s purpose and unrelated to the function or value of the bullshit. For Frankfurt, the bullshitter is different from the liar: whereas the liar recognizes the truth and takes steps to conceal the truth, the truth just does not play a role of any kind in the bullshitter’s decision of what to say. Frankfurt writes:

“Telling a lie is an act with sharp focus. It is designed to insert a particular falsehood at a specific point in a set or system of beliefs, in order to avoid the consequences of having that point occupied by the truth ... The liar is inescapably concerned with truth-values. In order to invent a lie at all, he must think he knows what is true. ... On the other hand ... [the bullshitter] is not constrained by the truths surrounding that point or intersecting it.” (Frankfurt, 1986, 2005).

The late philosopher Gerald Cohen’s 2002 paper “Deeper into Bullshit,” which Cohen later expanded into a 2012 book chapter entitled “Complete Bullshit,” raises the point that Frankfurt’s framework may be too narrow to include “a certain variety of nonsense...which is found in discourse that is by its nature *unclarifiable*, discourse, that is, that is not only obscure but which cannot be rendered unobscure.” Cohen (2002) suggests, with some later support from Eubanks and Schaeffer (2008), that much of the unclarifiable bullshit in the world emanates from academia, where unclarifiable bullshit “comes close to being celebrated for its very unclarity by some of its producers and consumers.” Cohen names this type of unclarifiable academic bullshit, and calls it “Cohen’s Bullshit.”

One of the classic pieces of evidence that the postmodern humanities may be mired deeper in bullshit than most communities is the physicist Alan Sokal's 1996 spoof article "Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity," which Sokal intentionally stuffed full of gibberish and "published as a non-spoof in the thereby self-condemning *Social Text*" (Cohen, 2002), a prominent cultural-studies journal. Sokal's hoax article is primarily a collage of the most incoherent quantum-physics bullshit spewed out by the major figures of postmodern cultural theory. For instance, Sokal (1996) quotes Gilles Deleuze and Félix Guattari on chaos theory, from their book *What is Philosophy?*:

"To slow down is to set a limit in chaos to which all speeds are subject, so that they form a variable determined as abscissa, at the same time as the limit forms a universal constant that cannot be gone beyond (for example, a maximum degree of contraction). The first functions are therefore the limit and the variable, and reference is a relationship between values of the variable or, more profoundly, the relationship of the variable, as abscissa of speeds, within the limit" (Deleuze and Guattari, 1994).

"And there's much more," wrote Sokal later, after he had revealed the hoax, rattling off some of the B-list French intellectuals he had cited in the paper: "Jacques Lacan and Luce Irigaray on differential topology, Jean-Francois Lyotard on cosmology, Michel Serres on nonlinear time—but let me not spoil the fun...look up the originals, and decide for yourself. You'll find that these passages are even worse in context than out of context."

Sokal's own interstitial commentary mostly just showers praise on the authors of the Cohen's Bullshit in all of the quotes. This authorial restraint is one of the most exquisite aspects of Sokal's expert handiwork, the author's consistent ability to write and write and write while saying nothing at all but still convincing the *Social Text* editors of the legitimacy of the work, to go on for almost 17,000 words without making a single substantive claim yet still get the paper accepted. Sokal has shown us that sometimes it is necessary, as a sort of experimental method, to create and submit some bullshit in order to expose a larger pattern of bullshit. Cohen calls Sokal's method "deliberate bullshit," but I would rather call it "bullshit on bullshit." I will employ this method in Chapter 6 of this dissertation, which investigates the *Wine Spectator* Award of Excellence program.

After the *Social Text* hoax was exposed, some readers commented that France was the implicit butt of Sokal's elaborate practical joke, given the high percentage of gibberish quotes in the article that came from French intellectuals. Cohen is more explicit about the point: he concludes his expanded 2012 essay "Complete Bullshit" by asking "the question of why so much of that kind of bullshit is produced in France." Cohen's three main explanations, in short, are (1) that the French intellectual and social elite are heavily concentrated in Paris, where they are subject to "intellectual authoritarianism"; (2) that

the French are “stylish,” and thus have a talent and taste for creative artistry in their verbal communications; and (3) that in part because “all French *lycée* students study philosophy,” there is a large French “lay audience for intellectual products” that is “interested in interestingness,” which is “quite different from being interested in truth. ... Theories go not because they are refuted but because they are passé.”

It is hard to imagine that anyone could take Cohen’s gratuitous rant against French academia very seriously, given the irony that its main premise, “so much of that kind of bullshit is produced in France,” is unclarifiable and untestable and pretty much your classic example of Cohen’s Bullshit, and in spite of the even greater amount of Cohen’s Bullshit that’s packed into the vague platitudes and anachronistic stereotypes behind Cohen’s various possible explanations for the supposed phenomenon. Perhaps Cohen relies on Voltaire’s *Candide* for historical support.

Regardless, today, given the recent proliferation in the United States of for-profit universities, college admissions scams, and vanity journals established by pharmaceutical companies; given the increasing diversity of the population and diminishing power of the old technocratic elite in France; and given the increasing proportion of academic papers that are published in the English language, it is hard to imagine that France could possibly be the world’s leading producer of academic bullshit, even if one were to include this dissertation in the calculation.

At a minimum, the U.S. now appears to be the global leader in the publication of deliberate bullshit, with more than a dozen gibberish papers published in recent years by researchers following in the footsteps of Sokal. One such project began when writer/mathematician James A. Lindsay and philosopher Peter Boghossian published a paper in *Cogent Social Sciences* entitled “The conceptual penis as a social construct” (Lindsay and Boyle, 2017). Following this success, Lindsay and Boghossian teamed up with Helen Pluckrose, editor of *Areo* magazine, and the trio got seven more deliberate bullshit papers accepted to journals in a variety of fields, including “Human reactions to rape culture and queer performativity at dog parks in Portland, Oregon” in *Gender, Place & Culture* (Wilson, 2018; Lindsay, 2018); in *Sex Roles*, a two-year “thematic analysis of table dialogue” to determine why men like to eat at Hooters (Schuessler, 2018; Lindsay, 2018); in the *Journal of Poetry Therapy*, “Moon meetings and the meaning of sisterhood: a poetic portrayal of lived feminist spirituality,” with some of the text written by a “teenage angst poetry generator” (Lindsay, 2018); and in a feminist social-work journal, “Our struggle is my struggle,” which, in the words of the *New York Times*, “simply scattered some up-to-date jargon into passages lifted from Hitler’s ‘Mein Kampf’” (Schuessler, 2018).

Frankfurt asks: “Why is there so much bullshit?...In public life...people are frequently impelled...to speak extensively about matters of which they are to some degree ignorant... Bullshit is unavoidable whenever circumstances require someone to talk without knowing

what he is talking about...Instances arise from the widespread conviction that it is the responsibility of a citizen in a democracy to have opinions about everything.” I would claim, further, that instances of a different kind of bullshit also arise from the widespread conviction that it is the responsibility of a consumer in a free market to be a connoisseur of everything.

To serve my empirical purpose of identifying bullshit in present-day consumer preferences, I err on the side of simplicity and start with a version of Frankfurt’s definition as my basis where it is the consumer, not the producer, who displays a lack of regard for truth-value. Analogously to Frankfurt, I define a “bullshit attribute” as any stated attribute of a good whose truth or falsity is physically unverifiable by the consumer, even after consuming the good. The statement may be true, false, or indeterminate, but the statement’s truth is irrelevant to the value derived by the consumer. To determine whether a statement is bullshit, Frankfurt looks to the intent of the producer of the statement. To determine whether a stated attribute is bullshit, I look to the intent of the consumer of the statement (and of the good).

One assumption usually made in the economics of consumer information, as in Akerlof (1980), is that the truth or falsity of information matters to the consumer. True information—for instance, a photograph of an item of furniture for sale online, or an auto mechanic’s report on the condition of a used car’s engine—is valuable when the truth describes some fact or prediction about a good that is capable of affecting the consumer’s experienced outcome. The furniture shopper armed with a photograph, like the used car shopper armed with a mechanic, can better evaluate the consumer’s own expected future experience of consuming the good and adjust willingness to pay according to these expectations.

In the modern era of experimental and behavioral economics (e.g. Kahneman and Tversky, 1977; V. Smith, 1982; Kőszegi and Rabin, 2007), a purely probabilistic statement can also be treated as information. But for the economic experiments to work, the probabilistic statement must still be about a material outcome (e.g., in a prospect theory experiment, a 30% likelihood of winning \$100). The truth-value of the probabilistic statement, therefore, is of material interest to the consumer, as it affects the subject’s chance of receiving \$100 in cash. I distinguish information from bullshit by whether or not the truth or falsity of the information materially affects the consumer’s outcome.

What kinds of stated properties of goods are *not* information because the truth or falsity of the statement does not affect the consumer’s outcome? Consider, as an example, the words “Anno 1366” printed on a Stella Artois beer label. Although this statement, if interpreted to mean that the company was founded in 1366, could be true or false, the truth or falsity of the statement does not affect the consumer’s experience. There is no



characteristic aroma or taste of beer made by a company founded in 1366, and even the world's most sophisticated chromatography equipment could not possibly verify the age of a brewery by analyzing its beer. The statement—regardless of its truth-value, if any—does not cast any new light on the consumer's future sensory experience, nor does it convey anything physical about the beer. The only way that the consumer can detect the attribute “Anno 1366” is to read the words themselves. Thus “Anno 1366” is a bullshit attribute of Stella beer. (If meant as a claim of age, it is also factually false: the Stella Artois brand launched in 1927. But even if Stella had launched in 1366, “Anno 1366” would still be a bullshit attribute because of the consumer's indifference to its truth-value.)

If a consumer is willing to pay extra for a bottle with this bullshit printed on the label, then this must be a willingness to pay for the words “Anno 1366,” not a willingness to pay for some characteristic that would be present if the words were true and absent if the words were false. The words themselves, not their truth, generate their own value. The words do not point toward an expected consumption experience: they *are* the consumption experience. This self-sufficiency, this closed ontological loop, is a basic characteristic of bullshit attributes.

Quandt (2007) writes:

“In some instances, there is an unhappy marriage between a subject that especially lends itself to bullshit and bullshit artists who are impelled to comment on it. I fear that wine is one of those instances where this unholy union is in effect. Lots of commentators describe wines; either because they want to sell wines (e.g., wine stores) or because they are professional wine writers whose business is to evaluate wines. Of course, neither group can do its job properly without imbibing substantial quantities of wine, which may perhaps explain in part the purple prose that flows from their pens. (Note that “purple” is good in wines but not so good in prose.) We, the wine-drinking public, are happy to read their evaluations, because we are largely ignorant of the quality of wines.”

For now I will leave aside Quandt's comment that “we are largely ignorant of the quality of wines.” I will confront this topic (and criticize Quandt's point here) in Chapter 4, “On quality.” The main point I take from Quandt here is that wine descriptions are particularly prone to bullshit because of their vested interests in the outcomes and their unavoidable tendency to drink wine while evaluating, which are points well taken.

At the heart of the discussion in Quandt (2007), though, are two points about wine critics: first, their numerical scores disagree substantially, suggesting that they are not measuring any objective quality of wine; and second, the adjectives they use to describe wine, such as “olive-tinged black currant,” do not actually convey any information. How is the consumer to understand the difference, asks Quandt, between:

“... scorched earth and spicy earth ... I have never eaten earth, and I have never been near scorched earth. ... Is new saddle leather different from any other new leather? ... I am not sure how zesty minerals differ from just minerals ... If I am given a choice between two wines that are described by identical attributes except for their textile quality, should I choose the one with silky tannins or velvety tannins? It is a problem to drive a person crazy” (Quandt, 2007).

If, as Quandt suggests, the value derived by the consumer from these attributes—scorched earth versus spicy earth, or zesty minerals versus minerals—does not depend on the truth or falsity (or, in this case, unverifiability) of the statement. So these adjectives are not information attributes: they are bullshit attributes.

A final note on use of the word “bullshit”: I do not use the word “bullshit” in some parts of the dissertation. This is for two reasons: first, chapters 1, 2, and 6 were published as journal articles before I had developed the ideas here. Although I believe that the content of chapters 1, 2, and 6 is relevant and complementary to the new material, I have left the published articles in their mostly-original forms rather than awkwardly trying to adapt their old text to my new framework. (In Chapter 6, “price signal effect” is roughly equivalent to bullshit.)

Second, chapters 3, 8, and 9 contain unpublished material that I will eventually submit to academic journals. The word “bullshit” is still considered unprintable by some editors. Therefore, in these chapters, I sometimes refer to bullshit attributes as “invisible attributes.” I also refer to “Veblen attributes,” defined in Chapter 3, a broader category that includes packaging (decorative) attributes and invisible (bullshit) attributes.

## Chapter 1: Do more expensive wines taste better?<sup>1</sup>

*Evidence from a large sample of blind tastings*

### Introduction

When symbolic content is an important part of consumption, the hedonic enjoyment of a good might become decoupled from its primary sensory characteristics. The symbolic content of a price tag has been emphasized in marketing research (e.g., Cialdini, 1998).

At the same time, when two goods with similar characteristics differ in price, a reasonable prior is that the more expensive good will, on average, be preferred to the cheaper one, if preferences are measured, for instance, by ratings on a hedonic pleasantness scale. People have been shown to expect a positive correlation between price and pleasantness (e.g., Rao and Monroe, 1989). Consistent with this expectation, a meta-analysis reports positive correlations between price and hedonic or so-called “quality” ratings for most, but not all, of 1,200 different consumer product markets, but also finds that the range of these correlations is very large (Tellis and Wernerfelt, 1987).<sup>2</sup>

For some goods, there is much heterogeneity in consumer tastes, making it harder to infer quality from revealed preferences. Nonetheless, a reasonable prior is that consumers on average will derive more enjoyment from the good with the higher price. Wine seems to be a good for which consumer tastes are highly heterogeneous (Amerine and Roessler, 1976; Lecocq and Visser, 2006). While individuals may frequently disagree over which wine they prefer, the above hypothesis suggests a positive correlation between the enjoyment of a wine and its price.

A number of studies have reported positive correlations between price and subjective appreciation of a wine for wine experts (e.g., Oczkowski, 1994; Landon and Smith, 1997; Benjamin and Podolny, 1999; Schamel and Anderson, 2003; Lecocq and Visser, 2006). Non-experts, however, may not be particularly sensitive to some of the refinements that

---

<sup>1</sup> Revised version of an article published in *Journal of Wine Economics*, Volume 3, Number 1, Spring 2008, Pages 1–9. Co-authored with Johan Almenberg (Department of Economics, Stockholm School of Economics and the Research Institute of Industrial Economics, Stockholm School of Economics), Anna Dreber (Department of Economics, Stockholm School of Economics, and Program for Evolutionary Dynamics, Harvard University), John W. Emerson (Department of Statistics, Yale University), Alexis Herschkowitsch (Fearless Critic Media), and Jacob Katz (Department of Statistics, Yale University). We thank Jacopo Anselmi, Zoe Chance, Shane Frederick, Richard Friberg, Barry Goldstein, Erik Grönqvist, Daniel Horwitz, Roy Ip, Magnus Johannesson, Thomas Pfeiffer, Hal Stubbs, Sue Stubbs and an anonymous referee for helpful comments and suggestions. Johan Almenberg thanks the Ragnar and Torsten Söderberg Foundations for financial support, and Johan Almenberg and Anna Dreber thank the Jan Wallander and Tom Hedelius Foundation for financial support. The Program for Evolutionary Dynamics is sponsored by J. Epstein. Article reprinted with permission from the American Association of Wine Economists.

<sup>2</sup> Because the word “quality” is so inconsistently used, and so rarely defined, in the wine economics literature, I avoid using the term hereafter.

are held in high esteem by wine aficionados. Weil (2001, 2005) uses the following experimental setup: two bottles of wine are poured into four containers.

Tasters are then given three of the containers and asked to distinguish which one differs from the other two. A random guess has 1/3 chance of being correct. In Weil (2001), the two wines are identical apart from year, but one wine is from a “good” vintage, and the other from a “bad” vintage.<sup>3</sup> The tasters get it right 41% of the time—only marginally better than a random guess.<sup>4</sup>

In Weil (2005), the wines are a reserve bottling and a regular bottling, from the same producer and year. The fraction of correct answers is merely 40%. Moreover, Weil finds that even when tasters can distinguish between the vintages, they are about as likely to prefer the good one as the bad one. And among those who can distinguish the reserve bottling from the regular bottling, only half prefer the reserve. In both cases, the wines differ in price by an order of magnitude. These experiments highlight the discrepancy between experts and non-experts and the subjectivity of the wine experience.

Extrinsic factors, such as peer consumption and marketing actions, can also influence how a good is experienced. The price tag may in itself be such a factor. Recent research has shown that individuals appreciate the same wine more when they think that it is more expensive (Brochet, 2001; Plassmann et al., 2008). In other words, the price of a good affects the experienced utility derived from that good. Thus, to test the conjecture that the price of wine and the enjoyment of its intrinsic qualities are positively correlated, we need to examine the enjoyment of wine when individuals are unaware of the price.

Blind tastings offer the opportunity to isolate the experience of the wine itself from psychological confounds related to its price, presentation or published expert ratings. In this paper, we use a large sample of more than 6,000 U.S. blind tasting data points in a series of 17 wine tastings conducted by Robin Goldstein and Alexis Herschkowitsh in a variety of different U.S. cities. These tastings also formed the basis for the recommendations in the book *The Wine Trials* (Goldstein, 2008a). We investigate the relationship between price and subjective appreciation of wines when the price is unknown to the taster. Subjective appreciation is measured by overall ratings assigned to wines by individual participants.

Our main finding is that individuals who are unaware of the price do not, on average, derive more enjoyment from more expensive wine. In fact, unless they are experts, they enjoy more expensive wines slightly *less*. Our results are robust to the inclusion of individual fixed effects, and are not driven by outliers: when we omit the extremes of the

---

<sup>3</sup> Weil uses pairs for which the famous wine critic Robert Parker has rated one of the bottles “average” to “appalling” and the other bottle “excellent” to “the finest”.

<sup>4</sup> All of the significant difference is driven by the testers’ ability to distinguish between the good and bad vintages from Bordeaux Pomerol.

price distribution our results are even more pronounced.

Our paper is organized as follows. In section 7.2, we describe our data. In section 7.3, we present our econometric model and report our results. We also perform a robustness check. We conclude in section 7.4, where we discuss some implications of our results and suggest directions for future research.

### Data

The data set contains 6,175 observations from 17 blind tastings organized by Goldstein. The blind tastings took place in a variety of U.S. cities between April 2007 and February 2008. In total, 506 participants tasted wine flights composed from 523 different wines. The wines were presented in a double-blind manner, so that neither the person serving the wine nor the person tasting the wine knew the identity, price, or any other characteristics of the wine aside from its color. Each taster assigned an overall rating to every wine tasted prior to discussing the wines with the rest of the group, and was not permitted to change his or her answer after discussion. The rating was the response to the question “Overall, how do you find the wine?” and the available answers were “Bad”, “Okay”, “Good”, and “Great”. In the data, these alternatives are coded from 1 to 4, with 1 corresponding to “Bad” and 4 corresponding to “Great”.<sup>5</sup>

The price per bottle ranged from \$1.65 to \$150. The prices are average retail prices and were obtained from [www.wine-searcher.com](http://www.wine-searcher.com). The wines represent a broad variety of types (e.g., red, white, rosé, sparkling), country origins, and grapes.

The participants were unpaid volunteers from 21 to 88 years of age. Selection bias is a concern with any voluntary subject pool, and we have no reason to think that this is an exception. It is quite likely that the sample contains an over-representation of highly educated individuals, and an over-representation of individuals working in the food and wine industries. Nonetheless, the size of the sample and the general diversity of the tasters lead us to hope that inference will not be too restricted. For a more extensive description of the blind tastings, see Goldstein (2008).

### Results

Throughout the regression analysis, we use both an ordered probit estimator and a linear estimator (OLS). In both cases, we consistently use robust standard errors. The ordered probit estimator is particularly well suited to an ordinal dependent variable, but we find that OLS also performs well, and yields estimates that are easier to interpret. In any case,

---

<sup>5</sup> Tasters ticked one of four boxes. In about 3% of the sample, tasters ticked in between two boxes, suggesting a rating somewhere in between the two responses. For simplicity, we dropped these observations from the regression. Including them makes no difference to our qualitative results, and the changes to the estimates are negligible.

the two models generate highly consistent results. The dependent variable is the overall rating, measured on a scale from 1 to 4, with 4 being the highest rating. The main independent variable is the price variable, expressed as the natural logarithm of the average retail price per 750 mL of the wine in US dollars.<sup>6</sup>

In Model 1, we regress the overall rating assigned to wine  $i$ , by individual  $j$ , on the price of the wine. About 12% of participants had some wine training, such as a sommelier course. In Model 2, we allow for the possibility that these “experts” rate wines in a different manner. We include a dummy variable for being an expert, as well as an interaction term for price and the expert dummy. In a linear regression, this allows both the intercept and the slope coefficient to differ for experts and non-experts.

The linear specifications of Models 1 and 2 can be written as:

$$y_{ij} = \beta_0 + \beta_1 \ln(P_i) + \varepsilon_{ij} \quad (1)$$

and

$$y_{ij} = \beta_0 + \beta_1 \ln(P_i) + \beta_2 EXPERT_j + \beta_3 \ln(P_i) * EXPERT_j + \varepsilon_{ij} \quad (2)$$

where  $P_i$  is the price of wine  $i$ , and  $EXPERT_j$  is a dummy variable indicating if taster  $j$  has wine training. If individuals found that more expensive wines tasted better, the correlation between overall rating and price would be positive. In our sample, this is not the case: for both ordered probit estimates and OLS estimates, the coefficient on price is negative.

In Model 1, the OLS coefficient is about  $-0.04$ , implying that a 100% increase in price is associated with a 0.04 reduction in the overall rating (Table 7.1, column (c)). The negative effect for more expensive wines is statistically significant.

For non-experts, the relationship between price and overall rating is negative; for experts, however, this is not the case. Our estimates of Model 2 show that the correlation between price and overall rating is positive—or, at any rate, non-negative—for experts (Table 1.1, columns (b) and (d)).

The price coefficient for non-experts is still negative, of about the same size as before, and with greatly improved statistical significance. The coefficient on the  $\ln(\text{price}) * \text{expert}$  interaction term is highly statistically significant (ordered probit  $p = 0.017$ ; OLS  $p = 0.015$ ). For experts, the net coefficient on price is the sum of the two, i.e., about 0.11 for the ordered probit and 0.09 for OLS. This net coefficient is significantly different from zero, but only at the 10% level (ordered probit  $p = 0.099$ ; OLS  $p = 0.095$ ), despite the large sample size.

---

<sup>6</sup> If we didn’t do this, we would be expecting a one-dollar increase to have the same effect at the \$5 price level as at the \$50 price level. We get similar qualitative results using the raw dollar prices rather than logarithms but the statistical significance of the coefficients is lower (but still significant).

Table 1.1. Ordered probit and OLS results.

<b>Dependent Variable: Overall Rating</b>				
	<i>Ordered probit</i>		<i>OLS</i>	
	<i>(a)</i>	<i>(b)</i>	<i>(c)</i>	<i>(d)</i>
ln(Price)	-0.047 (0.039)**	-0.061 (0.013)**	-0.038 (0.038)**	-0.048 (0.012)**
ln(Price)*Expert		0.171 (0.017)**		0.138 (0.015)**
Expert		-0.558 (0.001)***		-0.448 (0.001)***
Constant			2.297 (0.000)***	2.337 (0.000)***
<i>N</i>	5986	5972	5986	5972
<i>R</i> <sup>2</sup> /pseudo- <i>R</i> <sup>2</sup>	0.000	0.002	0.001	0.005

Robust *p* -values in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

Note: regressions with Expert variable exclude 14 observations from subjects who did not complete the expertise questionnaire.

In sum, we find a non-negative relationship between price and overall rating for experts. Due to the poor statistical significance of the price coefficient for experts, it remains an open question whether this coefficient is in fact positive.

How large are these price effects? The coefficients are of a moderate magnitude, but non-negligible, given that wine prices cover a large range—both in our sample and in the general wine market. Suppose we have two wines, A and B, and Wine A costs ten times more than Wine B in dollar terms. In terms of a 100-point scale (such as that used by *Wine Spectator*), the OLS estimation of Model 2 predicts that non-experts will assign an overall rating that is four points lower for wine A, whereas experts will assign an overall rating that is seven points higher.

In addition, the coefficient on the expert dummy is negative, quite sizeable, and statistically significant (OLS expert dummy coefficient:  $\square 0.448$ ;  $p = 0.001$ ). In other words, the OLS estimation of Model 2 consists of two linear relationships, one with a higher intercept (2.337) but a negative slope ( $\square 0.048$ ), and one with a lower intercept (1.889) but a positive slope (0.090). The point where the two lines cross each other is the price level at which experts and non-experts are expected to assign the same rating. If we take the model literally, this point occurs at the price of \$25.70, i.e.,  $\ln(\text{price}) =$  approximately 3.25. At this price, the model predicts that both groups will assign a rating of about 2.2. Below this price, the model predicts that experts will assign lower ratings to

a wine than non-experts, and vice versa.

We also test a third model, including individual fixed effects. The linear specification of Model 3 can be written as:

$$y_{ij} = \beta_0 + \delta_j + \beta_1 \ln(P_i) + \beta_3 \ln(P_i) * EXPERT_j + \varepsilon_{ij} \quad (3)$$

where  $\delta_j$  is a dummy for each individual taster. Including individual fixed effects has very little effect on the qualitative results and the minor differences only serve to reinforce our earlier conclusions, as both the negative effect for non-experts and the positive effect for experts become slightly stronger. These results are presented in Table 1.2. For each of the four regressions in Table 1.2, a Wald test rejects that the fixed effects are jointly equal to zero by a wide margin ( $p$ -value  $< 0.001$ ), suggesting that this is a suitable addition to the model.

Table 1.2. Individual fixed effects results.

	<b>Individual Fixed Effects</b>			
	<i>Ordered probit</i>		<i>OLS</i>	
	<i>(a)</i>	<i>(b)</i>	<i>(c)</i>	<i>(d)</i>
ln(Price)	-0.070 (0.007)***	-0.089 (0.001)***	-0.050 (0.009)***	-0.064 (0.002)***
ln(Price)*Expert		0.209 (0.011)**		0.151 (0.013)**
Individual fixed effects	Yes	Yes	Yes	Yes
Constant			2.487 (0.000)***	2.183 (0.000)***
<i>N</i>	5986	5972	5986	5972
<i>R</i> <sup>2</sup> /pseudo- <i>R</i> <sup>2</sup>	0.080	0.081	0.181	0.182

Robust  $p$ -values in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

Note: regressions with Expert variable exclude 14 observations from subjects who did not complete the expertise questionnaire.

To make sure that our results are not driven by wines at the extreme ends of the price distribution, we also run our regressions on a reduced sample, omitting the top and bottom deciles of the price distribution. Given the broad range of prices in the sample, this is an appropriate precaution. The remaining wines range in price from \$6 to \$15.



Using the reduced sample, we estimate Model 2 (Table 7.3, columns (a) and (c)) and Model 3 (Table 7.3, columns (b) and (d)). This produces consistent and even more pronounced estimates. The coefficient on price is still negative, and in each case larger than when using the full sample. The statistical significance of the coefficients improves further, and the R-squared is higher.

**Table 1.3. Reduced Sample, With and Without Individual Fixed Effects.**

*Excludes the top and bottom deciles of the price distribution.*

	<b>Reduced Sample, With and Without Individual Fixed Effects<sup>6</sup></b>			
	<i>Ordered probit</i>		<i>OLS</i>	
	<i>(a)</i>	<i>(b)</i>	<i>(c)</i>	<i>(d)</i>
ln(Price)	-0.225 (0.001)***	-0.173 (0.019)**	-0.182 (0.001)***	-0.122 (0.025)**
ln(Price)*Expert	0.523 (0.002)***	0.515 (0.006)***	0.421 (0.002)***	0.364 (0.009)***
Expert	-1.301 (0.000)***		-1.044 (0.000)***	
Individual fixed effects	No	Yes	No	Yes
Constant			2.622 (0.000)***	1.910 (0.000)***
<i>N</i>	4817	4817	4817	4817
<i>R</i> <sup>2</sup> / <i>pseudo-R</i> <sup>2</sup>	0.003	0.094	0.007	0.206

Robust *p*-values in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%

In sum, we use the reduced sample to check the robustness of our results with regard to mid-range price levels. Based on the above, we conclude that our results are not only robust but in fact even more pronounced when omitting observations at the extremes of the price distribution.

### Discussion

The pleasure we get from consuming wine depends both on its intrinsic qualities such as taste and smell and external attributes such as price and presentation. One may argue that the former influences our subjective appreciation through a bottom-up process, where the sensory apparatus plays a key role, and that the latter works through a top-

down process, where beliefs and expectations about quality are important determinants.<sup>7</sup>

In this paper we have explored the bottom-up effects by looking at how participants in blind tastings rate wines. We find that, unless they are experts, individuals who are unaware of the price enjoy more expensive wines slightly less.

There is a large relevant literature related to the influence of extrinsic signals on taste experience. Lee et al. (2006) look at how knowledge of a beer's ingredients (normal beer with added vinegar) can affect subjective appreciation. They show that the timing of the information plays a substantial role. One group of tasters is told about the vinegar, tastes the beer, and assigns ratings. A second group is told about the vinegar after tasting the beer, but before the ratings are assigned. On average, individuals in the first group assign significantly lower ratings, suggesting that informing participants about the vinegar influences the experience in itself.

Using functional magnetic resonance imaging (fMRI), McClure et al. (2004) find that having the subject's favourite brand's name on a drink makes it taste better than if it is unlabeled. In another fMRI study, Plassmann et al. (2008) test whether marketing actions such as changes in the price can influence the experienced pleasantness of a product such as wine. Testers are given different wines that they are told differ in price. In reality, some of the wines are the same but simply presented with different prices. Price cues are found to correlate positively with experienced pleasantness, measured through both subjective reports and fMRI scans.

Marketing provides one channel through which consumers can be influenced to buy certain wines. But it is not the only one: wine critics/experts may also play a role in affecting wine prices and shaping consumer preferences. For example, Hadj Ali et al. (2007) find a positive effect of wine critic Robert Parker's ratings on the price of Bordeaux wine.

There is, however, some research expressing scepticism towards wine ratings and their use for the average wine drinker. According to Quandt (2007), many wine ratings do not actually convey any information, nor is there substantial agreement in ratings by experts. Consistent with this view, Weil (2007) investigates whether wine descriptions by experts convey information to wine consumers. This is tested by having testers match wine descriptions to wines. In a similar setup to Weil (2001, 2005), tasters are asked to distinguish the odd one out of three different glasses of wine. Only about 50% of the participants in Weil (2007) can distinguish the odd one out, and of those who manage to do it, only about half can correctly match a wine critic's description of the wine with the wine itself—which is no better than a random guess.

Our results indicate another reason for why the average wine drinker may not benefit

---

<sup>7</sup> This, in turn, might depend on ulterior motives such as status concerns. Wine as a status signal, and the prospect that expensive wine could function as a positional good, is discussed in Goldstein (2008), chapter 5.

from expert wine ratings: he or she simply doesn't like the same types of wines as experts. This is consistent with Weil (2001, 2005), who finds that even among the subset of tasters who can distinguish between good and bad vintages, or reserve or regular bottlings, they are as likely to prefer the "worse" one as to prefer the "better" one.

These findings raise an interesting question: is the difference between the ratings of experts and non-experts due to an acquired taste? Or is it due to an innate ability, which is correlated with self-selection into wine training?<sup>8</sup> Investigating this further would be a fruitful avenue for future research.

In sum, in a large sample of blind tastings, we find that the correlation between price and overall rating is small and negative. Unless they are experts, individuals on average enjoy more expensive wines slightly less. Our results suggest that non-expert wine consumers should not anticipate greater enjoyment of the intrinsic characteristics of a wine simply because it is expensive or is appreciated by experts.

---

<sup>8</sup> For a further discussion, see Chapter 4 of Goldstein (2008).

## Chapter 2: Can people distinguish pâté from dog food?<sup>9</sup>

*An experiment to test the sensory impact of food labeling*

Considering the similarity of the ingredients of canned dog food to those of human foods, dog food could be a suitable and inexpensive substitute for pâté or processed blended meat products such as Spam or liverwurst. However, the social stigma associated with the human consumption of pet food makes an unbiased comparison challenging.

This chapter reports the results of a double-blind test in which subjects were presented with five unlabeled blended meat products, four of which were intended for human consumption and one of which was Newman's Own dog food. To prevent visual bias, we pulsed all five samples in a food processor, garnished them with parsley and salt, and served them on crackers.

After ranking the samples on the basis of taste, subjects were challenged to identify which of the five was dog food.

Although 72% of subjects ranked the dog food as the worst of the five samples in terms of taste (Newell and MacFarlane multiple comparison,  $p < 0.05$ ), subjects were not better than random at correctly identifying the dog food.

### Introduction

What qualifies as food fit for human consumption is culturally defined. In some cultures, grasshopper, snake, dog, and horse are on the menu. Elsewhere, these healthy protein sources provoke disgust. There has also been a substantial flexibility of diet within cultures over time. Lobster, once considered fit only for fertilizer and slave food in 18th-century North America, is consumed there today as an expensive delicacy. Such cultural evolution is ongoing, with comestible goods constantly moving into or out of fashion. We investigated the potential of canned dog food for human consumption by assessing its palatability alone.

The diet of domestic dogs in most of the world consists of scraps, the by-products of human food preparation and consumption. Indeed, the close overlap between the diet of *Canis familiaris* and *Homo sapiens* may have been crucial for its evolution as a human companion species (Bradshaw 2006). Commercialized dog food is a recent phenomenon, becoming popular only in relatively wealthy industrialized nations since the mid-20th Century (Michel 2006). Nonetheless, it has grown rapidly into a \$45 billion industry (Euromonitor International 2007). Intense competition for market share has kept the

---

<sup>9</sup> Co-authored with John Bohannon and Alexis Herschkowitsch. Abridged version published in Chance as Bohannon et al. (2010). Appears here in the original unabridged form in which it appeared as AAWE Working Paper #36 (Bohannon, Goldstein, and Herschkowitsch 2009; Bohannon, Goldstein, and Herschkowitsch 2010).

price of dog food low relative to comestible goods for human consumption, even those derived from very similar meat industry by-products such as liverwurst and Spam.

In spite of its attractive price, commercial dog food is left virtually untouched by human consumption. One valid concern is the risk of food poisoning. The discovery in 2007 that several brands of commercial pet food were contaminated with melamine, an industrial fire retardant that can cause renal failure, caused widespread concern (Barboza 2007). However, partly as a result of this scandal, "organic" pet foods have gained significant market share. For example, Newman's Own® Organics Premium Pet Food is made exclusively from "human grade" agricultural products.<sup>10</sup>

But even if dog food is safe for human consumption, it must overcome considerable prejudice. Part of the barrier is the perception that dog food is unpalatable. The pet food industry has invested decades of research and development to make their products more appealing to the humans who must purchase and handle their products (Bradshaw 1991). Human volunteers have been used to compare the taste qualities of pet food formulae (Pickering 2008). The aim has been to reduce feelings of disgust while owners serve the food to their pets, rather than to make it more palatable for human consumption, but the result is the same. The diet and lifestyle of dogs in the industrialized world has converged with that of humans (Schaffer 2009). Could dog food be approaching acceptance as comestible good fit for humans?

Assessing the intrinsic palatability of dog food is a first step in answering this question. Controlling for bias is a challenge. Expectation has a large effect on the hedonic tone of food. There are many levels at which expectation can have its effects, and many mechanisms have been proposed (Deliza and MacFie 1991). The effects can be subtle and depend on when information is gained relative to consumption (Lee, Frederick, & Ariely 2006). Measuring the hedonic tone free of bias requires a double-blind trial (Goldstein *et al.* 2008).

We predicted that in a double-blind taste test, subjects would be unable to identify dog food among 5 samples of meat products with similar appearance and texture, thus allowing them to assess palatability independent of prejudice. We hypothesized that, if the dog food were ranked favorably relative to human comestible goods with similar ingredients, it should be considered fit for human consumption.

---

<sup>10</sup> See <http://www.newmansownorganics.com/pet/faqs>.

## Materials and Methods

The dog food tested was Canned Turkey & Chicken Formula for Puppies/Active Dogs (Newman's Own® Organics, Aptos, CA).<sup>11</sup> The four meat products used for comparison were duck liver mousse ("Mousse de Canard," Trois Petits Cochons, New York, NY), pork liver pâté ("Pâté de Campagne," Trois Petits Cochons, New York, NY), supermarket liverwurst (D'Agostino), and Spam (Hormel Foods Corporation, Austin, MN).<sup>12</sup> Each product was pulsed in a food processor to have the consistency of mousse.

Samples were allocated to serving bowls, labeled A through E, garnished with parsley to enhance presentation, and chilled in a refrigerator to 4°C. To allow one researcher (Bohannon) to perform a double-blind trial, the preparation was carried out by the coauthors (Goldstein and Herschkowitsch).

The experiment was carried out between 7:00 PM and 10:00 PM on 31 December 2008 in Brooklyn, New York<sup>13</sup>. After we fully disclosed the aim of the experiment—to evaluate the taste of dog food—18 subjects volunteered. Subjects were college-educated male and female adults between the ages of 20 and 40.

The five sample dishes, A through E, were presented to subjects with a bowl of crackers ("Table Water Crackers," manufactured by Carr's of Carlisle, U.K.). The identity of the samples, unknown to the researcher serving the samples, was as follows:

- A: Duck liver mousse.
- B: Spam.
- C: Dog food.
- D: Pork liver pâté.
- E: Liverwurst.

Subjects were asked to rank the "tastiness" of the samples relative to each other on scale of 1 (best) to 5 (worst). They were instructed to taste all of the spreads, in any order and as many times as necessary, in order to make a sound judgment. After the rankings were recorded on data sheets, subjects guessed which of the five samples they believed was the dog food.

---

<sup>11</sup> Ingredients: Organic Turkey, Water Sufficient for Processing, Chicken Liver, Organic Chicken, Ocean Whitefish, Organic Brown Rice, Carrots, Flaxseed, Oat Bran, Tricalcium Phosphate, Dried Kelp, Guar Gum, Carrageenan, Potassium Chloride, Sea Salt, Minerals (Iron Amino Acid Chelate, Zinc Amino Acid Chelate, Cobalt Amino Acid Chelate, Copper Amino Acid Chelate, Manganese Amino Acid Chelate, Sodium Selenite, Potassium Iodide), Vitamins (Vitamin E, A, B12, D3 Supplements, Thiamin Mononitrate, Biotin, Riboflavin Supplement). Nutrition: 8% protein, 7% fat, 29 Kcal/oz.

<sup>12</sup> Ingredients: Pork, Ham, Salt, Water, Sugar, Sodium Nitrite. Nutrition: 12% protein, 27% fat, 85 Kcal/oz.

<sup>13</sup> For more details of the event, see <http://www.sciencemag.org/cgi/content/full/323/5917/1006b>

## Results

The dog food (sample C) was ranked lowest of the five samples by 72% (13) of subjects. The duck liver mousse (sample A) was rated as the best by 55% (10) of subjects.

Between these extremes, the majority of subjects ranked Spam, pork liver pâté, and liverwurst in the range of 2nd to 4th place (see Tables 2.1 and 2.2). The rankings were analyzed using the multiple comparison procedure described by Christensen *et al.* (2006). The absolute differences between summed rankings were compared to the threshold values for  $p=0.05$  and  $p=0.01$  levels of significance (see Tables 2.1 and 2.2).

Table 2.1. Pâté vs. dog food: Raw data.

A duck liver pate  
 B spam  
 C dog food  
 D pork liver pate  
 E liverwurst

Subject	Ranking of samples					Which is dog food?
	A	B	C	D	E	
1	1	4	3	2	5	E
2	1	4	5	3	2	D
3	5	2	1	4	3	E
4	1	2	5	3	4	B
5	2	4	5	3	1	B
6	1	5	4	2	3	E
7	2	4	5	1	3	E
8	1	2	5	4	3	E
9	1	4	5	3	2	D
10	2	1	5	3	4	A
11	3	2	5	1	4	C
12	4	2	5	1	3	E
13	1	4	5	4	4	E
14	1	5	4	3	2	C
15	1	4	5	3	2	C
16	3	2	1	5	4	E
17	3	4	5	2	1	B
18	1	3	5	2	4	B
sums:	34	58	78	49	54	

Table 2.2. Pâté vs. dog food: Distribution of rankings.

Ranking (n)		A	B	C	D	E
1st		10	1	2	3	2
2nd		3	6	0	4	4
3rd		3	1	1	7	5
4th		1	8	2	3	6
5th		1	2	13	1	1

Ranking (%)		A	B	C	D	E
1st		0.56	0.06	0.11	0.17	0.11
2nd		0.17	0.33	0.00	0.22	0.22
3rd		0.17	0.06	0.06	0.39	0.28
4th		0.06	0.44	0.11	0.17	0.33
5th		0.06	0.11	0.72	0.06	0.06

The aggregate taste rankings of the dog food were statistically significant against a null hypothesis of random rankings (see Table 2.3).

Table 2.3. Pâté vs. dog food: multiple comparison test.

**Critical Values for p = 18 panelists**

P=0.05	CV = 4.3553p <sup>0.5012</sup>	18.5422
P=0.01	CV = 3.6582p <sup>0.5011</sup>	15.5699

<b>Significance Critical Difference</b>	P=0.01 <b>24</b>	P=0.05 <b>18</b>
---	---------------------	---------------------

**Differences between sample rankings**

	A	B	C	D	E
A	*	<b>24</b>	<b>44</b>	15	<b>20</b>
B	*	*	<b>20</b>	9	4
C	*	*	*	<b>29</b>	<b>24</b>
D	*	*	*	*	5
E	*	*	*	*	*

Significant differences in **bold**



The ranking difference between dog food and Spam was greater than the  $p < 0.05$  threshold, and the difference was greater than the  $p < 0.01$  threshold for all other samples. Subjects' preference for the duck liver mousse was also highly significant. The only sample that was not ranked significantly differently than the duck liver mousse (at the  $p < 0.05$  level) was the pork liver pâté.

Only 3 of 18 subjects correctly identified sample C as the dog food (see Table 2.4). A Chi-Squared test did not support the hypothesis that the distribution of guesses was significantly different from random ( $X^2 = 0.433$ ,  $p = 0.9797$ ).

Table 2.4. Pâté vs. dog food: identification test.

Sample	Which sample is dog food? identified as		expected if random
	dog food	frequency	
A	1	0.06	0.2
B	4	0.22	0.2
C	3	0.17	0.2
D	2	0.11	0.2
E	8	0.44	0.2

**Chi-Squared test**

$X^2 = 0.433$

2-tailed P value = 0.9797

\* Not significantly different from random distribution

Discussion

Subjects significantly disliked the taste of dog food compared to a range of comestible meat products with similar ingredients. Subjects were not better than random at identifying dog food among five unlabeled samples. These two results would seem to be paradoxical. Why did the 72% of subjects who ranked sample C as worst in terms of taste not guess that sample C was dog food?

One possibility is that slight differences in appearance and texture skewed the guesses. While the distribution of guesses failed a Chi-Squared test of statistical significance, 44% of subjects incorrectly chose liverwurst (sample E) as the dog food. As the texture of samples had been equalized with a food processor, it is possible that subjects were attempting to discern which sample was dog food based on taste, not texture. The explanation we find more compelling, however, is that subjects were

primed to expect dog food to taste better than it does. As we assured subjects that the experience would not be disgusting, they might have excluded the worst-tasting sample from their guesses.

Regardless of the cause of the distribution of guesses, we can be confident that the comparison of taste was free of prejudice. Even with the benefits of added salt, a smooth texture, and attractive presentation, canned dog food is unpalatable compared to a range of similar blended meat products.

We conclude that, although human beings do not enjoy eating dog food, they are also not able to distinguish its flavor profile from other meat-based products that are intended for human consumption.

### Chapter 3: Hide the label, hide the difference?<sup>14</sup>

*Two experiments on consumer preferences for bullshit in lager beer*

“Business concerns...are capitalized on their capacity to produce earnings, not on their capacity to produce goods. ... Certain customary lines of waste and obstruction...are unavoidable so long as industry is managed by businesslike methods, and for businesslike ends. ... [T]he eventual elimination of salesmanship and sales-cost would lighten the burden of workday production for the underlying population by some fifty per cent. ... [A]ll this wasteful traffic of salesmanship is using up their productive forces, with nothing better to show for it than an increased cost of living ... [I]ts logical outcome is a growing distrust of the business men and all their works and words. But the underlying population is still very credulous about anything that is said or done in the name of Business, and there need be no apprehension of a mutinous outbreak, just yet.”

—Thorstein Veblen, *Engineers and the Price System* (1921)

#### Introduction

In chapter 3, I observe bullshit attributes within a larger class of what I call “Veblen attributes” in the U.S. market for imported European pale lager beer by using the Stella Artois brand as an example.

U.S. consumers are willing to pay more per beer for Stella than for Heineken, its biggest competitor. In their bottled and canned formats, Stella and Heineken have many easily verifiable similarities in their packages and basic verifiable information conveyed on their labels. For instance, both beers have 5.0% alcohol by volume. But do some of the other sensory attributes of the beer inside the two packages differ?

To answer this question, I use a blind-tasting method called a “triangle test,” and in an experiment with 138 subjects, I find that the sensory properties of Stella, Heineken, and Budvar (another European pale lager) are identical when the two beers are poured into glasses, without their packaging. I infer that the price difference between Stella and Heineken must be generated by a combination of verifiable packaging attributes (e.g. color, size, and shape of the bottles or cans) and unverifiable bullshit (e.g. “Anno 1366”). When the beers are sold in keg form, with essentially identical packaging, the Stella premium can be attributed wholly to packaging and bullshit, which in this chapter I collectively call the “Veblen attributes.”

Some bullshit attributes are also found outside the package, in the realms of marketing, social influences, and expert opinions. For instance, ad campaigns associate

---

<sup>14</sup> For their significant contributions to this experiment, I thank Seamus Campbell and Alexis Herschkowitsch, who helped run the tastings; the Green Dragon beer bar, for graciously volunteering their space in Portland for the experiment; and Johan Almenberg, who helped enormously with data analysis and interpretation. An earlier draft of this paper appeared as AAWE Working Paper #165.

Stella with the European aristocracy and Heineken with sex appeal, and Stella has a considerably higher Beer Advocate rating than Heineken. Given that beer ratings are meant to describe the sensory attributes of beers, and not of packaging, and that the sensory attributes of Stella and Heineken appear identical to consumers in a blind tasting, I consider Stella's higher ratings on Beer Advocate and other ratings sites to be bullshit attributes: they do not convey any truth about the consumer experience.

Next, starting with the premise that consumer heterogeneity in brand preferences for Stella, Heineken, and Budvar can only be explained by bullshit attributes, I ask two questions: First, do U.S. consumers have consistent and heterogeneous preferences for these three brands? Second, if so, what patterns can be observed in such preferences? In a multiple-choice survey experiment, conducted through Google Consumer Surveys, I split a group of about 3,000 U.S. consumers into two groups, asking each group a slightly different question: (V1) "Which beer would you be most likely to buy?" (V2) "Which beer would you prefer to drink?" The choices given are Stella, Heineken, and Budvar, plus the bestselling beer in America (Bud Light). I observe a variety of systematic differences between "buying" and "drinking" preferences, as well as demographic differences between consumers in their brand preferences including statistically significant gender effects. I account for the results in terms of differences between brands in their Veblen and bullshit attributes, and I supplement the discussion with examples from the advertising campaigns from Stella Artois and Heineken.

### **Experiment 1: Triangle test for sensory differentiation**

This chapter describes the results from a controlled field experiment conducted with 138 beer drinkers at a beer bar in Portland, Oregon. The method I use is called a "triangle test," pioneered by Amerine et al. (1965). The method is meant to measure whether two samples have identical or noticeably different sensory properties from the point of view of the human sensory system.

The triangle test I conduct is meant to measure whether three competing brands of lager beer differ noticeably in their sensory properties. Subjects are presented with three tasting glasses of draft beer, marked with the numbers 1, 2, and 3. Two of the glasses contain identical beers from the same keg (the "twins"). The third glass contains a different brand of beer from a different keg (the "singleton"). The subject's task is simply to identify which of the three glasses is the "singleton."

My experiment tests subjects' ability to discriminate the sensory differences between three European lager brands from three different countries: Heineken, from the Netherlands; Stella Artois, from Belgium; and Budvar, from the Czech Republic

(Czechia) and marketed in the U.S. as “Czechvar,” due to a marketing dispute with AB InBev (Swinnen, 2017).<sup>15</sup>

I test these three beers against each other, two at a time, in three different triangle tests (Heineken vs. Stella, Heineken vs. Budvar, and Stella vs. Budvar). Overall, I find that consumers do not perform better than chance at this task, suggesting that the sensory attributes of these three competing products, which I call the “useful attributes,” are functionally identical from the perspective of the U.S. beer consumer. European mass-market lager beer may thus be an example of a market in which competitive differentiation and brand loyalty is driven largely by Veblen attributes.

### Experiment 1: Motivation

In competitive consumer goods markets, consumers often display loyalty to specific brands. Thorstein Veblen, in *Theory of the Leisure Class* (1899), suggests that in a leisure economy where basic consumer needs are satisfied, firms seeking sustained profitability may increasingly compete on what Veblen calls “Salesmanship (includes, e.g., needless multiplication of merchants and shops, wholesale and retail, newspaper advertising and bill-boards, sales-exhibits, sales-agents, fancy packages and labels, adulteration, multiplication of brands and proprietary articles).”

Veblen’s “Salesmanship” attributes, which he calls “superfluities and spurious goods,” are the packaging and invisible attributes generated by the “business” activities of firms (perhaps roughly equivalent to today’s imperfectly overlapping non-cost-of-goods-sold categories such as “overhead,” “SG&A,” “sales and marketing,” “cost of sales,” and so on), as opposed to the “industrial” activity of manufacturing the visible attributes that satisfy the basic consumer need: “Production (and sales-cost) of superfluities and spurious goods ... are incurred solely for purposes of business, not for purposes of industry; they are incurred for the sake of private gain, not for the sake of productive work” (Veblen, 1899). I call such “superfluities” the “Veblen attributes” of a good, and I distinguish Veblen attributes from what I call “useful attributes” of goods. I define “useful attributes” as attributes that are:

- (1) Material;
- (2) Visible;
- (3) Functional.

---

<sup>15</sup> Czechvar is the US brand of the Czech beer brand Budějovický Budvar, which is not permitted to use the “Budweiser” or “Budvar” brand name due to a non-compete agreement with AB InBev (owner of the U.S. Budweiser brand). It is sold in the rest of the world as “Budvar.” See Swinnen (2017) for the back story.

I will now give each of these three properties (materiality, visibility, and functionality) special definitions. First, by “material,” I mean a three-dimensional object, or portion of an object, that exists in Newtonian space and follows the laws of motion and gravitation. Second, by “visible,” I mean an attribute whose existence can be verified with the naked eye (or other senses) of an ordinary consumer in the market for the good, upon inspecting (and if necessary, opening and/or consuming) the good, with his or her own sensory apparatus, unaided by any special machinery, information, or social influence.<sup>16</sup> Third, by “functional,” I mean an attribute that has a material impact on the satisfaction of the good’s intended function.<sup>17</sup>

Each of these three properties—materiality, visibility, and functionality—are necessary conditions for an attribute to be what I call “useful,” and collectively, the three properties are sufficient conditions for an attribute to be “useful.” With respect to the set of all attributes of a good, I define useful attributes and Veblen attributes as mutually exclusive and collectively exhaustive complements: that is, all attributes that are not useful are Veblen, and all attributes that are not Veblen are useful.

In the example of a stapler, the attributes that determine its ability to staple successfully, to handle a certain number of staples, to staple a certain number of pages, etc., would be “useful attributes”; whereas the color of the stapler, the brand name printed on the top of it, and the information printed on its box would be “Veblen attributes.”

What are the “useful attributes” of beer? By properties 1 and 2 (materiality and visibility), the useful attributes would be limited to the sensory attributes that could be detected by consumers. By property 3 (functionality), the useful attributes would be further limited to the attributes of the liquid that is drunk by the consumer, plus the attributes of the bottle or keg that make it functional and transportable as a beer container.

The “Veblen attributes” of beer might include a variety of different non-beer and non-bottle attributes, including brand name and reputation, label design, market prices and quantities, peer reviews, expert ratings, the country or region of origin, size of the producer, manufacturing process and ingredients (if their sensory differences cannot be

---

<sup>16</sup> Here I use an expansive definition of “visible” that includes that which is perceivable by the entire human sensory apparatus, rather than just the visual system.

<sup>17</sup> I recognize that this definition is unsatisfying in that it leaves room for dispute over what is an “intended function.” I choose to err on the side of intuitive simplicity rather than imposing a series of non-intuitive conditions. Resolving this dispute for all goods (e.g. wall art) would take another dissertation. In this dissertation, I examine material goods whose material functions are relatively obvious: the function of beer and wine is to be drunk; the function of food is to be eaten; the function of a restaurant is to serve meals. I limit the scope of this entire dissertation, including all discussions of “useful” or “Veblen” attributes, to material goods, i.e. goods that contain at least some material attributes as defined above.

distinguished by consumers—if they could be distinguished, these would instead be “useful” attributes), long-term health warnings, and statements of the producer’s philosophical values.

Beer brands spend a comparatively large portion of their expenses on marketing, advertising, and brand-building. In 2018, AB InBev, the world’s largest beer producer (and owner of Stella Artois, one of the three brands tested in this experiment), reported revenue of US\$56.4 billion and spent \$18.1 billion (32%) on “SG&A” (sales, general, and administrative) costs, including marketing (AB InBev, 2019). By comparison, Tyson Foods, the world’s largest chicken producer, reported 2018 revenue of \$40.0 billion and spent only \$2.1 billion (5%) on SG&A.

Unlike the world’s biggest chicken manufacturer, the world’s biggest beer manufacturer spends one-third of its money on SG&A, which Veblen would roughly call “Salesmanship.” Tyson, on the other hand, does not typically buy Super Bowl ads.

It is not hard to see why Veblen attributes might play an important role in the market for European lager beer. First of all, the market is mature and highly competitive, and the product is well-defined, having been in commercial production in Europe is more than 500 years (Swinnen, 2017). At first glance, the useful attributes of pale lager beer—which is by all accounts the dominant style of beer around the world—appear to be relatively homogenous, from a sensory perspective, like the useful attributes of water. Most lager beers are golden-colored, fizzy, and contain about 5% alcohol. If lager beers are served on draft (i.e. from a keg, not a bottle or can), in a generic beer glass, where the beers are dissociated from their packaging, then differences between the visible attributes of competing brands of beer may even become difficult for everyday consumers to distinguish. Might they sometimes be altogether impossible for consumers to distinguish?

I assume that the beer consumer is simultaneously motivated by two different sets of product attributes at once: useful attributes and Veblen attributes. It could be, for instance, that useful attributes serve to satisfy the inner areas of the nervous and hormonal system”), such as the midbrain, brainstem, and thalamic areas, (still often referred to using 19th-century slurs like “lower brain” or “reptilian brain”), plus maybe the skin, musculoskeletal system, and other human organs that respond to tactile sensory experience. Veblen attributes, on the other hand, could serve to satisfy higher (outer) brain areas such as the information-processing areas of the cerebral cortex. Cortical (invisible) and subcortical (visible) attributes could then be viewed as being bundled together in the packages purchased by consumers.

Since pulling apart packages of product attributes is much easier than pulling apart packages of neuronal assemblies in human brains, I adopt a blind tasting approach to observing the relative influence of visible and invisible attributes in beer. Blind tasting experiments have the potential to isolate visible attributes from invisible attributes.

### Experiment 1: Background

Although the Pepsi Challenge, first introduced in 1975, might have brought blind tasting into the public eye more than any other event in history, blind beer tasting in a scientific context was pioneered by Ralph Allison and Kenneth Uhl in the early 1960s. Allison and Uhl's results, published in the *Journal of Marketing*, suggest that brand loyalty had little to do with differences in taste for the differences in sensory attributes between beer products. In their study, Allison and Uhl (1964) selected five different lager brands that, according to expert tasters, contained objective perceptual differences. Next, they sent six-packs of identical 12-ounce beer bottles to more than 300 regular beer drinkers.

Each participant received two bottles each of three different brands. If subjects had previously indicated that one of the five brands was their regular brand, then this brand was contained in the six-pack. Participants were then asked to consume the beers at home, and rate each of them.

The main trick was that some of the six-packs had beers with labels, while others were unlabeled. When the beers were labeled by brand, participants rated the beers differently, and as expected, they rated their favorites higher than other beers. When beers were unlabeled, however, participants showed virtually no preferences for certain beers over others. In the blind tasting condition, no beer was judged by its regular drinkers to be significantly better than the other samples. In fact, regular drinkers of two of the five beers scored other beers significantly higher than the brand that they stated was their favorite (Allison and Uhl, 1964).<sup>18</sup>

Some more recent blind tasting studies of competing beer brands, e.g. Valenzi and Eldridge (1973), have essentially replicated the results of Allison and Uhl (1964). Others, e.g. Jacob, Olson, and Haddock (1971) and Mauser and Uhl (1978), have replicated the results in part (e.g. with respect to clusters of brands producing similar styles of beer) while narrowing them to some extent (e.g. finding some consumer discriminability between style categories such as light beer vs. heavy beer, or American beer vs. Canadian beer). All of these experiments solicited extensive batteries of sensory evaluations (on hedonic rating scales) from subjects, e.g. strength, aftertaste, bitterness, and overall pleasantness.

In blind wine tastings, meanwhile, Weil (2001, 2005) finds that tasters perform only marginally better than random at distinguishing between wines from "good" and "bad" vintages made by the same producer, or between reserve bottlings and regular bottlings from the same producer and year. Weil observes that amongst those who can distinguish between reserve and regular bottlings, only half prefer the reserve, even though the wines differ in price by an order of magnitude. Consistent with Weil's results is Goldstein et al.

---

<sup>18</sup> That brands can influence tasting ratings has been shown on a number of products. For example, McClure et al. (2004) show, with the help of functional magnetic resonance imaging (fMRI), that having the subject's favorite soft drink brand's name on a drink makes it taste better than if it is unlabeled.



(2008), which appears in this dissertation as Chapter 9, shows that in a large sample of blind wine tastings, tasters on average actually prefer expensive wines slightly *less* than cheaper wines, in contradiction to the commonly held belief amongst wine consumers, especially on the high end, that expensive wines tend to have hedonically superior useful attributes.<sup>19</sup>

This experiment, on the other hand, avoids the use of hedonic ratings, sensory evaluations, and preference elicitations altogether. Instead, I ask a simpler question: are there any useful differences at all between three competing brands of European lager beer that are sold at three different price points?

My triangle test aims to see whether consumers can differentiate between the useful attributes of Budvar, Heineken, and Stella. The simple question posed is whether or not, from the point of view of beer enthusiasts in a typical consumption setting (in a glass at a beer bar), these three competing European beer brands have sensory differences that might help explain the brands' price differences, or whether competition between the brands is occurring only on the basis of Veblen attributes. By removing all Veblen attributes from the brands, I hope to test for the existence or absence of useful differences between the three.

The rest of this chapter is organized as follows: first I discuss the experimental methods and data in more detail; then I present my results; and finally I discuss the findings and suggest some implications.

### Experiment 1: Methods and design

One challenge that arises in the attempt to separate out beer attributes into visible and invisible attributes is that when beer is bottled or canned, packages are differentiated by visible attributes such as size, shape, color, weight, and graphic design. Often the beer is consumed directly from these packages, so the consumption experience is never decoupled from the packaging attributes.

Some European beers are visibly differentiated by their physical features, such as the swing-stopper of a Grolsch bottle, the faux-porcelain of a Delirium Tremens bottle, or the Champagne cork in a bottle of Chimay Grande Réserve. Even some relatively inexpensive mass-market beer brands have recently begun to differentiate on some visible package attributes, such as wider necks, tighter freshness seals, or easier-to-open cans, and some brands may eventually introduce self-chilling beer cans. Even in the absence of

---

<sup>19</sup> Several studies show that when tasters know the price they tend to prefer the more expensive wine (Brochet 2001, Plassmann et al. 2008). Almenberg and Dreber (2011) look at wine and price information with a setup similar to that of Lee et al. (2006), and find that price information matters for an expensive wine but not a cheap wine, but only for women, and only when the information is given before the actual tasting experience.

brand information, the attributes of such packages are easily discriminated by consumers and thus fall into my “visible” category.

When beer is served on draft, on the other hand—and thus competing brands, after being ordered, are presented to consumers in identical unmarked beer glasses—the sensory attributes of packages are equalized, thus removing most packaging attributes from the demand calculation. (There are some exceptions, as when the consumer watches the bartender pour from a particular tap, a visual brand experience, or when a specialty beers are served in branded glasses).<sup>20</sup>

Although Allison and Uhl were careful to standardize their bottles, so bottle shape and size would not be likely to have biased their results unless there were variable bottle-beer sensory interactions, my experiment eliminates the sensory impact of bottles entirely by serving draft beer in identical tasting-sized glasses to consumers in a beer bar, thus simulating the real-life draft-beer purchase and consumption experience.

I also take a different, simpler approach to eliciting preferences that dispenses with hedonic ratings in favor of a narrow focus on addressing the question of whether there is any sensory variability at all in the functionality of certain competing brands’ products within the draft lager market segment—i.e. whether beer consumers have any preferences at all for beer visibles.

Rather than testing whether consumers prefer their favorite brand of beer to another brand, or asking consumers to rate beers on an array of sensory characteristics, I simply test whether or not beer consumers can tell competing brands apart from each other when their associated invisibles are concealed. In order to do this, I apply the triangle test introduced by Amerine et al. (1965), and applied to several wine studies since (e.g. Solomon 1990; Weil 2001, 2005, 2007), to three mass-market European lager beers that are readily available in the U.S. marketplace.

Triangle tests are frequently used in food science and sensory science as a means of verifying minimum threshold differences perceivable by consumers. The basic procedure is that three glasses of beer marked “1,” “2,” and “3” are presented to subjects simultaneously. Two of the glasses contain the same product (the “twins”), and one glass contains a different product (the “singleton”). Subjects visually inspect and taste all three beers and are simply asked to indicate which one is the singleton. If subjects cannot correctly pick the singleton at a rate above chance (33%), then I will infer that they do not perceive a visible difference between the two beers.

---

<sup>20</sup> Belgian beer brands are the pioneers and leaders in differentiation by distinctive beer glasses at bars, with Kwak perhaps the paragon. This might be related to why Belgian beers, anecdotally, have been able to stake out higher (and more differentiated) price points on premium draft lists than have competing premium domestic craft beers, which are generally served in identical pint glasses in pubs, or brandy snifters in pubes. The intuition that greater price variation depends on greater package variability is consistent with my later finding in Chapter 7 that bottle prices are more variable than keg prices. (All kegs look alike.)

I used three beers in this experiment: Budvar (A), Heineken (B), and Stella Artois (C). These are shown in their packaged (bottled) and unpackaged (draft) form in Figure 3.1.

In my triangle test, three samples are poured from the tap and presented to subjects in three 4-ounce clear beer tasting glasses: two glasses of one brand (the twins), and one glass of a different brand (the singleton). Subjects are provided with simple forms on which they are asked to pick out the singleton.

Figure 3.1. Three competing brands of European lager beer, in bottle and draft forms.



The triangle test compares tasters' ability to discern different content against a random guess. With the design described above, a random guess has a one in three chance of being correct. Weil (2001, 2005, 2007) applies this test to different categories of wines. I apply the test to different brands within the same category of beer: pale European lager. I repeat the test three times, allowing me to pit each of the three lagers against each of the others in pairwise comparisons (A vs. B, B vs. C, A vs. C). In round 1, beer A is poured into two glasses and beer B into one glass. In round 2, beer B is in two glasses and beer C is in one glass. In round 3, beer C is in two glasses and beer A in one glass. As such, each beer appears once as a twin and once as a singleton, and each beer is pitted against each other beer once.

I was allowed by the bar to install one keg in each of three of their tap lines. With the help of my colleagues Seamus Campbell and Alexis Hersckowitsch (both of whom also helped to run the experiment), I purchased the three kegs at a local liquor store and arranged to have them delivered to the Green Dragon bar, where they were temporarily installed into the bar's tap lines.

I surveyed 138 individuals, all aged 21 and over. At the request of the Green Dragon managers, in order not to cannibalize sales, each taster only participated in one triangle test. For a more elaborate description of the beer tasting setup, see Campbell and Goldstein (2010).

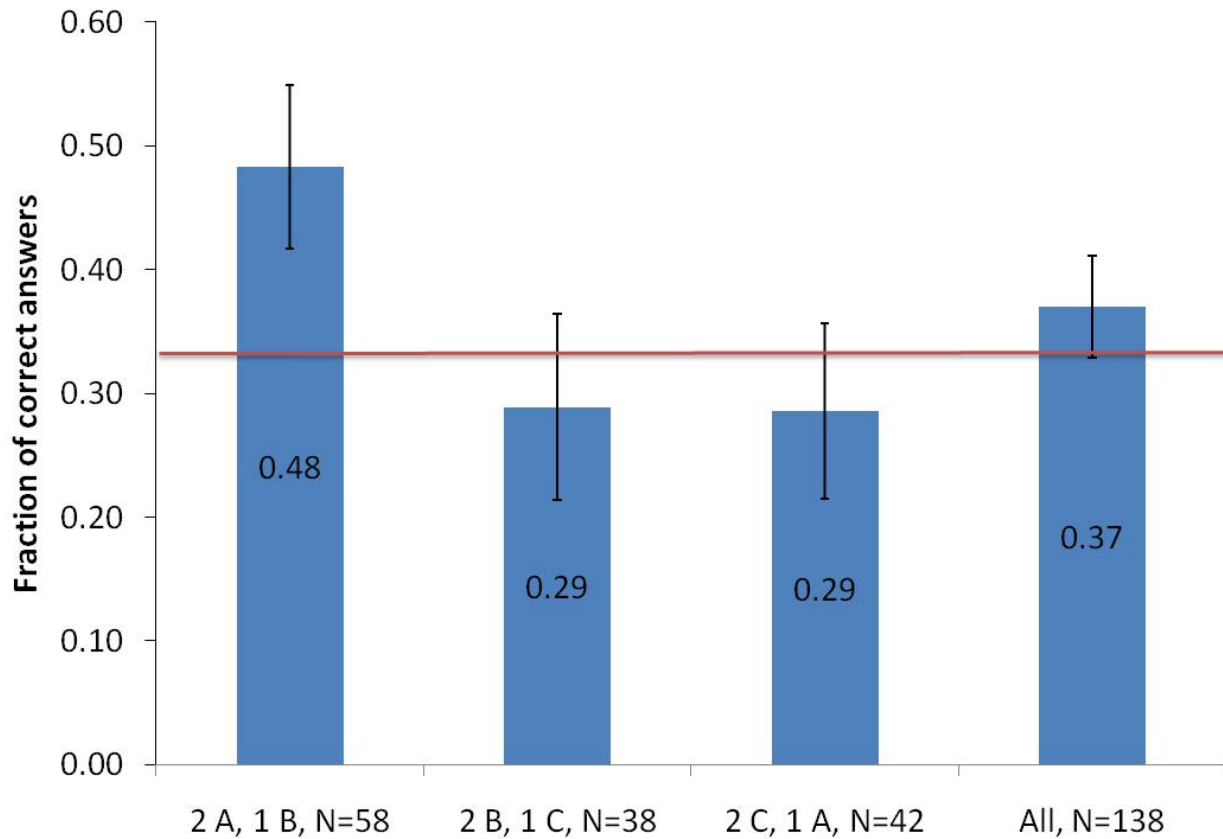
### Experiment 1: Results

I find that adult beer consumers are by and large unable to distinguish between the visible (i.e. sensory) attributes of European lager beers in a triangle test. In two of three tastings, participants are no better than random (33%) at telling the lagers apart. In the third tasting, subjects are slightly better than random, but still fewer than half are able to guess the singleton correctly. Across all three tastings together ( $n=138$ ), subjects do not perform statistically better than chance. I thus provide evidence that for lager beer—a product category for which consumers typically display strong brand loyalty and claim to distinguish functional differences in flavor between brands—the visible differences between different brands' products seem to be negligible.

I analyze each tasting separately as well as jointly by pooling the three tastings. Results are presented in a similar manner to that of Weil (2001, 2005): to test whether subjects perform better than chance, I compare the share of subjects who correctly identify the singleton with the share who would get it right if they had just guessed, i.e. one-third. In Figure 3.2, the share that would be correct if all guesses were completely random is indicated by the red line. Figure 3.2 shows that in two out of three rounds (rounds 2 and 3), tasters perform slightly worse than random chance at correctly identifying the singleton.

Figure 3.2. Experiment 1: Triangle test results, fraction of correct answers in each of the three settings separately and jointly.

*Error bars are standard error of the mean.*



In one of the three rounds (round 1), tasters perform slightly better than random chance (about 48%, compared to 33% for a random guess). But even in this round, the majority of tasters still guess wrong. When the three tastings are pooled, subjects' overall accuracy is not significantly different from the expected accuracy of random guesses—neither statistically nor in terms of the magnitude (about 37% versus 33% for chance).

This result is supported by a binomial test for whether the fraction of correct answers in each of the three settings separately and jointly differs significantly from a random guess. This is not the case in rounds 2 and 3 ( $p=0.73$  and  $p=0.62$ ) or the joint sample ( $p=0.32$ ), but it is the case in round 1 ( $p=0.017$ ).

### Experiment 1: Discussion

If this sample of U.S. beer consumers can be taken as reasonably representative of beer consumers in general, then my results suggest that when tasting blind, beer drinkers are unable to distinguish between the useful attributes of different European lager brands. Consumer loyalty to specific brands is thus unlikely to be grounded in sensory differences between the beer flavor, aroma, color, texture, etc. A more likely explanation for such brand loyalty is that product differentiation in the European beer market primarily reflects investment in Veblen attributes, e.g. marketing, sales promotions, advertisements, country of origin, and connected effects such as social desires and associations.

This inference would be consistent with, and perhaps help explain, the earlier observation that beer companies spend a relatively high percentage of their money on marketing and advertising. In this sense, beer brands might be “identity brands” (Holt 2004), i.e. consuming a particular brand of beer forms part of the identity of some consumers, and in this case, the reason for why a person identifies with a certain beer may be derived not from the direct sensory inputs provided by that beer upon consumption (a so-called “bottom-up” sensory process, in the terminology of psychology and neuroscience), but rather from the aspects of experience generated by packaging and invisibles such as brand reputation (a “top-down” cognitive process).

On the other hand, when consumers in *non-blind* tasting settings report sensory differences between brands in reviews and hedonic ratings, these differences may be explained by the sensory expectation effects generated by brand cues.<sup>21</sup> When beer is consumed in the usual setting, i.e. casually at a bar or in the home, brand cues are typically present during the consumption experience from start to finish. Whether from the label on a bottle, the logo on a can, or the tap medallion at a bar, consumers usually know what brand of beer they are drinking before they take the first sip, and top-down cues thus shape the sensory experience of consumption on a fundamental level.

This hypothesis is supported by a result in Lee et al. (2006), in a field experiment at a pub, where the addition of vinegar to beer affects the tasting experience negatively if the unusual additive is known to subjects before they taste the beer, but positively if it is not known. This interesting result suggests that cognitive information can affect the sensory experience of beer (as reported on a hedonic liking scale) when the information is introduced prior to drinking it, but not when it is introduced afterward, a sort of sensory expectation effect.

In sum, these findings add to the growing body of research on how consumers respond to sensory products in the absence of brand cues. My results suggest that U.S. consumers are largely unable to distinguish between competing brands of European lager

---

<sup>21</sup> See Ariely and Norton (2009) for more studies on “conceptual consumption” in general as well as the subclass “consuming expectancies.”

beer in blind tastings. Consumer brand loyalty in the beer market is thus unlikely to be grounded in the visible attributes of beer. European lager beer, at least in the U.S. market, could be an example of a product category in which Veblen attributes drive all consumer differentiation.

## **Experiment 2: Stated brand preferences for buying or drinking**

The results of the triangle test suggest that packages of beer from three competing European lager beers, from three different countries, served on draft—Heineken (from the Netherlands), Stella Artois (from Belgium), and Budvar (from Czechia)—have identical sensory beer attributes from the perspective of consumers at a beer bar in Portland, Oregon. However, the U.S. prices and market shares of these three European lager brands differ substantially. Any differences amongst the prices of, and consumer preferences for, these three brands may therefore be attributed to differences in Veblen attributes—and not to differences in useful beer attributes.

To what extent do U.S. lager beer consumers differ in their preferences for these three competing packages of apparently interchangeable liquids? Can some patterns be observed in their preferences? The first goal of this chapter is to verify that U.S. consumers do in fact differ in their preferences for the non-beer (packaging and invisible) attributes of Heineken, Stella, and Budvar in the absence of price information. The second goal is to see what differences emerge in preferences for non-beer attributes amongst gender, age, regional, and household income groups.

This chapter is organized as follows. First, I describe the experimental design, the survey form, and the method. Next, I summarize the data, report the main survey results, make some demographic comparisons, and observe some response-time effects. Finally, I provide some context by comparing the Heineken, Stella, Budvar, and Bud Light advertising strategies.

### Experiment 2: Methods and design

The main data set in this chapter comes from a nationwide survey of about 3,300 Internet users around the United States that I conducted on October 10–13, 2018, through the Google Consumer Surveys system. Google Consumer Surveys function as a paywall where the Internet user pays with his or her time, as in watching a TV commercial. Google then charges the surveyor and distributes a portion of the income to the Internet content provider that runs the paywall.

My one-question survey pops up on an Internet users' screen before visiting a premium website. Surveys are distributed on laptop and mobile browsers. This Google

Survey suffers from many of the same flaws as any online survey, e.g. the lack of a direct (financial or other) incentive for subjects to give an honest answer, and the general problem of eliciting stated preferences versus revealed preferences.

My survey asked about 3,300 U.S. consumers which beer they preferred of the same three European lager choices offered in the experiment in chapter 4: Heineken, Stella Artois, and Budvar.<sup>22</sup> In order to more realistically simulate a choice that might be made in a store, I also added a fourth non-European choice to the survey: Bud Light, the most popular beer brand in America. The multiple-choice survey form is shown in Figure 3.3.

3,307 subjects were split randomly into two treatments: about 55% were asked which beer they would “be most likely to buy” (the “buy” treatment, n=1,805), whereas 45% were asked which beer they “prefer to drink” (the “drink” treatment, n=1,502). The difference in sample size between the two groups was unintentional; there is some random variation in the number of subjects for each Google survey. (Each treatment was set up as a separate survey, but the two surveys ran simultaneously.)

In both treatments, the order in which the four brands (Heineken, Stella, Budvar, and Bud Light) appear was also randomized. No graphics or additional text were shown in the window, other than Google’s standard fine print. Given the strong possibility of thoughtless answers, my approach was to maximize signal to noise in responses. For instance, this meant using as few words as possible and including a “no opinion” option that was as easy to click as one of the other answers, thus reducing (if not fully eliminating) the incentive for users to click a random answer just to get past the survey window. Demographic (age and gender) and some (state-level) geographical information were collected. Response time (in milliseconds) was also measured.

The logic behind the treatment variable (“buy” vs. “drink”) is as follows. Assuming that every consumer is to some extent budget-constrained, asking consumers what they would be likely to “buy” *implies* that the choice is subject to the consumers’ own budget constraints. Changing the question from “which beer would you be most likely to buy” to “which beer would you prefer to drink” removes this implicit budget constraint.

Bud Light is the cheapest of the four beer options in my survey. In Sacramento, California, for instance, Bud Light is about 40% cheaper than Heineken, the next-cheapest survey beer. Following are the minimum prices per 12-ounce serving of beer sold by the case for the four survey beers at a Total Wine store in Sacramento, CA, in January 2019: Bud Light, \$0.61; Heineken, \$0.99, Stella, \$1.08; Budvar, \$1.79.<sup>23</sup>

---

<sup>22</sup> The experimental cue included both the names “Czechvar” (the U.S. name for the brand) and “Budvar” (the European name). Hereafter I refer to the brand only as “Budvar.” For more on the Czechvar/Budvar story, see Swinnen (2017) and footnote 12 of Chapter 4.

<sup>23</sup> To be clear, prices were not indicated on the survey. Some consumers are expected to be familiar with the approximate prices or relative prices of Budvar, Heineken, Stella, or Bud Light, and others are not.



Figure 3.3. Experiment 2: survey form.

*A survey question window pops up with the following text, with five check-boxes where the user can click or tap to choose.*

**“Buy” treatment (1,805 subjects)**

**Which beer would you be most likely to buy?**

- Bud Light
- Heineken
- Stella Artois
- Budvar (Czechvar)
- No opinion

*(Order of 4 beers rotated; “No opinion” always displayed last)*

**“Drink” treatment (1,502 subjects)**

**Which beer would you prefer to drink?**

- Bud Light
- Heineken
- Stella Artois
- Budvar (Czechvar)
- No opinion

*(Order of 4 beers rotated; “No opinion” always displayed last)*

One hypothesis that I formed before conducting the survey was that subjects would choose Bud Light (the cheapest beer) more often in the “buy” variant (with the implicit budget constraint in place). This effect, if observed, would also confirm that the “buy”/“drink” treatment is working as intended. As reported below, this effect is indeed present and statistically significant in the data.

Subjects were asked only one question (either the “buy” or “drink” question), and given no further information. In both treatments, subjects were shown the four choices as text strings (no pictures), and they are also offered the choice of “no preference” in order to remove consumers not in the beer market from the data set. The task was short and simple.

### Experiment 2: Data

Median response time was 9.0 seconds for “buy” and 8.3 seconds for “drink.” Frequency of response time for both variants appears to be similar and roughly normally distributed, as shown by the Epanechnikov kernel density plot in Figure 3.4.

Figure 3.4. Experiment 2: log response time, “buy” vs. “drink” treatments.

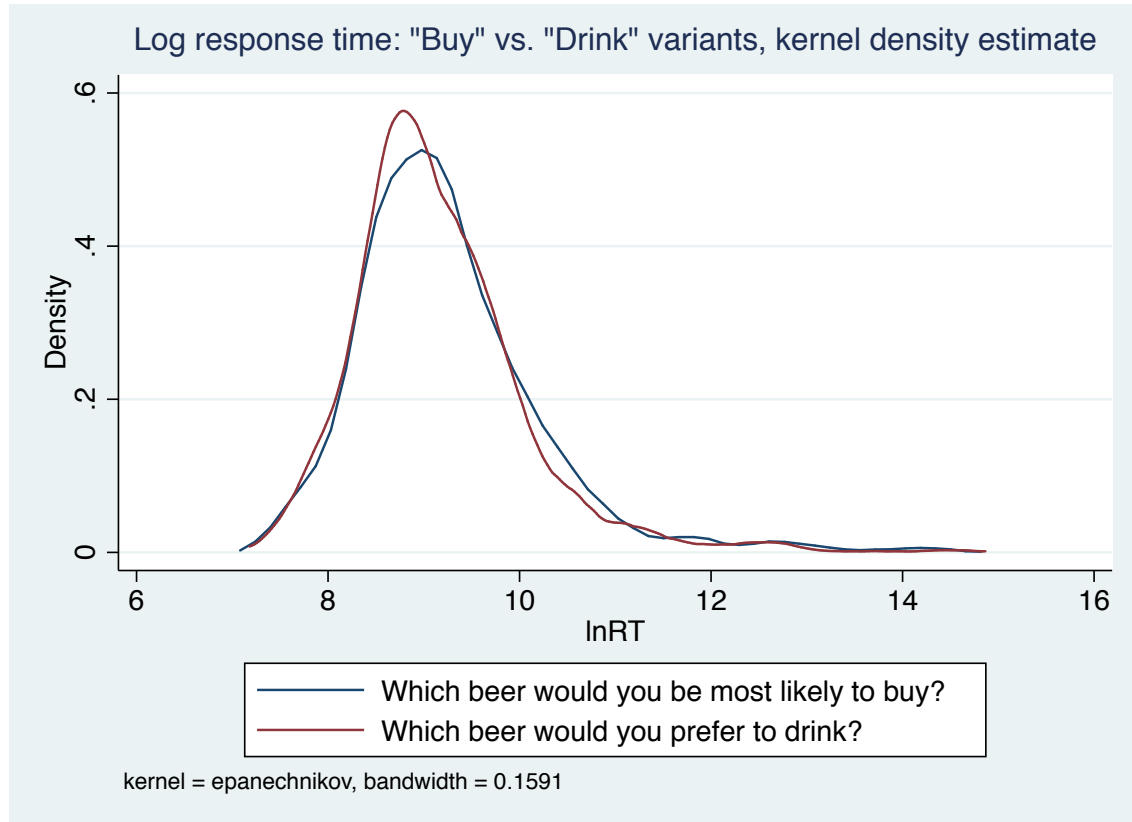


Table 3.5 summarizes the data collected in the survey. As discussed above, the main purpose of the survey is to elicit preferences amongst Stella, Heineken, and Budvar using Bud Light as an option for consumers who preferred the light American style (or American brand), thus (along with the “no opinion” option) giving me even cleaner sample of people with some taste for some European beer, and assuring that light beer drinkers are not forced to choose from amongst three beers they would be unlikely to buy or drink. Table 3.5a shows the number of consumers and breakdowns of the overall subject pool by gender and U.S. region (Northeast, South, Midwest, and West). Table 3.6a shows how many subjects were in each variant.

Table 3.5. Experiment 2: data overview.

Table 3.5a. Categorical variables.

Variable	Obs	Pct of sample
male	1608	49%
female	1232	37%
no gender	467	14%
Total n=	3307	100% <i>Including “buy” and “drink”</i>
Variable	Obs	Pct of sample
northeast	455	14%
south	1004	30%
midwest	1113	34%
west	709	21%
no region	26	0.1%
Total n=	3307	100% <i>Including “buy” and “drink”</i>

Table 3.5b. Treatment variable.

Two treatments

“Which beer would you be most likely to buy?” (V1)

“Which beer would you prefer to drink?” (V2)

*Percent of subjects choosing a beer, not including subjects who choose “no opinion”*

Variable	Obs	Pct of sample*
V1. “Buy”	864	55%
V2. “Drink”	673	45%
All	1537	100% <i>Not including “no opinion”</i>

Table 3.6b and 3.6c show male-female differences in whether or not subjects have any opinion (i.e. choosing one of the four beers vs. “no opinion”). In both the “buy” and “drink” variants, more males than females (54% of males vs. 41% of females in “buy,” 50% of males vs. 37% of females in “drink”) state a preference for one of the beer brands. Pearson chi-squared tests verify that these male-female differences are statistically significant (chi-squared = 24.19 for “buy” and chi-squared = 25.20 for “drink,”  $p < 0.001$  in each case). This finding supports the common-sense intuition that American males are more likely than American females to be beer consumers.

Table 3.6. Experiment 2: any beer opinion vs. no opinion overall and by gender.

3.6a. Opinion vs. no opinion.

Variable	Obs	Pct of sample
opinion	1537	46%
no opinion	1770	53%
Total n=	3307	100% <i>Including “buy” and “drink”</i>

3.6b. “Buy”: opinion vs. no opinion by gender.

“Buy” by gender	No opin	Opin	Total
Female	59.32%	40.68%	100.00%
Male	45.79%	54.21%	100.00%
Total	51.76%	48.24%	100.00%

Pearson chi2(1) = 24.1876 Pr = 0.000

3.6c. “Drink”: opinion vs. no opinion by gender.

“Drink” by gender	No opin	Opin	Total
Female	62.62%	37.38%	100.00%
Male	49.59%	50.41%	100.00%
Total	55.16%	44.84%	100.00%

Pearson chi2(1) = 25.1983 Pr = 0.000

## Experiment 2: Main results

I report choice results for the subject pool as a whole and for individual demographic groups, comparing patterns of “buy” and “drink” choices in each case. I also observe some response-time effects.

Overall results shown below are not weighted to be representative of all U.S. consumer demographics. Given that the choice is limited to four brands that do not collectively represent or approximate the U.S. beer market as a whole, the survey is not primarily intended to represent or estimate the true U.S. population means for beer preferences. Rather, I hope to explore some of the more interesting amongst the demographic breakdowns and some response-time trends. On top of all other limitations is the potential selection bias that comes from any group of Internet users, which overrepresents people who own computers or smartphones, who may have higher incomes than people who do not.

In the “buy” variant, I find that the first choice of both males and females is Bud Light, America’s bestselling beer. Overall, more subjects choose Bud Light than any of the three European lagers. In support of the idea that replacing “buy” with “drink” removes the budget constraint, about 15% fewer subjects choose Bud Light, the cheapest option, in the “drink” variant than in the “buy” variant. Budvar is chosen by the fewest subjects, and is chosen more often by males than females (by a four to one margin). Basic results in the “buy” vs. “drink” variants are shown in Table 3.7.

Table 3.7. Experiment 2, summary of results: “buy” vs. “drink.”

<u>“Buy” treatment (amongst subjects choosing Euro lagers)</u>			
Which beer would you be most likely to buy?			
	N=	Pct.	
○ 1. Stella Artois:	<b>130</b>	<b>47%</b>	
○ 2. Heineken:	<b>118</b>	<b>42%</b>	
○ 3. Budvar:	<b>30</b>	<b>11%</b>	
<u>“Drink” treatment (amongst subjects choosing Euro lagers)</u>			
Which beer would you prefer to drink?			
	N=	Pct.	Diff. vs. “Buy”
○ 1. Stella Artois:	<b>230</b>	<b>54%</b>	<b>+7%</b>
○ 2. Heineken:	<b>156</b>	<b>36%</b>	<b>-6%</b>
○ 3. Budvar:	<b>43</b>	<b>10%</b>	<b>-1%</b>

Table 3.8 reports the main results of choices by brand, combining “buy” and “drink” variants. First, Table 3.8a compares Bud Light responses with European lager responses (grouping Stella, Heineken, and Budvar together). Overall, about 42.9% of the 1,537 subjects choosing any beer choose Bud Light, 30.5% choose Stella, 19.3% choose Heineken, and 7.2% choose Budvar. A multinomial logit test, shown in Table 5.5b, rejects the null hypothesis ( $p < 0.001$ ) that subjects are choosing randomly amongst these four brands (i.e., where the probability of each brand choice is 25%). I infer that U.S. consumers in the market for European beer have well-defined and heterogeneous relative preferences with respect to the choice between these four beers.

Table 3.8. Experiment 2: percent of all subjects with opinions choosing each brand. Combines “buy” and “drink” treatments.

Table 3.8a. Brand preferences amongst subjects choosing any beer.

All beers	Freq.	Percent	Rank
Bud Light	660	42.94	1
Stella	469	30.51	2
Heineken	297	19.32	3
Budvar	111	7.22	4
All beers	1,537	100.00	

Table 3.8b. Multinomial logit test against null hypothesis of random choices.

Multinomial logistic regression	Number of obs	=	1537
	LR chi2(0)	=	0.00
	Prob > chi2	=	.
Log likelihood = -1894.5651	Pseudo R2	=	0.0000

All beers	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Bud_Light	(base outcome)					
Stella Artois	-.342	.060	-5.66	0.000	-.460 - .223	
Heineken	-.799	.070	-11.43	0.000	-.935 - .662	
Budvar	-1.783	.103	-17.38	0.000	-1.984 -1.582	

Table 3.9a shows the heterogeneity in beer preferences amongst subjects choosing European beers. About 53.5% of the 877 subjects choosing any European beer choose Stella, 33.9% choose Heineken, and 12.7% choose Budvar. A multinomial logit test, shown in Table 3.9b, rejects the null hypothesis ( $p < 0.001$ ) that subjects choosing any European beer are choosing randomly amongst these three brands (i.e., where the probability of each brand choice is 33%). I infer that U.S. consumers in the market for European beer have heterogeneous preferences for Stella, Heineken, and Budvar, in spite of my finding above that U.S. consumers cannot differentiate amongst the sensory beer attributes of these three brands.

Table 3.9. Experiment 2: percent of European beer drinkers choosing each brand. Combines “buy” and “drink” treatments.

Table 3.9a. Brand preferences amongst subjects choosing any European beer.

European	Freq.	Percent	Rank
Stella	469	53.48	1
Heineken	297	33.87	2
Budvar	111	12.66	3
Total	877	100.00	

Table 3.9b. Multinomial logit test against null hypothesis of random choices.

Multinomial logistic regression	Number of obs	=	877
	LR chi2(0)	=	0.00
	Prob > chi2	=	.
Log likelihood = -844.56764	Pseudo R2	=	0.0000

Choose Euro beer	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
Stella_Artois	(base outcome)				
Heineken	-.457	.074	-6.16	0.000	-.602 - .312
Budvar	-1.441	.106	-13.65	0.000	-1.648 -1.234

Next, I compare preferences in “buy” vs. “drink” treatments. First I observe the difference in preferences for the cheapest option between treatments to verify that imposing the “buy” condition acts like a mental budget constraint, as was intended.

If the “buy” condition functions as an implicit budget constraint, as intended, then more people should choose to “buy” than to “drink” Bud Light, which is the cheapest beer in the survey by a fair margin.<sup>24</sup> As shown in Tables 3.10 and 3.11, I observe this expected implicit budget constraint effect in the overall subject pool (i.e. including those with “no opinion”), with 21.6% all of subjects choosing to “buy” Bud Light but only 18.0% of all subjects choosing to “drink” it (Pearson chi-squared = 6.76,  $p = 0.009$  against the null hypothesis that relative preferences for Bud Light are equally likely in the “buy” and “drink” treatments). This result appears in Table 3.10. I also observe this expected implicit budget constraint effect in the pool of subjects choosing any beer (i.e. excluding those with “no opinion”), with 45.1% of subjects choosing to “buy” Bud Light but 40.1% of subjects choosing to “drink” it (Pearson chi-squared = 3.89,  $p = 0.049$  against the null hypothesis that preferences are equal in both treatments). This result appears in Table 3.11.

Table 3.10. Experiment 2, budget constraint effects: “Buy” vs. “Drink” preferences for Bud Light amongst all subjects (including “no opinion”)

*All numbers are percentages.*

Variant	Other	Bud Light	Total
V1. Buy	78.39	21.61	100.00
V2. Drink	82.02	17.98	100.00
Total	80.04	19.96	100.00

$$\text{Pearson chi2}(1) = 6.7646 \quad \text{Pr} = 0.009$$

---

<sup>24</sup> It is not necessary that all, or most, consumers know that Bud Light is the cheapest option in order to find implicit budget constraint effects generated by some subset of consumers that do.



Table 3.11. Experiment 2, budget constraint effects: “Buy” vs. “Drink” preferences for Bud Light amongst subjects choosing any beer

*All numbers are percentages.*

Variant	Euro beer	Bud Light	Total
V1. Buy	54.86	45.14	100.00
V2. Drink	59.88	40.12	100.00
Total	57.06	42.94	100.00

$$\text{Pearson chi2}(1) = 3.8911 \quad \text{Pr} = 0.049$$

Tables 3.12 and 3.13 report implicit budget constraint effects of choices across all brands by “buy” vs. “drink” condition. Table 3.12 reports percentages of all subjects (including those with no opinion), and Table 3.13 reports percentages of subjects choosing any beer. In both groups, I observe statistically significant differences in consumer choices amongst the four beers, against a null hypothesis that subjects answer the same way under the “buy” and “drink” conditions (Pearson chi-squared = 11.65,  $p = 0.020$  for all subjects; Pearson chi-squared = 8.63,  $p = 0.035$  for subjects choosing any beer).

Given that consumers make different choices under the “buy” and “drink” conditions, and that some of these differences are statistically significant, I infer that this budget-constraint treatment exerts a nonzero influence on preferences amongst these four beers.

Table 3.12. Experiment 2: “Buy” vs. “Drink” preferences for all beers, all subjects (including “no opinion”)

*All numbers are percentages.*

Variant	Bud Light	Budvar	Heineken	No opin.	Stella	Total
V1. Buy	21.61	3.77	9.25	52.13	13.24	100.00
V2. Drink	17.98	2.86	8.66	55.19	15.31	100.00
Total	19.96	3.36	8.98	53.52	14.18	100.00

$$\text{Pearson chi2}(4) = 11.6538 \quad \text{Pr} = 0.020$$

Table 3.13. Experiment 2: “Buy” vs. “Drink” preferences for all beers amongst subjects choosing any beer (excluding “no opinion”)

*All numbers are percentages.*

Variant	Bud Light	Budvar	Heineken	Stella	Total
V1. Buy	45.14	7.87	19.33	27.66	100.00
V2. Drink	40.12	6.39	19.32	34.18	100.00
Total	42.94	7.22	19.32	30.51	100.00

Pearson chi2(3) = 8.6290 Pr = 0.035

Table 3.14. Male-female differences in beer preferences (includes “buy” and “drink”).

Number and % choosing	Female	Male	Total
None	752	769	1,521
Pct	61.04 *	47.82	53.56
Bud Light	247	328	575
Pct	20.05	20.40	20.25
Heineken	82	175	257
Pct	6.66 **	10.88	9.05
Stella	139	262	401
Pct	11.28 **	16.29	14.12
Budvar	12	74	86
Pct	0.97 ***	4.60	3.03
Total	1,232	1,608	2,840
Pct	100.00	100.00	100.00

\* Likelihood is statistically significant differences in a Pearson chi-square test ( $p < 0.05$ ).

\*\* Likelihood is statistically significant differences in a Pearson chi-square test ( $p < 0.01$ ).

\*\*\* Likelihood is statistically significant differences in a Pearson chi-square test ( $p < 0.001$ ).

### Experiment 2: Gender effects

Preferences for Heineken vs. Stella appear to vary by gender. In the “buy” variant, more men than women choose Heineken, whereas more women than men choose Stella. In the “drink” variant, these patterns invert for both men and women: more men choose Stella, and more women choose Heineken.

Table 3.14 reports overall male-female differences for both variants combined, amongst all subjects, including subjects with no opinion. For the analysis of male-female differences, the number of observations is reduced by about 14% (to 2,840 from 3,309) because Google Surveys does not have gender information about some subjects. The most useful take-away from Table 3.14 is that including consumers without beer opinions, equal proportions of males and females (about one in five of each) choose Bud Light, but that considerably higher proportions of males than females choose each of the three higher-priced European beers.

Next, Table 3.15 summarizes buy-drink differences by gender and brand preference. Table 3.15a summarizes choices, for all consumers with opinions, between Bud Light vs. any of the three European beers. Table 3.15b summarizes choices from amongst the three European beers, out of all subjects choosing any European beer. Also indicated in Table 3.15 is the average number of seconds (“Secs”) for responses of each type.

Two main findings are easy to see in Table 3.15: (1) a high proportion of people choosing Bud Light are female; and (2) a high proportion of people choosing Budvar are male. First, as shown in Table 3.15a, I find that Bud Light, the most common and inexpensive beer, is preferred more often by females than by males. More females (55%) than males (42%) choose to “buy” Bud Light, and more females (47%) than males (36%) choose to “drink” Bud Light.

Table 3.16 shows that this male-female difference is statistically significant in a Pearson chi-squared test (chi-squared = 11.38,  $p = 0.001$  for “Buy Bud Light”; chi-squared = 7.53,  $p = 0.006$  for “Drink Bud Light”).

Second, as shown in Table 3.15b, I find that within the group of European beer drinkers, Budvar, the most uncommon and expensive beer, is preferred more often by males than females by a large margin. About three times as many males (16%) as females (5%) choose to “buy” Budvar, and about two and a half times as many males (13%) as females (5%) choose to “drink” Budvar. Table 3.17 shows that this male-female difference is statistically significant in a Pearson chi-squared test (chi-squared = 6.74,  $p = 0.009$  for “Buy Budvar”; chi-squared = 6.88,  $p = 0.009$  for “Drink Budvar”).

Table 3.15. Main results by gender: “Which beer would you prefer to “buy” or “drink”?”

*Nationwide Google survey of about 3,300 U.S. consumers conducted in August 2018.*

*"No opinion" option always appears last. Results do not include "no opinion" (55%).*

*Four possible responses: Bud Light, Heineken, Stella, or Budvar. Order of answers randomized.*

3.15a. Results for all consumers with opinions

	<u>Version 1: "Buy"</u>			<u>Version 2: "Drink"</u>			<u>Buy-Drink Differences</u>	
	Pct*	Avg Secs**	n=	Pct	Avg Secs	n=	Pct*	Avg Secs
<b>BUD LIGHT</b>								
Male	42%	11.4	159	36%	9.9	157	-6%	10.7
Female	55%	9.9	119	47%	9.8	113	-8%	9.8
<b>EUROPEAN***</b>								
Male	58%	10.9	218	64%	10.1	276	6%	10.5
Female	45%	11.2	96	53%	10.2	127	8%	10.7

3.15b. Results within consumers choosing one of the three European beers

	<u>Version 1: "Buy"</u>			<u>Version 2: "Drink"</u>			<u>Buy-Drink Differences</u>	
	Pct	Avg Secs	n=	Pct	Avg Secs	n=	Pct	Avg Secs
<b>HEINEKEN</b>								
Male	39%	11.3	85	30%	11.0	82	-9.3%	-0.3
Female	34%	10.8	33	38%	12.7	48	3.4%	1.9
<b>STELLA</b>								
Male	45%	11.4	98	57%	9.8	157	11.9%	-1.6
Female	60%	11.3	58	57%	9.7	73	-2.9%	-1.6
<b>BUDVAR</b>								
Male	16%	10.9	35	13%	10.1	37	-2.6%	-0.8
Female	5%	11.2	5	5%	10.2	6	-0.5%	-1.0

\* Percent choosing option. Figure in "Differences" calculated as (% buy - % drink).

\*\* Omits outlying response times (>33 secs). Response time difference = (RTbuy - RTdrink).

\*\*\* “European” = any of the other three options: Heineken, Stella, or Budvar.

Note: Numbers may not add due to rounding.

Table 3.16. Experiment 2: Chi-squared tests for gender diffs. in pct choosing Bud Light.  
*Results show percentages of all subjects with beer preferences.*

<b>“Buy”</b>	Other	Bud Light	Total
Female	44.17%	55.83%	100.00%
Male	57.88%	42.12%	100.00%
Total	52.79%	47.21%	100.00%

Pearson chi2(1) = 11.3845 Pr = 0.001

<b>“Drink”</b>	Other	Bud Light	Total
Female	52.92%	47.08%	100.00%
Male	63.74%	36.26%	100.00%
Total	59.88%	40.12%	100.00%

Pearson chi2(1) = 7.5314 Pr = 0.006

Table 3.17. Experiment 2: Chi-squared tests for gender diffs. in pct choosing Budvar.  
*Results show percentages of all subjects choosing any of the three European beers.*

<b>“Buy”</b>	Other	Budvar	Total
Female	94.34%	5.66%	100.00%
Male	84.26%	15.74%	100.00%
Total	87.39%	12.61%	100.00%

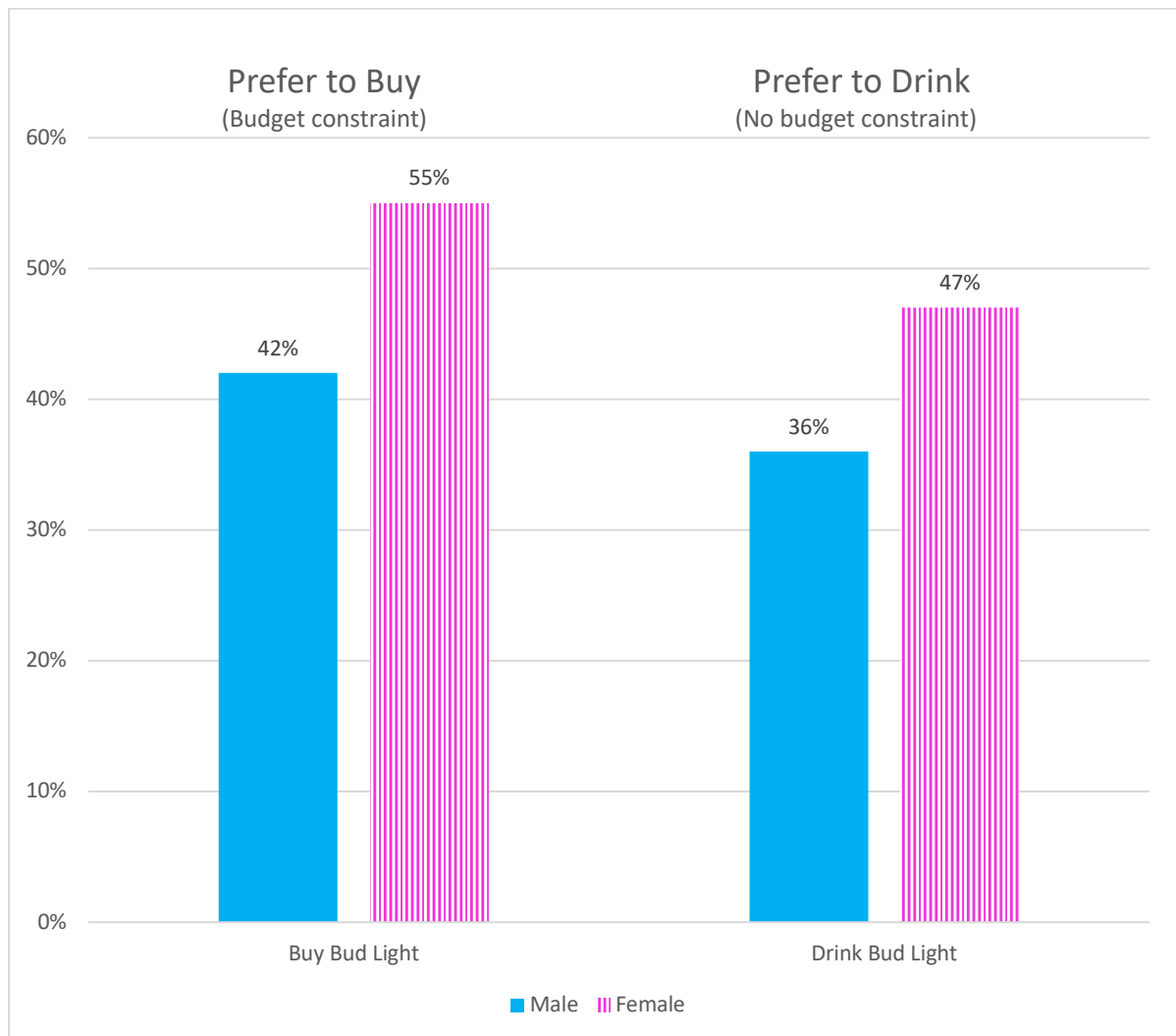
Pearson chi2(1) = 6.7412 Pr = 0.009

<b>“Drink”</b>	Other	Budvar	Total
Female	95.28%	4.72%	100.00%
Male	86.59%	13.41%	100.00%
Total	89.33%	10.67%	100.00%

Pearson chi2(1) = 6.8774 Pr = 0.009

Although more women than men prefer Bud Light in each variant, the budget constraint effect of the “Buy” condition is observed in roughly equivalent proportions in both males and females: both sexes make a shift of similar percent magnitude (about 15%) away from Bud Light (the cheapest option) when moving from the “Buy” variant to the “Drink” variant. The shifts (in overall percent) from Bud Light to European are from 42% to 36% for males and from 55% to 47% for females. These budget constraint effects are shown in the rightmost columns of Tables 3.15a and 3.15b, labeled as “Buy-Drink Differences,” and are illustrated visually by Figure 3.18.

Figure 3.18. Experiment 2: Budget constraint effects by gender: “Buy” Bud Light vs. “Drink” Bud Light.



### Experiment 2: Regional effects

Table 3.19 shows patterns of differences in responses amongst subjects from four different U.S. regions. Out of all subjects, including those responding with “no opinion,” subjects in the Northeast and West regions (coastal areas) show a lower likelihood of choosing Bud Light (16% in the Northeast and 15% in the West) than the South and Midwest (22% and 23%, respectively). Heineken and Stella are correspondingly more popular in the Northeast and West than in the South and Midwest.

Given the smaller subject pools being compared here, and the fact that some subjects were not coded with regional attributes, I did not run statistical tests on these effects or the age effects that I report next.

Table 3.19. Experiment 2: U.S. regional differences in beer brand preferences.  
*Includes “buy” and “drink” variants combined.*

beerchoice	Northeast	South	Midwest	West	Total
No opinion	233	530	591	395	1,749
Pct	51.21	52.79	53.10	55.71	53.31
Bud Light	72	223	261	104	660
Pct	15.82	<b>22.21</b>	<b>23.45</b>	14.67	20.12
Heineken	53	95	73	76	297
Pct	11.65	9.46	6.56	10.72	9.05
Stella	76	126	152	111	465
Pct	<b>16.70</b>	12.55	13.66	<b>15.66</b>	14.17
Budvar	21	30	36	23	110
Pct	4.62	2.99	3.23	3.24	3.35
Total	455	1,004	1,113	709	3,281
Pct	100.00	100.00	100.00	100.00	100.00

### Experiment 2: Age effects

Age effects are shown in Table 3.20. I merely observe a couple of broad trends. First, in spite of (or perhaps because of) its implicit claims of enhancing male sexual success, Heineken appears to appeal more to older generations of beer drinker. Table 5.15 pinpoints the modal Heineken drinker as a 55-to-64-year-old, and the second-most support for Heineken comes in the group of consumers 65 and older. This may be more of a cohort effect rather than an effect of aging: that is, in the United States, Heineken built up a market presence many years ago.

Table 3.20. Experiment 2: Age effects in beer brand preferences.

*Includes “buy” and “drink” variants combined.*

age	Bud Light	Budvar	Heineken	Stella	Total
18–24	<b>50.00</b>	5.56	16.67	27.78	100.00
25–34	42.80	<b>8.00</b>	17.60	31.60	100.00
35–44	42.60	5.42	18.05	<b>33.94</b>	100.00
45–54	41.39	7.33	20.15	31.14	100.00
55–64	44.96	5.04	<b>22.09</b>	27.91	100.00
65+	45.27	6.47	20.40	27.86	100.00
Unknown	50.00	33.33	16.67	0.00	100.00
Total	43.59	6.52	19.48	30.40	100.00

My 75-year-old father doesn’t drink beer—ever—but I still I asked him what beer he would choose if he was on his way to a party, in an arbitrary city away from home, where the host had asked him to pick up a six-pack.

“Heineken, I guess,” said my father, who also happens to be the least pretentious person I’ve ever met. Nonetheless, when I asked him why, he just shrugged his shoulders and said, questioningly: “I don’t know, really. Maybe because it’s respectable?”

My discussion of Stella Artois advertising below suggests that rejecting mere respectability, and shooting for something more—“Class,” or perhaps even the faux-crystal “Chalice” that is promoted by Stella in its marketing—might be what the Stella drinker is going for when he spends an extra 10% on his beer. This 10% surcharge need not be reflected in the price per bottle, which in many supermarkets is set equal to



Heineken's price. Rather, the price premium is implemented via a quantity penalty (which may be noticed by some consumers and not by others).

The Stella bottle, unlike the Heineken bottle, comes in 11.2-ounce (330ml) bottles—almost an ounce less beer per bottle than Heineken's or Bud Light's 355ml. Think about that next time you're back three deep at an overcrowded urban lounge, trying to get the bartender's attention, and thinking about ordering a Stella.

Table 3.19 pinpoints the modal Stella drinker as a 35-to-44-year-old, thus implicating my own 40ish demographic as the archetypal Stella drinkers, the assholes who want to pay a little bit more for a little bit less beer, poured a little bit more slowly, into a Chalice.

### Experiment 2: Response-time effects

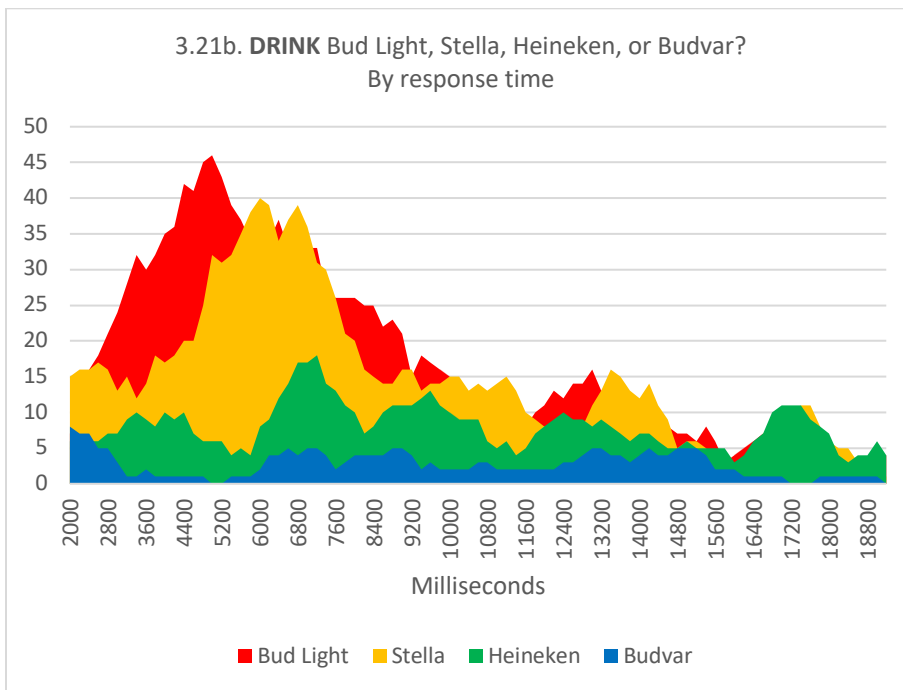
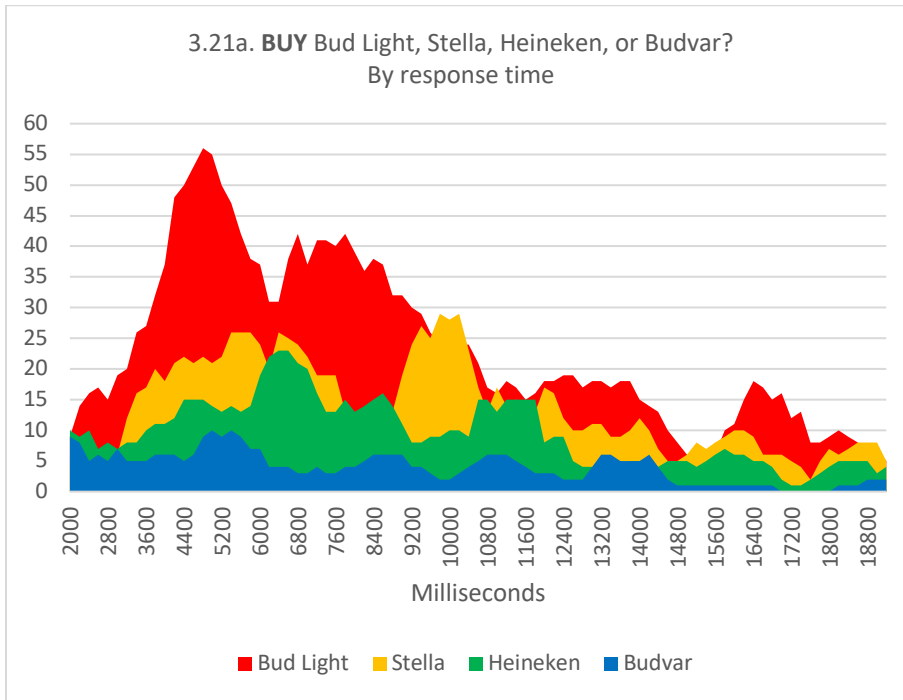
In the final section of my empirical results, I report some response-time effects, using Google's data on how many seconds each subject takes to enter a response to the survey.

Figures 3.21 and 3.22 plot choices for Bud Light vs. European beer by response time in milliseconds, in absolute number of answers per 300-millisecond interval bin (i.e. bin 1 = 2000-2299 ms, 2300-2599 ms, etc.).

In both "Buy" and "Drink" variants, Bud Light peaks at a modal response time of about five seconds. In both variants, European beer overtakes Bud Light at about six seconds, and peaks at a modal response time of about seven seconds. The biggest difference between the "Buy" and "Drink" variants is that in the "Drink" variant, European beers far outshine Bud Light from 6 to 8 seconds, with a steep climb to a much more dramatic seven-second peak. As is illustrated by the comparison between Figures 3.20a and 3.20b, much of the overall budget constraint effect (i.e. overall difference between "Buy" and "Drink" results) is accounted for by the increased percentage of subjects choosing European beers around seven seconds after the question is asked—with the mass concentrated about two full seconds after the modal Bud Light response time of five seconds.

One possible interpretation of these response-time differences is that a preference for drinking Bud Light, the cheapest and most popular option, may be more instinctive and less elaborative, akin to Kahneman (2011)'s proposed "System 1," whereas the choice of a premium option is roiled around for an extra couple of seconds by something like Kahneman's "System 2," the cognitive processing pathways, before coming out with the more honorable answer. From this perspective, Bud Light might be more of a five-second, System-1 decision, whereas the European beers might be more of a seven-second, System-2 decision. Figure 3.21, aside from proving that Google-colored mountains can spring out from anywhere, puts the previous two figures into relative context by showing choice frequencies (the Y-axis, in raw responses) over time (the X-axis, in bins of 200 milliseconds) plotted as non-additive areas (essentially a stylized histogram).

Figure 3.21. Experiment 2: Budget constraint (“Buy” vs. “Drink”) effects on preferences for Stella, Heineken, or Budvar, by response time.



Figures 3.21a and 3.21b show clear differences in the response-time characteristics of the subjects choosing each of the four answers, and they may jointly provide the best support in this chapter for the System-1/System-2 processing explanation.

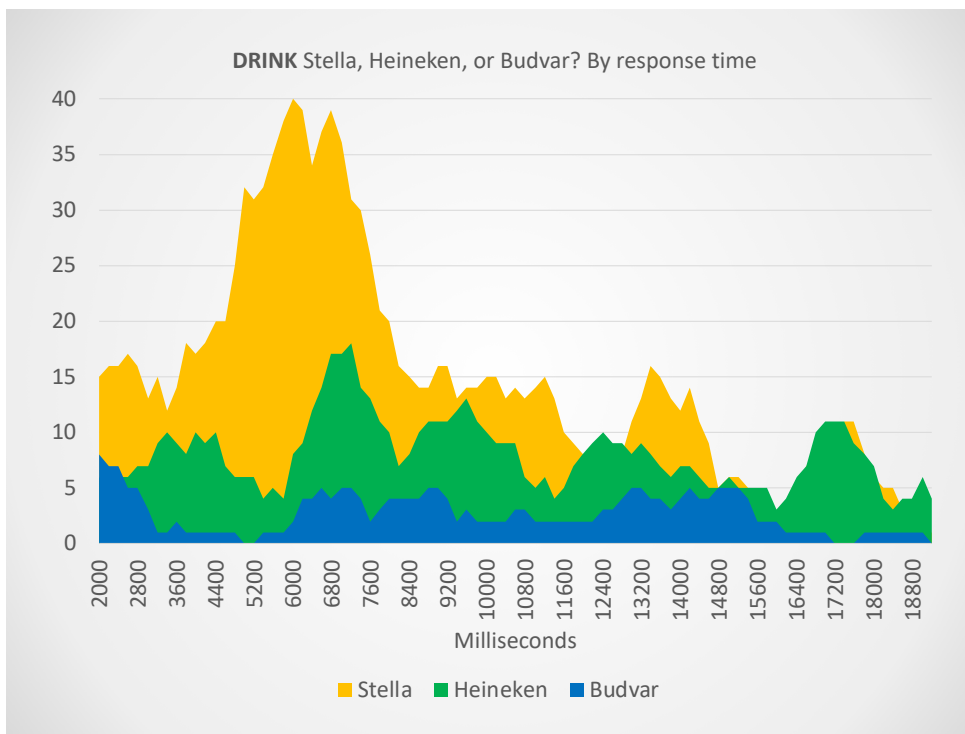
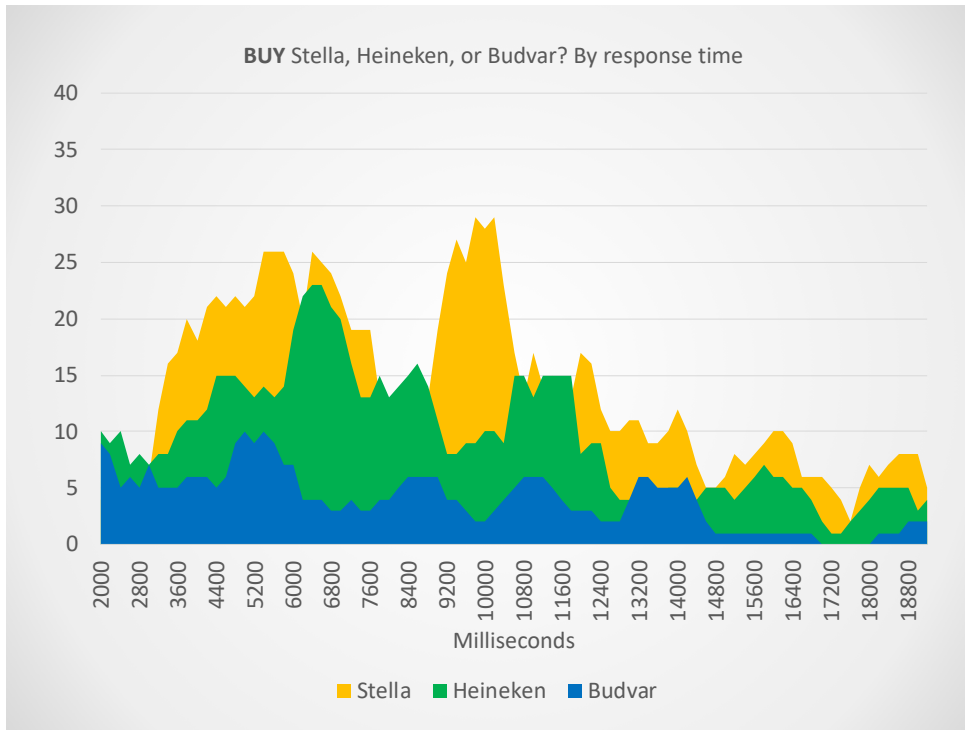
In “Buy” choices, the dominance of Bud Light of the first five seconds of decision-making is made clear, with Stella not peeking over Bud Light until about 10 seconds. the “Drink” division, golden Stella leaves its European competitors in the dust, with its steepest ascent beginning at five seconds, just as Bud Light begins to fall. Stella surpasses Bud Light at six seconds and sustains its dominance to the seventh second. The late resurgence of the Dutch brand in the geezer league after 16 seconds will not be enough to catch the Belgian hare.

What is most notable, perhaps, about the comparison between Figures 3.21a and 3.21b is what I would call the “Stella pop” peaking at seven seconds, a pop that is present in the “Drink” variant without the budget constraint, but absent in the “Buy” variant with the constraint. Thus the seven-second European-beer effect described above and illustrated in Figure 3.21b—the effect of giving the decision a couple of additional seconds of thought, the contemplative extra moment of the educated elite—might be driven mainly by Stella Artois, the beer of upper-class pretensions, and not by Heineken (beer of ubiquity) or Budvar (beer of unquity).

Another way of explaining these response-time effects, suggested by Julian Alston, might be that the response time reflects the time consumers take to follow a decision-tree heuristic (for instance, a decision tree that starts with the choice of “any beer” vs. “no opinion,” then (if “any beer”) moves onto a choice between “any European beer” vs. “Bud Light,” then finally moves onto the Budvar/Heineken/Stella choice. Although this is a compelling approach to understanding the response-time results, it would be difficult to impose any kind of empirical validity test on any competing interpretations of these response-time results.

Finally, Figure 3.22 shows the same effects as Figure 3.21, but results are limited to subjects who chose one of the three European lagers (Stella, Budvar, or Heineken). Amongst European lager drinkers, it appears that aspirations for Stella peak around six seconds in the “drink” condition, whereas in the “buy” condition, the same peak is not seen. Maybe six seconds is when reality kicks in for the “buy” subjects. They say that’s how long an orgasm lasts.

Figure 3.22. Experiment 2: Budget constraint (“Buy” vs. “Drink”) effects on preferences for the three European lager brands, by response time.



### **Discussion: how do Stella and Heineken differ in their Veblen attributes?**

The sensory similarities amongst Heineken, Stella, and Budvar (5% alcohol, pale golden color, carbonation level) were introduced in the first part of this chapter. The results of the triangle test in that chapter suggested that useful beer attributes cannot account for differences, if any, in consumer preferences amongst Budvar, Heineken, and Stella. Thus far this chapter has focused on identifying differences amongst the three brands in U.S. consumer preferences for their Veblen attributes.

Almost any real beer drinker faced with a choice amongst Stella, Heineken, and Budvar has some instinct on which direction to go. It might involve some thought. There may be a few instincts, even: a three-second instinct, a five-second instinct, and a seven-second instinct. For regular beer drinkers, people who go to beer bars on occasion and talk about their favorite beers—which is a damn lot of people—the instinct rarely seems to be indifference.

Given that media marketing is a major expenditure of beer brands, comparing the content of some examples of advertisements for the four beer brands considered in this chapter may provide some useful context for understanding variation in their target audiences. I thus conclude Chapter 3 by observing some differences between Stella’s and Heineken’s approaches to marketing through the lens of advertising.

#### Stella Artois advertising

Stella Artois, originally launched in 1927 as a Christmas beer (in spite of its claim, on the packaging, to date from 1366), has been a part of AB Inbev, the world’s largest beer conglomerate, since 2008, and it is sold in more than 100 countries. However, Stella’s brand message diverges from Heineken’s themes of recognizability, uniformity, globalization, and heterosexual male conquest.

Figures 3.23–3.27 give a view onto the differentiated consumer mind-space targeted by Stella Artois advertisements, which focus not on ubiquity or machismo but rather on social class. Stella is presented not as the beer for everyone, but rather as the beer for the upper class. This message is conveyed simply and directly in the three examples of U.S. Stella print ads that are shown in Figure 3.23, suggesting that Stella’s marketing angle is perhaps more straightforwardly Veblen-esque than any of the other brands in this experiment. Each of the Stella Artois ads shows one or more bottles of Stella, a short three-to-four-word slogan set in large Copperplate Gothic type, and little else. The slogans in Figure 3.23a, 3.23b, and 3.23c, respectively, are “Definition of Class,” “That Kind of Class,” and “What Class Are You In?”.

Figure 3.23. Stella Artois advertisements, United States, 2018.

Figure 3.23a.



Figure 3.23b.



Figure 3.23c.



Source: James Montefusco II, <https://www.behance.net/gallery/14917619/Mixed-Media-Mock-Stella-Artois>

Figure 5.24 shows an early-1980s Stella print ad with the headline “Come along, gentlemen. Haven’t you got mansions to go to?”, suggesting that the theme of the upper class is nothing new for the brand’s positioning. Even in this vintage campaign, the costliness of making Stella is associated with the upper-class lifestyle and the brand’s “reassuringly expensive” price. The point is made in detail with an extensive 14-paragraph text block:

“Who could possibly afford to drink Stella Artois and keep a stately home up to scratch in this day and age? For ours is, after all, a costly beer to brew. ... Not a drop sees the light of day until it has matured for six weeks or more. (Twice as long, in fact, as most rival brews.) ... Stella simply wouldn’t be Stella if we sidestepped any stage of the process. Especially the final separation. When you part with the ludicrous price of a pint. *Stella Artois. Reassuringly expensive.*”

Figure 3.24. Stella Artois advertisement (early 1980s).

**“Come along, gentlemen. Haven’t you got mansions to go to?”**

Gentlemen? Undoubtedly. But mansions? Surely the landlord is a little wide of the mark there. Who could possibly afford to drink Stella Artois and keep a stately home up to scratch in this day and age? For ours is, after all, a costly beer to brew.

We send our buyers far and wide, at great expense, to corner the best raw ingredients. No job lot could ever fit our bill. We use only Europe’s most fragrant hops. (For which we pay through the nose.) And the current asking price for the Continent’s finest barley?

(Frankly, that is something we’d rather you didn’t ask.) Then there is the inordinate amount of time we take to brew our beer. Not a drop sees the light of day until it has matured for six weeks or more. (Twice as long, in fact, as most rival brews.)

You could be forgiven for thinking there’s madness in our method. We beg to differ. Stella simply wouldn’t be Stella if we sidestepped any stage of the process. Especially the final separation. When you part with the ludicrous price of a pint. **Stella Artois. Reassuringly expensive.**

Source: [Best Copy Ads](#)



Round Creative (2018) claims that this “high-class” angle for Stella was first launched in the early 1980s with an “advertising genius” named Frank Lowe, who rescued the brand from an earlier pitch that had apparently been relatively ineffective:

“Before [Lowe] took charge of the account in 1981, the branding theme for Stella Artois was strength; with imagery including phonebooks torn in half and taglines like ‘Stella’s for the fellas who take their lager strong.’ The wrong message for the wrong market, it failed to help Stella take a decent market share. Lowe created the masterful ‘Reassuringly Expensive’ campaign which took the lager’s big negative and flipped it completely on its head. By suggesting that the higher price was a mark of quality, Lowe cleverly positioned Stella as the upmarket choice” (Round Creative, 2018).

Figures 3.25 and 3.26 illustrate the presentation and packaging of Stella’s notion of “class.” First there is the unique shape of the Stella glass, a faux-crystal goblet with a gold rim that holds less beer (11.2 oz) than a standard American pint glass (16 oz) or a standard American can or bottle (12 oz). Figure 3.25 shows a Stella billboard in Los Angeles bearing the motto “It’s a Chalice, Not a Glass.”

Figure 3.25. Stella Artois billboard, Los Angeles, CA.



Source: [Daily Billboard Blog](#)



Figure 3.26. Stella Artois advertisement.

## Stella Artois 9-Step Pour



### 1. THE PURIFICATION

Rinse a clean chalice. A chilled chalice sustains the head.



### 2. THE SACRIFICE

Sacrifice the first drops to ensure the freshest taste.



### 3. THE LIQUID ALCHEMY

Hold the chalice at a 45 degree angle as you pour. The angle creates the perfect combination of foam and liquid.



### 4. THE CROWN

Straighten the chalice to form the perfect head.



### 5. THE REMOVAL

Close the tap as you remove the glass.



### 6. THE SKIMMING

Use the skimmer to trim the head.



### 7. THE JUDGEMENT

Three centimetres of foam.



### 8. THE CLEANSING

Rinse the outside of the glass for a brilliant chalice and stunning presentation.



### 9. THE BESTOWAL

Present your Stella Artois on a fresh coaster.

Source: [theversatileagent.com](http://theversatileagent.com)

This billboard ad dovetails with the Stella bottle, which only contains 11.2 oz of beer. Figure 3.26 shows a ClearChannel billboard in Los Angeles showing the Stella glass with the words Figure 3.26 describes, in excessive detail for a light comedic effect, the “Stella Artois 9-Step Pour.” The theme tying together the nine images, headlines, and text blocks in Figure 3.26 is the extraordinary amount of labor and waste that goes into the pouring of a Stella into the Stella chalice. For instance, in “The Sacrifice,” some beer must first be wasted by pouring it outside the Stella Chalice “to ensure the freshest taste.” “The Skimming” is the removal of some foam from the top of the glass. The visual presentation of non-beer attributes is emphasized in the two final steps, “The Cleansing” (“Rinse the outside of the glass for a brilliant chalice and stunning presentation”) and “The Bestowal” (“Present your Stella Artois on a fresh coaster”).

In February 2019, Stella Artois aired a widely viewed Super Bowl TV commercial (Figure 3.27) featuring Sarah Jessica Parker as her *Sex and the City* character, Carrie Bradshaw, and Jeff Bridges in character as “The Dude” from *The Big Lebowski*.

Figure 3.27. Stella Artois 2019 Super Bowl TV commercial.



Source: [Eater](#)

The plot is that Carrie and the Dude each swap out their usual orders (a White Russian and a Cosmopolitan) for Stella. The motto is “Change Up The Usual.” Although the setting for Stella’s Super Bowl ad is unwaveringly upper-class—a restaurant with tuxedoed bartenders and a string quartet—Stella’s latest marketing push suggests that the brand is also aiming to satisfy aspirational leisure consumers’ desire for novelty, perhaps in contrast to the universal recognizability of its European lager archrival, Heineken.

### Heineken advertising

According to Heineken’s U.S. website landing page, “25 million Heineken’s [*sic*] are served daily across 192 countries. That’s a big family. Our Global Brew Masters ensure every Heineken beer has the special perfectly-balanced taste the world over. From Seville to Selangor, you’ll enjoy the same great taste of an ice-cold Heineken” (Heineken, 2019). As shown in Figure 3.28, globalization, uniformity, and recognizability have been long-running themes in Heineken’s marketing.

Figure 3.28. Heineken.com website (United States, 2019).

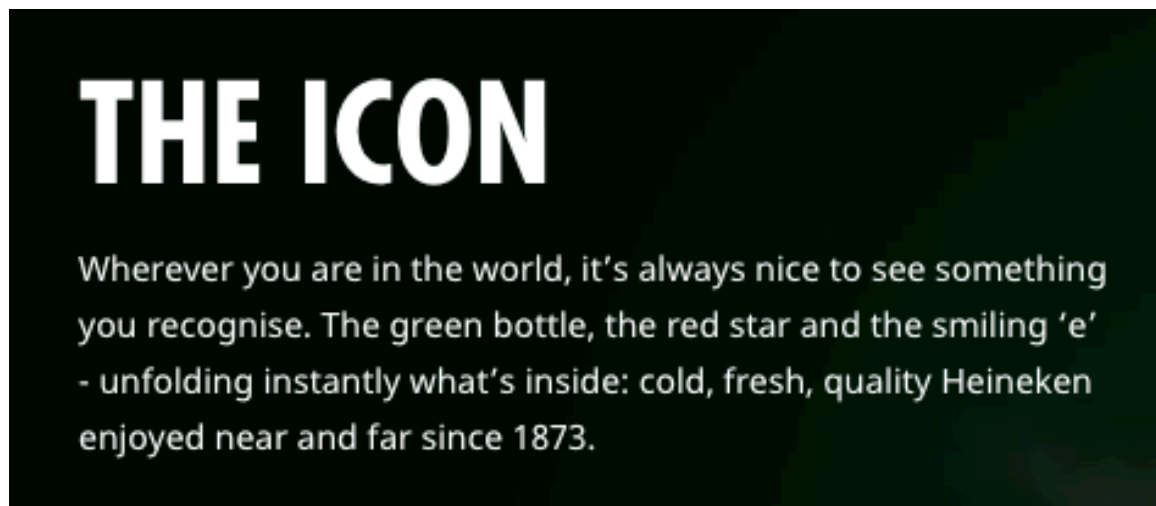


Figure 3.29 shows a 2012 advertising campaign from Germany that is built around images of four identical Heineken bottle caps from London, Berlin, New York, and Tokyo, again reinforcing the idea of recognizability. The campaign includes giant banner ads with relief bottle caps in airport terminals, conveying a focused message that Heineken is available, and identical, everywhere in the world. [Heslinga et al. \(2010\)](#) summarizes the pitch as follows: “It does not matter where in the world you are and what you do because when you see the green bottle of Heineken you are or you feel completely at home again.”



Figure 3.29. Heineken ad campaign by the Bernstein Agency (Germany, 2012).



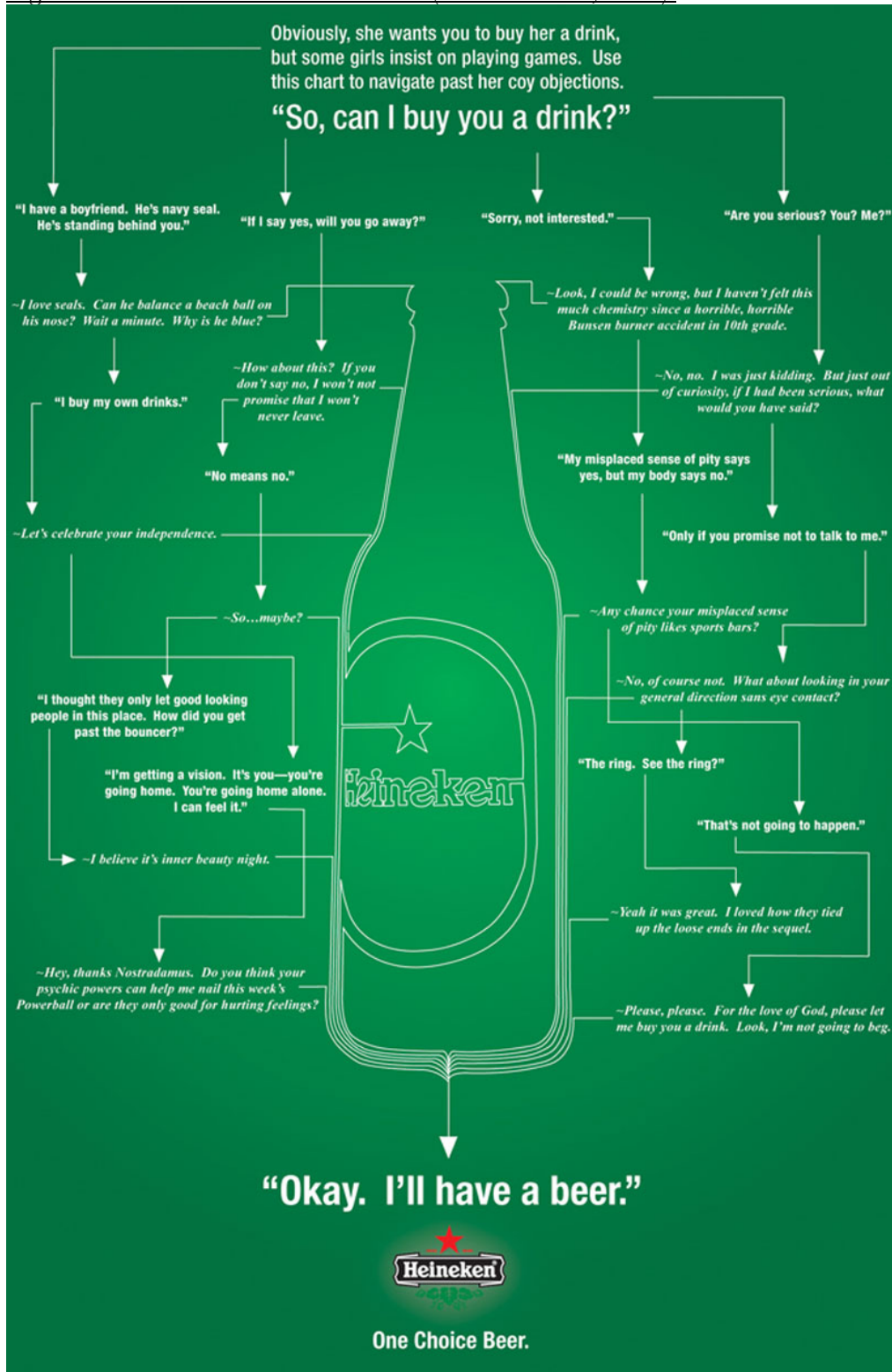
Source: [Bernstein.de](http://Bernstein.de)

Figure 3.30. Heineken advertisement (United Kingdom, 1960s).



Source: [Findery.com](http://Findery.com)

Figure 3.31. Heineken advertisement (United States, 2007).



Source: [Mybeerpix.com](http://Mybeerpix.com)

Figures 3.30 and 3.31 show examples of another consistent element of Heineken's marketing angle: the presentation of the Heineken drinker as a womanizer. Figure 5.3 shows a vintage U.K. Heineken ad from the 1960s that offers a light-hearted endorsement of what in present-day America might be classified as sexual assault. The ad shows two beautiful young women dressed as Heineken bottles stepping through the snow at a European ski resort, with the headline: "Grab a cold Heiney!"

Incredibly, by the early 2000s, the theme of light-hearted sexual assault had not only been carried forward in Heineken ads, but had taken on an even more politically incorrect form. The 2007 advertisement shown in Figure 5.21 consists of a full-page flow chart with instructions for how to pick up a woman (or "girl," as Heineken puts it) at a bar. Heineken's instructions to the man begin as follows: "Obviously, she wants you to buy her a drink, but some girls insist on playing games. Use this chart to navigate past her coy objections." For instance, when the woman says "No means no," the man is instructed to respond: "So...maybe?"

By the 2010s, at least some of Heineken's U.S. advertising had finally shifted to the endorsement of a slightly subtler and more sophisticated form of womanizing: co-branding with James Bond, playboy par excellence. Figure 3.32, for example, shows a 2012 U.S. ad that features the *Skyfall* 007 logo and a black silhouette of Bond, packing heat, in front of a giant Heineken star, with a headline that reads: "Are you Bond enough?"

Figure 3.32. Heineken advertisement (United States, 2012).



Source: [pricechopperblog.com](http://pricechopperblog.com)

### Stella vs. Heineken advertising: conclusions

I conclude this comparative branding review with Table 3.33, which shows a one-line summary of the marketing themes of the four beers in my choice experiment, as I infer them from the material discussed above as well as Budvar and Bud Light advertising not discussed here. Table 3.33 also shows “buy” and “drink” results from the two experimental treatments.

Table 3.33. Comparison of “buy” and “drink” experiment results and marketing themes inferred from the advertising programs of Heineken, Stella, and Budvar (Czechvar).

<u>Brand</u>	<u>Retail price<sup>1</sup></u>	<u>% Likely to “Buy”<sup>2</sup></u>	<u>% Prefer to “Drink”<sup>2</sup></u>	<u>Marketing Theme 1</u>	<u>Marketing Theme 2</u>	<u>Marketing Theme 3</u>
Heineken	\$0.99	19%	19%	Globalization	Uniformity & ubiquity	Sexual conquest
Stella	\$1.08	28%	34%	The upper class	“Chalice” glass	Expense & labor
Budvar	\$1.79	8%	6%	Authenticity	Rarity & individuality	Unique place of origin
Bud Light	\$0.61	45%	40%	Comedy	Fantasy	Sports

<sup>1</sup> Lowest price per 12-ounce serving of beer sold by the case for the four survey beers at Total Wine, Sacramento, CA, in January 2019.

<sup>2</sup> Out of subjects choosing any beer.

Although the three brands compete in the same larger market for European lager, I find that the three brands’ marketing themes, and by extension their target audiences, are different from each other in important ways. This is consistent with my findings in the survey experiment.

Some tastes we’re born with. Others develop from within, through trial and error, partying and blind tasting, engagement and learning, sensitization and habituation. There are all the things we learn from our parents, their ways, their tastes. And then there are a bunch of tastes we all have, maybe even most of them, that seem to come from nowhere and everywhere at the same time. They just buzz around through the cosmos, bounce off things until they hit you. That’s what it is to be a human being these days: that’s just our lives in the radio kingdom.

To me, the most inescapable conclusion from the results reported in this chapter is that American consumers do have non-random beer brand preferences, and that they are measurable, demographically patterned, and responsive to budget constraints. Given the

conclusion from Chapter 4 that Stella, Heineken, and Budvar have indistinguishable sensory beer attributes from the perspective of U.S. consumers, I conclude that this particular set of non-random beer brand preferences are preferences not for useful attributes but for Veblen attributes—and within Veblen attributes, particularly, perhaps, the bullshit.



## Chapter 4. An impostor at the Great American Beer Festival

*A lager drinker shows up to sign books in craft beer heaven*

Before I even sat down next to the other authors at the book-signing table, I had already started to realize that I was in deep shit.

I'd been invited to Denver for the Great American Beer Festival (G.A.B.F.), America's biggest and oldest beer convention, to promote a new book called "The Beer Trials," of which I was co-author and which the G.A.B.F. organizers were under the impression that I had co-written. However, I knew little about beer, except that I liked to drink a lot of it and that it tended to make me fat. My favorite type of beer was cold beer. I had also picked up a few tidbits from having read *The Beer Trials* at least once, all the way through, from cover to cover.

I was listed as the co-author of *The Beer Trials* because the book had been my idea. It was a sequel to *The Wine Trials*, a book that I'd written two years earlier on a topic that I had actually been trained in.

The co-author of *The Beer Trials*, who in *Trials* circles was generally referred to as "the author of *The Beer Trials*," was named Seamus Campbell. Seamus was not only a homebrewer and a professionally certified beer judge: he was also a Certified Cicerone, meaning that he had made it through an incredibly rigorous and competitive process that involved being tested on things like sorting five identical-looking beer samples into order by their comparative levels of diacetyl, dimethyl sulfide, or acetaldehyde.

There were only a few hundred Certified Cicerones in the world at the time, and Seamus was one of them. This fact had often given me solace when I was confronting the realities of life as an impostor co-author, but it was of little help in Denver.

In terms of color, weight, sugar levels, carbonation levels, and every other possible sensory attribute of beer, Seamus and I had approximately opposite beer preferences. Almost every beer I liked, Seamus disliked, and vice versa. *The Beer Trials* contained only Seamus's beer opinions and none of my own, which was a good thing given not only my lack of education but also because my tasting notes, the few times I had tried, said things like "good; tastes like beer," and "bad; doesn't taste like beer."

Unfortunately for me, Seamus was unable to attend the G.A.B.F. signing due to a prior commitment to serve as a judge at a World of Warcraft convention in Rome. In every way—intellectually, morally, emotionally—this was an abhorrent decision on Seamus's part. The depth of my predicament hit me when, just before sitting down at the book-signing table, I noticed that there was already a half-hour-long queue of conference attendees waiting to pose for pictures with the authors of a book called *Yeast: The Practical Guide to Beer Fermentation*.

I wondered what Seamus was doing right now. I pictured his tall, lanky frame standing between two pimply-faced adolescent Italian gamers, resolving their heated Warcraft avatar infringement dispute with his characteristic calmness, intelligence, and grace. In my head I cursed every one of those fuckers.

I slithered into my assigned seat at the book-signing table, where several distinguished-looking beer authors, all of whom were male and at least fifteen years older than me, were crammed into a row of small folding chairs behind a long white plastic slab. Ready on the table in front of my seat was a laser-printed, lucite-encased place card with my name on it. A stack of copies of *The Beer Trials* and a few black ballpoint pens were also laid down in front of me.

Sitting on one side of me at the signing table was a courtly man in his sixties, wearing glasses, almost completely bald but perfectly fit, with excellent posture and the dimpled grin of a geeky child. He had a name like the hero of a comic novel: Stan Hieronymous. Like the *Yeast* authors, Stan Hieronymous was being hounded by adoring admirers and Instagrammers.

Although Hieronymous was there to promote his newest title, *Brewing With Wheat: The 'Wit' and 'Weizen' of World Wheat Beer Styles*, many of his fans had also brought along their dog-eared copies of his older classic, *Brew Like A Monk: Trappist, Abbey, and Strong Belgian Ales and How to Brew Them*.

On the other side of me sat a big, jolly man, around the same age as Hieronymous, with the crinkles of knowledge, a hoarse voice, and a great rocking belly laugh. His name was Jay Brooks, and I later discovered that he was the world's best-known beer blogger. At the moment of my arrival, Hieronymous and Brooks had been in the midst of trading hints, across my empty chair, about the difficulties of brewing göse.

I soon deduced that göse was a type of beer. Otherwise, I recognized few of the words that were used by Hieronymous and Brooks during their discussion, other than the prepositions and linking verbs. At first I thought they were talking about geuze, a Belgian style of beer that I had tried once or twice and that might have been mentioned in *The Beer Trials*. But I ultimately became convinced that göse and geuze, although quite similar, were completely different.

At first, the three of us exchanged pleasantries and little more, with Hieronymous and Brooks consumed by autographs. In a rare ebb of traffic, Hieronymous briefly flipped through my book, sometimes nodding, sometimes frowning, and offering what I can only assume to be an incredibly insightful counterpoint about one of our stated views on the behavior of volatile esters in Bavarian Hefeweizen foam. I agreed that it was an important point he had raised. I was spare with my words.

After an awkward pause, I could tell that Hieronymous was on to me, so I came clean and told him that Seamus had basically written the book. He laughed, agreed that World

of Warcraft was a poor decision on Seamus' part for a multitude of reasons, and promised not to tell anyone else.

We chatted for a little while about where to get green chile cheeseburgers in Hieronymus's hometown of Albuquerque. In spite of my beer ignorance, I did consider myself to be one of the world's top cheeseburger theorists, and I found Hieronymus's opinions on ingredients and prep (e.g. American, not cheddar; griddle, not charcoal) to be spot-on in the sense that they lined up with the correct opinions on the matter (i.e. mine).

At some point, when the crowds were finally starting to thin, I finally shared an intimate moment with Brooks, who I think had detected me as a beer fraud from the moment he first looked me in the eye. Brooks, like Hieronymus, handled this fact in the most unpretentious and polite way possible. We soon discovered that we both lived in the broader Bay Area, and before long he had invited me to come drink beer with him sometime in Marin County.

The unwarranted display of genuine humility and grace on the part of two of the world's foremost beer authors was highly suspicious in that it was almost the opposite of the vibe in the nouveau-riche pissing contest of psycho-sensory elitism that is the dominant culture in wine writing. What made this dissociation even more unlikely was that the top beer critics applied more empirical rigor to their work than the wine critics did. The beer geeks had *more* reason to be condescending than the wine geeks did, not less. Why were they so damn nice?

When it comes to science, beer geeks are the geekiest geeks of the booze world. This is in part because many beer geeks also brew their own beer at home. Wine geeks, on the other hand, typically acquire expertise in drinking and describing wine, but not in making it. Wine writing thus tends to be fluffier, from a technical standpoint, than beer writing.

There is a good historical reason for this: the first people in modern history to categorize, evaluate, and write about wine were not winemakers themselves, but rather middlemen—British wine merchants, a.k.a. *négociants*—who would buy wine from French farmers and sell it at a markup in the United Kingdom, with fancy labels to indicate its provenance and quality and justify the markup.

Wine *négociants* seem to have existed as early as the 1300s, but the business really started to take off around the time of the famous 1855 classification of the red wines of Bordeaux's *Médoc* region, which formally distinguished aristocratic Bordeaux wines from plebian Bordeaux wines and brought the term "grand cru" into being.

From the beginning, the *négociant* project was a brand-building enterprise. It didn't matter whether or not the middlemen knew the first thing about planting vines or fermenting juice. The role of the early *négociants*, like that of modern-day wine experts and *sommeliers*, was to be tastemakers and communicators, lyricists and storytellers, minstrels of farmer mystique.

Beer never grew up around a *négociant* system. Unlike grapes, which must be juiced after picking and subsequently refrigerated, barley and hops are shelf-stable and easily shipped around the world. Beer can be made anywhere, regardless of climate or terrain.

Since it's a lot easier and cheaper to import beer ingredients than the beer itself, the idea of shipping beer from one country to another never caught on like it did with wine. Quality classifications, ranking systems, and higher-priced premium brands were practically non-existent in the beer world even a century after they were introduced to the wine world. Beer was simply a local, undifferentiated commodity.

Many of today's beer pundits, like Hieronymous and Brooks, were already brewing illegally before homebrewing was even legalized in America in 1979.<sup>25</sup> They brewed garage beer for themselves, their friends, and their families.

Initially, this was a heartfelt reaction to the overconcentration and homogenization of American beer. legalization of homebrewing went into effect just as the number of licensed commercial breweries in the US was hitting an all-time low of 44. Suddenly, the underground homebrewing community had a chance to come out of hiding and crack open the white market.

To make homebrew that's even drinkable, you need a pretty good understanding of the chemistry and microbiology of fermentation, which puts the beer critics into stark contrast with the wine critics. Even the exalted Master of Wine and Master Sommelier exams, whose difficulty has periodically been fodder for reality TV shows and documentary films, would never ask candidates to sort samples by microbial content.

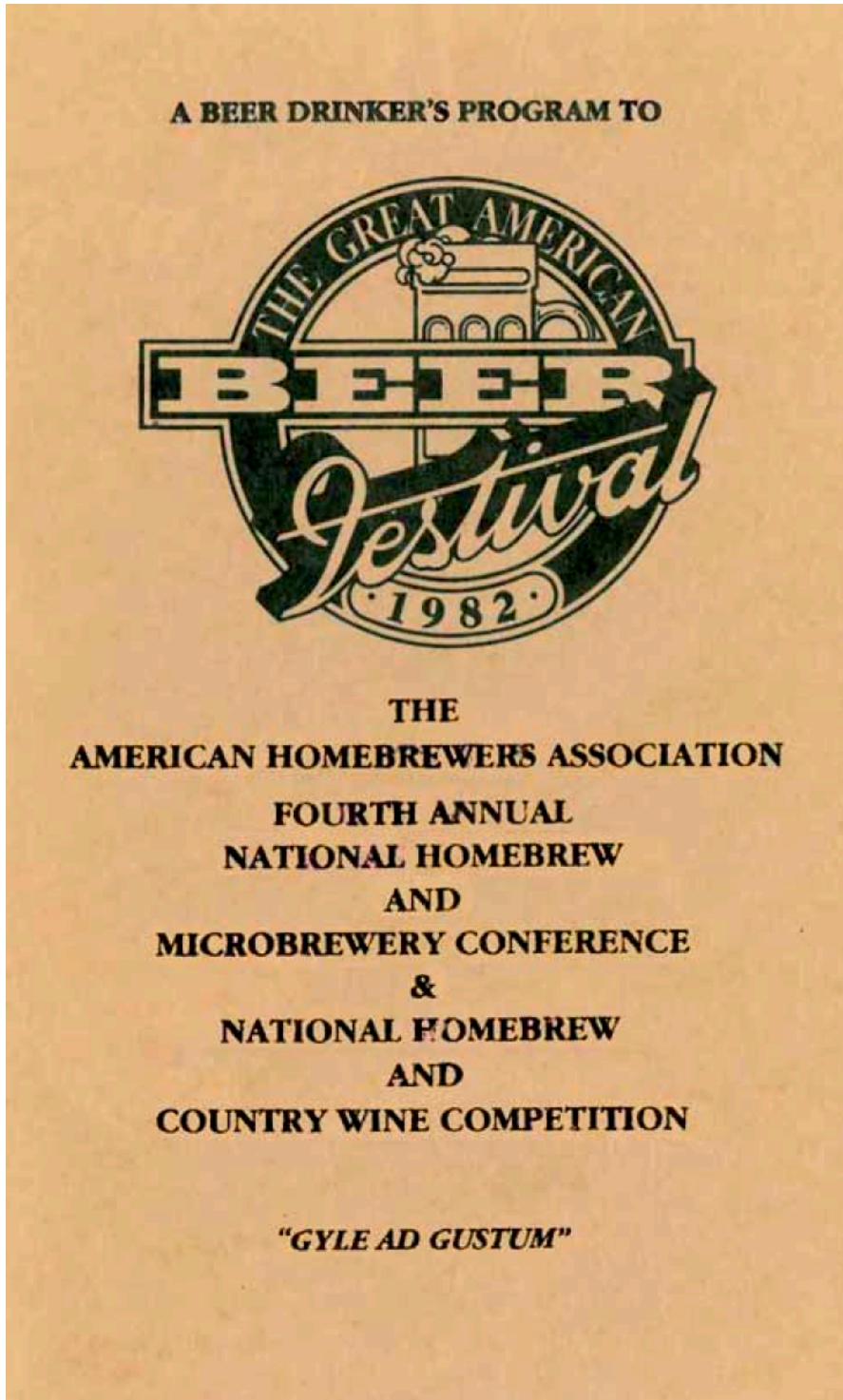
One of the leaders of the homebrewing community from the start was an unassuming nuclear engineer and beer hobbyist named Charlie Papazian, who had been teaching classes in homebrewing at the Community Free School of Boulder, Colorado, since before legalization. When his craft was legitimized, Papazian wasted no time in founding a trade organization, conference, and homebrewing competition.

Three years later, in 1982, he christened his fourth annual conference in Boulder, Colorado, as the first Great American Beer Festival (Figure 4.1). There were 24 commercial breweries and 47 beers represented. The largest brewery to show up was Coors, which came with its premium-angled Killian's Irish Red. Other participants included West Coast mega-micro pioneers Anchor and Sierra Nevada; midsized traditional American lager brands like Yuengling, Genessee, Rolling Rock; and a handful of smaller operations.

---

<sup>25</sup> Legalization, signed into law by President Carter in 1978, removed the federal ban on homebrewing but did not require individual states to legalize it. Only in 2013 did homebrewing become legal in all 50 states. Mississippi and Alabama were the last two states to legalize.

Figure 4.1. Program for the first Great American Beer Festival, the fourth annual conference organized by Charlie Papazian (1982).



In the three decades since, Papazian and his organization, which is now the Brewers' Association, have grown the GABF into one of America's highest-profile mega-conferences, with almost 50,000 attendees per year. Not only is the festival a major tourist draw for the city of Denver, but it has also helped turn the American West—especially California, Oregon, Washington State, and Colorado—into America's craft-beer epicenter.<sup>26</sup>

Thanks in large part to Papazian and Jackson, the craft-beer segment grew out of nowhere to capture the public imagination with increasing speed, fervor, and media coverage to match, throughout the 1990s and 2000s. By 2013, craft beer had captured 14.3% of the total \$100 million American beer market, with sales growing at 20% per year and stealing increasing share from large breweries each year.

There are now more than 6,000 breweries in America—128 times as many as we had in 1979—and 99% of them are craft breweries. In just over three decades—three percent of the lifespan of the global beer industry—the US went from having one twenty-eighth of the number of breweries in Germany alone to having not only the most breweries of any country in the world, but almost twice as many breweries as all of Europe combined.

From any sane economic standpoint, in terms of the efficient allocation of society's resources, this is a preposterously inefficient oversupply of producers. But we're American: when we do something, we do the fuck out of it.

There are some benefits to this oversupply of producers. First, the oversupply increases the beer consumer's opportunity to drink beer close to its production source, which is a good thing for freshness. Second, the oversupply of firms and resulting fierceness of competition between them. In some economic situation, fierce competition keeps prices close to marginal cost, which is a good thing for consumers.

On the other hand, when marginal cost is high—as in the example of 6,000 small, inefficient producers of the same commodity pepper the land as densely as gas stations, but each of them make their own product from scratch so don't benefit from scale economies, competition alone can't save consumers from high prices.

The end result of the massive structural change in the US beer industry is that Americans are paying more for their beer than ever before. On average, craft beer costs 83% more than normal beer,<sup>27</sup> and the most expensive beer on the market costs about 20 times what the cheapest beer does. Still, this is a far cry from the ratio of most expensive wine to cheapest wine, which is more than 1,000:1. One can pay \$2,000 for a brand-new

---

<sup>26</sup> The Brewers Association's definition of "craft beer" is beer produced by a brewery with production of fewer than six million barrels per year. The definition also requires that no more than 25% of a craft brewery be owned by a larger-than-craft brewery, and that most of its beer be brewed with barley malt, rather than corn, rice, or other grains.

<sup>27</sup> This is calculated from 2013 statistics by dividing US craft beer market share by sales (14.3%) by US craft beer market share by volume (7.8%).

bottle of Napa Valley Cabernet from a vineyard owned by a dot-com IPO veteran when there are other bottles from the same appellation available for five dollars. Many wine critics and writers—in contrast to their counterparts in the beer world—continue to defend and justify these kinds of price multiples, even when blind tastings reveal that the experts don't prefer more expensive wines to cheaper ones when their labels are hidden.<sup>28</sup>

Another difference in pomposity between beer and wine critics is that the beer critics admit that they like to drink. Some of them may even admit, albeit while drunk, that they like to get drunk. That's why beer critics actually drink beer when they taste beer. Wine critics don't swallow, they spit, and this makes all the difference.

Beer critics swallow because they're normal people. It's normal human behavior to swallow the liquid that you intentionally sip.<sup>29</sup> It's also normal human behavior to swallow the wine that you drink. Wine geeks don't swallow precisely because the behavior is too normal for them. They don't want to behave like normal people. That's what it is to be a snob.

Wine critics spew Cabernet-saliva cuvée into their effete little silver-plated spittoons, sometimes in a long rainbow stream that you can see from halfway across a large tasting room. Some of them are actually willing to argue that swallowing the liquid isn't actually the point of drinking wine—and that people who actually consume wine, including the alcohol, are missing the whole point of wine. Beer geeks, on the other hand, generally seem to recognize that the alcohol is part of the point of drinking beer, and that the beer can't be fully appreciated without the alcohol. Even the founders of the big-time microbreweries, e.g. Jim Koch of Samuel Adams, are basically just chill guys who like to get wasted.

What makes the G.A.B.F. a particularly impressive organizational feat, even when compared with another convention of similar enormity, is that virtually all of its attendees, at any given time, are drunk. By now the corollary of this statement should already be clear to the reader, which is that wine critics, compared with beer critics, tend to be less fun at parties.

There can be certain downsides to mass inebriation, such as the possibilities of chaos, death, and destruction. The fact that the GABF staff is able to throw a four-day party for 50,000 wasted people—mostly men—in a single giant room that contains the world's biggest all-you-can-drink beer buffet, and still end each day with no bigger problems than a few bros winding up the drunk tank with puke all over their polo shirts, is a testament to the staggering organizational prowess of Papazian and his crew.

It was generally only the tipsiest of attendees that would stop, between the two beer-theory legends that sandwiched me, at the deserted island that was my own signing

---

<sup>28</sup> See, e.g., Goldstein et al. (2008).

<sup>29</sup> Listerine and post-toothbrushing rinses are amongst the exceptions to this rule.

station. Often, customers' first impression would be that I was a handler for one of the other authors. I was frequently asked where the bathrooms were, or if they could pay me for their copies of *Yeast*.

When a potential customer did realize that I was actually there to sign books, he or (very occasionally) she would generally pick up a copy of *The Beer Trials* and turn it over in his or her hands a couple times, handling it as one might a suspect foreign object, like a rock that a guy at a swap meet tells you came from the Mars rover.

This passing curiosity would typically turn to revulsion after they started flipping through the reviews and noticing that our reviews (as per my instructions to Seamus, in one of my only legitimate contributions to the ratings section of the book) included beers like Bud Ice, Natural Light, Pabst Blue Ribbon, and Miller Genuine Draft; or that we awarded 6 points out of a maximum 10 to Steel Reserve and noted in our review that it fit more snugly and easily into a brown paper bag than its competitors. By the end of the one-hour signing session, I knew exactly what it felt like to be a Costco food-sample server hawking samples of diet carrot juice one aisle over from the free meatballs.

The occasional customer whose book I would end up signing usually fell into one of two categories: either the customer was (1) collecting a complete set of signed books by all five or six authors at the table; or (2) just wanting a signed book to give as a gift, and lacking the necessary stamina that would have been required to wait in the Hieronymous or Brooks queue.

When the session finally ended, after my two compadres finished posing for their last few photos and I carried my largely unsold pile of books back over to the cash register, the authors invited me to go taste with them. This was exactly what I had been hoping for all along, but I hadn't had the audacity to suggest it myself. So we set out in formation, onto the convention floor, ready, at last, to drink.

The proportions of the 290,000-square-foot Colorado Convention Center were staggering. There were 455 American breweries pouring 2,248 beers out of the 3,523 beers that were submitted to the competition. Of these, only 2.2% would receive gold medals: one in each of 79 beer categories. There were also silver, bronze, and "Pro-Am" medals, but the total of 239 medals still represented only 6.8% of the beers entered.

Unsurprisingly, it was hardest to win medals in the big categories like American-Style IPA, which had 150 competitors for three medals, than it was in the more specialized categories like Classic Irish-Style Dry Stout (23 entrants), European-Style Dunkel (19), or Gluten-Free Beer (13). Regardless, this was the most selective blind tasting competition I had ever seen. There were 151 BJCP-certified judges from 10 countries on hand to evaluate the beers, and the process was laborious, meticulous, and tightly controlled.

When Brooks, Hieronymous, and I first entered the tasting hall, we each got a small (but not that small) plastic beer-tasting glass, about the size of a large urine cup. We had



the next four-and-a-half hours to drink an unlimited number of the thousands of different beers on the floor.

A printed guide to the medal winners was released in the middle of the festival, resulting in an instant rush on the 79 gold-medal beers that depleted most of their supply almost instantly.<sup>30</sup> After the flop (i.e. the revelation of medal winners), a regular conference attendee could only hope to taste only a few gold-medal beers at best, and maybe a slightly bigger handful of silvers and bronzes.

The first major advantage to tagging along with Hieronymus and Brooks was that they seemed to know the fastest routes to the booths of the gold-medal winners that might still have beer left. Better still, even the brewers who had officially run out of supply would start scampering around when they saw Hieronymus and Brooks approaching, and they'd almost always manage to dig up a stash of something rare.

As a result of this situation, for the first time in my life, I was treated to a roll call of America's best beers according to the palates of America's best beer experts. I was on Seamus's turf now, and I hoped he was somehow sensing this from across the Atlantic and feeling threatened, right there in Italy in his Warcraft referee suit.

We tasted the gold-medal beers in the most prestigious categories, the small-batch cult brews that ran out first. We moved quickly and efficiently across the floor, sampling beers from Pizza Port of Carlsbad, Great Divide of Boulder, Mad River of Blue Lake, Russian River of Santa Rosa, Cigar City of Tampa, Bear Republic of Cloverdale. After five or ten tastes, I started to realize that the most prestigious craft beers in America share two main things in common. The first was that most of them came from California, Colorado, Washington, or Oregon.<sup>31</sup> The second was that none of them tasted anything like beer.

When I say "beer," I mean pale lager beer. Walk into a bar almost anywhere in the world and ask for "a normal beer"—without any further adjectives or qualifications—and you'll almost certainly be served some form of pale lager. The beer will be cold, fizzy, and light gold in color, and it will pour off a white, foamy head.

Ask for "a beer," and you might get Höfbräu in Munich, Mythos in Athens, Tsingtao in Guangzhou, Tecate in Tijuana, Estrella in Barcelona, Pilsener in Quito, Singha in Bangkok, Gambrinus in Prague, Molson in Toronto, Kronenbourg in Nice, Victoria in Melbourne, or Quilmes in Buenos Aires. All of these are really the same product, more or less: normal pale lagers, fundamentally similar in color, flavor, and texture to each other and each containing approximately 5% alcohol by volume, plus or minus half a percent.

The first thing that hits you about how different craft beer is from normal beer is the aromatics. Whereas beer just smells like beer, craft beer gives off a bouquet of overwhelming aromas that you would never normally associate with beer, such as wood,

---

<sup>30</sup> I say "most," rather than "all," for reasons that will become clear later in this chapter.

<sup>31</sup> Perhaps coincidentally, or perhaps not, these states are also the four cannabis powerhouses of America.

molasses, smoke, caramel, coffee, and sometimes vinegar. Craft beer is often cask-aged and thus flatter than beer. It's typically served warmer, so that its aromas come out. Craft beer is also often at least somewhat darker, thicker, sweeter, or more bitter, than normal beer. Many craft beers have several or even all of these features, but almost every craft beer is extremely intense in at least one of these respects.

The way in which craft beer deviates most consistently from beer, however, is its high alcohol content. At press time, for instance, the ABVs of the seven top-rated beers out of thousands listed on the Beer Advocate ratings website were 8%, 13.4%, 11%, 11%, 8%, 14.1%, and 15%. Three of these seven beers actually have more alcohol than a typical wine. In most parts of the world, these beverages would not even be referred to as "beer." In Germany, for instance, the specialty category called "strong beer" rarely exceeds 6.5% alcohol.

In current US craft beer circles, on the other hand, a beer with 5% alcohol is considered almost pathetic. This is part of why the US is possibly the only country on Earth in which "a beer" doesn't automatically mean pale lager. If you walk into a craft beer bar in a hipster neighborhood in just about any city in the US and ask for "a beer," here is what will likely happen.

First, the bartender, who will typically be extremely friendly and have an enormous beard, will point to a lineup of anywhere from 20 to 100 or more taps, and say "well, it depends on what you like!" If you repeat that you'd simply like "a normal beer," then he will apologize gently, and say that he doesn't sell normal beer: he sells craft beer.

If you say: "Okay, I'll have whatever the normal craft beer is," then you'll probably end up with an IPA with between 6% and 7% alcohol. If you know what you're doing, then you'll stop the bearded bartender at this point and ask for an imported lager. Then, if you're lucky, you'll get a slightly skunky Carlsberg or Beck's poured into a pint glass, where it will lose its foam within seconds.

At the Great American Beer Festival, amongst America's best 2,248 beers, it was almost impossible to find a normal one. As I continued to roam the hall with Hieronymous and Brooks—both of whom, to my pleasant surprise, actually seemed genuinely interested in soliciting my naive beer opinions—I was presented with a strange dilemma: should I pretend to appreciate, or at least enjoy somewhat, the craft beers that they and the other judges—at this competition and in decades' worth of the world's best-read beer tomes—have deemed to be America's best?

Did I dare to disrespect the artisanal architects of America's heroic yuppie rescue from the mega-brewery era by suggesting that what has replaced Big Beer isn't really beer? Or if I spent a few years learning about beer like them, would I end up sipping my beer from a brandy snifter?

I don't really remember which approach I ended up taking, because there was too much beer. In fact, I don't really remember much that happened, period, because after 10 or 15 samples of thick brandy-snifter beer, the world was spinning. I do remember lots of flushed faces, laughter, beer at Rock Bottom Brewery, a beer dinner at a restaurant called Euclid Hall, and then more beer after that. We met the Canadian beer-drinking champion, talk weed strains with California's most famous microbrewers.

The second night, before I blacked out, Brooks and Hieronymus had already turned in for the night. Fortunately, later that night I was adopted by a whole other group of buzzed beer geeks. I played drinking games with the "Beer Wench," a blogger who was getting simultaneously pawed at by the 95% of Rock Bottom. At some point in the night, I bluffed my way through an interview with the roving Beer TV crew. The whole time, start to finish, I never saw a single person spit out a single sip of beer.

The next day, as I was wandering around the world's largest all-you-can-drink beer buffet by myself, I noticed something I hadn't seen the day before: a gigantic MillerCoors booth. It was right in the middle of the convention floor. Just because the craft beer lobby ran the joint apparently didn't mean that the mass-market breweries couldn't show up. And they could sure as hell afford a booth—a big and elaborate one, in fact, complete with three Miller Lite bikini girls in the flesh.

As I approached the big MillerCoors banner, I could see that there was almost nobody else in the area. From a probabilistic standpoint, this was a virtual impossibility; given the sardine-like crowds on almost every other square foot of the hall and the presence of several hot, almost-naked girls in the middle of what may be the single biggest sausage fest in America.

At first I walked toward the booth cautiously, glancing around to make sure there weren't any authors nearby. But there was nobody nearby, so I picked up the pace. With every step, the kind of fresh, cold, low-rated beer that might actually help with my throbbing headache was becoming more and more of a reality.

I finally arrived at the pouring station and made eye contact with the MillerCoors rep, who, unfortunately, was a dude. He was a tall, clean-shaven man, on the young side—about my age—and I felt immediate compassion for him. Just like me at my own book-signing station the day before, my compadre at MillerCoors was trying to make the best of what amounted to a virtual boycott of his product.

Coors is originally from Denver: corporate holding groups notwithstanding, it was still to some extent a local company. The MillerCoors rep was just standing there, trying to smile, glancing from side to side occasionally, trying to ignore the occasional snickers as people would walk by and gawk for a second, occasionally taking a snide Instagram picture, before moving on.

The rep looked to be neither shocked nor thrilled by the situation. He had a look of studied neutrality on his face, gazing out at nothing in particular, like a driver at the wheel of a minivan. He must have been used to this kind of treatment at beer festivals. Just to give my compadre some much-needed love, I looked down at my official program to see if MillerCoors had won any medals at the GABF. Was there some kind of “normal beer” or “shitty beer” category, I wondered, where Miller might have been competitive?

It turned out that there was: Miller Lite had won a gold medal in the “American-Style Light Lager” category, which, as I scanned the other competitors—Bud Light, Olympia—I saw was the clearinghouse for all of the trash that normal, everyday Americans drink the most, by far. I realized that I might have stumbled upon the last remaining publicly available sample of gold-medal beer in the entire festival.

But then, as I continued to scan the winners, I saw another gold for MillerCoors, in a category that sounded way more to the liking of beer critics—crafty, almost—under the company’s “Tenth and Blake” subsidiary. And then another gold. And another.

I counted them all up and discovered that MillerCoors had claimed six GABF gold medals, plus three silvers, plus two bronzes: a total of 11 medals. It turned out that this was more gold medals, and more medals overall, than any other beer company had won at the 2010 Great American Beer Festival. MillerCoors was the most decorated beer company at the entire festival, the winner of the whole damn thing. Nobody gave a shit.

I struck up a conversation with my MillerCoors compadre, who was more than happy to pour me as much gold-medal beer as I could possibly drink—which was plenty. It was my kind of gold-medal beer. I learned that since 1996, MillerCoors, including medals previously won as Miller Brewing Company and Coors Brewing Company, had won a combined 232 medals at the GABF. By at least one definition, then—and possibly the most scientific one, given that it was the direct result of the evaluations of expert judges in the world’s most rigorous and prestigious blind beer tastings—MillerCoors was the best brewery in America.

“I’ll toast to that,” I said to my lonely friend. He poured one for himself too. I held up my tasting glass, clinked my plastic urinal cup against his, and took a deep, long swig of cold, fresh, gold-medal-winning Miller High Life, lowest of all brows, the cheapest beer in America. My headache was gone in a single instant, and I knew then that Miller High Life was all I had ever wanted or needed. It was the King of Beers, and even the kings of beer agreed.

## Chapter 5. Craft beer is wine

*A question, a poem, and a song*

### What is craft beer?

The results reported in Chapters 3 and 4 suggest that there is some bullshit even in normal beer, whether imported or domestic. Take Pabst Blue Ribbon (PBR), for instance: there is plenty of bullshit here. You picture the private-equity hustlers in LA who worked a leveraged buyout, cut a deal with Miller to brew it in Milwaukee, and got a bunch of hipsters in Portland to drink PBR. You suspect, in the bottom of your heart, that there is bullshit here somewhere.

But in the end, even in the craftiest of normal beers, the bullshit is not found in such copious quantities as it is in other segments of the beverage industry, such as wine. First, normal beer is not that expensive. Second, the price differences between normal beer brands are relatively minor. Third, as shown in Chapter 3, most normal beers are similar to each other in their sensory qualities. Consumer expectations are well understood and well satisfied. Normal beer is gold, clear, fizzy, crisp, refreshing, more bitter than sweet, and lightly alcoholic. It has about 5% alcohol by volume. At a grocery store, it costs a dollar or two per serving. Normal beer is best served cold and fresh on a sweaty afternoon, at the beach or on the porch, to a group of people you love.

Then there is craft beer. Its bottle or can might be spray-painted with profanity-laden Hentai, or laminated with a digital transfer of a mango with giant boobs shooting up a school. Its liquid might trickle flat from the spout of a gnarly barrel or spew sparkling from the neck of a Champagne bottle. It can be thick and black like ink or pale and milky pink. It can be a Quad as strong as wine or a stout infused with oyster brine.

There are only three rules that every craft beer must obey: (1) it must be somehow stronger and less refreshing than normal beer; (2) it must be different from all other craft beers; and (3) by decree of the Brewers' Association, it must be brewed in smaller quantities than normal beer—currently the maximum is 6 million barrels per year (this number has been periodically adjusted to allow Jim Koch, creator of Samuel Adams, not to get kicked out of the club he helped create)—and therefore be more expensive than normal beer.

In Harry Frankfurt's differentiation of bullshit from lying (see Introduction), he writes:

“The mode of creativity upon which [bullshit] relies is less analytical and less deliberative than...lying. It is more expansive and independent, with more spacious opportunities for improvisation, color, and imaginative play” (Frankfurt, 1986, 2005).

If craft beer is a low-end segment of the fine wine market, then the segment benefits from fewer barriers to entry (as a producer or consumer) and more tolerance for error in production than other wine segments. To make normal or craft wine you must hire a winemaker, but to make craft beer you can download an app. You can end up with a batch of soy-sauce must in your bathtub and call it cask-conditioned triple bock. Craft beer is best served rotten and warm, in a snifter, paired with foie gras sliders, on a cold and rainy evening, to a group of middle-aged white men. Craft beer may thus be a partial substitute for single malt whiskey or counterfeit wine.

The first craft beer I have just randomly withdrawn from my refrigerator, for instance ( $n=1$ ), describes itself as a “New England-style India Pale Ale” made by Heretic of Fairfield, California. The name of the beer is “Make America Juicy Again.”

Figure 5.1. Reagan and Heretic seal.



Figure 5.2. Craft beer name and BA logo.



On the right side of its label, in the background, is an original work of stencil art depicting former U.S. President Ronald Reagan as the Devil (Figure 3.34). The stencil art is terrifying in that (1) it is zoomed in just on the Devil/Reagan’s maniacal face; (2) the whole face is in a shade of gray light enough that you can miss the Devil/Reagan on the first sip and be surprised by him halfway through your beer, and (3) the only exception to (2) is the Devil/Reagan’s demonic right eye, which has a dark gray pupil and warm gray white and penetrates your soul like a laser.

Barely violating the right edge of Reagan’s devil goatee is the corporate “Heretic” logo, set in a werewolf font and encircled by a ribbon of four slogans: “fiercely independent,” “genuine craft brewery,” “family-owned,” and “served fresh.” As shown in Figures 5.2 and 5.3, the Brewers’ Association certified craft logo, for which consumers are willing to pay extra (e.g. in Hart, 2018), shows a silhouette of an upside-down beer.

Figure 5.3. Launch of the Brewers’ Association Supporter Seal and website.

## Brewers Association Launches Supporter Seal and Website



Source: Baker (2019).

Whatever your definition of bullshit, some bullshit, like craft beer bullshit, is plain to the naked eye. It is expansive and independent, with opportunities for improvisation, color, and imaginative play. To summarize:

There was an old warehouse in Portland  
that was crafted, tricked out, and LEED-certed.  
They brewed up some brews and they priced them to move  
at twelve dollars per snifter-sized serving.

Behind the steel taps stood the stalwart,  
the Czar of this barley-wine barnyard.  
He mumbled and peered through his Taliban beard  
and slung beers from a red leather gauntlet.

He handed a plate to an orderly harlot  
with bats on her breasts and ripe cheeks of bright scarlet.  
Her beer-braided placenta with deer brains and fennel  
was sprayed with aged tears from the cask of Aunt Charlotte.

The harlot felt just like a princess that day  
as the Czar sang of hop kingdoms far away  
and helped her to pick between Sour White Prick  
and Mom's Lyme Disease Tick IPA.

Reversing the progress of centuries,  
and creating deliberate scarcity,  
he fights for the cause of re-making the flaws  
that were cured by our modern technology.

It is rated by scholars and critics.  
Its varieties grow by the minute.  
Its label's oration of vast information  
has value with no upper limit.

It is juicy and boozy and fine.  
It grows soft and complex over time.  
It is sniffed in a glass, and reviewed by an ass:  
This is not a craft beer, it's a wine.



Song: “It Was Beer”

Intro

I cry for normal beer. There’s less and less each year.  
It used to flow like water, now it flows like frozen tears.

Verse A1

There was once a day when all the beer was cheap.  
There was once a day when it wasn’t judged by geeks.  
There was once a monk who had a great idea.  
There was once a monk who brewed a normal beer.

Chorus

It was beer, it was beer, it was normal beer.  
He saw God. God was clear: God sent normal beer.  
It was beer, it was beer, it was normal beer.  
It was beer, it was beer, it was beer!

Verse B1

It is always Craft Beer Week in California  
without a normal beer tap to be found.  
Of our abnormal beers we are very proud—  
of their dark and musty barrels aging underground.

Verse C1

It is flat! It is warm! Brown and bitter, sweet and strong.  
Tastes like soy sauce, smells like Brett—  
makes you woozy, not refreshed.  
Swirl it in a brandy snifter, scratch and sniff it, say it’s different—  
It’s not beer, it’s not beer, it’s not beer!

Verse B2

In the California valley of the sunshiney days,  
you can have your beer in thousands of abnormal ways.  
But ask your bartender for a normal beer,  
and he’ll say: which IPA?

### Verse C2

IPA for the critics and the hipsters of the Bay.  
IPA for the Tesla dealers north of San Jose.  
IPA for the surfers and the sommeliers.  
IPA for the MBAs.

### Verse C3

IPA in the night, IPA in the day.  
IPA for the starlets and the douchebags of LA.  
IPA for the sea lions of Monterey.  
IPA for the MBAs.

### Verse B3

We've got Racer, Bear Republic, Pliny Elder, Pliny Younger.  
We've got gluten-free pork belly marijuana nitro porter.  
We were freed by bearded bartenders with tats of Che Guevara  
from the shackles of our ordinary lager!

### Verse A1

There was once a day when all the beer was cheap.  
There was once a day when it wasn't judged by geeks.  
There was once a monk who had a great idea.  
There was once a monk who brewed a normal beer—

### Chorus

It was beer, it was beer, it was normal beer.  
He saw God. God was clear: God sent normal beer.  
It was beer, it was beer, it was normal beer.  
It was beer, it was beer, it was beer!

### Coda

It was fresh, it was gold, it was cold. It was beer.  
It was crisp, it was foamy, it was clear. It was beer.  
It was the same every day, every year. It was beer.  
It was just a normal beer.

## **Chapter 6: Many laws of many prices?**

*In a consumer survey of about 3,300 U.S. consumers, I observe and describe multi-modal price distributions in “stated tendency to pay” for wine and beer.*

This chapter asks two questions about present-day U.S. retail prices: (1) How much do American consumers tend to pay for common retail packages of beer and wine? (2) How does this tendency to pay vary between individuals?

These two questions could be approached with many different empirical methods. One technique would be to observe mean prices and price distributions in large data sets of retail prices, such as supermarket scanner data. An alternative technique, which I introduce in this chapter, is to survey a substantial number of U.S. consumers and simply ask them to report how much they typically pay for retail beer and wine packages. I call this method “stated tendency to pay” (hereafter abbreviated “STTP”).

My survey elicits STTP from an initial pool of 3,086 U.S. consumers. In this chapter, I analyze STTP frequency distribution for wine and beer, comparing the distributions and illustrating differences through comparative means, histograms, and kernel density plots. I formally define several different forms of multi-modality, and I apply these definitions to the STTP frequency distributions in my survey results. I also observe gender, age, household income, population density, regional, and response time effects on STTP. Finally, I speculate on some effects that may derive from natural occurrences of prices and numbers in real-world environments, and I look to Benford’s Law for help explaining these distributions.

### Methods and design

In spring 2017, I surveyed 3,086 U.S. internet users online about their STTP for a six-pack of beer or a 750ml bottle of wine. All subjects were surveyed through the Google Surveys interface (formerly known as Google Consumer Surveys) during the four-day period from April 11, 2017, to April 14, 2017.

Google Surveys share many of the drawbacks and advantages common to all survey techniques in which consumers self-report information about their spending habits. Two of the biggest drawbacks of my technique are (1) the potential inaccuracy of some consumers’ estimates of their own spending patterns, and (2) the lack of a direct incentive for honesty or disincentive for dishonesty. Amongst the benefits of my technique are the demographic information provided by Google Consumer Surveys, including gender, age group, region, and income bracket; the breadth of the subject pool, which covers all 50 states and a diversity of demographic profiles; the simplicity, consistency, and ease of use of the Google Surveys interface; the recording of each subject’s response time in milliseconds; and the short time commitment necessary for subjects to participate.

My survey asked subjects only one simple question of about 10 words, using the Google Consumer Surveys interface. Approximately half of subjects were randomly assigned to each of two variants of the survey (“beer” and “wine”). In the “beer” variant, I asked subjects the following question: “How much do you typically pay for a six-pack of beer at a store?” (n=1,570). In the “wine” variant, I asked: “How much do you typically pay for a bottle of wine at a store?” (n=1,516).

The survey was designed to take less than 20 seconds for the typical subject to complete, and to be as easy as possible to understand for a large and diverse subject pool. For these reasons, the survey did not specify the bottle volume in either wine or beer variants (the assumption being that subjects would generally interpret the question to mean the standard sizes of 750ml for wine and 6 x 12 ounces for beer). Subjects were presented with a blank response field and were required to type in a number as a response. Only nonnegative numerical responses were accepted. To reduce the amount of potential noise from consumers who did not wish to answer or who were not beer or wine buyers, I also offered subjects an easy way of opting out of the survey by entering a response of “0.”

A total of 3,086 U.S. consumers participated in the survey between April 11 and April 14, 2017.

### Data

Other than not allowing negative responses, I chose not to limit the numerical values that would be accepted as responses. A subject could respond to the question by entering any nonnegative integer. I chose not to set bounds for responses so as to avoid the potential bias of “anchoring” responses to the stated high/low range. (For instance, if I limited six-pack prices to between \$5 and \$25, then subjects who desired to submit a relatively “neutral” answer might be biased, perhaps unconsciously, toward the midpoint of \$15). One problematic side effect of not bounding responses was some unreasonably low and high data points (e.g. “\$0.0000001” or “\$1,000,000”) submitted by subjects who completed the survey (in order to access the premium content) without answering the question in good faith.

In order to eliminate the noise and bias that would have resulted if non-sensical responses were included in an analysis of the survey data, I dropped a considerable number of outlying STTP responses at both the high and low ends of the numerical response spectrum. First, in analyzing both the “beer” and “wine” variants, I dropped all responses of “0” (about 23% of “beer” responses and 20% of “wine” responses), which indicated that the subject did not want to participate in the survey or was not in the market for wine or beer. Next, I dropped the top 9% and bottom 3% of nonzero numerical responses.

I chose these cut-off points of 9% at the top and 3% at the bottom by imposing a rough plausibility condition on responses, beginning with the “beer” variant. To do this, I began by sorting each of the 1,205 nonzero beer responses into percentile bins (ordered from 1 to 100, where bin 1 represents the 1st percentile, or bottom 1%, of responses, and bin 100 represents the 100th percentile, or top 1%, of responses).

If all responses were unique numbers, this process would assign 12 or 13 responses to each of 100 different percentile bins. However, my data set contains many non-unique responses: for instance, 91 “beer” subjects entered responses of exactly \$7.00. My percentile process handles such cases by forcing all equal numerical responses into the same percentile bin. Thus the 91 responses of \$7 are all assigned to the 36th percentile, representing the fact that approximately 35% of all responses fall below \$7. The next higher response, \$7.25, is assigned to the 43rd percentile, representing the fact that approximately 42% of all responses fall below \$7.25. Percentile bins 37 through 42 are thus left empty. For the whole set of “beer” responses, the percentile sorting process thus results in a total of 36 (out of a possible 100) percentile bins being populated with anywhere between 1 and 145 responses.

After sorting the responses into percentile bins, I dropped all data points in percentiles whose price ranges I determined to be “implausible” with respect to the actual upper and lower bounds of prices for six-packs of beer in the U.S. retail market (excluding special cases such as marketing giveaways or going-out-of-business sales where goods may occasionally be liquidated below wholesale cost, which would not reasonably correspond to any consumer’s “typical” STTP for a six-pack of beer).

Implausible responses might have been entered for a variety of different reasons. Some subjects might have entered intentionally dishonest responses for the purposes of avoiding deliberation and proceeding as quickly as possible to their premium content. Others might have been dishonest merely for the purpose of intentionally causing trouble, perhaps in some cases because they were annoyed by the presence of an online paywall preventing them from accessing the content to which they had intended to click through.

On the bottom end of the “beer” data, I dropped the lowest 3% of nonzero responses (percentile bins 1 through 3), corresponding to the 52 responses between \$0.0000001 (an obviously nonsensical answer) and \$2.00 (an STTP below the wholesale cost of any six-pack on the U.S. market). There were no responses between \$2.00 and \$2.99, the lowest value in the 5th percentile, which I allowed as “plausible.” On the top end, the 92nd percentile included responses up to \$55, a price well above that of any six-pack in the U.S. retail marketplace. (The only possible exception might be Sapporo Space Barley, made with ingredients from the International Space Station, but this six-pack was only sold to 250 customers in the world by lottery and thus could not be purchased “in a store”

as per the survey question). I thus dropped the 108 top-end responses in bins from the 92nd to 100th percentiles.

The results of this sorting and dropping process for the “beer” variant of the survey are shown in Tables 6.1 (with summary statistics) and 6.2 (with detail by percentile). The process reduced the initial pool of 1,570 subjects in the “beer” variant to 1,045 subjects (67%), who reported paying between \$2.99 and \$20 per six-pack.

Table 6.1. National consumer survey results: “How much do you typically pay for a 6-pack of beer at a store?”

Total subjects assigned to “beer” group	n = 1,570 obs.
Drop top 9% (pctiles 92-100) of nonzero answers (\$21 and over)	108 obs.
Drop bottom 3% (pctiles 1-3) of nonzero answers (\$2 and under)	52 obs.
Drop all \$0 answers	365 obs.
-----	
Total observations dropped (33% of all answers, 12% of nonzero)	525 obs.
-----	
<u>Remaining observations in “beer” group used in analysis</u>	<u>n = 1,045 obs.</u>

Beer group	Obs	Mean	Std. Dev.	Min	Max
All	1570	648905	22900000	0	900000000
All nonzero	1205	845461	26100000	.0000001	900000000
-----					
Drop top 9%	108	9433072	87000000	21	900000000
Drop bot 3%	52	1.383	0.585	.0000001	2
Drop \$0	365	0	0	0	0
-----					
Tot dropped	525	1940518	39500000	0	900000000
-----					
Used	1045	7.91	2.65	2.99	20
-----					

Table 6.2. National consumer survey results: "How much do you typically pay for a 6-pack of beer at a store?" Responses in US dollars, by percentile.

*Raw survey data. Red indicates outlying percentiles dropped from analysis. Percentiles calculated with respect to all 1,205 nonzero answers.*

Pctile bin	n=	min	max	mean
Dropped 1	28	.0000001	1	0.90
3	24	1.34	2	1.95
Used 5	21	2.99	3	3.00
7	46	3.37	4	3.94
10	2	4.49	4.5	4.50
11	16	4.55	4.99	4.88
12	90	5	5	5.00
19	2	5.1	5.18	5.14
20	26	5.49	5.99	5.87
22	131	5.99	6	6.00
33	38	6.35	6.99	6.88
36	91	7	7	7.00
43	4	7.25	7.35	7.28
44	12	7.4	7.6	7.50
45	36	7.8	7.99	7.98
48	145	8	8	8.00
60	11	8.09	8.69	8.43
61	42	8.78	8.99	8.98
64	75	9	9	9.00
70	21	9.59	9.99	9.97
72	117	10	10	10.00
82	16	10.49	11	10.93
83	66	11.99	12	12.00
88	1	12.25	12.25	12.25
89	13	12.5	13	12.96
90	11	13.5	15	14.68
91	12	15.99	20	19.40
Dropped 92	13	21	55	34.38
93	11	57.35	87	70.94
94	13	88	200	120.73
95	12	234	599	403.58
96	11	600	642	603.82
97	15	699	800	743.67
98	9	850	1200	968.55
99	12	1343	12345	5347.42
100	12	14553	90000000	85000000
All nonzero	1205	.0000001	90000000	845460
Used	1045	2.99	20	7.91

Next, I applied the same process described above to the wine data, removing the zero responses and the same percentages of high (9%) and low (3%) responses from the “wine” data. This reduced the initial pool of 1,516 subjects in the “wine” variant to 1,060 subjects (70%), who reported paying between \$1.29 and \$35 per bottle. Results of the sorting and dropping process for the “wine” variant of the survey are shown in Tables 6.3 (with summary statistics) and 6.4 (with detail by percentile).

Table 6.3. National consumer survey results: “How much do you typically pay for a bottle of wine at a store?”

Total subjects assigned to “wine” group	n = 1,516 obs.
Drop top 9% (pctiles 92-100) of nonzero answers (\$36 and over)	107 obs.
Drop bottom 3% (pctiles 1-3) of nonzero answers (\$1 and under)	53 obs.
Drop all \$0 answers	296 obs.
-----	
Total observations dropped (33% of all answers, 12% of nonzero)	456 obs.
-----	
<u>Remaining observations in “wine” group used in analysis</u>	<u>n = 1,060 obs.</u>

Wine group	Obs	Mean	Std. Dev.	Min	Max
All	1516	345450	12800000	0	500000000
All nonzero	1220	429265	14300000	.01	500000000
-----					
Drop top 9%	107	4894297	48300000	36	500000000
Drop bot 3%	53	.908	.243	.01	1
Drop \$0	296	0	0	0	0
-----					
Tot dropped	456	1148443	23400000	0	500000000
-----					
Used	1060	12.25	6.31	1.29	35
-----					



Table 6.4. National consumer survey results: “How much do you typically pay for a bottle of wine at a store?” Responses in US dollars, by percentile.

*Raw survey data. Red indicates outlying percentiles dropped from analysis. Percentiles calculated with respect to all 1,205 nonzero answers.*

Pctile bin	n=	min	max	mean
Dropped 1	53	.01	1	0.91
Used 5	27	1.29	2	1.89
7	6	2.19	2.99	2.83
8	18	3	3	3.00
9	17	3.99	4	4.00
10	63	5	5	5.00
16	38	5.62	6	5.99
19	41	6.95	7	7.00
22	6	7.5	7.99	7.82
23	70	8	8	8.00
28	4	8.45	8.99	8.73
29	43	9	9	9.00
32	225	9.99	10	10.00
51	16	10.99	11	11.00
52	105	11.25	12	11.99
61	14	12.49	13	12.95
62	24	13.99	14	14.00
64	120	14.5	15	15.00
73	1	15.08	15.08	15.08
74	12	15.99	16	16.00
75	19	17	18	17.70
76	7	18.99	19.99	19.32
77	100	20	20	20.00
85	9	21	24	22.44
86	42	25	25	25.00
89	25	26	30	29.68
91	8	30.98	35	33.37
Dropped 92	10	36	40	39.21
93	17	44	50	48.18
94	8	54	66	59.50
95	13	69	80	74.69
96	14	90	100	98.43
97	9	110	250	155.67
98	12	300	838	592.58
99	12	886	3500	1664.42
100	12	4111	50000000	4360000
All nonzero	1220	.01	50000000	429265
Used	1060	1.29	35	12.25

## Results

To give an overview of STTP responses, I first recalculate percentiles using the data that remain after dropping \$0 responses, the top 9% of nonzero responses, and the bottom 3% of nonzero responses from both the “beer” and “wine” groups. I call this the “final data set.” These results are shown as Table 6.5 (beer) and Table 6.6 (wine).

Mean STTP responses in the final data set are \$7.91 for a six-pack of beer and \$12.25 for a bottle of wine. Beer STTP shows relatively less variance than wine STTP. The standard deviation for beer STTP is \$2.65 (34% of mean price), and the standard deviation of wine STTP is \$6.31 (52% of mean price).

Median STTP is \$8 for a six-pack of beer and \$10 for a bottle of wine. For both beer and wine, the median responses are also the modal responses (with the maximum response frequencies). 225 subjects, or 21% of all “beer” subjects, reported paying \$8 (the median and modal responses) per six-pack of beer. 145 subjects, or 14% of all “wine” subjects, reported paying \$10 (the median and modal responses) per bottle of wine.

Compared with the frequency distribution of beer STTP, the frequency distribution of wine STTP has a longer right tail. The median value for beer STTP (\$8) is close to the mean, whereas the median wine STTP is notably lower than the mean. Table 3.5 shows STTP by percentile for beer (recalculated for the final data set), and Table 3.6 shows price responses by percentile for wine.

Intuitively, the “low” cut-off point of \$1.29 per bottle for wine, like the “low” cut-off point of \$2.99 for a six-pack of beer, seems plausible (if barely so) in the sense that it is slightly (but not much) higher than the minimum wholesale cost per bottle for the most inexpensive wines. It is worth noting, however, that some responses above the “high” cut-off point, such as \$36, are certainly “plausible” in that they are well below the maximum price of a retail bottle of wine.

This dropping of some potentially good-faith responses is a result of having generated the cut-off points for both groups based on the plausibility of “beer” responses rather than “wine” responses. I chose to proceed in this order (first, generating the cut-offs based on beer, and second, applying them to wine) because proceeding in the opposite order would have resulted in a considerable number of implausible “beer” responses being included in the data set, which would have biased the means and other statistics for the “beer” group.

I also note that dropping the 92nd percentile of “beer” data resulted in dropping a few “plausible” responses (at the low end of the 92nd percentile), such as \$21 per six-pack, so there is reason to believe that a rough parity between the “high” cut-off points for wine and beer STTP might nonetheless have been achieved by this method

Table 6.5. National consumer survey results: "How much do you typically pay for a 6-pack of beer at a store?" Responses in US dollars, by percentile.  
*Percentiles recalculated after dropping \$0 responses, top 9% of nonzero responses, and bottom 3% of nonzero responses.*

Percentile	N	min	max	mean
1	21	2.99	3	3.00
3	46	3.37	4	3.94
7	8	4.49	4.75	4.64
8	10	4.99	4.99	4.99
9	90	5	5	5.00
17	4	5.1	5.49	5.32
18	24	5.5	5.99	5.90
20	131	5.99	6	6.00
32	1	6.35	6.35	6.35
33	10	6.49	6.98	6.63
34	27	6.99	6.99	6.99
36	91	7	7	7.00
45	9	7.25	7.49	7.38
46	9	7.5	7.85	7.58
47	34	7.89	7.99	7.99
50	145	8	8	8.00 (median value)
64	9	8.09	8.59	8.37
65	44	8.69	8.99	8.97
69	75	9	9	9.00
76	21	9.59	9.99	9.97
78	117	10	10	10.00
89	6	10.49	10.99	10.83
90	10	11	11	11.00
91	66	11.99	12	12.00
97	6	12.25	12.99	12.78
98	11	13	14	13.23
99	10	15	18	15.40
100	10	18.79	20	19.88
Total	1045	2.99	20	7.91

Table 6.6. National consumer survey results: "How much do you typically pay for a bottle of wine at a store?" Responses in US dollars, by percentile.

*Percentiles recalculated after dropping \$0 responses, top 9% of nonzero responses, and bottom 3% of nonzero responses.*

Wine STTP	N	min	max	mean
1	27	1.29	2	1.89
3	6	2.19	2.99	2.83
4	18	3	3	3.00
5	17	3.99	4	4.00
7	63	5	5	5.00
13	38	5.62	6	5.99
16	1	6.95	6.95	6.95
17	40	6.99	7	7.00
20	2	7.5	7.5	7.50
21	74	7.95	8	8.00
27	1	8.45	8.45	8.45
28	46	8.5	9	8.99
32	225	9.99	10	10.00 (median value)
53	16	10.99	11	11.00
55	105	11.25	12	11.99
65	14	12.49	13	12.95
66	24	13.99	14	14.00
68	4	14.5	14.99	14.87
69	116	15	15	15.00
79	1	15.08	15.08	15.08
80	12	15.99	16	16.00
81	19	17	18	17.70
82	1	18.99	18.99	18.99
83	106	19	20	19.96
93	51	21	25	24.55
97	3	26	28	27.33
98	22	29.99	30	30.00
100	8	30.98	35	33.37
Total	1060	1.29	35	12.25

A reasonable alternative method would have been to set different “high” and “low” percentile cut-off points for wine and beer.<sup>32</sup>

Although my method has its drawbacks, I chose to sacrifice some potentially meaningful data at the high end of wine responses in order to make the “wine” and “beer” data sets as comparable as possible. I made this choice in light of the fact that comparisons between wine and beer results are particularly important for the analysis that will follow.

Given that median prices are equal to modal prices, the simplest answer to the first question initially posed in this chapter—how much do American consumers typically pay for beer and wine?—is \$8 per six-pack of beer (\$1.33 per beer) and \$10 per bottle of wine.

These two common alcoholic beverage packages contains a similar total quantity of alcohol: approximately 100 milliliters. Assuming the median price of \$8 for beer and \$10 for a bottle of wine, the typical price paid for wine—as measured by price per liter of pure alcohol—is about one-third above the typical price paid for beer. This result is calculated as follows. Assuming that an individual bottle or can holds 12 ounces and 5% alcohol by volume (ABV),<sup>33</sup> a six-pack of beer would therefore contain 2.13 liters of beer, or 106.5 ml of pure alcohol, at an STTP of \$75.15 per liter of pure alcohol. Assuming that a typical bottle of wine contains 750 ml of wine at 13% ABV, it would contain 97.5 ml of pure alcohol at an STTP of \$102.56 per liter of pure alcohol—a premium of about 36% over beer.

An important caveat to this result, however, is that the wine STTP premium (over beer) is not observed at the low end of the market. Given that my survey data set includes some implausible responses, the practical U.S. minimum prices of beer and wine cannot be determined from these data. To obtain a rough estimate of the actual bottom end of

---

<sup>32</sup> One additional justification for extending the 9% “high” cut-off to wine responses was the disproportionately high percentage of subjects answering \$36 or more per bottle of wine came from 18-to-24-year-old subjects, whose average incomes are the lowest of any age group, and who also (amongst all age groups) had the highest proportion of implausible responses on the low end. Beyond the apparent contradiction of the youngest or lowest-income consumers reporting unusually high prices paid for wine, it is also worth noting here is the fact that the U.S. minimum drinking age, in all 50 U.S. states, is 21. It is thus technically illegal under state law, in every one of the 50 U.S. states, for three-sevenths of the 18-to-24-year-old group—subjects from 18 to 20—to buy alcohol, even if anecdotally, many Americans under 21 do purchase alcohol some way or other. I chose not to omit 18-to-24-year-olds from my experiment altogether, because I hypothesize that 21-to-24-year-old consumers (and perhaps even under-21 consumers) may exert a nontrivial influence on the marketplace and may have preferences that are different from those of consumers in other age groups. Google Surveys does not provide information about whether consumers within the 18-to-24-year-old group are over or under age 21. In order to incorporate the opinions of under-25 consumers without subjecting means and other results to excessive noise, given that under-25 consumers appear (again anecdotally) to give more obviously dishonest answers than other age groups, I thus choose a broad approach to eliminating outliers.

<sup>33</sup> This may slightly overestimate the median (or maybe even the mean) ABV of mass-market beer in the United States, but 5% is the mean ABV in a sample of 300 beer prices I took at Total Wine in Sacramento, CA, in January 2019.

the retail market, as a point of reference, the most inexpensive bottle of wine available at U.S. retail stores, I refer to Total Wine, a major national retail chain with branches in several U.S. states, including California, Oregon, Nevada, and Virginia.

It is easily verified that the most inexpensive way to buy wine (in terms of price by volume) is the 5-liter box. The most inexpensive 5-liter boxes of wine listed on wine-searcher.com are priced at \$9.49. A variety of brands are priced at this same price point, including Peter Vella Blush and Franzia Chillable Red. At the \$9.49 price point, all competing wines contain 9% alcohol by volume, for a price per liter alcohol of \$21.09. Slightly more expensive per unit liquid, but less per unit alcohol, is a group of wines sold by Total Wine with more typical ABVs (for wine), such as Franzia Merlot (\$11.99, 12.5% ABV = \$19.18 per liter pure alcohol).

Cheapest for the 750ml bottle package format is Trader Joe's Charles Shaw ("Two-Buck Chuck," \$1.99 in some U.S. states, 12% ABV = \$22.11 per liter pure alcohol). In other words, buying 12–13% ABV wine in a 750ml bottle (as supposed to buying similar wine in a 5-liter box) adds approximately \$3 per liter of pure alcohol at the very bottom end of the price range in the marketplace.

As for beer, the bottom-end price on Total Wine (shared by several competing brands) is \$17.99 for a 36-pack (432 oz): \$0.50 per can. This is equivalent to 0.64 liters of pure alcohol, so the lowest-end price is \$28.15 per liter pure alcohol. The implication is that the cheapest wine is cheaper than the cheapest beer, if measured on a per-liter pure alcohol basis. Specifically, the cheapest wine in a 5-liter box is 32% cheaper per liter of pure alcohol than the cheapest case of beer (\$19.18 vs. \$28.15), and the cheapest wine in a 750ml bottle is 21% cheaper per liter of pure alcohol than the cheapest case of beer (\$22.11 vs. \$28.15). In terms of six-packs of beer—the common package type about which I survey U.S. consumers—the lowest price observed is \$3.99 (\$0.67 per can), or \$37.64 per liter pure alcohol.

Coming back to my survey results, if one accepts the assumptions I used to drop outlying responses, then the median and modal STTP per liter of pure alcohol for both beer and wine are several times higher than these bottom-end prices that can be observed by surveying the marketplace. U.S. consumers in my survey spend 36% *more* per liter pure alcohol on wine (\$102.56 per liter pure alcohol, or about 3.6 times the bottom-end price per liter pure alcohol of a 750ml bottle of wine) than U.S. consumers in my survey spend on beer (\$75.15 per liter pure alcohol, or almost exactly twice the bottom-end price per liter pure alcohol of a 6-pack of beer).

### Premiumization

The observation that median and modal U.S. consumer STTP for wine is 36% higher per unit volume than STTP for beer, taken together with the fact that wine is cheaper per

liter of pure alcohol on the bottom end, might imply that wine is more “premiumized” than beer, i.e. that wines, compared with beers, have more premium or value-added attributes that lead consumers (even in their typical everyday purchases) to pay relatively higher multiples of the product’s base price for premium versions. The hypothesis that wine is more premiumized than beer appears to have some support in the common-sense intuition that high-end wines, especially vintage wines, trade at much higher prices than high-end beers.

In my interpretation of Veblen (1899), premiumization is one of the central predictions of the theory of leisure demand. Veblen observes that in a leisure economy where disposable income is increasing, staples are not scarce, and basic consumer wants are met, consumers are often willing to pay premia for goods that comport with the “honorable” or “respectable” tastes set by a society’s elites.

Notably, Veblen mentions wine, and not beer, in his list of archetypal forms of “conspicuous leisure.” In this sense, he would likely have predicted the same phenomenon that I observe here: that the frequency distribution of wine prices would have lower kurtosis, and/or would be skewed more rightward, than the frequency distribution of beer prices would be. But what about segmentation? Would the distributions of wine versus beer prices also differ in their smoothness, or in their number of peaks and valleys? On this point, Veblen does not say much.

One conjecture of my own, which prior to this investigation was driven more by intuition than by empirical evidence, is that premiumization is commonly marked not just by kurtosis or skew, but also by a multi-modal price frequency distribution. In this view, premium “segments” arise not in a continuous fashion, but rather at multiples of the bottom-end price. The kurtosis and skew of frequency distributions in premiumized markets may thus be generated by a bumpy spreading-out of locally modal values (i.e. local maxima in the price frequency distribution) at popular price-multiple points along the right tail of the distribution, rather than a uniform flattening and/or skewing of a continuous-looking frequency distribution curve.

To see if this might be the case, and in an attempt to better understand the phenomenon of premiumization generally, I will report results below that compare the distributions of wine STTP to the distributions of beer STTP in a variety of different ways. In my view, one of the best ways to examine premiumization, particularly if price multi-modality plays a role, is to examine the shapes of price distributions graphically. When it comes to multi-modal price distributions, between-means comparisons may often be less useful than graphical illustrations such as price frequency histograms or kernel density plots—even if such graphical methods suffer from major quantitative drawbacks such as the lack of an industry-standard approach to verifying the statistical significance of multi-modal effects.

In the sections of Chapter 4 that follow, I first make observations about the STTP frequency distributions for wine and beer, and then I look for some finer-grained differences in premium preferences between consumers in different demographic groups.

### Multi-modality

I begin by looking at a series of histograms and kernel density plots that plot the frequency distributions of STTP for beer and wine from my survey. First, I state some formal definitions. I define a histogram-generating function that I call *Freq*, which takes as its inputs (1) some set  $X$  of three or more numerical values  $x, x \in X \subset \mathbb{R}, |X| \geq 3$ , for which the distribution is to be described; and (2) a set of three or more “bins” that specify the upper and lower bounds for the sub-intervals on which frequencies are to be calculated.

Bins. Let *bins* be defined as a set of  $n$  mutually exclusive and collectively comprehensive partitions of  $X$ ,  $bin_i \subset bins, n \geq 3$ . I denote each such partition with the index number  $i$  in the form  $bin_i$ , such that  $bins \equiv bin_1 \cup bin_2 \cup \dots \cup bin_n$ . I denote a partition  $bin_i$  by specifying the ordered pair of its unique lower and upper bounds  $(bin_{i,min}, bin_{i,max})$ . As a measure of the interval covered by the partition, I define *bandwidth* =  $bin_{i,max} - bin_{i,min}$ , where  $bin_{i,min} < bin_{i,max}$ . In this dissertation I generally use uniform or log-uniform bandwidth, but I do not impose any kind of uniformity condition on the partitioning of bins in general.

I require that *bins* be ordered sequentially by index number, such that for all  $i$ ,  $bin_{i,max} < bin_{i+1,min} < bin_{i+1,max} < bin_{i+2,min}$ , and so on. Formally this condition would be an order on *bins* where  $bin_1 < bin_2 < \dots < bin_n$  and  $bin_i < bin_j \Leftrightarrow bin_{i,max} < bin_{j,min}$ . This orders *bins* sequentially and also implies the mutual exclusivity condition that  $bin_1 \cap bin_2 \cap \dots \cap bin_n = \emptyset$  (an empty set).

Freq function. Given  $X$  and *bins*, the frequency function is generated for each  $bin_i$  by counting the number of values of  $x$  that fall within the minimum-maximum range that defines the bin, divided by the total count of values in  $X$ . This is equivalent to constructing a percent frequency distribution table. Formally, I use the notation  $1_x$  (with a lower-case  $x$ ) to mean an indicator or “counting” function that adds 1 to the sum for each value of  $x$  that falls within the *bin* interval:

$$Freq(X, bin_i) = \frac{\sum_{x=bin_{i,min}}^{x=bin_{i,max}} 1_x}{\sum_{x=\min(x)}^{x=\max(x)} 1_x}, \text{ where}$$

$$i \in \mathbb{N}; x \in X \subset \mathbb{R}, |X| \geq 3; bin_i \subset bins, |bins| \geq 3;$$

$$bin_{i,min} \in X \leq bin_{i,max} \in X.$$



*Freq* example. As an example, consider the hypothetical set (“Data Set 1”) of 10 US dollar-denominated retail prices  $X: \{x_1:3, x_2:4, x_3:4.5, x_4:4.5, x_5:5, x_6:5, x_7:5.5, x_8:6, x_9:7, x_{10}:7.5\}$ , and  $bins: \{bin_1:[3,3.5], bin_2:[4,4.5], bin_3:[5,5.5], bin_4:[6,6.5], bin_5:[7,7.5]\}$ . For this data set, these  $bins$  could also be equivalently specified by  $\{bin_1:[3,4), bin_2:[4,5), bin_3:[5,6), bin_4:[6,7), bin_5:[7,8)\}$ .

To generate the *Freq* outputs for Data Set 1, we must first verify that there are at least three prices and three bins, i.e.  $|X| = \geq 3$  and  $|bins| = \geq 3$ . These conditions are both satisfied, as  $|X| = 10$  and  $|bins| = 5$ . For Data Set 1, there would be 12 total possible bins (two endpoints and 10 points in between), so the output of *Freq* would be:

$$Freq(X, bin_1) = \frac{1}{12}$$

$$Freq(X, bin_2) = \frac{3}{12}$$

$$Freq(X, bin_3) = \frac{3}{12}$$

$$Freq(X, bin_4) = \frac{1}{12}$$

$$Freq(X, bin_5) = \frac{2}{12}$$

*Mode.* I construct indicator variables for two types of modal bin characteristics. The first type I call “local modality.” This is measured by the *Localmode* variable, which assigns 1 to the maximum value in the set and 0 to all other values. I call the second type of modal bin characteristic “global modality.” This is measured by the *Globalmode* variable, which assigns 1 to the maximum value of each local maximum in a histogram of given bin size and 0 to all other values. I go to the trouble of adopting my own terminology with “Mode,” as I do with “Frequency,” not because I object to any of the prevailing usages in the statistical literature, but rather because many different usages exist in the literature, each suited to their own purposes. Rather than adopting one of these usages and attempt to defend it, I state my own definitions. I classify various types of modality for the express purpose of generating useful variables that can best help to describe, analyze, and predict real-world retail prices using data collected from stores.

One characteristic shared by most of these retail price data sets is that at their finest level of granularity, their frequency distributions, at the finest level of granularity (i.e. the

smallest possible “bin” size—see “Frequency”) are rendered naturally discontinuous by the natural occurrence of numerical prices in the marketplace: for instance, few retail stores would be likely to have any items selling for \$100.13.

The retail price frequency distribution, if viewed as a curve rather than a histogram, is thus likely to be nonlinear and discontinuous, and might be non-differentiable. Like *Freq*, *Globalmode* (as well as *Localmode*) is undefined on sets containing fewer than three unique values (and thus three unique frequency bins). One implication of this is that the degenerate case of a uniform distribution is excluded from my definitions of modality, whereas in classical definitions, a “mode” can generally be defined on a uniform distribution or on a set with one or two members.

Local mode. When I refer to the “local modes” of  $X$ , I mean the bin(s) that generate the local maxima of  $Freq(X, bins)$ ; see “Frequency” for this definition.

Formally, I define local modes as the bins for which the *Localmode* function, defined as follows, generates a value of 1:

Let  $Localmode(Freq(X, bin_i)) = 1$   
if any of conditions (1) through (4) hold:

- (1)  $bin_i$  is a local mode limited to a single bin, with a count greater than the counts of the bins on either side of  $bin_i$ , such that  $Freq(X, bin_{i-1}) < Freq(X, bin_i)$ , and  $Freq(X, bin_i) > Freq(X, bin_{i+1})$ ; or
- (2)  $bin_i$  is a local mode spanning multiple contiguous bins that are equal, such that  $Freq(X, bin_{i-q-1}) < Freq(X, bin_{i-q})$ , and  $Freq(X, bin_{i+r}) > Freq(X, bin_{i+r+1})$ ; where  $Freq(X, bin_{i-q}) = Freq(X, bin_{i-q+1}) = \dots = Freq(X, bin_{i+r-1}) = Freq(X, bin_{i+r})$ , where  $r, q \in \mathbb{Z}$ ;  $r, q \geq 0$ ; and either  $r > 0$  or  $q > 0$ ; or
- (3)  $bin_i$  is the smallest bin (which by construction is  $bin_1$ ), and its value is higher than the next-to-smallest bin, such that  $i = 1$ , and  $Freq(X, bin_1) > Freq(X, bin_2)$ ; or
- (4)  $bin_i$  is the largest bin, and its value is higher than the next-to-largest bin, such that  $i = j$ , where  $\forall bin_k \subset bins, k \neq j : bin_j > bin_k$ ; and  $Freq(X, bin_j) > Freq(X, bin_{j-1})$ .
- (5) Let  $Localmode(Freq(X, bin_i)) = 0$  for the remaining values of  $bin_i$  that do not satisfy any of conditions (1) through (4).

Note that *Localmode* will always equal zero for a perfectly uniform distribution of  $X$ , because the percent frequency distribution cannot rise or fall as  $x$  increases. An ideal bell-shaped Gaussian-type distribution with a finite number of values will generate *Localmode* = 1 for only one value of  $X$  (the mean value).

Global mode. I define the indicator function *Globalmode*, shown in Equation Mode.2, to mark the bin(s) for which  $Freq(x, bin_i)$  generates the highest value(s) of any bin:

$$Globalmode(Freq(x, bin_i)) = 1, \text{ if } (Freq(x, bin_i) = \max (Freq(x, bins))).$$

Let *Globalmode* = 0 for all other values of  $bin_i$ .

Like *Localmode*, *Globalmode* can generate outputs of 1 for more than one bin, in the case that the global mode is also a local mode satisfying condition (2) of the *Localmode* function.

Number of modes. In evaluating the number of modes in a frequency distribution, I count the number of locally modal “peaks” across all bins. Given that local modes can span multiple bins, the number of locally modal peaks is not equal to the number of locally modal bins; that is, unimodality does not necessarily imply that  $Localmode(Freq(X), bin_i)=1$  for only one  $bin_i$ . For contiguous local modes spanning multiple bins that meet condition (L.2), such as  $bin_2$  and  $bin_3$  in the example above, their contiguous span of bins where  $Localmode(Freq(X), bin_i)=1$  is counted as a single mode.

One mode (unimodality). I call the distribution of a data set  $X$  “unimodal,” with respect to a set of bins, if the data set generates exactly one local mode. Data sets with normal (Gaussian) distributions, for example, would satisfy unimodality if the bins were of uniform bandwidth.

Two modes (bimodality). I call the distribution of a data set  $X$  “bimodal,” with respect to some set of bins, if  $Localmode(Freq(X), bins)$  generates exactly two local modes. The example set above, given the bins above, would thus be described as “bimodal” in my terminology.

More than two modes (multi-modality). For distributions that have than two local modes, I use the form “ $n$ -modal,” as in “3-modal” or “5-modal.”

Maximum number of modes. Given *bins*, the maximum number of possible modes is, for even numbers of bins,  $bins/2$ ; and for odd numbers of bins,  $(bins + 1)/2$ . I do not take the time to prove this intuition, as it would be severely off topic.

Coming back to Data Set 1, applying all of the definitions above to the hypothetical data set used in the “Mode” and “Frequency” sections, the price frequency distribution for Data Set 1 would be said to be “bimodal,” with 2 “local modes” at [\\$4, \\$6) and [\\$7, \\$8), and a “global mode” at [\\$4, \\$6). The local (and global) mode at [\\$4, \\$6) spans two bins— $bin_2$  [\\$4, \\$5) and  $bin_3$  [\\$5, \\$6)—but I consider it one mode with respect to  $n$ -modality.

Example. Let  $X$  represent Data Set 1:

$X: \{x_1:3, x_2:4, x_3:4.5, x_4:4.5, x_5:5, x_6:5, x_7:5.5, x_8:6, x_9:7, x_{10}:7.5\}$ ;  
 $bins: \{bin_1:[3,4), bin_2:[4,5), bin_3:[5,6), bin_4:[6,7), bin_5:[7,8)\}$ .

First, it is easy to verify that there are at least three prices and three bins in these sets:  
 $|X| = 10 \geq 3$ ;  $|bins| = 5 \geq 3$ .

For Data Set 1, the outputs of *Localmode* and *Globalmode*, calculated using the *Freq* function as defined in the “Frequency” section, would be:

$$Freq(X, bin_1) = \frac{1}{12} \rightarrow Localmode(X, bin_1) = 0;$$

$$Globalmode(X, bin_1) = 0.$$

$$Freq(X, bin_2) = \frac{3}{12} \rightarrow Localmode(X, bin_2) = 1, \text{ by condition 2};$$

$$Globalmode(X, bin_2) = 1.$$

$$Freq(X, bin_3) = \frac{3}{12} \rightarrow Localmode(X, bin_3) = 1, \text{ by condition 2};$$

$$Globalmode(X, bin_3) = 1.$$

$$Freq(X, bin_4) = \frac{1}{12} \rightarrow Localmode(X, bin_4) = 0;$$

$$Globalmode(X, bin_4) = 0.$$

$$Freq(X, bin_5) = \frac{2}{12} \rightarrow Localmode(X, bin_5) = 1, \text{ by condition 4};$$

$$Globalmode(X, bin_5) = 0.$$

## Histograms

To illustrate the relationship between bin size and multi-modality, Figures 6.7.1b–6.7.6b show for beer, and Figures 6.7.1w–6.7.6w show for wine, a sequence of histogram plots generated by the the same STTP frequency distribution (i.e. the price response sample from my survey after removing zeroes and outliers, leaving 1,060 wine and 1,045 beer responses) and varying only the bin size, which progressively decreases (as one reads downward), resulting in a finer-grained histogram and an increasing number of modes.

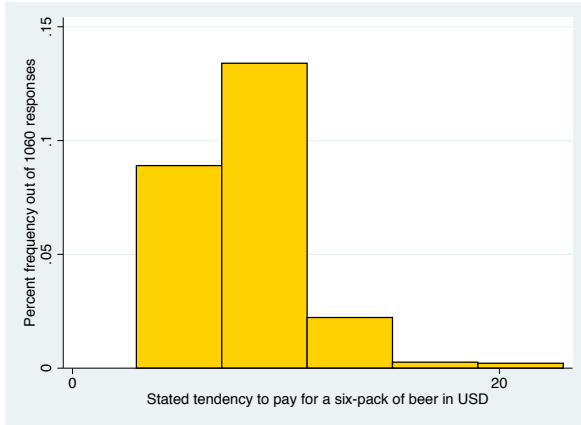
Any price frequency distribution, if given some large enough bin size or bandwidth, can be sufficiently smoothed, in a price frequency histogram or kernel density plot (see below). as to appear unimodal (i.e. with frequency or “density” values increasing monotonically between the minimum-valued bin and the modal bin, and decreasing monotonically between the modal bin and the maximum bin). Recall that my equations for modality are undefined are undefined in the degenerate case of one or two bins.

One way of viewing the price distribution of my survey unimodally is shown in Figures 6.7.1b and 6.7.1w, on the top of the beer and wine columns. These are the histograms that are the most “smoothed” because of using the largest bin sizes of all (analogous to the widest bandwidth in a kernel density plot). Given that the beer distribution is approximately half as wide across prices as the wine distribution, I select bin sizes for these side-by-side illustrations that are double the dollar values for wine as for beer, thus generating histograms with similar numbers of bins and roughly equalizing the granularity of the beer and wine histograms and making them easier to compare.

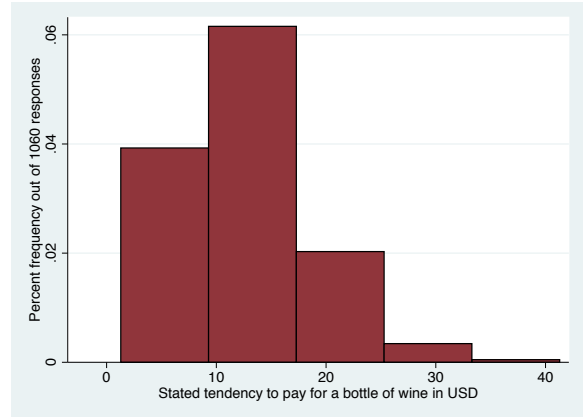
The widest bin size I show is \$4 per six-pack for beer and \$8 per bottle of wine. For example, this would mean that all responses between \$8.01 and \$12 for beer, and between \$8.01 and \$16 for wine, would be represented by one bar in the histogram (in these examples, each is the modal bar, covering about 13% of beer responses and 6% of wine responses). These bin sizes result in 5 bins for both beer and wine.

Figures 6.7b and 6.7w. Beer: Stated tendency to pay, frequency distribution (after dropping outlying answers and \$0 answers), with increasing bin sizes

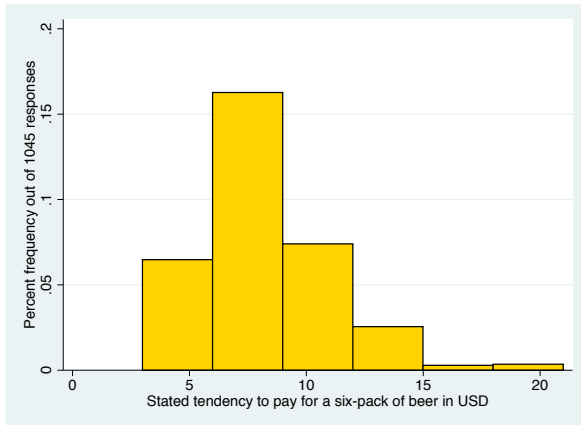
6.7.1b. Beer: Bin = \$4, 5 bins, 1 modal value



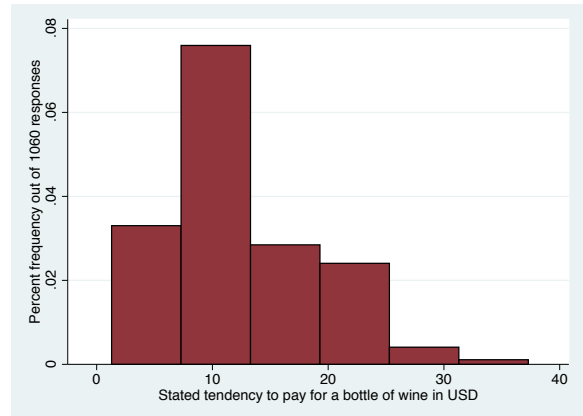
6.7.1w. Wine: Bin = \$8, 5 bins, 1 modal value



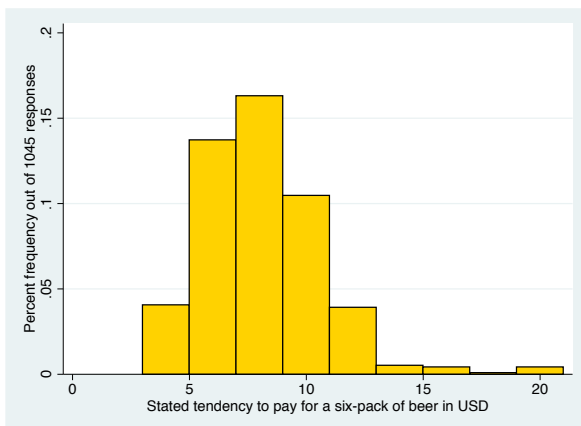
6.7.2b. Beer: Bin = \$3, 6 bins, 2 modal values



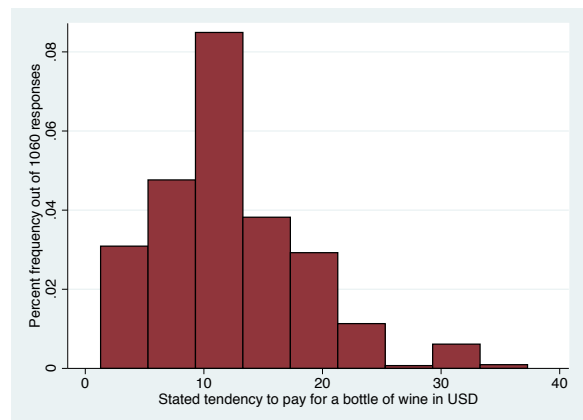
6.7.2w. Wine: Bin = \$6, 6 bins, 1 modal value



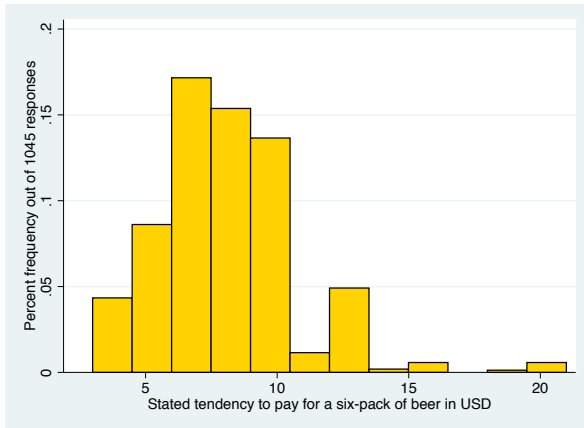
6.7.3b. Beer: Bin = \$2, 9 bins, 2 modal values



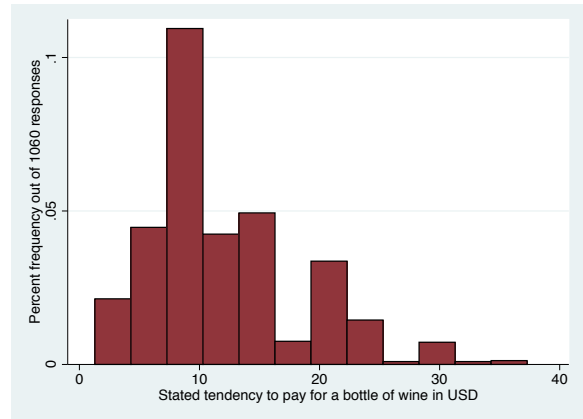
6.7.3w. Wine: Bin = \$4, 9 bins, 2 modal values



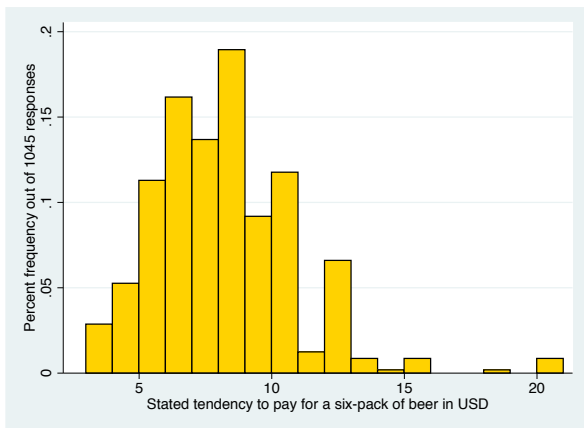
6.7.4b. Beer: Bin = \$1.50, 12 bins, 4 modal values



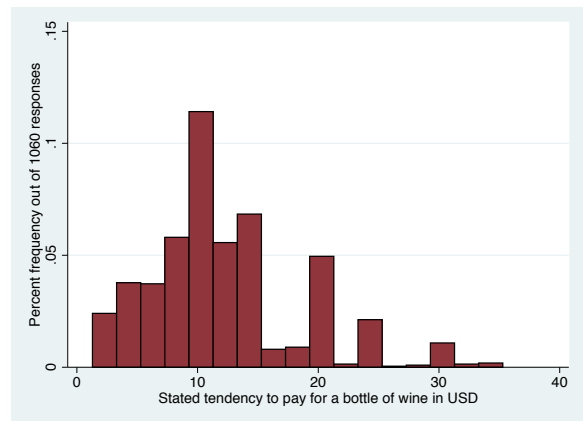
6.7.4w. Wine: Bin = \$3, 12 bins, 5 modal values



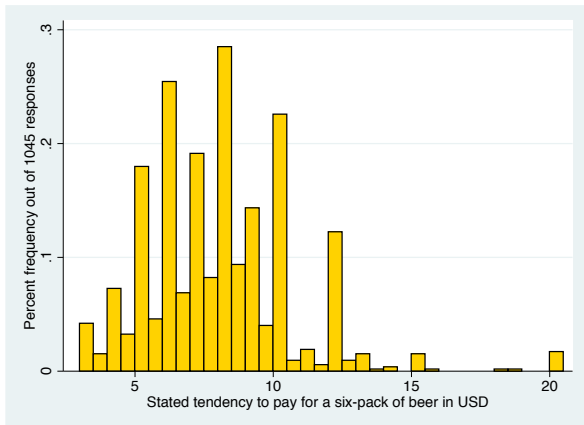
6.7.5b. Beer: Bin = \$1, 18 bins, 7 modal values



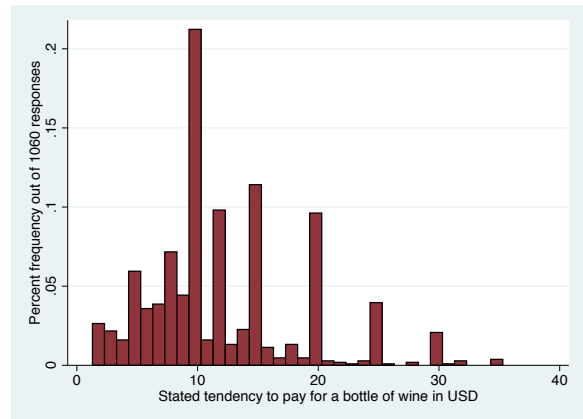
6.7.5w. Wine: Bin = \$2, 17 bins, 7 modal values



6.7.6b. Beer: Bin = \$0.50, 35 bins, 15 modal values



6.7.6w. Wine: Bin = \$1, 34 bins, 13 modal values



In figures 6.7.1b and 6.7.1w, the shapes of the beer and wine distributions are similarly unimodal with some rightward skew, with the main difference being the relatively lower kurtosis for wine.

Moving down to Figures 6.7.6b and 6.7.2w, bin sizes of \$4 and \$6 for beer and wine, respectively, generate 6 bins for each. The wine distribution remains unimodal but the beer distribution becomes bimodal, with a slightly higher proportion of responses around \$21 as around \$18. Clearer bimodality arises in both beer and wine in Figures 6.7.3b and 6.7.3w, where bin sizes of \$2 and \$4 generate 9 bins for beer and wine, and both beer and wine having a second, smaller modal value near the high end of the price bounds. For beer, this is around \$20, and for wine, it is around \$30. Here we see the first suggestion of a concrete high-end market segment where some prices are more popular than others, and demand seems to surround those secondary modes.

The sequence of progressively decreasing bin sizes shown in Figures 6.7.4b–6.7.6b and 6.7.4w–6.7.6w illustrates the emergence of more locally modal values as the frequency histograms become increasingly granular.

Table 6.8 summarizes the absolute and relative bin sizes and numbers of modes for wine and beer given four of the six relative bin sizes shown in Figures 6.7.1b–6.7.6b and 6.7.1w–6.7.6w.

Table 6.8. Relationship between bin size and number of modes, STTP for wine and beer.

<u>Relative bin size</u>	<u>Beer: bin size</u>	<u>STTP modes</u>	<u>Wine: bin size</u>	<u>STTP modes</u>
$(0.5)^0 = 1$	\$4 = $2^2$	1	\$8 = $2^3$	1
$(0.5)^1 = 0.5$	\$2 = $2^1$	2	\$4 = $2^2$	2
$(0.5)^2 = 0.25$	\$1 = $2^0$	7	\$2 = $2^1$	7
$(0.5)^3 = 0.125$	\$0.5 = $2^{-1}$	15	\$1 = $2^0$	13

### Kernel density plots

In general, the overall shape and “smoothness” of a price frequency histogram depends on the size of price bins selected for the X-axis of the histogram. The bin size in a histogram corresponds to the “bandwidth” in a kernel density plot: essentially, is a measure of the degree of smoothing, where higher bin sizes or bandwidth generate a more “smoothed” histogram or density plot.

A variety of different estimators can be used to generate kernel density plots. I use the Epanechnikov kernel function (Epanechnikov, 1969), which is a popular estimator because it is the most efficient in minimizing the mean integrated squared error (see, *e.g.*



Baum, 2013). All visual kernel density plots shown below are generated by Stata software, whose documentation describes the process as follows: “Kernel density estimators approximate the density  $f(x)$  from observations on  $x$ ” (Stata, 2019).

Histograms do this, too, and the histogram itself is a kind of kernel density estimate. The data are divided into non-overlapping intervals, and counts are made of the number of data points within each interval. Histograms are bar graphs that depict these frequency counts—the bar is centered at the midpoint of each interval—and its height reflects the average number of data points in the interval.

Figures 6.9a, 6.9b, and 6.9c below illustrate the same approximate relationship between relative smoothness and number of modal values that is illustrated by the histograms in Figures 6.7.4b–6.7.6b and 6.7.4w–6.7.6 above, but with continuous curves that are estimated using an Epanechnikov kernel density function with progressively decreasing bandwidths (1, 0.5, and 0.25 for beer; 2, 1, and 0.5 for wine).

The smoothest plot with the widest bandwidth, Figure 6.9a, shows unimodal distributions for both beer and wine. In Figure 6.9b, beer and wine each have six modes (\$6, \$8, \$10, \$12, and \$15, and \$20 for beer; \$10, \$12, \$15, \$20, \$25, \$30 for wine), and in Figure 6.9c, beer and wine each have roughly nine modes. Consistent with the relationships shown in Table 3.8, the wine distribution spreads out across the price spectrum roughly twice as “widely” as the beer distribution. This spread also extends in the other direction to the low values, which is consistent with the fact that on the marketplace, the cheapest wine (~\$2) is cheaper than the cheapest six-pack of beer (~\$4), whereas the modal wine (\$10) is more expensive than the modal beer (\$8).

In the sense that premiumization refers to the spreading-out of value-added “premium” product versions at progressively higher price points, I would thus characterize wine as being approximately twice as “premiumized” as beer. This relative measurement of premiumization between two distributions, which is roughly equivalent to the kurtosis ratio of the distributions, is one simple way that premiumization might be characterized. The number of modes, meanwhile, might be viewed as a relative measure of “segmentation,” which, interestingly, occurs in roughly equal proportions for wine and beer, in spite of their difference in premiumization.

I infer from these data that from the point of view of consumers’ own STTP, wine is more premiumized than beer, but that segmentation of wine and beer is approximately equal. These inferences could be tested against on a set of actual prices (e.g. wine and beer scanner data), and comparing the differences between STTP and actual prices could yield interesting insights into consumer perceptions.

Figure 6.9a. Beer vs. wine STTP distribution, large bandwidth (most smoothed).

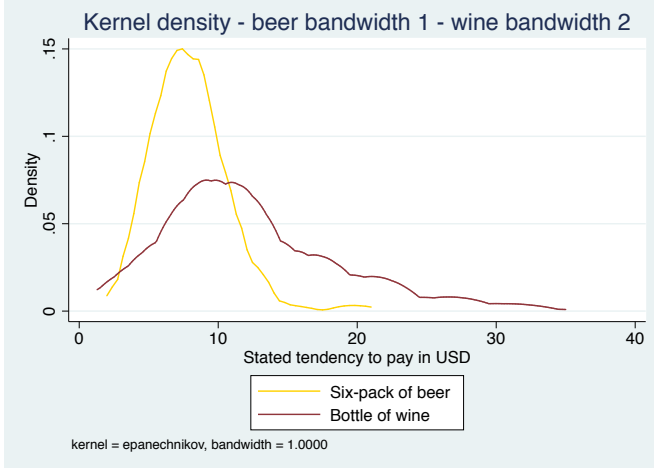


Figure 6.9b. Beer vs. wine STTP distribution, medium bandwidth (more smoothed).

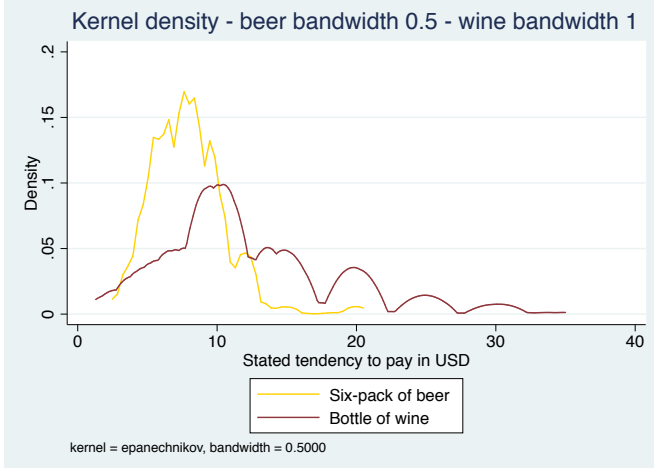
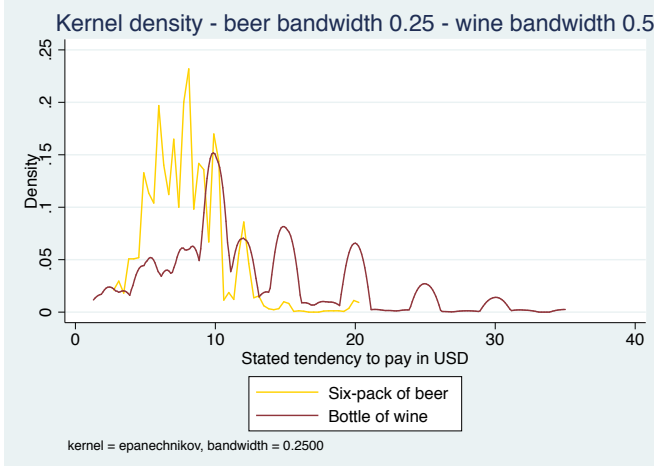


Figure 6.9c. Beer vs. wine STTP distribution, small bandwidth (least smoothed).



### Gender effects

Figures 6.10b and 6.10w show the STTP frequency distribution with an Epanechnikov kernel density plot for male (blue) vs. female (pink) subjects, with bandwidth of 0.25 for beer and 0.5 for wine. For beer, males and females show seven distinct local modes, with few if any male-female gender differences. For wine, two differences are noticeable: there is one bottom-end local mode around \$3 that is observed in males and not females, and there is one mid-level local mode around \$8 that is observed in females and not males.

Figure 6.10b. Gender effects on STTP distribution for beer.

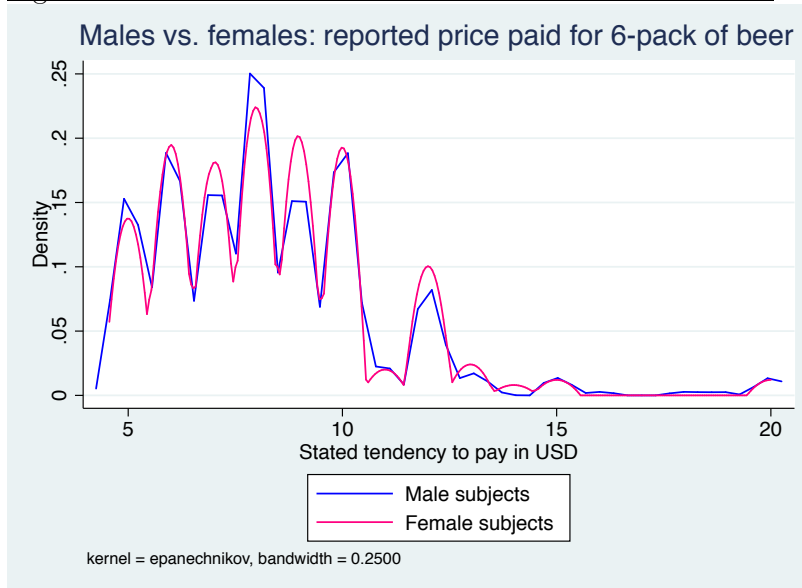
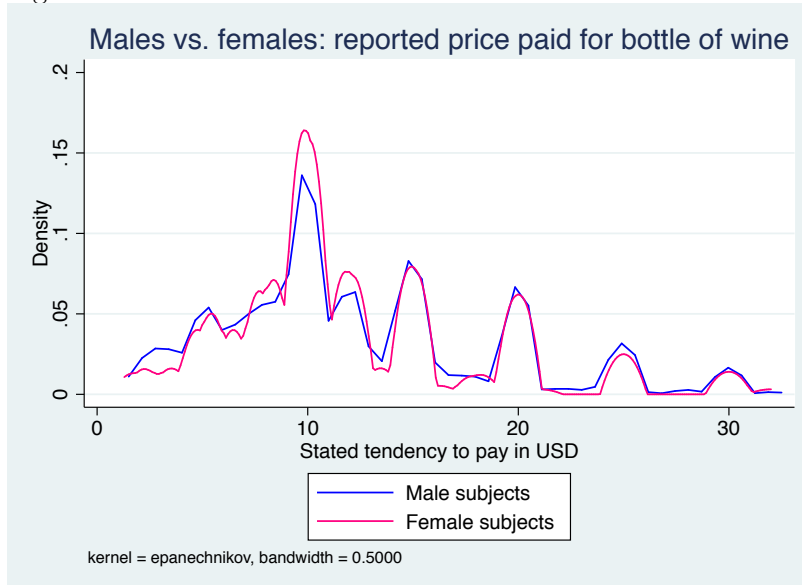


Figure 6.10w. Gender effects on STTP distribution for wine.



I do not report any statistical test results for these male-female prices: I am merely reporting the results that would be generated by the *Freq* function specified in Section 6.5.2 if generalized to measure modes on a continuous curve, in this case the kernel density curve. On a continuous curve rather than a histogram, *Freq* would simply be defined in terms of maxima and minima. The Epanechnikov kernel density function is an example of a quasi-continuous-looking price frequency curve.

Response-time effects.

Table 6.11 and Figures 6.12b and 6.12w show the effects of slower vs. faster response time on price responses in the survey. Google Surveys reports the time, in milliseconds, between when the survey question appears on the subject’s screen and the subject completes his or her response.

Table 6.11. Response time summary (in milliseconds, 1000 ms = 1 sec) on mean, minimum, maximum, and median STTP for beer and wine.

Time (ms)	Obs	Mean	Std. Dev.	Min	Max	Median
Beer	1045	28402.07	113564.2	3905	2963820	15581
Wine	1060	21299.95	39903.61	2783	809358	13344

Perhaps the most interesting aspect of Figures 6.12b and 6.12w is that there appear to be response time effects on STTP for beer but not on STTP for wine. A t-test verifies that the mean STTP differs significantly ( $p < 0.001$ ) between faster-than-median and slower-than-median responders. More interesting than this mean difference, perhaps, is the modal difference shown in Figure 3.12b: the modal STTP for slow responders is \$5, whereas the modal STTP for fast responders is \$8.

For wine, modal STTP is \$10 for fast and slow responders, although the density around the mode of \$10 is higher for slow responders than fast responders. Fast responders are more likely to report paying \$25 or \$30 per bottle of wine, but the shape of the distribution does not differ for fast vs. slow responders in the way that it does for beer responses. I do not currently have any good explanation for these differences, although many potential explanations are imaginable: for instance, stated tendency to pay more may correlate with impulsivity.

Figure 6.12b. Response time effects on STTP distribution for beer.

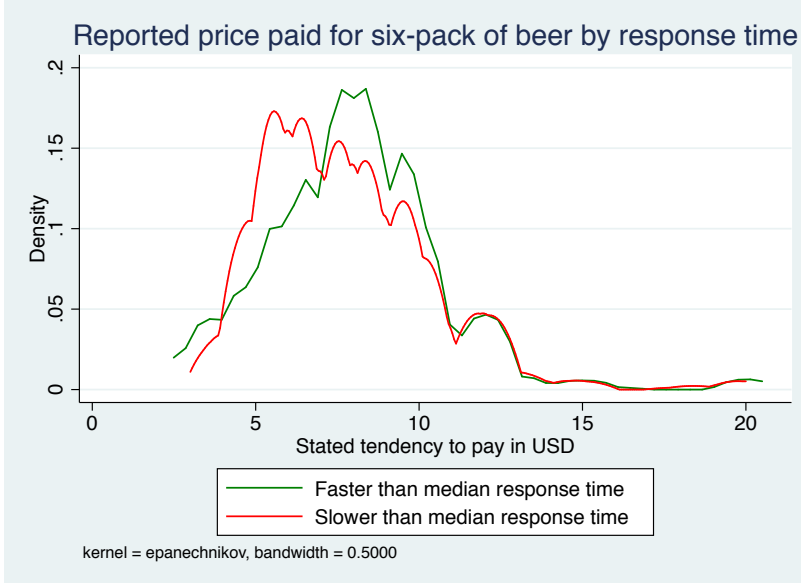
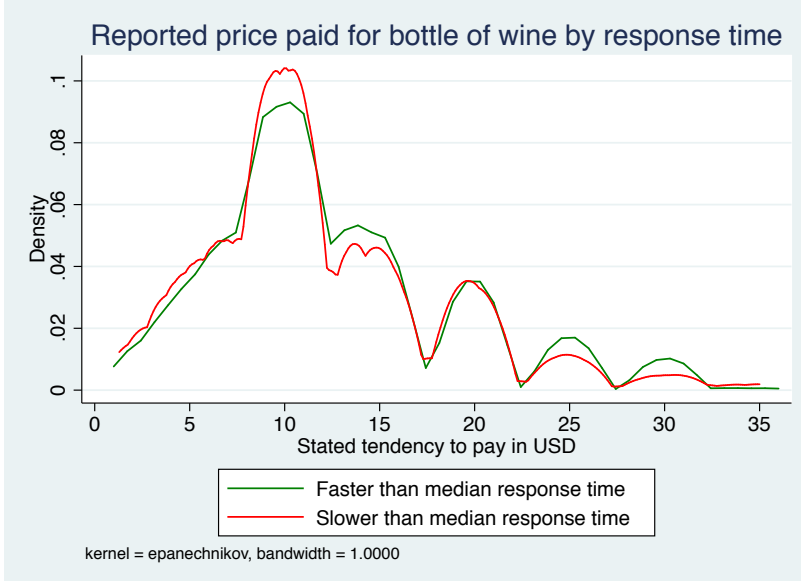


Figure 6.12w. Response time effects on STTP distribution for wine.



### Age effects

Tables 6.13b and 6.13w show the differences between mean, minimum, and maximum price for six different age groups. Not all subjects were identifiable as belonging to certain age groups, so about 200 observations are left out from age groups for both beer and wine.

Table 6.13b. Age effects on STTP means for beer.

Beer by Age	Obs	Mean	Std. Dev.	Min	Max
18-24	130	7.848	2.484	3	15
25-34	165	7.686	2.575	3	20
35-44	163	7.885	2.787	3	20
45-54	155	8.061	2.436	3	15
55-64	147	<b>8.245</b>	3.058	3	20
65+	79	8.115	2.780	3	20
All*	1045	7.905	2.652	2.99	20

\* Includes 206 subjects with unknown ages.

Table 6.13w. Age effects on STTP means for beer.

Wine by Age	Obs	Mean	Std. Dev.	Min	Max	W/B ratio**
18-24	112	12.446	5.910	2	30	1.586
25-34	197	12.266	6.838	1.5	35	<b>1.596</b>
35-44	157	12.459	5.983	1.5	35	1.581
45-54	162	12.402	6.381	2	35	1.539
55-64	149	<b>12.617</b>	6.657	1.5	32	1.530
65+	75	11.346	5.419	2	29.99	1.398
All*	1060	12.255	6.312	1.29	35	1.550

\* Includes 208 subjects with unknown ages.

\*\* (Wine STTP / Beer STTP for age group)

In both beer and wine groups, the 55–64 age group reported highest STTP. Figures 6.14b and 3.14w use kernel density plots to compare STTP for subjects aged 55–64 vs. subjects aged 18–54. (65-and-over subjects are omitted.)

Figure 6.14b shows that the difference in comparative means is driven by the existence of a much larger segment at \$12 for 55–64-year-olds than for younger consumers. Otherwise, the distributions look similar. Figure 6.14w shows that consumers 55–64 report buying more \$20 and \$25 wine than other age groups, but interestingly, 55–64-year-olds also report buying more \$5 wine, and less \$15 wine, than other age groups.

Figure 6.14b. Age effects on STTP distribution for beer.

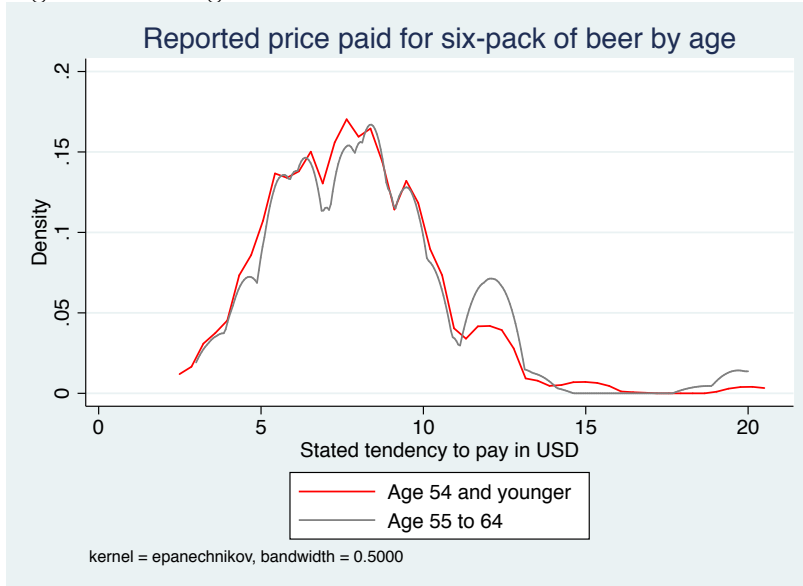
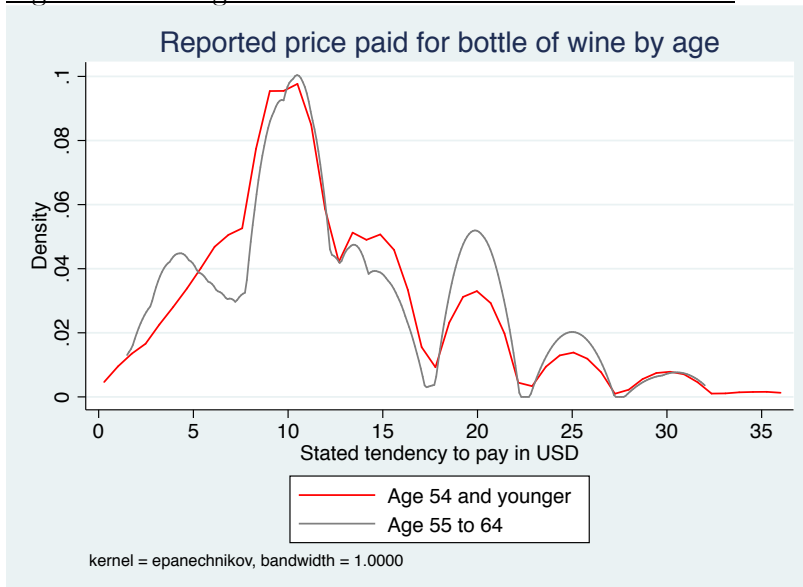


Figure 6.14w. Age effects on STTP distribution for wine.



### Income effects

Tables 6.15b and 6.15w show the differences between mean, minimum, and maximum price for five different household income groups. Not all subjects were identifiable as belonging to certain age groups, so about 40 observations are left out of income groups for both beer and wine.

Table 6.15b. Income effects on STTP means for beer.

Beer by inc.	Obs	Mean	Std. Dev.	Min	Max
< \$25K	75	8.565	2.974	4.99	20
\$25-50K	562	<b>7.601</b>	2.601	2.99	20
\$50-75K	267	7.931	2.648	3	20
\$75-100K	65	8.363	2.245	3	13
\$100K +	40	<b>9.173</b>	2.600	3.99	15.99
All*	1045	7.905	2.652	2.99	20

\*\* Includes 36 subjects with unknown household income.

Table 6.15w. Income effects on STTP means for wine.

Wine by inc.	Obs	Mean	Std. Dev.	Min	Max	W/B ratio**
< \$25K	86	11.821	6.480	2.99	30	1.380
\$25-50K	504	<b>11.815</b>	6.198	1.29	35	1.554
\$50-75K	282	12.271	6.000	1.5	35	1.547
\$75-100K	91	13.312	6.757	2	30	<b>1.592</b>
\$100 +	57	<b>14.377</b>	7.488	3	30	1.567
All*	1060	12.255	6.312	1.29	35	1.550

\* Includes 40 subjects with unknown household income.

\*\* (Wine STTP / Beer STTP for income group)

Predictably, in both beer and wine groups, the highest household income level (\$100,000 per year and above) is associated with the highest STTP (\$9.17 for beer, and \$14.38 for wine). More surprisingly, for both beer and wine, the second-lowest household income level (\$25,000-50,000 per year) is associated with the lowest STTP. This could reflect in



part the fact that people earning less than \$25,000 could include students who are partially supported by parents, loans, or scholarships. Compared with other people their age, college and graduate students come from relatively wealthier backgrounds, and may be spending money other than earned income.

Figures 6.16b and 6.16w use kernel density plots to compare STTP for wine and beer amongst subjects with household income above vs. below \$75,000 per year.

Figure 6.16b. Age effects on STTP distribution for beer.

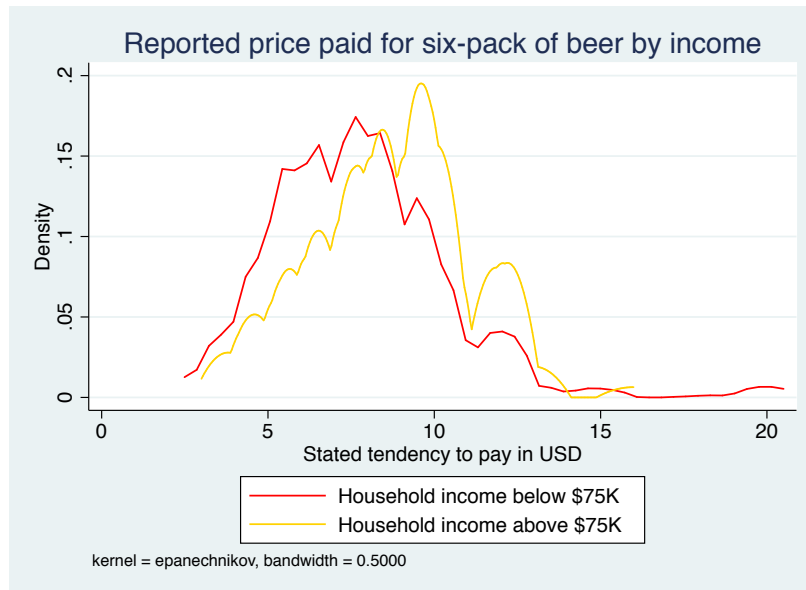
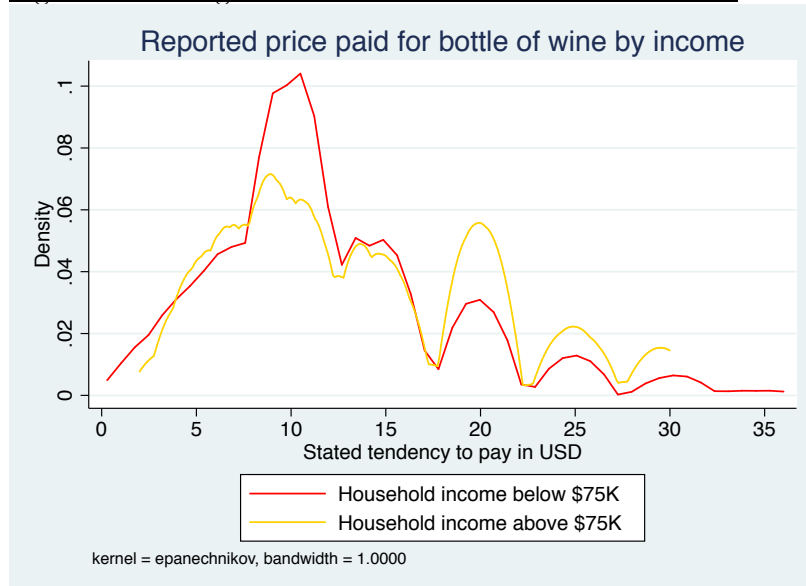


Figure 6.16w. Age effects on STTP distribution for wine.



For beer, the shift looks similar to the response-time shift: an overall shift in the whole distribution, from a distribution centered around a modal response of \$10 (for consumers with household incomes above \$75,000) to a distribution centered around a modal price of \$8 (for consumers with household incomes below \$75,000). For wine, however, the shift looks more like a categorical leap, with many consumers abandoning the \$10 price point in favor of the \$20 price point (and to a lesser extent, the \$25 and \$30 price points). This result brings further focus to the nature of the wine “premiumization” described more generally above.

#### Population density and regional effects

Tables 6.17b and 6.17w show regional effects, comparing STTP for urban, suburban, and rural population density areas. I observe an interesting inversion in order between beer and wine for urban versus rural consumers. Urban consumers report paying the most for beer and the least for wine. Rural consumers report paying the least for beer and the most for wine. In other words, the wine and beer order flips between urban and rural. Although there is no easy explanation for this phenomenon, it may be anecdotally observed that the “craft beer” movement is still mostly concentrated in urban areas. In many rural areas where craft beer might not be readily available, it might thus be possible to find premium-priced wine but not premium-priced beer.

Table 6.17b. Population density effects on STTP means for beer.

<b>Density:Beer</b>	Obs	Mean	Std. Dev.	Min	Max
Urban	325	<b>8.194</b>	2.752	2.99	20
Suburban	483	7.801	2.435	3	20
Rural	198	<b>7.748</b>	2.983	3	20

Table 6.17w. Population density effects on STTP means for wine.

<b>Density:Wine</b>	Obs	Mean	Std. Dev.	Min	Max
Urban	372	<b>11.849</b>	5.985	1.5	30
Suburban	481	12.420	6.257	1.29	35
Rural	176	<b>12.573</b>	7.095	2	35

Finally, Tables 6.18b and 6.18w show regional effects, comparing STTP for U.S. states in the Northeast, South, West, and Midwest. For both wine and beer, Northeasterners reported paying the most, and Midwesterners reported paying the least.

Table 6.18b. Regional effects on STTP means for beer and wine.

<b>Region:Beer</b>	Obs	Mean	Std. Dev.	Min	Max
Northeast	173	<b>8. 296</b>	2. 799	3	20
South	349	7. 816	2. 720	2. 99	20
West	204	8. 157	2. 446	3	20
Midwest	318	<b>7. 623</b>	2. 593	3	20

Table 6.18w. Regional effects on STTP means for beer and wine.

<b>Region:Wine</b>	Obs	Mean	Std. Dev.	Min	Max
Northeast	<b>191</b>	<b>13. 346</b>	6. 099	2	32
South	361	12. 765	6. 576	1. 29	35
West	246	11. 981	6. 182	1. 5	35
Midwest	261	<b>11. 026</b>	6. 026	1. 29	32

### **The natural occurrences of prices**

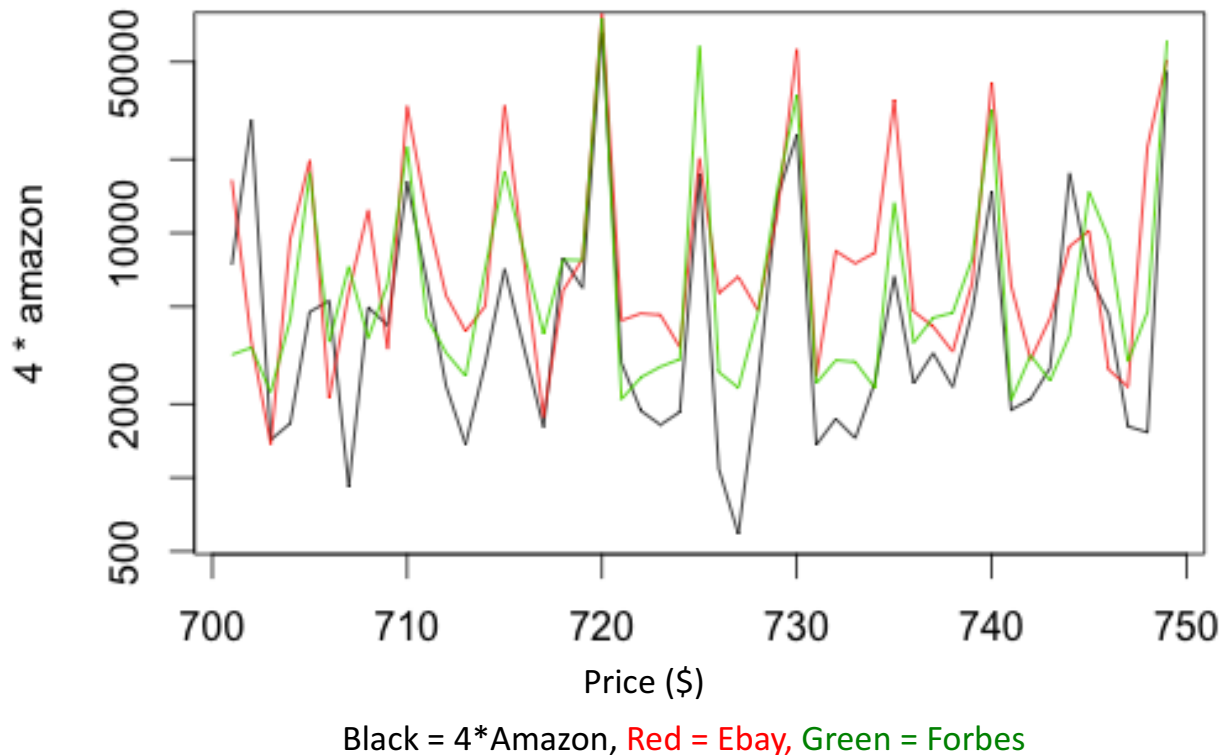
Above, I have attempted to show several different ways of approaching the multi-modality of beer and wine STTP. But what do price distributions look like in general? Can any baseline effects be established that underlie the price distributions of all goods? Do these have some things in common with the multi-modal effects that have been shown above in survey responses? What parts of the multi-modality can be attributed to the natural gravitation of survey respondents to common numbers like \$10, \$12, and \$15 versus less common numbers like \$11 and \$13?

### Frequency distributions of prices at online retailers

One initial way to approach this question is simply to look at frequency-distribution plots of the incidence of all prices occurring in whatever context on e-commerce sites. A simple

analysis of a relevant data set is shown in Fig. 6.19, which plots the frequency of integer prices (for any good in any category) on Amazon, Ebay, and Forbes between \$701 and \$749 in mid-2012, indicate the gravitation of prices toward round “reference numbers,” resulting in multi-modality in almost all price data sets.

Figure 6.19. Integer part of prices of goods on Amazon, Ebay, and Forbes between \$701 and \$749.



I chose the range from \$701 to \$749 completely arbitrarily, for illustrative purposes only. The main intuition guiding my choice was that this range was likely to be relatively unperturbed by noise that might result from proximity to zero or a power of 10 (e.g. \$999), and thus that the [\$701, \$749] range might reveal a clearer signal of the predictable or systematic digit effects that ripple out through the larger numbers indefinitely.

I cut off the endpoints of \$700 or \$750 because I expected them to be disproportionately higher than all other frequencies in the interval, and I wanted to compare the smaller frequency values within the interval against each other, not against their larger endpoints.

I hypothesize that prices generally gravitate toward human reference numbers based on three human effects: (1) organizing numbers around powers of 10; (2) organizing numbers that bisect the powers of 10 (i.e. numbers ending in 5); (3) rounding numbers after performing these bisections.

Consider the similarity between the price frequencies in the arbitrary “cut” of the price range from \$701 to \$749 (in one-dollar increments) shown in Figure 6.19. Here in three different data sets we see a similar nonlinear price frequency distribution with an wavelength of \$5, amplitude varying between 5,000 and 50,000.

Looking at the three data sets as approximations of the same curve, such curve would have a wave-like or spike-like form, with a set of global maxima that can be obtained as follows: Beginning with (700, 750), since the range does not include the endpoints (which are rounder base-10 numbers), we first bisect the interval, then round down to obtain 720, the global maximum of (700,750). Next we round up to obtain 730, local maximum #2 (ranked by amplitude). We see smaller peaks of roughly the same amplitude at the other two multiples of 10 (710 and 740) and the midpoint of the range, 725. Finally, we see significantly smaller peaks at the numbers ending with the digit 5.

### Benford’s Law

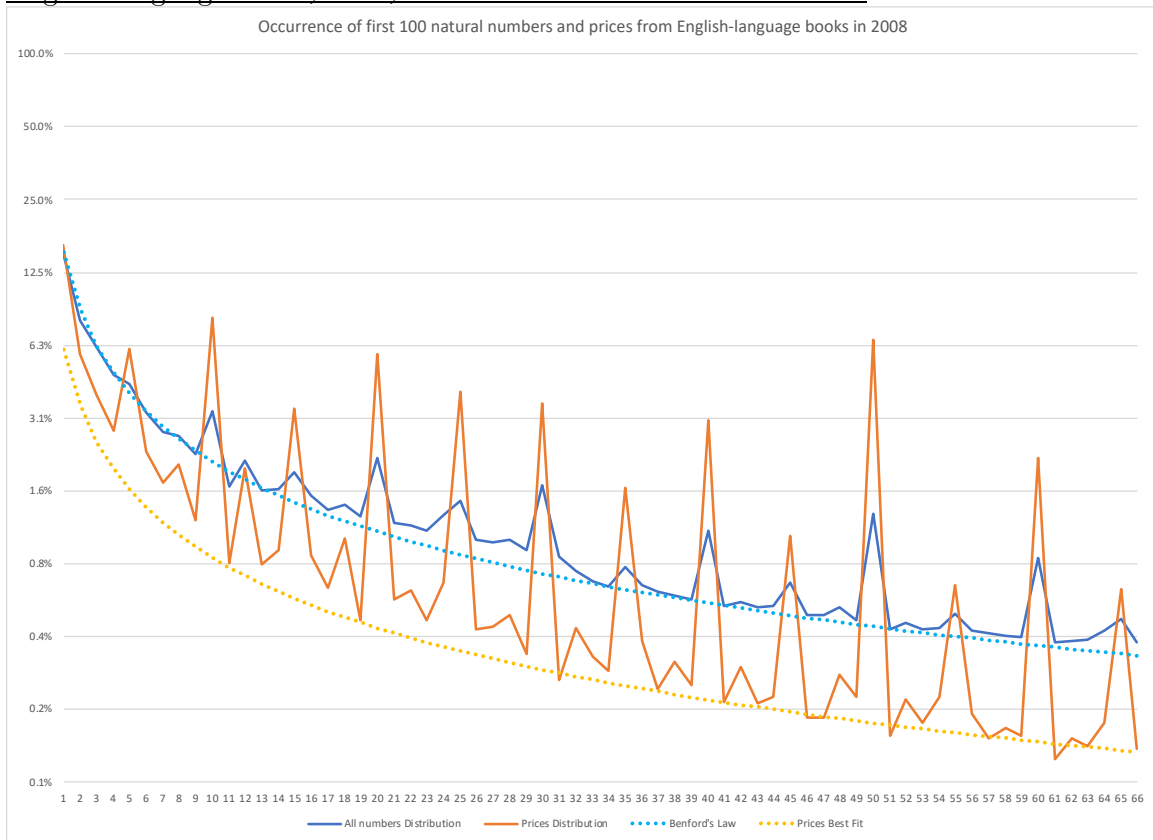
The “first-digit law,” or “Benford’s Law” (Benford, 1936), also known as the “first digit law,” holds that the probability that the base-10 expansion of a random integer has leftmost nonzero significant digit  $d$  is given by the following:

$$\Pr(d) = \text{Log}_{10}\left(1 + \frac{1}{d}\right), \text{ where } 1 \leq d \in \mathbb{N} \leq 9$$

How well do price incidence data from online marketplaces conform to Benford’s Law? Figure 3.6 plots the frequency distribution of the occurrence of each natural number (e.g. “68”) and prices (e.g. “\$68”) between 1 and 100 in the text of all English-language books published in 2008 and scanned by Google nGrams. In Figure 3.20, these natural occurrences are plotted against a curve generated by Benford’s Law (transformed by a scalar and a constant).

Figure 6.20 provides some visual evidence that Benford’s Law provides a baseline below which the occurrence of numbers never falls, but that it cannot explain the spike-like pattern that seems to be related to base-10 bisections (with multiples of 2, 5, and especially  $2 \times 5 = 10$  attracting far more than their Benford baseline occurrences).

Figure 6.20. Occurrence of the first 100 natural numbers and prices in the text of English-language books, 2008, and fit line based on Benford’s Law.



The general numbers and the prices appear to follow the same general patterns and baseline curve. The baseline price frequency curve may even be a linear fraction of the general number frequencies. Interestingly, however, the “spikes” at multiples of 2 and 5 are highly exaggerated on the price curve, spiking up above the general number curve at those round values. This suggests that base-10 effects are stronger on price frequencies than on number frequencies in general. (In other words, “\$50” is much more common than “50,” whereas “32” is much more common than “\$32”.)

Self-similarity of price frequency distributions

I conjecture that these patterns may demonstrate self-similarity when multiplied by powers of 10 (scaling up or down, perhaps, by some blanket multiplier). Some support for this conjecture is shown by the plot in Figure 6.21 of the frequency of prices on the English-language Internet as measured by millions of results or “hits” generated by a Google search for each price including a U.S. dollar sign before the number (e.g. “\$68”).

The number of hits for each integer price level was collected in the range from \$3 to \$30. The gray, yellow, and blue curves represent the frequencies of each base number, 10 times the number, and 100 times each number, respectively.

It is not difficult to see evidence of self-similarity from this pattern in Fig. 6.21. Can a function be defined analytically, perhaps recursively, to generate the correct order as measured against real-life data sets? The function would be self-similar, nowhere differentiable, everywhere discontinuous. It would be a castle splayed across the x-axis like a cartoon crown, its walls penetrated by infinitely ragged complexity.

Figure 6.21. Frequency of prices on the English-language Internet.

### 6.21a. Natural multi-modality? Beer and wine STTP and Google price frequencies



#### Neurobiological basis for base-10 and number cardinality effects

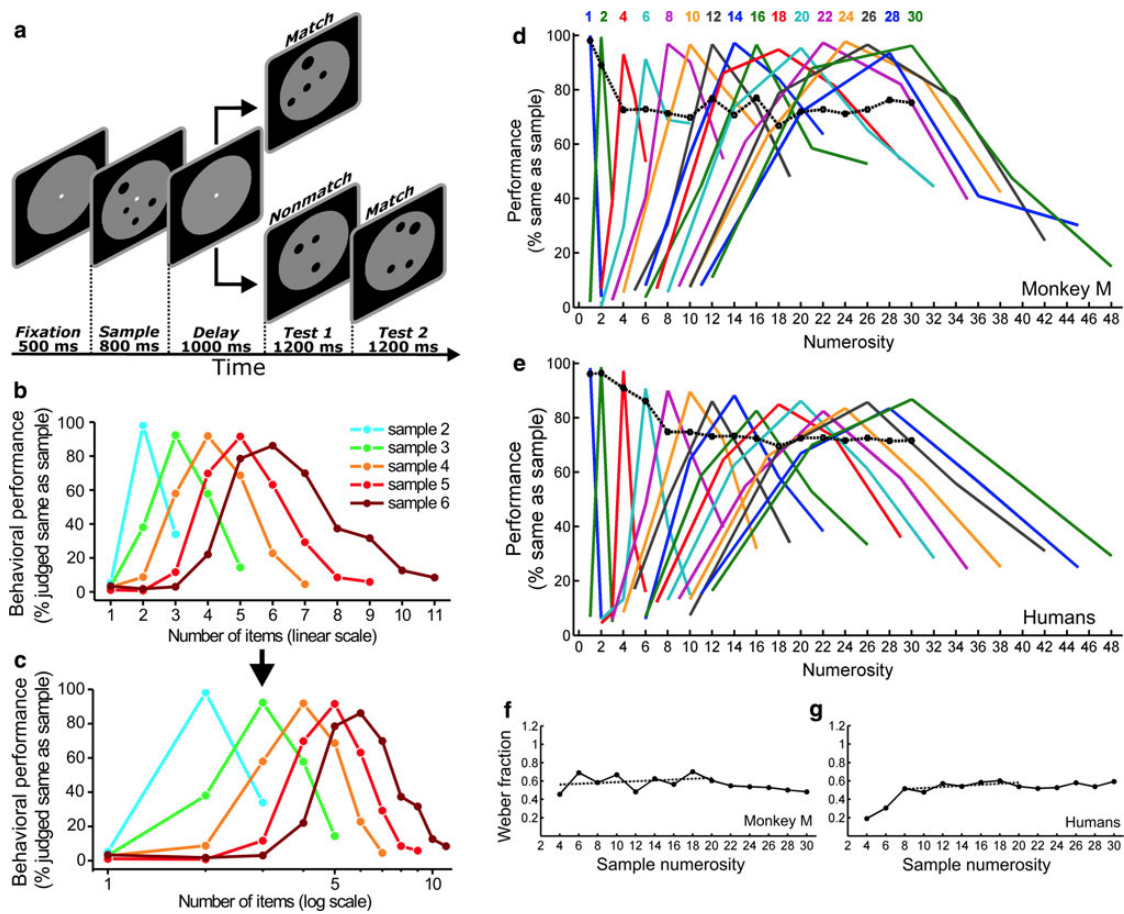
I hypothesize that human salience effects are efficiently adjusted this way for maximum sensitivity across ranges of most interest. In cognitive neuroscience, the human faculty for processing numerical information in this manner is known as the “Approximate Number System” (ANS), which is thought to develop during early childhood (Starr et al., 2018; see also Dehaene, 1997). Transduction efficiency is maximized by taking “cuts” of round numbers. Rounding and bisection heuristics may be evolutionary adaptations in the cerebral cortex—perhaps not only in human beings, but also in other primates.

Nieder et al. (2012), as shown in Fig. 6.22, provides evidence from numerosity experiments with macaque monkeys that primates beyond the human species may have

this same wiring efficiencies, and perhaps thus the same tendencies toward bisection and rounding. The mammalian brain has processing networks of neurons where each individual neuron is trained to respond to groups of a particular number of similar objects. For instance, fascinatingly, one set of neurons spiked more rapidly when five objects (five of anything) were placed in front of the monkey; a different set of neurons responded instead to six of anything; and so on. Monkey see, monkey count.

For more math with monkeys, see Chapter 8.

Figure 6.22. Cardinal number processing in the brain.



Source: Reprinted from Nieder et al., 2012.



## Chapter 7. The stench of cheap perfume<sup>34</sup>

*A half-blind tasting method for comparing placebo and nocebo responses to wine bullshit*

In this chapter, 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 we measure consumers’ placebo (positive) and nocebo (negative) responses to wine prices and labels, in a laboratory setting, controlling for sensory characteristics, without using deception? (2) If consumers’ placebo and nocebo responses to prices and labels are measurable, how do their magnitudes compare? (3) Do placebo and nocebo responses to prices and labels behave like gains and losses in prospect theory?

I control for the sensory characteristics of the wine using an experimental setup that does not require deception: 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 down 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 participated in a half-blind tasting. Subjects were sorted into two treatment groups. In the “placebo” treatment, the identical wines are worth \$50, generating a positive bias and swaying tasters toward the exposed bottle. In the “nocebo” treatment, the identical wines are worth \$5, generating a negative sensory bias and swaying tasters toward the mystery bottle. Deception is not used with either treatment; the (non-blind) wines’ actual prices are shown. I find statistically significant placebo and nocebo response, in spite of the small sample size. Comparing the two treatments, 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).

### Motivation

The challenge for any experimental economics method aimed at separating placebo and nocebo responses to packaging and invisible attributes is how to avoid deception in the experimental set-up. Plassmann et al. (2008)’s fMRI work with Antonio Rangel at Cal Tech is, for me, the canonical statement to date on placebo and nocebo responses to consumer prices. What was revolutionary about Plassmann et al. (2008) was the

---

<sup>34</sup> Submitted to the *Journal of Wine Economics*, September 2019.

observation of significant positive and negative price effects not only in differences between subjects' self-reported hedonic ratings, but also in the activation of subjects' brain areas associated with pleasure and reward (through fMRI-measured blood flow).

In Plassmann et al. (2008), subjects were deceived about the prices of wines they tasted (for instance, by being told a \$40 wine was worth \$5, or vice-versa). Deception is generally allowed in fMRI studies such as Plassmann's, and in other areas of neuroscience as well. The judgment of the formidable *Proceedings of the National Academy of Sciences*, which published the Plassmann paper, was that her deception methodology was completely warranted given the necessity of generating a signal strong enough that it could be detected by fMRI machines. The subsequent fanfare the Plassmann paper received vindicated *PNAS's* good judgment that deception was a useful methodology. In the end, a whole new field of research was created by the unique insights yielded by that paper. For instance, it was Plassmann et al. (2008) that sparked my own first interest in economics, a decade ago, while also rekindling an undergraduate interest in neuroscience that I had left behind still another decade earlier.

The field of economics diverges from neuroscience, psychology, and other fields in generally frowning upon the use of deception in experiments. The justification generally given for this policy, is a "public good" argument, vaguely analogous to the policy against excessive antibiotics: that the scientific profession doesn't want to pollute its collective subject pool with a bunch of skeptical subjects who expect to be lied to, and will thus show up more suspicious and less gullible to their next experiment.

The experimental economist Julian Jamison and his colleagues explain the principle as follows: "This rule exists 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" (Jamison et al., 2008; see also Levitt and List, 2007).

In an experiment that purposefully deceives participants in order to test the effects of deception on future subject behavior, Jamison et al. find some support for the rule as applied to laboratory experiments: "significant differences in the selection of individuals who return to play after being deceived as well as (to a lesser extent) the behavior in the subsequent games" (Jamison et al., 2008).

On the other hand, a movement against the anti-deception principle appears to be underway in some sub-fields of economics. For instance, the use of deception was the focus debate topic for the February 2019 issue of *Food Policy*. Jayson Lusk, in that issue, argues that the field is ready to move past this norm:

“Deception in economic experiments became taboo because of the context in which early experiments were conducted. Several decades ago, when experimental economics began gaining traction as a distinct method and sub-discipline, the typical experiment involved student subjects who interacted in an induced value or other abstractly framed decision environment in a computer lab on campus. In this environment, the ban on deception had a rationale explained in early texts on the subject. Given that the empirical research suggests spill-over or public-good type effects from deception in experiments appears to be relatively small (Ortmann and Hertwig, 2002), profession-wide policies to protect a common pool resource of trust in the profession are hard to justify. My guess is that the psychology discipline is no less trusted generally than economists despite the fact that their discipline permits deception in experiments” (Lusk, 2019).

As Lusk (2019) suggests, there may be a place for both deceptive and non-deceptive research in economics, as there is in the other social sciences. But for now, the anti-deception policy remains an industry norm. 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 chapter, I propose a novel non-deceptive methodology for eliciting placebo and nocebo responses to non-sensory attributes such as price. My method and experimental design are intended to satisfy the demands of even hard-line anti-deceptionists. I call my method “half-blind tasting.”

The larger motivation for the half-blind tasting methodology comes from an important line of research that precedes it by more than a decade. The results of Plassmann et al.’s (2008) fMRI experiment on the influence of wine price information, Lee, Frederick, and Ariely’s (2006) experiment on information about vinegar in beer, and others have suggested that visible packaging information (a cheap- or expensive-looking wine bottle and label) and invisible cognitive information (a low or high price), after being revealed to a consumer and subsequently processed by language, number, and other cognitive and emotional areas in the brain, can significantly add or subtract from the desirability of the experiences associated with products. Non-sensory information may even impact a consumer’s primary sensory experience (i.e., where a sensory experience is generated by a non-sensory input).

The initial question I pose above is connected to some larger and perhaps more interesting ones, such as these: When people spend more money for a wine, do they get their money’s worth, even if they couldn’t tell the difference between the cheaper and the more expensive wine in a blind tasting? How might the presence or absence of packaging and invisible attributes influence pairwise choice, hedonic ratings, and tendency to pay in a blind wine tasting experiment? Finally, might parallels to prospect theory be drawn by viewing the revelation of high and low prices as gains and losses?

In the remainder of this chapter, I describe the experimental design for “half-blind pairwise choice,” I report the results of my first modest-sized pilot experiment aimed at testing the method, and I discuss some possible improvements and implications.

### Design

The experimental design I propose is what I refer to as a type of “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 for half-blind pairwise choice 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 in 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. As a reference point, I also elicit how much subjects typically spend on wine, and for other comparisons, I ask for gender, age group, frequency of drinking 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 used to elicit preferences is shown in Figure 7.1. Subjects were not debriefed after participating.

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.

Figure 7.1. 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: \$\_\_\_\_\_

Data

My experiment included 53 subjects, who were randomly divided into two groups based on what time in the afternoon they took the test. 18 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 the nocebo response by disproportionately choosing the brown paper bag over the exposed bottle. There were fewer subjects in the “placebo” version than in the “nocebo” version due to budget constraints.

## Results

The main results for the pilot study are shown in Table 7.2. 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 top row (“All wine drinkers”) of each condition (Low Price “Nocebo” and High Price “Placebo”), in their pairwise-choice answers, 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 (“Mean WTP” on Table 7.2), 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 top 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.<sup>35</sup>

Table 7.2 also reports some means comparisons between groups based on gender, self-reported frequency of wine drinking, and STTP (“stated tendency to pay”; see Chapter 3), 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 female subjects (19), and the considerable selection bias in the subject pool (U.C. 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).

---

<sup>35</sup> 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 7.2. Results of Berkeley Half-Blind Pairwise Choice Experiment.

7.2a. Comparison 1: Cheap wine (Nocebo response)

Wine with \$5 price label vs. identical wine with unknown price

	n=	<u>Known price = \$5</u>		<u>Unknown price</u>		Price signal effect size
		Prefer in pairwise choice	Mean stated WTP	Prefer in pairwise choice	Mean stated WTP	
All wine drinkers	35	<b>0.257***</b>	\$6.75	0.742***	\$8.89	<b>-\$2.19**</b>
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 on wine	12	0.250**	\$9.00	0.750**	\$10.67	-\$1.67

\*\*\* 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.

7.2b. Comparison 2: Expensive wine (Placebo response)

Wine with \$50 price label vs. identical wine with unknown price

	n=	<u>Known price = \$50</u>		<u>Unknown price</u>		Price signal Effect size
		Prefer in pairwise choice	Mean stated WTP	Prefer in pairwise choce	Mean stated WTP	
All wine drinkers	18	<b>0.592</b>	\$15.89	0.421	\$11.44	<b>\$4.78***</b>
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 on wine	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.

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.

### 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 in order to observe the effects of price and label separately.

Amongst all objections to my design, one that was raised by Matthew Rabin at Berkeley stands out as the most serious. The problem is 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 would have reduced power to untenably low levels. Not a single one 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.

### Discussion and future directions for research

The main purpose of this chapter 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 53-subject pilot experiment and mean differences showing the expected signs of effect, I conclude from the experiment that the method is worth exploring further: subjects as a



whole exhibit both placebo and nocebo responses that are strong enough, 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 different price points could therefore 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 7.3.

In Kahneman and Tversky’s prospect theory—as elaborated, for example, in Kahneman and Tversky (2013)—subjects in a lab experiment defy expected-value rationality by over-protecting against losses in a systematic (nonlinear, in a roughly 2:1 loss:gain slope ratio if linearized) fashion that is illustrated, in idealized form, by the familiar prospect theory curve shown in Figure 7.4.

Comparing the upward-sloping curves and their kinks in Figures 7.3 and 7.4, 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 7.4 would be analogous to the wine-price axis in Figure 8.3. The “value” (Y) axis in Figure 7.4, meanwhile, could be viewed as analogous to the percent choosing, the Y-axis in Figure 7.3.

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.

Figure 7.3. Idealized form for placebo/nocebo curve constructed from half-blind tasting results at a number of different price points.

*Percent of consumers preferring a wine labeled with price to an identical unlabeled wine, based on half-blind tasting results at several different price points (populated with hypothetical data)*

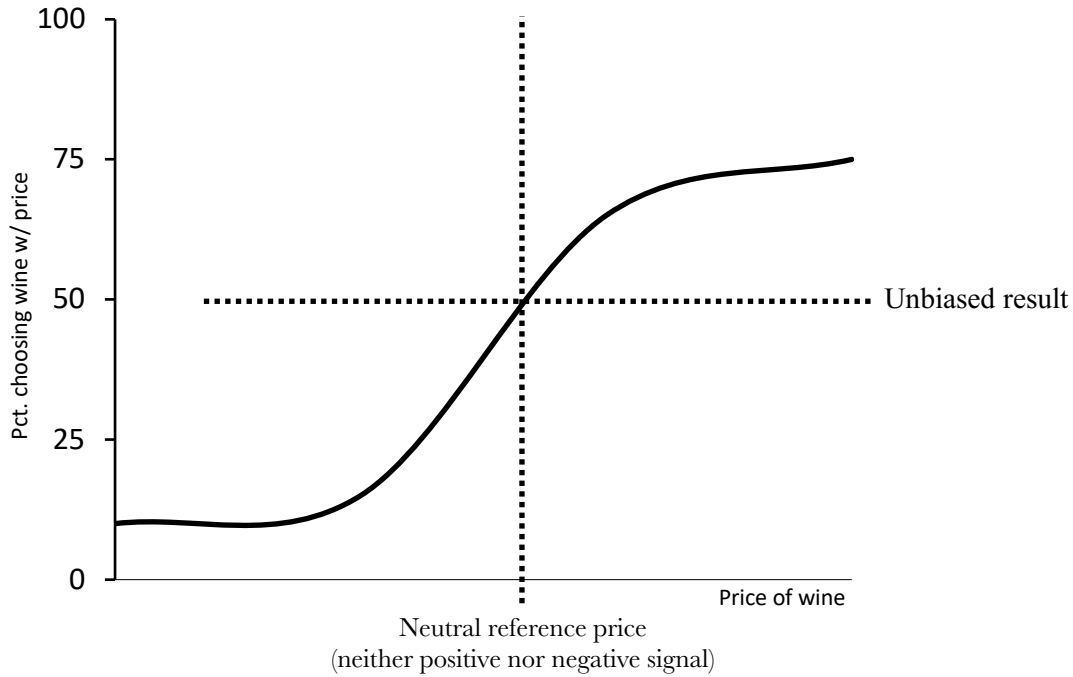
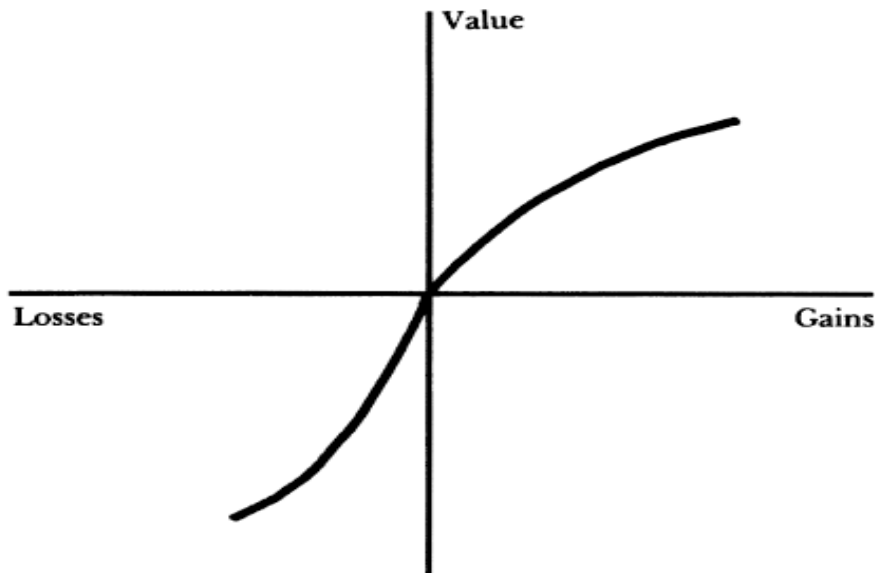


Figure 7.4. Kahneman and Tversky's prospect theory curve.



Source: Reprinted from Kahneman and Tversky (2013).

## Chapter 8: Capuchin monkeys do not show human-like price effects<sup>36</sup>

*Four lab experiments to test whether demand responses to price bullshit are uniquely human*

Congratulations! You have just won a bottle of wine. You have two options: a 2001 pinot noir that costs \$10 or another pinot noir from the same year that costs \$50. Which do you choose? You probably chose the more expensive wine. Indeed, when given a choice like this, people tend to pick the most expensive options, whether those options involve alcohol (e.g., Jacoby et al., 1971; Plassmann et al., 2008), meat (Makens, 1964; Bello Acebrón, and Calvo Dopico, 2000), or even cassette players (Dodds and Monroe, 1985, see Rao and Monroe, 1989 for a review).

What's more surprising, however, is that our preferences for more expensive goods seem to hold even in cases where the price is arbitrary. For example, Plassmann et al. (2008) allowed participants to sample the same wine when it was labeled as either inexpensive (\$5 or \$10) or expensive (\$45 or \$90). Participants reported greater experienced pleasure for wines that were labeled as more expensive, even though what they actually drank was the same in both cases. These results suggest that merely labeling a good as more expensive seems to affect the subjective utility a person experiences from that good.

One might be tempted to write off such findings as the result of a strange demand characteristic; perhaps participants self-report that they enjoy expensive options more in order to signal that they're the kind of person who prefers expensive things. However, some evidence suggests that pricing effects may run deeper than mere demand characteristics. First, Plassmann et al. (2008) found that price affected participants' preferences at the neural level; they found that activation in the medial orbitofrontal cortex (mOFC)—an area of the brain thought to encode the subjective reward utility of different stimuli (see review in Levy and Glimcher, 2012)—was higher when participants thought the wine was expensive than when they thought the wine was cheap.

These results suggest that participants actually *experienced* the wine as tasting better when it was labeled as expensive than when it was labeled as inexpensive. Second, pricing influences how effective people think a good will be. Shiv et al. (2005) allowed participants to pay different prices for energy drinks and observed how well they

---

<sup>36</sup> Co-authored with Rhia Catapano, Nicholas Buttrick, Jane Widness, and Laurie R. Santos (all of Yale University Department of Psychology, Comparative Cognition Laboratory). Revised from its original form, published in *Frontiers in Decision Neuroscience* as Catapano et al. (2014). We are grateful to the following students for their help with this project: Kimberly Ashayeri, Melissa Baranay, Cliff Bielinski, Elizabeth Brim, Linda Chang, Nicki Cohen, Lindsay Davis, Daniel Friedman, Jasmeet Jernaill, Mia Kunst, Joyce Li, Emily Rosenberg, and John Mark Taylor. This study was supported by a McDonnell Scholar Award to Laurie Santos and by Yale University.

performed on a set of mental acuity puzzles. People who had paid more for the drink showed greater energy-boosting effects than those who got the drink more cheaply (for a similar finding, see Waber et al., 2008). This result further suggests that pricing effects appear to go beyond mere self-reported differences in preferences; simply changing the price individuals pay for a drink affects not only how well they think it works, but also its actual effectiveness.

Although much work has shown that price affects people's expectations about a good (see also Rao and Monroe, 1989 for a review), less work has explored how such effects emerge in the first place. One possibility is that our expectations concerning price information stem from our experience with how markets operate. As any economics major knows, markets tend to conform to the rules of supply and demand. People prefer products that are particularly good or effective, and thus demand for such effective products should increase. Companies, therefore, will likely end up charging more for products that are especially effective due to the higher demand for such products. This relationship means that better tasting and more effective products are likely to be more expensive.

One possibility, then, is that our experience with markets causes us to develop an association between price and value; we come to implicitly assume that expensive products must actually be valuable because otherwise sellers would have to lower their prices. In this way, one could explain the expectations we described above as an extension of our experience with human-like markets.

It is also possible that our experience with markets allows us to develop more explicit theories about how markets work: we may come to develop rich beliefs about the connection between price and value based on our own understanding of markets. A second possibility, however, is that our expectations about the connection between price and value have nothing to do with our experience in markets. Instead, our preference for more expensive items may stem from more domain-general processes, ones that are not specific to monetary values or markets.

Are our expectations regarding price and value motivated by a domain-general mechanism or by experience with human-like markets? One way to distinguish between these alternatives is to test a population that does not have rich experience with human markets. Here, we test just such a population—capuchin monkeys (*Cebus apella*).

Although turning to capuchins may seem at first glance a strange way to test the mechanisms underlying pricing effects in our own species, there are several reasons why this population is well-suited for this question. First, researchers have successfully used capuchins as subjects in economics studies examining the origins of judgment and decision-making biases (Chen et al., 2006; Egan et al., 2007, 2010; Addressi et al., 2008; Lakshminarayanan et al., 2008, 2011).

In many of these studies (e.g., Chen et al., 2006), monkeys were trained to trade tokens with a human experimenter for different kinds of food. Monkeys were then allowed to enter a market in which they had to choose between different experimenters who sold different goods at different prices. Using this market method, researchers have observed that capuchins appear to have many of the same strategies and biases as humans (see review in Santos and Chen, 2009). Like humans, capuchins exhibit endowment effects (Lakshminarayanan et al., 2008), loss aversion (Chen et al., 2006), reflection effects (Lakshminarayanan et al., 2011), and choice-induced preference reversals (Egan et al., 2007, 2010). Given that capuchins show many of the same economic biases as humans, it makes sense to examine whether this species shows pricing biases as well.

In addition, recent work suggests that capuchins seem to understand some aspects of price in the context of their experimental market. Chen et al. (2006) tested whether monkeys' choices in their market obeyed the tenets of standard price theory (see Becker, 1962). Capuchins were asked to allocate a set of tokens across two different kinds of food (e.g., apples and grapes) at a cost of one token per food item. Chen and colleagues then introduced a compensated price shift, in which the price of one of the goods dropped (e.g., a subject now received two apples per token rather than one). The researchers then tested whether monkeys switched their consumption after this compensated price shift; did monkeys buy more of the cheaper good after the price change? Chen and colleagues observed that subjects attended to price information, buying more of the cheaper good after the price shift. These results suggest that monkeys' choices in this market obey standard price theory, and thus that monkeys attend to price information in their market in some of the same ways as human consumers do in real markets.

Because capuchins appear to understand price information in token markets, this species can provide a particularly useful test case for distinguishing between the two different accounts of pricing biases described above. Although capuchins seem to understand certain aspects of pricing information in a token economy (e.g., Chen et al., 2006), they lack human-like experience with how price works in real markets. The capuchin token economies differ greatly from those of human participants, particularly with regard to the connection between a good's potential value and its price. If human-like market experiences are indeed necessary for the development of an association between price and value, then capuchins should not show the same kinds of pricing effects as humans do.

In the current studies, we developed a series of experiments to determine whether capuchins show human-like price biases. Capuchin subjects were taught the price of two novel foods in the context of their token economy (see Chen et al., 2006). We then assessed monkeys' preferences for the two goods in the absence of tokens (i.e., during free

choice). If monkeys exhibit human-like pricing effects, then they should prefer the more expensive good to the cheaper good when they have a chance to freely choose without paying with their tokens.

## **Experiment 1**

Experiment 1 began by teaching capuchins prices for two new foods: differently colored flavored ices. We then allowed monkeys to freely choose between the two colors and tested whether they spontaneously preferred the good that we had told them was more expensive.

We tested seven brown capuchin monkeys (AG, AH, FL, HR, MD, MP, NN) from the population at the Comparative Cognition Laboratory at Yale University (New Haven, CT). All monkeys had participated in a variety of experiments involving making decisions in their token economy (Chen et al., 2006; Lakshminarayanan et al., 2008, 2011). All studies were approved by the Yale Institutional Animal Care and Use Committee.

### Experiment 1: Design

Testing was conducted in a cubical testing chamber (75 x 75 x 75 cm), which monkeys entered via a sliding door attached to their main large social enclosure. Only one monkey was allowed into the testing area at a time. Monkeys were free to walk into an adjacent section of the enclosure during testing where no other monkeys were present. Two panels on opposite sides of the testing enclosure allowed participants to interact with the experimenters. Each of the panels had two trading holes (5 x 9 cm), spaced such that the participants could reach through one but not both of the openings at the same time (approximately 25 cm apart).

### Experiment 1: Materials and Methods

During testing, subjects were allowed to choose different foods presented on a small table with a sliding component that was hooked to the outside of the testing enclosure. The sliding component had two trays each positioned to line up with a trading hole when slid up to the chamber, allowing the subject easy access to their contents. Monkeys were presented with 12 tokens (1-inch-diameter aluminum disks) which they could use to “purchase” small Flavor-Ice ice chunks (2.5 x 1 cm) of different colors (orange and blue).

Before testing began, each monkey was given one piece of each ice color in order to familiarize them with the taste. The order of color presentation was counterbalanced across monkeys. After familiarization with the different flavors, monkeys began the study.

All monkey subjects began on an initial preference phase, followed by a price learning phase, and then a preference assessment phase (see Figure 8.1 for more details).

In the initial preference phase, subjects were given a choice between equal quantities of the blue and orange ice. The goal of this phase was to assess monkeys' initial preference between the two colors of ice. We also wanted to be sure that monkeys did not have strong aversion to one ice color; those monkeys who showed significant preferences were removed from the study.

Note that because we did not introduce any tokens during this phase, monkeys had the chance to sample the ices before learning about their prices. During each trial, monkeys were given a choice between the two ices. To do so, the experimenter slid the sliding component away from the testing chamber, placed each piece of ice on its respective tray, and then pushed the sliding component back to allow the subject to choose. The subject then selected one of the two ices; after the subject made its choice, the sliding component was withdrawn and the other ice was removed. We presented subjects with two sessions of 12 trials each, with ice placement counterbalanced across trials. Sessions were run on separate days in order to prevent ice satiation.

After the initial preference phase, subjects moved on to the price learning phase. The goal of this phase was to use the token economy to teach monkeys that one of the two goods could be bought at a discount relative to the other. On each trial, monkeys had a chance to give a token to an experimenter who would return either one piece of one color of ice (the expensive ice), or three pieces of the other color of ice (the cheap ice). Note that the expensive color of ice was priced at three times the price of the cheap color of ice. We chose this difference in price because previous work has shown that monkeys are able to distinguish items that are three times the value of other items (e.g., vanMarle et al., 2006). We conservatively chose which ice was cheaper based on monkeys' initial preferences; monkeys who indicated an initial preference for orange ice (AG, FL, MD) were taught that orange ice was cheap while those that indicated a preference for blue ice (AH, HR, MP, NN) were taught that blue ice was cheap. Testing proceeded as in previous token trading procedures (e.g., Chen et al., 2006; Lakshminarayanan et al., 2008).

Specifically, subjects could "purchase" one of the two ices by placing a token into the hand of an experimenter. At the start of each trial, the experimenter placed his hand open to receive the monkey's token. At the same time, he displayed a small dish holding the amount of ice to be traded. Upon receiving a token from the monkey, the experimenter brought the dish up to allow the monkey access. Once the monkey had eaten all the ice, the dish was reloaded, and the next trial began. Each monkey received two sessions of 12 trials with the order of the first color counterbalanced across sessions.

Figure 8.1. Description of different experimental conditions and phases

Expt #	Phase	Purpose	Tokens Used?	Choice Available?
Expt 1	Initial Preference Phase	To assess initial preferences between two goods	X	✓
	Price Learning Phase	To teach an initial set of prices for two goods	✓	X
	Preference Assessment Phase	To test whether the changed prices affect preference	X	✓
Expt 2: Exposure	Exposure Phase	To expose monkeys to two goods	X	X
	Price Learning Phase	To teach an initial set of prices for two goods	✓	X
	Preference Assessment Phase	To test whether changed prices affect preference	X	✓
Expt 2: No Exposure	Price Learning Phase	To teach an initial set of prices for two goods	✓	X
	Preference Assessment Phase	To test whether difference in prices affects preference	X	✓
Expt 2: Price Shift Control	Initial Purchasing Preference Phase	To assess initial preferences between two goods when prices are equal	✓	✓
	Price Shift training Phase	To teach a new set of prices	✓	X
	Price Shift Assessment Phase	To assess whether new prices shift purchasing decisions	✓	✓
Expt 3: Price Shift Control	Initial Purchasing Preference Phase	To assess initial preferences between two goods when prices are equal	✓	✓
	Price Shift Training Phase	To teach a new set of prices	✓	X
	Price Shift Assessment Phase	To assess whether new prices shift purchasing decisions	✓	✓
Expt 3: Price/Delay Conditions	Price/Delay Learning Phase	To teach an initial set of prices (in tokens or delay) for two goods	✓	X
	Price/Delay Preference Assessment Phase	To assess whether differences in price affect preference	X	✓
Expt 4: Price/Delay Conditions	Price/Delay Learning Phase	To teach an initial set of prices (in tokens or delay) for two goods	✓	X
	Price/Delay Preference Assessment Phase	To assess whether differences in price affect preference	X	✓
	Price/Delay Manipulation Check Phase	To assess whether monkeys attend to brand characteristics	✓	✓

Following the price learning phase, monkeys moved on to the preference assessment phase. The goal of this phase was to see if learning the price of the two ices affected the monkeys' preferences for each of the two colors. To test this, we gave the monkeys a free choice between the two colors (i.e., they could take the ices without having to purchase them using their tokens). The preference assessment phase was identical to the initial preference phase; monkeys again received a free choice between the two colors on the sliding trays. Importantly, the experimenter presenting the trays to the monkey was blind to which ice color had been cheap and which had been expensive, and thus could not influence the monkeys' choices. Assuming monkeys were indifferent between the two ice colors initially, we could test whether monkeys show human-like price effects by examining whether they reliably chose the color shown to be more expensive in the price learning phase when choosing in the preference assessment phase.

### Experiment 1: Results

We first tested to see if monkeys had an initially strong preference for one of the two colors of ice in the initial preference phase. Two of the monkeys showed a strong and significant initial preference (HG: 12.5%,  $p = 0.0003$ , JM: 25.0%,  $p = 0.02$ ) and thus were dropped from further testing. All other monkeys did not show a significant



preference (FL: 29.2%,  $p < 0.064$ ; AG: 37.5%,  $p = 0.31$ ; AH: 41.7%,  $p = 0.54$ ; HR: 50.0%,  $p = 1.00$ , MD: 37.5%,  $p = 0.31$ ; MP: 29.2%,  $p < 0.064$ ; NN: 41.7%,  $p = 0.54$ ; all tests exact binomial probability estimates against chance), suggesting that these subjects had no initial preference between the two colors and thus could be used as subjects in the subsequent phases.

Monkeys then went on to learn about the price of the two ices during the price learning phase. Note that in this phase monkeys had no choices—they received equal numbers of trades across the cheap and expensive goods. After learning about the prices, we again tested the monkeys' preferences in the preference assessment phase. Specifically, we compared monkeys' preference for the expensive color in the test condition with their preference for that color before they learned the price information. To do this, we used a repeated measures ANOVA with the color used as expensive as a between subject variable (orange or blue) and time of choice as a within subject variable (before price information and after price information). We observed no main effect of color [ $F(1, 5) = 0.328$ ,  $p = 0.591$ ], suggesting that monkeys had no strong preference for one color over another. We also, however, observed no main effect of timing [ $F(1, 5) = 0.527$ ,  $p = 0.50$ ]. Monkeys' preferences did not change after learning that one item was more expensive. Non-parametric tests confirmed this finding that there was no effect of price training (Wilcoxon signed rank:  $Z = 0.94$ ,  $p = 0.35$ ). We also observed no interaction between color and price [ $F(1, 5) = 0.004$ ,  $p = 0.95$ ].

### Experiment 1: Discussion

In Experiment 1, we taught monkeys that one color of ice was three times more expensive than the other, and explored whether monkeys subsequently preferred the more expensive kind of ice when they had free access to both options. In contrast to what's often been observed in humans (Plassmann et al., 2008), monkeys showed no preference for the more expensive ice. Learning which kind of ice was more expensive in the price learning phase did not seem to affect monkeys' preferences in the preference assessment phase. This result suggests that learning that a food is expensive doesn't seem to make monkeys like it more.

## **Experiment 2**

There are a few problems with Experiment 1. The first concerns whether monkeys noticed the differential pricing of the two goods. Previous work has shown that this population of capuchins understands the "price" of different goods when such goods are sold in different amounts for a single token (Chen et al., 2006). Experiment 1 assumed

that similar presentations of different amounts of food would teach monkeys the specific price of each good. It is possible, though, that monkeys did not attend to this information.

### Experiment 2: Design.

In Experiment 2, we add a control condition to test that monkeys attended to the pricing information, testing the monkeys on a price shift condition similar to that used in Chen et al. (2006)'s original study.

A second potential flaw in Experiment 1 is that sampling a food may establish a preference that doesn't change once new price information is learned. In Experiment 1, we first exposed monkeys to both flavors of ice in order to obtain a baseline for their preferences. Unfortunately, it is possible that this experience allowed monkeys to establish a preference that could not be changed by price information. In Experiment 2, we directly explore this possibility by varying whether monkeys had previous experience with the foods whose prices were being manipulated. This manipulation allowed us to examine whether prior experience anchored preferences and contributed to the fact that we did not observe a human-like pricing bias in Experiment 1.

### Experiment 2: Materials and Methods.

We tested 8 capuchins (AG, AH, FL, HG, HR, JM, MP, NN) from the same colony. All but two monkeys (HG and JM) had previously participated in Experiment 1. One monkey who was tested in Experiment 1 (MD) was not included in this study due to a disinterest in entering the enclosure for testing during the period when this study was run.

We used the same testing enclosure as in Experiment 1, but with a couple of key differences. Instead of using flavored ice as in Experiment 1, Experiment 2 used pieces of differently flavored Jell-O brand gelatin (roughly 1 cm across and 0.65 cm deep). To standardize the shapes of the gelatin, we made each piece using a standard mold. We used six different color/shape combinations (pink watermelon squares, purple grape crescents, green lime stars, blue blueberry hearts, red strawberry clovers, and yellow lemon triangles). Hereafter each gelatin will be referred to by color. Gelatin colors and experimenter were counterbalanced across subjects.

The procedure of Experiment 2 was similar to that of Experiment 1 except that (1) we explicitly varied the exposure monkeys had to the gelatin, and (2) we included a condition to directly test whether monkeys encoded the price information in this study. Each monkey was run on three separate conditions: first, an exposure condition and a non-exposure condition (presented in a counterbalanced order) followed by a price shift control condition (see Figure 8.1 for more details). In both the exposure and non-exposure conditions, monkeys would ultimately get a choice between two colors of gelatin, one which was shown to be expensive and one which was shown to be cheap.

In the exposure condition, monkeys began with the exposure phase, consisting of two sessions. The goal of this phase was to allow the monkeys to systematically taste each of the two gelatin colors. On each trial, monkeys interacted with an experimenter who handed them a single color of gelatin. During each trial, the experimenter began by positioning herself outside either side of the test enclosure, displaying a small dish holding the gelatin to be delivered. The experimenter then brought the dish up to the trading hole, allowing the monkey full access to its contents. Once the monkey had eaten the gelatin, the dish would be reloaded, and the next trial would begin. Each monkey received two sessions of 12 trials (six for each color gelatin) with the order of the first color gelatin given counterbalanced across sessions.

After subjects completed the exposure phase, they moved on to a price learning phase, identical to that of Experiment 1, in which subjects were taught the prices for two colors of gelatin (either blue vs. green or red vs. yellow depending on the counterbalance). In this phase, capuchins participated in a series of five sessions of 12 trials. Each monkey was taught that one color was cheap (i.e., they received three pieces of that color gelatin for a single token), while the other color was expensive (i.e., they received only one piece of gelatin for a token).

After learning the price of the new goods, capuchins then were given a preference assessment phase, similar to the one used in Experiment 1, in which subjects were allowed to freely choose between the two colors of gelatin for which they had just learned prices in the price learning phase. Monkeys received 5 sessions of 12 trials. On each trial, monkeys interacted with an experimenter, blind to the nature of the cheap and expensive gelatin goods, who held two small dishes, each containing a colored gelatin piece. The experimenter then simultaneously brought each dish to a separate trading hole, and allowed the monkey to choose one. After the monkey chose the color to consume, the other dish was withdrawn, and the trays were reloaded. Again, the goal of this phase was to determine which of the two flavors the monkeys preferred when they got to freely choose, and to see if price information affected that preference.

In the non-exposure condition, monkeys proceeded through exactly the same phases as in the exposure condition except that we did not include the exposure phase; monkeys only went through a price learning phase and a preference assessment phase. In this way, monkeys in the non-exposure condition had never been exposed to the two novel gelatin colors when they first learned about the gelatins' prices in the price learning phase. If previous exposure to the novel foods had anchored the monkeys' initial preferences in Experiment 1, then wouldn't expect a similar effect here since the goods were totally novel at the point the monkeys learned their prices.

After subjects completed both the exposure condition and the non-exposure condition, they then moved on to the price shift control condition. The goal of this

condition was to be sure that monkeys learned the different prices. We used a version of the price shift studies used in Chen et al. (2006): monkeys were presented with an initial purchasing preference phase (to test the monkeys' initial preferences between two colors of gelatin when those colors were the same price), a price shift training phase (in which we taught the monkeys a new price for one of the two goods), and a price shift assessment phase (to see whether the monkeys responded rationally to this shift in price and allocated their budget accordingly).

The initial purchasing preference phase involved assessing the monkeys' initial preferences between the two colors of gelatin being offered when monkeys got equal amounts of the two colors for a single token. Monkeys received five sessions of 12 trials each, with one session run per day. At the beginning of each session, the subject monkey was given 12 tokens with which to buy the gelatin. Experimenters positioned themselves on opposite sides of the testing chamber. One experimenter consistently dispensed pink gelatin, while the other dispensed purple. Both experimenters began the session with their backs facing the monkey.

Then, on a synchronized count, both turned around, offering one hand toward a trading hole to receive a token from the monkey, and displaying the gelatin that they offered in the other hand. The monkey was then able to choose the color he preferred by depositing a token in one of the experimenters' hands. The chosen experimenter then presented his tray up to the trading hole so that the monkey could reach the gelatin while the other experimenter would turn his back around in order to indicate that they were no longer available for trade. After the gelatin was completely consumed, experimenters would reload their trays, switch sides and proceed with the next trial. Gelatin colors were counterbalanced across monkeys.

After we had taught monkeys that the two kinds of gelatin were equal in price, subjects moved into the price shift training phase. Here the monkeys learned that the prices that they were originally exposed to had changed. Specifically, the good that the subject had liked least (purple: AG, AH, HG, HR; pink: FL, JM, MP, NN) was discounted such that one token went from buying a single piece of gelatin to buying three pieces. As in the initial assessment, monkeys were given 12 tokens to allow them to purchase gelatin from the two experimenters. However, in this case monkeys did not have a choice; on each trial, only one experimenter was available to trade.

Both experimenters began each trial with their backs to the testing chamber, but only one experimenter turned at a time, thus providing only one person with whom the monkey could trade. One experimenter offered three pieces of the "discounted" gelatin for a single token, while the other offered the original price of one gelatin piece per token. Each monkey received three sessions of this training before moving onto the price shift assessment phase.

In the price shift assessment phase, we tested whether monkeys responded to the change in price they just witnessed by switching their consumption to the cheaper good (see Chen et al., 2006, for another version of this price-shift test). The assessment was similar to the initial purchasing preference phase except that monkeys had a choice of the two gelatin colors now at the new prices. If monkeys accurately attend to and track price, they should switch their consumption to the less costly good in this assessment phase. Each monkey completed 5 sessions of 12 trials each.

### Experiment 2: Results.

We first analyzed how monkeys performed on the price shift assessment as compared to the initial purchasing preference phase. Did subjects successfully understand the more expensive price and therefore choose the cheaper good when they had to pay for it? To test this, we ran a repeated measures ANOVA with price (initial purchasing preference phase with equally priced goods vs. price shift assessment where one good was more expensive) as a within-subjects factor and color of good chosen to be cheap (pink or purple) as a between-subjects factor.

We observed no effect of color [ $F(1, 6) = 0.25, p = 0.879$ ]. We did however observe a significant main effect of price [ $F(1, 6) = 123.85, p < 0.0001$ ]. Although monkeys overall didn't have a preference for either color initially [ $t(7) = 0.944, p = 0.377$ ], monkeys preferred the cheaper good after the price change [81.0% preference for the cheap good:  $t(7) = 5.89, p < 0.0006$ ]. All monkeys spent more on the cheap color after the price shift than before it (paired sign test:  $p = 0.0078$ ).

We also observed a small interaction of color and price [ $F(1, 6) = 7.67, p = 0.03$ ]; monkeys showed more of a shift toward the cheaper good when pink was the cheap good than when purple was the cheap good. Overall, these results demonstrate that subjects recognized the price at which each color gelatin was being offered and attended to this information in their choices. Importantly, all monkeys consumed more of the cheaper good after the price shift as one might expect given standard price theory (see Chen et al., 2006).

We then tested whether monkeys chose the expensive good above chance in the preference assessment phases of the exposure and non-exposure conditions. As in Experiment 1, we observed no effect of price in Experiment 2; monkeys chose the expensive good on average 53.1% of trials [One-sample t-test:  $t(7) = 0.431, p = 0.68$ ]. This chance-level performance was true for both the blue/green color combination [48.5% choice to expensive,  $t(7) = 0.26, p = 0.80$ ] and the red/yellow color combination [57.7% choice to expensive,  $t(7) = 0.718, p = 0.50$ ]. Despite the fact that monkeys robustly understood price in the price shift control, that information didn't seem to affect

their preferences or how much they valued each kind of gelatin when they got to freely choose one.

We also looked to see whether initial experience affected the magnitude of price effects using a repeated measures ANOVA with exposure level (exposure condition vs. non-exposure condition) as a within subject factor and the color combination subjects had experience with (blue/green vs. red/yellow) as a between subjects factor. We observed no effect of experience [ $F(1, 6) = 1.198, p = 0.316$ ]; subjects showed just as strong a preference for the expensive good in the exposure condition (Mean 57.9% preference for the expensive good) as they did in the non-exposure condition (48.3%).

### Experiment 2: Discussion.

We had two goals in Experiment 2. Our first was to confirm that monkeys attended to the prices that we had presented. To do this, we performed a price shift like that of Chen et al. (2006), changing the price of one color of gelatin to a “sale price” that was three times cheaper than the price of the other gelatin. Our subjects overwhelmingly purchased more of the cheaper gelatin, as they had done in previously published studies (Chen et al., 2006).

This result indicates that capuchin monkeys do attend to the price of the goods presented in this token exchange market. In addition, this result also demonstrates that monkeys use price as a factor in their purchasing decisions in this experimental market. The second goal of Experiment 2 was to examine the role of prior experience in the monkey’s preferences—specifically, to find out whether prior experience may have moderated any effect that price may have on preferences. Regardless of exposure, capuchins did not prefer the more expensive good; as in Experiment 1, no monkey showed a preference for either the cheap or expensive good. This new result indicates that monkeys’ lack of a preference for the more expensive good in Experiment 1 is not because of anchoring due to prior exposure. Instead, the results of Experiment 2 suggest that regardless of whether monkeys have previous experience with a particular food, capuchins base their preferences on their subjective experience with a food, rather than any external price information.

### **Experiment 3**

A possible issue with Experiments 1 and 2 is that both studies used foods that were different colors and—perhaps more importantly—different flavors. It is possible that price information did not influence monkeys preferences in these studies because monkeys may have had slight (although not statistically significant) preferences based on the colors and

flavors of the foods we offered them. These initial preferences may have overshadowed any changes in valuation that occurred due to differences in price. The results of Experiments 1 and 2 therefore leave open the possibility that monkeys may prefer the more expensive of two perfectly equal goods.

To examine this, Experiment 3 tested monkeys' preferences for differently priced yet perceptually identical foods, just as has been done in human pricing effect experiments (e.g., Plassmann et al., 2008).

### Experiment 3: Design.

In experiment 3, we used two of the same kind of food (pieces of Kix cereal) to ensure that the two goods were perfectly equal. However, to be sure that the individual foods were distinguishable in some way, we paired the pieces of cereal with novel "brand" logos. In this way, Experiment 3 was able to allow monkeys to choose between foods which would be identical in perceptual experience (i.e., taste) yet could have different prices.

Experiment 3 also aimed to explore whether any other factors could affect monkeys' preferences for different foods. Given that monkeys failed to show pricing effects in Experiments 1 and 2, we hoped to find another factor that could affect monkeys' preferences even if this species lacks pricing effects. One external factor that is unrelated to price but appears to play a role in humans' reward preferences is the wait time that comes with different rewards. Although we tend to dislike waiting for a reward (e.g., Berns et al., 2007) and find long wait times very costly, we also find rewards more enjoyable if we have to wait longer for them (e.g., Alessandri et al., 2008a). This so-called "delay justification effect" has also been observed in 7-year-old children, who also prefer a stimulus that usually follows a delay to a stimulus that usually does not follow a delay (Alessandri et al., 2008b). Although there is some controversy about the mechanisms underlying these delay justification effects (see Festinger, 1957 vs. Zentall, 2010), it seems clear that our tendency to overvalue stimuli that are associated with longer delays may be part of a larger, more general tendency to prefer stimuli associated with more cost or effort, regardless of whether such effort comes in the form of extra waiting (e.g., Alessandri et al., 2008a), more difficult work (e.g., Festinger and Carlsmith, 1959), or even more embarrassment (e.g., Aronson and Mills, 1959).

Interestingly, humans are not the only species to experience delay and other forms of effort justification effects. Recent comparative work suggests that some non-human species also prefer a stimulus associated with additional cost or effort (Clement et al., 2000; Kacelnik and Marsh, 2002; Friedrich and Zentall, 2004; Gipson et al., 2009), particularly in cases when delays are involved (DiGian et al., 2004; Pompilio and

Kacelnik, 2005; Zentall and Singer, 2007; Wanat et al., 2010; see Zentall, 2010 for a review).

For example, DiGian et al. (2004) presented pigeons with two stimuli that predicted an immediate reward: one stimulus was available immediately and one stimulus appeared only after a 6 s delay. When given a choice between these two reinforcing stimuli, pigeons reliably preferred the stimulus that appeared after a delay even though it predicted the same kind of reward as the stimulus that was available immediately. This result suggests that pigeons find a stimulus more rewarding if they have to wait longer for it. In this way, other species appear to value goods more highly the longer they have to wait for them.<sup>37</sup> We therefore wanted to see whether capuchins might use delay as an extrinsic factor that mediated their preferences for different kinds of foods, even though they don't incorporate price information into their preferences.

Experiment 3 tested whether varying either a food's wait-time or its price would affect monkeys' preferences when freely choosing between different options. The similarity between these two dimensions—delay and price—allowed us to set up two methodologically identical studies testing each of these factors.

We taught monkeys about two novel brands' prices (one piece for one token or three pieces for one token) or wait times (available after 30 s or immediately available) through repeated exposure, and then allowed the monkeys to choose between the two brands. If longer delays affected capuchins' preferences more so than higher prices, then monkeys should prefer brands that come after a long delay even though they show no preferences across differently priced brands. This type of effect would imply that capuchins' preferences in a token economy can in fact be influenced by extrinsic properties, but that price information is not one of these properties.

On the other hand, if the capuchins show neither delay nor price effects, this finding would suggest that capuchins may evaluate goods based entirely on their subjective experience with them, rather than any additional extrinsic features.

### Experiment 3: Materials and Methods.

Subjects were 7 adult brown capuchin monkeys (AH, FL, HG, HR, JM, MD, NN) from the same colony. Two monkeys who had previously participated in both Experiments 1 and 2 were excluded from Experiment 3: one low-ranking monkey (AG) was not included

---

<sup>37</sup> Nonetheless, there does exist a number of studies suggesting that human and non-human animals do not always show delay and other effort justification effects (Vasconcelos et al., 2007; Arantes and Grace, 2008; Shibasaki and Kawai, 2008, 2011; Vasconcelos and Urcuioli, 2009). For example, in one experiment, Shibasaki and Kawai (2008) observed that people prefer stimuli that follow low-effort tasks than those that follow high-effort tasks. These authors found a similar pattern in Japanese macaques (Shibasaki and Kawai, 2011) in a task in which different stimuli appeared after different numbers of touchscreen presses. Like humans, macaques appeared to show a reverse effort justification effect, statistically preferring the stimulus associated with the least effort over the one associated with the most.



due to social problems in the colony during the time of testing, while another monkey (MP) was not included due to a disinterest in entering the enclosure for testing during the period when this experiment was run.

Experiment 3 used pieces of Kix cereal (General Mills, US)— a familiar food for these subjects—as a reward. To differentiate between the cheap and expensive versions of the cereal, we created three pairs of “brands” which were denoted by three easily distinguishable pairs of symbols: red flower vs. yellow star, green clover vs. yellow moon, and orange balloon vs. blue horseshoe. These brand symbols (approximately 13 13 cm) were displayed prominently on the white shirt of the experimenter that consistently offered that brand. The brand symbol was also displayed on the container (a 3 oz clear plastic drinking cup with the front half cut out to make food easily reachable) from which the cereal was dispensed.

Each experimenter presented the cups to the monkey on white foam-core platforms that were covered in white duct tape for ease of cleaning. Each cup always contained a single piece of cereal, but each platform could hold one to three cups depending on condition. We attached the cups to the platform using Velcro to keep them stable. Each monkey always saw the same experimenter associated with each brand, but the brand and experimenter were counterbalanced across monkeys.

Each monkey participated in three conditions: a price shift control condition, a delay condition and a price condition (see Figure 8.1 for more details). The price shift control condition was administered first, with the delay and price conditions presented afterwards in a counterbalanced order.

The price shift control condition, which was nearly identical to the price shift control condition used in Experiment 2, was used to ensure that monkeys could attend to price information in the context of the branded cups used in Experiment 3. If monkeys are able to accurately track the prices of brands, then—as in Experiment 2—they should buy more of the cheaper brand after the price shift. After all of the monkeys had completed the price shift control, demonstrating that they were paying attention to the price of the cups and using this price information to make decisions, they then moved on to either the price condition or the delay condition (in a counterbalanced order).

As in the non-exposure condition of Experiment 2, the price condition consisted of two phases: a price learning phase (where monkeys were exposed to the prices of two new brands), and price preference assessment phase (where monkeys were allowed to choose between the two brands in the absence of price). The goal of the price learning phase was to expose the monkeys to the prices of two new brands of cups. One of the two brands was cheap (i.e., one token could purchase three cups with one piece of Kix each), while the other was expensive (i.e., one token could purchase only one cup with one piece of

Kix). Note that this is identical to the procedure used for different goods in Experiments 1 and 2, except that we used different brands instead of different foods.

Monkeys each received four sessions, each consisting of 16 trials each. In each trial, the two branded experimenters dropped a token into the enclosure in unison, before moving to opposite sides of the enclosure. Then one of the two experimenters turned around, offering to trade the contents of her cup(s) in exchange for a token. After the monkey paid one of the two experimenters and received the food, the two experimenters switched sides and repeated the process.

Monkeys then moved on to the price preference assessment phase. The goal of this phase was to determine whether capuchins changed their preference for the two brands based on the price information they had just been taught. Each price preference session began with four price reminder trials identical to the original price learning trials. We added these reminder trials to be sure that monkeys accurately remembered which brand was which.

After these reminder trials, a different experimenter (who was blind to condition) gave the monkeys a choice between the two different branded cups in the absence of any tokens. Importantly, in this case, each branded cup only contained a single piece of Kix. If monkeys had developed a preference for a specific brand based on the pricing information, then they should selectively choose that preferred brand when given a free choice to eat food from either brand. Monkeys received four sessions of 16 trials each. If capuchins use price as an indicator of quality, then they should selectively prefer the previously expensive brand (the one for which they had previously been offered one cup for a single token) to the previously cheap brand (the one for which they previously received three cups for a single token).

The delay condition mirrored the price condition with one key difference: instead of varying the price of the two brands, we instead varied the time the monkeys needed to wait in order to receive each brand. Like the price condition, the delay condition involved two phases: a delay learning phase and a delay preference assessment phase. The delay learning phase served to teach the monkeys that one brand was associated with a delay and one was not. Specifically, one of the two brands (the “expensive delay” brand) was associated with a 30 s delay while the other brand (the “cheap delay” brand) was given immediately.

Each monkey received four 16-trial sessions. On each trial, the two branded experimenters dropped a token into the enclosure in unison before moving to opposite sides. Then, one of the two experimenters turned around, offering to trade with the monkey. When the monkey gave a token to the experimenter holding the cheap brand, she moved her cup toward the trading hole immediately, allowing the monkey to immediately retrieve the piece of Kix in the cup. In contrast, when the monkey gave a

token to the experimenter holding the expensive brand, she waited 30 s before moving her cup toward the trading hole, thus requiring the monkey to wait before retrieving the piece of Kix in the cup.

Following the delay learning phase, monkeys moved on to the delay preference assessment phase. The goal of delay preference assessment phase was to determine whether the relative cost of the two brands in terms of time had affected the monkeys' preferences between the two. The structure of these sessions was identical to the sessions presented in the price preference assessment phase except that we varied the brands' delay times rather than prices. During each trial, an experimenter who was blind to which brand had previously been associated with the delay turned to face the enclosure, offering a cup from one of the brands at each of the two trading holes.

If capuchins come to value brands that are associated with a greater wait time, then they should selectively prefer the brand that was previously associated with the delay over the brand that was previously available immediately. On the other hand, if capuchins do not use increased delay as an indicator of quality, then they should show no preference in the delay preference assessment phase.

### Experiment 3: Results.

We first explored whether monkeys had an initial preference for one of the two brands in the price condition when they initially encountered them in purchasing preference phase. None of the seven monkeys showed a statistically significant preference across the two brands (percent choice to the brand that would later be made cheap: HR: 48%,  $p = 0.90$ ; FL: 48%,  $p = 0.90$ ; NN: 42%,  $p = 0.26$ ; MD: 38%,  $p = 0.059$ , JM: 41%,  $p = 0.17$ ; HG: 47%,  $p = 0.71$ , AH: 50%,  $p = 1.00$ ).

After the price shift, however, all the monkeys developed a significant preference for the cheaper of the two brands: (HR: 98%,  $p < 0.0001$ , FL: 97%,  $p < 0.0001$ ; NN: 94%,  $p < 0.0001$ ; MD: 98%,  $p < 0.0001$ ; JM: 95%,  $p < 0.0001$ ; HG: 95%,  $p < 0.0001$ ; AH: 70%,  $p = 0.0016$ ). A paired t-test revealed that monkeys as a group preferred the cheaper brand when using their currency [ $t(6) 9.83$ ,  $p < 0.0001$ ].

Again, this result suggests that monkeys do use price information when distinguishing between brands, actively shifting their consumption to the cheaper brand when they need to pay tokens to obtain it.

Given that all monkeys correctly paid attention to the price of these brands of cups, we then explored whether teaching monkeys that the brands had different prices had an effect on their preferences for each brand. Unfortunately, three subjects did not fully complete testing due to social problems in the enclosure. We therefore ran our analysis on just the four monkeys that completed all the tests.

As in previous experiments, we presented monkeys with two new brands and tested whether monkeys chose the more expensive one above chance in the price preference assessment phase. We again saw no effect of price on monkeys' preferences. As in all previous experiments, no monkeys showed a significant preference for the expensive brand over the cheap brand (percent choice to expensive brand: HR: 50%,  $p = 1.0$ , FL: 48%,  $p = 0.90$ ; NN: 48%,  $p = 0.90$ ; MD: 53%,  $p = 0.71$ ).

We then tested whether monkeys showed a preference for delay—whether they chose the expensive delayed brand above chance in the delay preference assessment phase. Interestingly, we also saw no significant preferences for the brand with the expensive delay. One monkey, HR, did show a significant preference, but her pattern of performance went in the opposite direction than we hypothesized—HR significantly preferred the immediately available cheap brand (20%,  $p < 0.0001$ ). All other monkeys chose between the immediate and delayed brand at chance (FL: 47%,  $p = 0.71$ ; NN: 48,  $p = 0.90$ ; MD: 55%,  $p = 0.53$ ).

### Experiment 3: Discussion.

Experiment 3 had two main goals. The first goal was to examine whether using two goods that are experientially identical but perceptually distinguishable would cause monkeys to show a price preference that they do not otherwise show.

In contrast to Experiments 1 and 2, which used different flavors of the same food, Experiment 3 used identical cereal pieces as the reward for both the expensive and cheap options. In spite of this change, monkeys still showed no preference between the two brands during price preference assessment trials. Importantly, all monkeys successfully switched to the cheaper good when it went on “sale,” again indicating that they were attending to the price information we had shown them. This finding implies that the monkeys' lack of preference in the first two experiments is likely not due to differences in the two foods being offered, as monkeys' failure to use price as an indicator of value persists even when identical items are offered as goods.

The second goal of Experiment 3 was to explore whether any factors could affect monkeys' preferences for different foods. To this end, we explored whether the amount of time monkeys were required to wait for one brand over the other affected their preferences. Rather than teaching monkeys different prices, the delay condition of Experiment 3 taught the monkeys that they had to pay different costs in terms of time. We found that monkeys did not show a preference for the good previously associated with a delay. This finding shows that the use of delay as a cue to quality may not be as robust as some previous studies had suggested (DiGian et al., 2004; Pompilio and Kacelnik, 2005; Zentall and Singer, 2007; Wanat et al., 2010; see Zentall, 2010 for a review). In this context, it seems that capuchins use neither price nor delay as cues to quality.

## **Experiment 4**

Our use of Kix cereal in Experiment 3 had one important drawback. This cereal was already quite familiar to our subjects, as it had previously been used in a number of experiments in the lab. Although Experiment 2 established that prior exposure to a food does not affect monkeys' preferences in terms of price effects, we still worried that it might be difficult for monkeys to think of this food differently based on its brand since it was so familiar to them previously.

### Experiment 4: Design.

Experiment 4 used the same approach as Experiment 3, but with a novel kind of food: Crunch Berries cereal (Quaker Oats, US).

The Crunch Berries offer a couple of major advantages over Kix. First, this cereal is available in four colors (red, green, blue, and purple) giving us the ability to make the two different brands perceptually different and thus easier to discriminate. However, all four of colors of Crunch Berries taste the same; in this way, there is no reason that subjects should form a strong preference for one color over another. Second, our subject monkeys had no outside experience with Crunch Berries, and thus we were able to ensure that the monkeys' only knowledge of differences between the colors and brands was acquired during the exposure they received during training and testing.

In addition to using a different food reward, Experiment 4 also aimed to be sure that monkeys were able to distinguish between the two brands and make choices based on them. We therefore added more reminder trials at the beginning of each testing session to be sure that the monkeys remembered which brand was cheap vs. expensive and which brand required a delay. To determine whether the monkeys were accurately tracking brands in both the price and delay conditions, we also added a set of manipulation checks to ensure the monkeys were accurately tracking which brand was associated with a higher price/delay.

### Experiment 4: Materials and Methods.

Subjects were 6 adult brown capuchin monkeys (FL, HG, HR, JM, MD, NN). Five of the monkeys tested (FL, HG, HR, JM, NN) had previously participated in Experiment 3. Two monkeys (HG, MD) who participated in Experiment 3 were excluded from Experiment 4 due a disinterest in entering the enclosure for testing during the period when this study was run. One monkey (MP) participated in Experiments 1 and 2, but not in Experiment 3.

We used four Crunch Berry cereal colors (red, blue, green, purple) as goods; these colors all tasted the same so monkey should not have had any preferences for colors based

on the flavor. We again associated each reward color with an individual “brand” using symbols: red Crunch Berries with a red flag, blue Crunch Berries with a blue sun, purple Crunch Berries with a purple teardrop, and green Crunch Berries with a green snowflake. The brand used in each condition and the experimenter associated with each brand were counterbalanced across monkeys. As in Experiment 3, brands were displayed both on the experimenter’s shirt, and the cups containing the Crunch Berries.

In Experiment 4, each monkey participated in 2 conditions: a price condition and a delay condition (see Figure 8.1 for more details).

As in previous experiments, the price condition began with two phases: a price learning phase (where monkeys were exposed to the prices associated with two brands), and a price preference assessment phase (where monkeys were allowed to choose between the two brands in the absence of price). However, at the end of the price condition, we added an additional phase, a price manipulation check phase.

In the price learning phase, monkeys were taught that one brand was expensive (for each token the monkeys got only one cup/Crunch Berry), while the other brand was cheap, (for each token the monkeys got three cups/Crunch Berries). The price learning phase consisted of one session with 16 trials with procedures identical to the price learning phase of Experiment 3.

After monkeys completed the price learning phase, they moved on to the price preference assessment phase, where we aimed to determine whether the prices of the two brands impacted preferences. The price preference assessment was performed over four sessions, each consisting of a set of 8 reminder trials (identical to the original price learning phase trials), and 16 test trials. Note that we doubled the number of reminder trials from Experiment 3 to ensure that the monkeys would remember the prices of the two brands.

After reminder trials were complete, a different experimenter (blind to condition) gave the monkeys a choice between the two brands of cups in the absence of any tokens. As in Experiment 3, monkeys chose between the two brands of cups now offering one Crunch Berry each. Again, if monkeys developed a preference for the more expensive Crunch Berry brand after learning its price, then they should selectively choose the expensive brand over the cheap brand in these trials.

We then moved on to the price manipulation check phase. This phase served to reaffirm that the monkeys were paying attention to the price of the two Crunch Berry brands. Specifically, we expected that when the monkeys had to pay for the two brands using their tokens, they would take price information into account and therefore choose the cheap brand (which gave them the most food).

The price manipulation check phase consisted of four sessions, each containing 8 reminder trials (identical to the original price learning phase trials) and 16 manipulation

check test trials. In these test trials, the monkeys had to use their tokens and choose between the expensive brand (which gave only one cup/Crunch Berry per token) and the cheap brand (which offered three cups/Crunch Berries per token).

These trials began when the two branded experimenters dropped a token into the enclosure before moving to opposite sides. After a synchronized count, the two experimenters turned around, allowing the monkey to choose to trade the token for either the cheap brand or the expensive brand. Monkeys who attend to price should choose to buy the cheap brand more often than the expensive brand since the cheap brand gives them more food overall.

As in Experiment 3, the delay condition mirrored the price condition (a delay learning phase, a delay preference assessment phase, and a delay manipulation check phase) with only one key difference: instead of varying the price of the two brands, we instead varied the delay time the monkeys needed to wait in order to receive each brand: one brand was associated with a 30 s delay (the “expensive delay”), while the other was given immediately (the “cheap delay”).

#### Experiment 4: Results.

We first tested whether monkeys showed a preference in the price preference assessment phase. As in all previous experiments, we saw that monkeys as a group did not show a preference for the expensive good [ $t(5) = 1.32, p = 0.25$ ]. Two monkeys showed preferences for the color associated with a cheap price (percent choice of expensive: HR: 31%,  $p = 0.004$ ; NN: 36%,  $p = 0.03$ ) and all other monkeys showed no preference (FL: 48%,  $p = 0.90$ ; HG: 42%,  $p = 0.26$ ; JM: 61%,  $p = 0.10$ ; MP: 48%,  $p = 0.90$ ).

In the price manipulation check phase, monkeys as a group showed a preference for the cheap good [ $t(5) 18.81, p < 0.0001$ ]. Indeed, all monkeys individually showed a preference for the cheap good (Preference for expensive symbol: FL: 6%,  $p < 0.0001$ ; HR: 0%,  $p < 0.0001$ ; HG: 6%,  $p < 0.0001$ ; JM: 0%,  $p < 0.0001$ ; NN: 0%,  $p < 0.0001$ ; MP: 15%,  $p < 0.0001$ ). Comparing monkeys’ performance across the two phases also revealed a significant effect of phase [ $t(5) 8.86, p = 0.0003$ ], suggesting that although monkeys attend to which good is more expensive when they must spend their tokens, the same subjects do not prefer the expensive brand when they can freely choose between the two brands.

We then explored how monkeys performed on the delay preference assessment. Overall, monkeys showed no preference for the delayed good [ $t(5) = 0.74, p = 0.49$ ]. One monkey showed a significant preference in the opposite direction (preference to delayed good: NN: 36%,  $p = 0.03$ ), but all other monkeys did not show any statistically significant preference (FL: 61%,  $p = 0.10$ ; HR: 53%,  $p = 0.71$ ; HG: 50%,  $p = 1.0$ ; JM: 44%,  $p = 0.38$ ; MP: 39%,  $p = 0.10$ ).

Even though monkeys did not show a preference for the delayed good, as a group they showed a significant preference for the immediate reward during the delay manipulation check phase [ $t(5) = 5.12, p = 0.004$ ]. Individually, four monkeys showed a significant preference for the immediate reward (preference to the delayed reward: FL: 4%,  $p < 0.0001$ ; HR: 19%,  $p < 0.0001$ ; JM: 23%,  $p = 0.0002$ ; MP: 8%,  $p = 0.0001$ ), but two other monkeys' performance was not statistically significant (HG: 38%,  $p = 0.11$ ; NN: 35%,  $p = 0.06$ ).

Comparing monkeys' performance across the delay preference assessment phase and the delay manipulation check phase revealed a significant effect of phase [ $t(5) = 3.27, p = 0.02$ ]. Monkeys attended to which brand was the delayed brand in the manipulation check, yet they still formed no preference for the brand previously associated with delay when they had the chance to get both brands immediately.

#### Experiment 4: Discussion.

We had two main goals in Experiment 4. Our first goal was to examine whether the use of a familiar reward in Experiment 3 had prevented monkeys from showing pricing and delay effects. Experiment 4 dealt with this issue by testing whether monkeys showed delay and pricing effects for food rewards that were both unfamiliar and easily distinguished (differently colored Crunch Berries). Despite this methodological change, we observed the same pattern of performance as in Experiment 3: capuchins showed no preference for the more expensive brand in either the price or delay condition.

The second goal of Experiment 4 was to determine whether the monkeys were in fact tracking the prices and delays that were associated with the two different brands. To assess this, we added a manipulation check in both the price and delay conditions. These manipulation checks revealed that monkeys were accurately tracking which brand was expensive in terms of both delay and price, yet the same subjects' preferences were unaffected by these cues.

### **General discussion**

Across four experiments, capuchins did not moderate their preferences with regard to price. In Experiment 1, capuchins showed no preference for the more expensive piece of ice. In Experiment 2, capuchins showed no preference for the more expensive color of gelatin, regardless of prior exposure. In Experiments 3 and 4, capuchins showed no preference for the more expensive brand of cereal.

Across several studies, however, capuchins consistently passed manipulations checks showing that they understood the price associated with each of the two goods involved.



Taken together, these findings imply that capuchins do not show a human-like pricing effect—learning the price of a good does not change the capuchins’ preference for that good. These results also suggest that capuchins’ preferences may not be affected by other factors—such as delay information. Across Experiments 3 and 4, capuchins failed to update their preferences for different foods based on the delay they were required to wait for that food. This lack of a preference for delayed rewards conflicts with previous studies demonstrating that human and non-human species do—at least in some cases—prefer stimuli associated with additional time or effort (DiGian et al., 2004; Pompilio and Kacelnik, 2005; Zentall and Singer, 2007; Wanat et al., 2010; see Zentall, 2010 for a review). It is worth noting, however, that the general preference for longer delays is likely to be less robust than the preference for higher-priced goods—a growing body of studies demonstrate that human and non-human participants do not always show effort and delay justification effects (Vasconcelos et al., 2007; Arantes and Grace, 2008; Shibasaki and Kawai, 2008, 2011; Vasconcelos and Uruioli, 2009).

Note that there are many situations in which humans and animals tend not to show delay justification effects; indeed, the literature on discounting effects in humans and animals suggests that both of these populations often prefer cases in which delays are shorter rather than longer (see review of this work in Stevens, 2010). For this reason, it is possible that our subjects might show other effects on value manipulation in future studies despite not showing the delay effects we hypothesized here. Nonetheless, the results of the current study still provide hints that delay and other effects on preferences may be less robust than previously thought.

Another potential problem with our study concerns how monkeys were presented with price information in their experimental token economy. As in previous studies (e.g., Chen et al., 2006; Lakshminarayanan et al., 2011), we communicated the price of a good to monkeys by changing the amount of that good that monkeys received for a single token. While this way of indicating price information has been validated in previous work (see Chen et al., 2006 for evidence that monkeys obey the tenets of standard price theory when tested using this method), it also resulted in a methodological worry: when learning the price of different goods, monkeys always received more of the “cheaper” good than of the more expensive one. In this way, monkeys inadvertently wound up having more experience with cheap vs. expensive items.

We attempted to deal with this potential confound in several ways. First, we ran all preference assessment phases on different days than we taught monkeys the price of the different goods. In this way, we hoped that if monkeys became satiated on the cheaper good during the price learning phases, this would not extend to their preference choices since preference assessments were run on separate days than exposure to cheap and expensive goods. Some of our studies used foods that are very familiar to monkeys (e.g.,

Kix) so that differential exposure in our studies would be trumped by the monkeys' previous experience with these foods. Nevertheless, it is worth noting that this differential exposure across cheap and expensive conditions is one confound in our studies that could potentially have influenced the lack of pricing effects we observed.

Based on the control and experimental conditions used, we can rule out a number of reasons for why the monkeys might not be showing a human-like pricing effect. First, our control conditions demonstrate that monkeys' lack of preference is not due to an inability to understand price. In Experiment 2, subjects showed that they understood the price of the two goods; in the price shift control condition, subjects spent more of their tokens on a good that was "on sale" than on an equivalent good that was not. This replicates previously published findings (e.g., Chen et al., 2006) showing that capuchins can both track price and use price information in their purchasing decisions.

Other control conditions revealed that capuchins successfully track the price and delay associated with different brands of goods. In Experiment 4, capuchins chose to buy a cheap brand over an expensive brand, and an immediate brand over a delayed brand, suggesting that subjects successfully use both price and delay information in their purchasing decisions even though they do not use these factors to form their preferences.

Finally, our control conditions rule out the possibility that capuchins failed to show pricing effects because their prior experiences anchored their subsequent preferences. In Experiment 2, we established that capuchin monkeys failed to show a price effect regardless of whether they had previous experience with the foods serving as goods.

Overall, then, our findings suggest that capuchin monkeys perform very differently than humans when interacting with differently priced goods. Although humans regularly prefer goods that are higher in price, capuchin monkeys appear to show no such effect.

This pattern of performance is relatively surprising for two reasons. First, our results suggest that capuchins fail to fall prey to arbitrary price information when deciding between different goods. Our failure to observe pricing effects in capuchins is also surprising in light of the fact that this species exhibits a number of other classic judgment and decision-making biases, such as the endowment effect (Lakshminarayanan et al., 2008), loss aversion (Chen et al., 2006), and the reflection effect (Lakshminarayanan et al., 2011). Indeed, to our knowledge, the price effect is the first judgment and decision-making heuristic to have been studied in non-human primates and not observed. Our results therefore suggest that pricing effects may rely on mechanisms that are distinct from those involved in these other biases. Indeed, our results suggest that pricing effects may be due to cognitive mechanisms or specific experiences that are uniquely human.

One uniquely human experience that could give rise to human-specific pricing effects is our species' practice of participating in markets in which there is often an association between price and value. In a free market, companies can only charge what people are

willing to pay for their goods. As such, in most human markets, there will often be an association between a good's price and its actual quality. Humans may thus generalize these experiences to falsely believe the price of an item is always indicative of its quality.

Under this potential explanation, we might not expect capuchin monkeys to show a similar effect since the markets they trained in are not markets that have associations between price and quality. In this way, our findings have narrowed down the kinds of human-specific experiences that likely lead to price effects.

The goal of these studies was to gain insight into the mechanisms underlying pricing effects in humans. Although we know much about how and when these effects occur, little work to date had addressed where these effects come from in the first place. By comparing our own biases to those of capuchin monkeys, we hope to have shed light on the mechanisms underlying human pricing effects.

Indeed, we have observed that—in contrast to other decision-making biases—pricing effects may be uniquely human. Our results therefore hint that monkeys may choose between goods simply based on their experience with different items rather than using the sorts of arbitrary factors that humans use.

## Chapter 9: The costs and benefits of excellence

*A bullshit-on-bullshit investigation of wine awards*

This chapter estimates the relationships between New York City restaurant prices, *Zagat* food ratings, and the three levels of *Wine Spectator* “Awards of Excellence,” which are given each year to about 4,000 restaurants that are determined by the expert critics at the world’s most popular wine magazine to be the world’s best wine restaurants (*Wine Spectator*, 2019). In order to evaluate the market role played by the *Wine Spectator* awards, we first undertake an investigative experiment to test the hypothesis that awards are “non-rationed,” i.e. available to any restaurant, or non-restaurant, for the \$250 entry fee, plus associated labor, materials, and postage costs.

Our analysis in this chapter draws on four different data sources: (1) one datum that we collect from an investigation into the standards for the *Wine Spectator* Awards of Excellence; (2) data on 1,709 restaurants in New York City from the *Zagat Survey*, including average meal price, location, cuisine type, and food, décor and service rating; and (3) data on which of these 1,709 restaurants won *Wine Spectator* Awards of Excellence at each of the three levels (a total of 141, or 8%, of restaurants).

We begin by describing and summarizing these three data sets. Next we run regressions using two hedonic price models. In Model 1, price per meal is the dependent variable; and location, cuisine, *Zagat* ratings of food, décor, and service, and *Wine Spectator* awards are independent variables. In Model 2, food rating is the dependent variable, and price becomes an independent variable. We report regression results for full and simplified versions of both models. For Model 2, we also report separate results for high-priced, midrange, and low-priced restaurants. We find statistically significant main effects for *Wine Spectator* awards and *Zagat* ratings, as well as effects for location, number of cuisines served, and some cuisine types.

### Wine Spectator experiment

#### Wine Spectator experiment: motivation

Each year, the world’s leading wine magazine, *Wine Spectator*, devotes one of its 12 annual print issue to announcing its Awards of Excellence. Information about award-winning restaurants is also published on *Wine Spectator*’s website, and award winners are sent a certificate suitable for framing (Figure 9.1).

Figure 9.1. *Wine Spectator* Award of Excellence certificate on display at Gordon Ramsay Steak, Baltimore, Maryland.



Source: <http://www.twipu.com/HorseshoeBmore>

*Wine Spectator* was originally founded as a San Diego tabloid in 1976. It was acquired by Marvin Shanken (Figure 9.2) and is now a business division of M. Shanken Communications of New York City.

Figure 9.2. Marvin Shanken (left), publisher of *Wine Spectator*.



Source: The Internet.



M. Shanken also publishes *Cigar Aficionado* and the newly launched, pompously spelled, and audaciously named *Whisky Advocate*.

### **“Whisk(e)y.”**

The name *Whisky Advocate* is audacious, to say the least, given that Robert Parker’s *Wine Advocate*, one of *Wine Spectator*’s main competitors in the wine-ratings space, has nothing to do with this new M. Shanken project. Shanken’s spelling is worth noting: it’s “Whisky,” not “Whiskey.” The Irish, notes Micallef (2018), “spell whiskey with an ‘e’ between the ‘k’ and the ‘y’ while their Scottish counterparts leave out the ‘e.’” However, this is not why common Americans (unlike Shanken) spell it “whiskey,” because the United States destroyed branded Irish Catholic whiskey with Prohibition. It was not until the Volstead Act that global Scotch whisky sales exceeded Irish whiskey sales for the first time. The Bureau of Alcohol, Tobacco and Firearms (BATF) initially adopted the Scotch Protestant “whisky,” but now uses “whiskey.” Micallef writes: “Since Irish whiskey was more popular, was considered by many to be of a higher quality than Scotch whisky and sold at a roughly 25% premium, many American distillers, tried to associate their products with Irish whiskey and adopted the Irish spelling. The practice of spelling whiskey with an e stuck, even after Irish whiskey sales declined” (Micallef, 2018). To sum up, “whiskey” was first used by the Irish to prove they weren’t Scotch, then by the Scotch to pretend they were Irish, before finally being abandoned by Shanken in favor of “whisky” to prove that he’s Scotch and not Irish, or, at the least, definitely not a guy from New Haven who moved to Miami.

*Wine Spectator* charges a fee to restaurants in exchange for being considered for an Award of Excellence. Although M. Shanken’s events business may be its most profitable unit,<sup>38</sup> the Awards of Excellence program was, and continues to be, a nontrivial moneymaker for the company. Restaurants applying for awards also incur some costs of assembling and

---

<sup>38</sup> The New York Wine Experience, launched in 1981, draws 6,000 attendees per year, according to <https://www.nywineexperience.com/tickets.html>. There’s the option of buying tickets to one of two evening “Grand Tastings,” featuring 269 wines rated 90+ points; these tickets are sold at a normal rate of 2.5 hours for \$375 @ \$150/hour, or a bulk/alcoholic’s (“VIP”) rate of 3.5 hours for \$475 @ \$135/hour. There is also the 2.5-day pass for \$2,495 that includes tickets to both evening tastings plus two days of morning and afternoon seminars by highly-rated winemakers, two midday lunches, and one evening banquet that honors the restaurants winning Wine Spectator Grand Awards of Excellence. It’s not clear what percentage of the 6,000 attendees buy which type of ticket, but if one third (2,000) bought each of the three types, this might be around \$7 million.

mailing the materials. Award winners are notified of their awards several months after applying.

In 2008, the year of our investigative experiment, about 4,000 restaurants won an Award of Excellence. At that time, the entry fee was \$250 per restaurant, translating to roughly \$1 million in aggregate revenue for M. Shanken Communications, not including revenues from the fees submitted by the unobserved number of applicants who fail. In 2008, restaurants also incurred postage and labor costs in preparing and mailing in the application materials.

The number of *Wine Spectator* Awards of Excellence has fluctuated in the years since 2008. In 2018, 3,759 restaurants received awards. Applications can now be submitted online. As of 2019, the regular entry fee payable to M. Shanken was \$425 per restaurant, with an early-bird fee of \$350, translating to roughly \$1.5 million in aggregate revenue for the company (*Wine Spectator*, 2019).

The background research for this paper began in the early spring of 2008, when we undertook an investigative experiment that explored the standards for Awards of Excellence. Our original motivation was the search for a better understanding of *Wine Spectator's* institutional definition of “Excellence,” given that failure rates were not (and still are not) disclosed by the magazine (*Wine Spectator*, 2008, 2019).

We developed a hypothesis that *Wine Spectator* awards are “non-rationed,” meaning that there is a theoretically unlimited supply of awards in the marketplace, in which case “Excellence” would simply mean the ability to complete the form correctly and submit the fee. We sought to test this against the null hypothesis that awards are “rationed,” i.e. that awards are measurements of restaurants’ actual wine cellars and wine service, as evaluated by *Wine Spectator's* expert judges.

We split our overall null hypothesis that Awards of Excellence are non-rationed into three null sub-hypotheses, any of which, if met, would indicate that awards are rationed to some degree. From narrowest to broadest, these three null sub-hypotheses are:

- (1) A restaurant must have a good wine list to win a *Wine Spectator* Award of Excellence.
- (2) A restaurant must have a wine cellar to win a *Wine Spectator* Award of Excellence.
- (3) A restaurant must exist to win a *Wine Spectator* Award of Excellence.

Note that each of our three null hypotheses are stated as “musts”: that is, they are null hypotheses about the *necessary* conditions for awards. A single counter-example, generated by one data point, could therefore suffice to reject each of these three null hypotheses.

We do not mean to imply that these three conditions are collectively *sufficient* for awards: a restaurant that exists, has a wine cellar, and has a good wine list might fail, for



instance, if it neglects to state on the application that it meets the prerequisites for awards or to fill out the application forms correctly. However, if all three null hypotheses are rejected, we can infer that anyone that is capable of filling out the forms correctly and paying the entry fee, as we did, is capable of purchasing an award. The need to comply with paperwork in order to complete the purchase is a condition on purchases in many different markets, so we do not consider the paperwork requirement to be a form of rationing; rather, we classify paperwork requirements, which accompany many types of large purchases, as transaction costs.

If we can reject null hypotheses (1), (2), and (3), we will reject our overall null hypothesis and proceed under the assumption that *Wine Spectator* awards are non-rationed and to model them as attributes that can be purchased by restaurants, not as signals of other restaurant attributes such as a good wine cellar, a wine cellar, or existence.

#### *Wine Spectator* experiment: design.

To test null hypotheses (1), (2), and (3) together, we created an imaginary restaurant, located in Milan, Italy, called “Osteria L’Intrepido.” We submitted an application for a *Wine Spectator* Award of Excellence in spring 2008. Our application form is shown in Figure 9.4 (Goldstein, 2008).

In compliance with the instructions on the form, we listed an address in Milan (the residence of Giuliano Stiglitz; more on Stiglitz below), and we obtained a digital voice mailbox in Milan. Stiglitz left a message on the voice mailbox saying that the restaurant was closed for vacation. We made one posting about the restaurant on Chowhound.com, a food enthusiasts’ website. We complied with the other criteria stated in the application materials on *Wine Spectator*’s website, including the requirement that the restaurant list at least 100 wines on its wine list.

Our application included a “Reserve Wine List” of Osteria L’Intrepido’s special cellar selections. The reserve list, entitled “I rossi italiani ‘riserva’ della nostra cantina” (reserve red wines from our cellar), included 15 wines priced between 80 and 300 euros. We constructed Osteria L’Intrepido’s Reserve Wine List by searching *Wine Spectator* archives for the lowest-rated Italian wines in *Wine Spectator* history and choosing from amongst them. A list of the wines on the reserve wine list and summary statistics of their prices and ratings are shown in Table 9.5a, and *Wine Spectator*’s (2019) published guide to the meaning of ratings is reprinted as Table 9.5b.

Figure 9.4. Osteria L'Intrepido Application for the *Wine Spectator* Award of Excellence.

NEW 2008 WINE SPECTATOR 2008 RESTAURANT AWARDS PROGRAM LISTING FORM 10769

Please review carefully the information we have on file for you (in the left hand column) and correct any errors and FILL IN ANY BLANK AREAS in the right hand column. Please remember that this data is used for both editorial overview and publication in *Wine Spectator* and on our website. PLEASE SIGN AND RETURN THIS FORM ASAP along with the materials needed as shown in the highlighted box below. Please DO NOT copy this listing form for use with any other entries!

Wine Spectator Restaurant Awards Program  
 PH: (212) 684-4224 ext. 781 FAX: (212) 481-0724 EMAIL: restaurantawards@mshanken.com

**Please Note:** We DO NOT HAVE your LISTING FORM for 2008. We HAVE your MENU for 2008. We HAVE your WINE LIST for 2008. We HAVE your ENTRY FEE for 2008.

**\*\* DATA CURRENTLY ON FILE \*\***

1. Restaurant: **Osteria L'Intrepido**
2. Hotel / resort:
3. Contact / Title: **Stiglitz G.S., Owner**
4. Mailing Address: **Viale Filippetti 33**
5. **Milan, 20122**
6. **ITALY**
- 7.
8. Street Address: **Viale Filippetti 33**
9. **Milan, 20122**
10. **ITALY**
- 11.
12. Sommelier:
  
13. Wine Director:
14. General Manager:
15. Owner(s): **Stiglitz G.S.**
  
16. Contact Email (to receive confirmation of receipt of materials and other information from *Wine Spectator*):  
**lintrepido@gmail.com**
17. Reservations **(39) 0-24-074-6174**
18. Business **(39) 0-24-074-6174**
19. FAX **(39) 0-24-074-6174**
20. Website:  
**http://www.**
  
21. Total number of selections on wine list: **0**
22. Total number of bottles in wine inventory: **0**
23. Corkage fee per bottle if allowed (in US\$): **Not Allowed**
24. Are you a private club (yes / no) **No**
25. Dinner entree price range in US\$ **\$0 - \$0**  
 (if prix fixe only, then prix fixe range)
26. Prix fixe menu offered? (yes / no) **No**
27. Prix fixe menu only? (yes / no) **No**
28. Cuisine type:
29. Chef:
30. Credit cards (check all that apply):
31. Meals offered (check all that apply):  Lunch  Dinner
32. Days closed: **None**
  
- By: GS Stiglitz Print Name/Title: Owner General Manager Date: 14 Febb 2008

**\*\* LIST CHANGES OR CORRECTIONS HERE \*\***

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_
6. \_\_\_\_\_
7. \_\_\_\_\_
8. \_\_\_\_\_
9. \_\_\_\_\_
10. \_\_\_\_\_
11. \_\_\_\_\_
12. Augusto CRAZIA
  
13. Augusto CRAZIA
14. Luca GAMBERINI
15. GS STIGLITZ  
 Please send correspondences to GS Stiglitz  
 But not publish name of owner in magazine
16. \_\_\_\_\_
17. \_\_\_\_\_
18. \_\_\_\_\_
19. \_\_\_\_\_
20. http://www.osterialintrepido.wordpress.com
  
21. 256
22. 2,100
23. \$ \_\_\_\_\_ per bottle
24.  YES  NO
25. \$ 30 to \$ 42 (secondi piatti eur. 20-28)  
 for a la carte entrees (or prix fixe if prix fixe only)
26.  YES  NO
27.  YES  NO
28. Milanese
29. Paolo GAGGINO
30.  AMEX  MC  VISA  DV  None
31.  Lunch  Dinner
32. Closed Monday and Sunday night

Table 9.5a. Wines on Osteria L'Intrepido's "reserve wine list."

<b>Vintage</b>	<b>Wine</b>	<b>Producer</b>	<b>Region</b>	<b>Tasting notes from <i>Wine Spectator</i> review</b>	<b>Rating</b>	<b>Price, EUR</b>
1998	Amarone Classico	Tedeschi	Veneto	Stale black licorice, slightly frothy	65	80
1998	Amarone Classico "La Fabriseria"	Tedeschi	Veneto	Unacceptable, cloying, smells like bug spray	60	185
1993	Amarone Classico "Gioé"	S. Sofia	Veneto	Too much paint thinner and nail varnish	69	110
1985	Barbaresco Asij	Ceretto	Piedmont	Earthy, swampy, gamy, harsh and tannic	64	135
1990	Barolo	Vajra	Piedmont	Musty, lacking in charm or much fruit	64	140
1982	Barolo Riserva	Giacosa	Piedmont	Sharp and harsh; aggressive tannins	72	250
1994	Barolo Zonchera	Ceretto	Piedmont	Disjointed, green, coarse, chewy, astringent	74	120
1996	Brunello di Montalcino Riserva	Soldera	Tuscany	Turpentine, hard, acidic, disappointing	74	235
1982	Brunello di Montalcino "La Casa"	Caparzo	Tuscany	Smells barnyardy and tastes decayed	67	200
1993	Brunello di Montalcino	Caparzo	Tuscany	Too much new wood, lacking in concentration	80	180
1995	Brunello di Montalcino Riserva	Caparzo	Tuscany	Light-bodied, diluted finish, disappointing	81	135
1995	Cabernet Sauvignon "I Fossaretti"	Bertelli	Tuscany	Something wrong here. Metallic and odd. Corky.	58	120
1976	Sassicaia	San Guido	Tuscany	Memorably bad vintage. Oxidized, bitter.	65	250
1980	Sassicaia	San Guido	Tuscany	Light, watery and diluted vanilla. Modest.	77	280
1995	Sassicaia	San Guido	Tuscany	Rich in currant, blackberry, dried herbs, leather.	88	300

<b>Summary</b>	<b>Rating</b>	<b>Price, EUR</b>
Count ( $n=$ )	15	15
Min	59	80
Max	88	300
Median	69	180
Mean	70.53	181.33
Standard deviation	8.46	68.33

Table 9.5b. *Wine Spectator* guide to ratings points.

95-100 Classic: a great wine.

90-94 Outstanding: a wine of superior character and style.

85-89 Very good: a wine with special qualities.

80-84 Good: a solid, well-made wine.

75-79 Mediocre: a drinkable wine that may have minor flaws.

50-74 Not recommended.

Source: *Wine Spectator* (2019)

To be clear, only the wines' vintages, names, producers, regions, and prices—not their *Wine Spectator* scores and descriptions—appeared on the list we submitted to *Wine Spectator* as part of the application for Osteria L'Intrepido. (We wanted to test whether *Wine Spectator*'s expert judges would bother to look up their own ratings of the most prized and expensive wines on our list.) The *Wine Spectator* ratings and tasting notes shown in Figure 9.5a were taken directly from the magazine's own online ratings archive and added afterward. All 15 wines on Osteria L'Intrepido's Reserve Wine List were rated by *Wine Spectator*. Ratings listed in Figure 9.5a are of the same vintages that were listed on the reserve list.

The mean *Wine Spectator* rating of the 15 wines on Osteria L'Intrepido's reserve wine list was 70.5 points, approximately four standard deviations below *Wine Spectator*'s overall mean rating of about 86 points, as measured in 2001 (Miu, 2001).<sup>39</sup> Only one of the 15 wines was rated above *Wine Spectator*'s mean rating.

All wine prices on the reserve list were at least 80 euros, more than two standard deviations above the mean price of *Wine Spectator*-rated wines (as measured several years earlier).<sup>40</sup> The mean wine price on the reserve list was 181.3 euros.

The application for the Award of Excellence also required a menu and a cover letter. We constructed an imaginary menu of typical bumbling Italian nouvelle fusion cuisine and wrote a cover letter in bumbling English. The application asked for a website, so we published a simple WordPress website on wordpress.com (Figure 9.6) that included the restaurant's address, voice mailbox, and menu. The restaurant's front page stated (in Italian) that the restaurant would be open through Christmas 2007 and New Year's 2008. No photographs were submitted with our application or posted elsewhere.

---

<sup>39</sup> In a sample of 5,908 *Wine Spectator* ratings, the mean rating was 85.91 points with a standard deviation was 3.83 points (Miu, 2001).

<sup>40</sup> Assumes 1.1 USD = 1 EUR. In Miu (2001), the mean wine price of *Wine Spectator*-rated wines was \$23.37 with a standard deviation of \$30.46 (Miu, 2001). However, this study was of 2000 data, and prices have not been adjusted for inflation.

Figure 9.6. Osteria L'Intrepido website, 2008.



### Wine Spectator experiment: datum

The first response we received from *Wine Spectator*, several months after we submitted our award application, came in the form of a voicemail that was left in our digital mailbox on May 22, 2008. It was an English-language message left by a representative from the ad sales department of *Wine Spectator*. She congratulated us for winning an Award of Excellence, but she did not mention which level of award we had won. She asked if we might be interested in purchasing an additional print advertisement to promote our award in the issue in which it would be announced (August 2008). The prices of these print advertisements, according to published rates at the time, ranged from \$3,000 to \$8,000. We did not respond to *Wine Spectator's* sales call or purchase an ad, nor did we contact the magazine for any other reason after submitting our application.

In early August 2008, the issue of *Wine Spectator* that revealed Osteria L'Intrepido's Award of Excellence was published.

Figure 9.7. *Wine Spectator* Award of Excellence for Osteria L'Intrepido.



On August 15, 2008, in a presentation of the results in Goldstein et al. (2008) at the annual meeting of the American Association of Wine Economists in Portland, Oregon, we announced Osteria L'Intrepido's award to the wine economics community, and wine economists at the meeting spread the word to journalists and the public. On the same day of the presentation, we replaced Osteria L'Intrepido's Wordpress website with a blog entry describing the experiment (Goldstein, 2008), the application materials shown in Figures 9.4–9.6, and an audio recording of the *Wine Spectator* voicemail. By this point, Osteria L'Intrepido's award certificate, stamped with Marvin R. Shanken's signature, had already been mailed to Stiglitz in Milan (Figure 9.7).

After the announcement at the AAWE meeting in Portland and posting of the article, *Wine Spectator* Editor-in-Chief Thomas Matthews published a response to what he called an “elaborate hoax” and “publicity-seeking scam” and pointed out, in his defense, that “the wine list earned an Award of Excellence, the most basic of our three award levels” (Matthews, 2008).

This gave us a working definition of basic “Excellence” from the point of view of the world's leading wine magazine. The full price of Excellence in 2008, we determined, was \$250 paid directly to *Wine Spectator* plus the material and labor costs of the application process. In the case of Osteria L'Intrepido, the task of applying for and receiving the award totaled to about 5 hours of labor, which we valued at about \$10 per hour in the restaurant industry, plus about \$50 in postage costs, for a total of \$350 in costs.

We can now venture answers to the three original questions posed by this investigation into the *Wine Spectator* standards of Excellence:

(1) *Must a restaurant have a good wine list to win an Award of Excellence?*

No. Results from the “reserve wine list” portion of our investigative experiment provide evidence that a good wine list is not a necessary condition for winning a *Wine Spectator* Award of Excellence. The single counter-example of Osteria L'Intrepido, whose “reserve list” was made up of 15 wines with a mean price of 181 euros per bottle and a mean *Wine Spectator* rating four standard deviations below *Wine Spectator*'s mean for all wines. This provides sufficient grounds for rejecting the Null Hypothesis 1: that a good wine list is necessary condition for winning a *Wine Spectator* Award of Excellence.

(2) *Must a restaurant have a wine cellar in order to win an Award of Excellence?*

No. The Milan address for Osteria L'Intrepido that we provided on the application form we sent to *Wine Spectator* was the home address of Giuliano Stiglitz, who worked in mobile advertising sales for Orange Telecom at the time. We listed “GS Stiglitz” as “Owner” of Osteria L'Intrepido on the application form shown in Figure 9.4 above. Giuliano Stiglitz

claims to be a distant relation to the Nobel Prize-winning economist Joseph Stiglitz but cannot prove it.<sup>41</sup>

At the time of the application, Giuliano Stiglitz's apartment did contain some wine. There was week-old half-empty box of Montepulciano d'Abruzzo, and there were two bottles of Calabrian plonk that had been given to Stiglitz by his kindly 98-year-old grandfather, Giovanni La Manna, a retired member of the Italian Parliament. La Manna was a member of the Communist Party, who drank plonk as a matter of principle.<sup>42</sup> All of the wine at Stiglitz's apartment was stored on the kitchen counter, next to the sink.<sup>43</sup>

In sum, there was no wine cellar at Osteria L'Intrepido, and the total wine inventory at the listed address amounted to a total of approximately 2.5 liters. Thus we can conclude that a wine cellar is not a prerequisite for the basic Award of Excellence. Observations of Stiglitz's apartment provide sufficient evidence to reject Null Hypothesis 2: that a wine cellar is a necessary condition for winning a *Wine Spectator* Award of Excellence.

(3) *Must a restaurant exist to win an Award of Excellence?*

No.

This preliminary investigative experiment was undertaken for the purpose of better understanding the process by which the world's top 4,000 or so wine restaurants are selected each year by *Wine Spectator's* expert judges, so as to determine whether *Wine Spectator* Awards of Excellence are rationed or non-rationed goods. We reject all three Null Hypotheses by counter-example, and we conclude that the *Wine Spectator* Award of Excellence is simply a non-rationed good sold to restaurants, or non-restaurants, at a 2008 price of \$250, with a total cost of approximately \$350 to the restaurant including labor, materials, and postage. Proceeding on the basis of this single but meaningful datum, we move on to investigating the relationship between awards and meal price at New York City restaurants. To do this, we collect data from the *Zagat Survey* restaurant guide, described in Section 9.5.

---

<sup>41</sup> In the Italian language, the leading initial "G" in "GS Stiglitz" could theoretically stand either for "Giuseppe" (Italian for "Joseph") or for "Giuliano." On the flip side, however, Joseph Stiglitz's middle name, "Eugene," is inconsistent with the middle initial "S."

<sup>42</sup> La Manna's Communist Party membership casts even more doubt upon the Giuliano-Joseph family connection.

<sup>43</sup> As the only person in possession of hard evidence that any of this really happened—in the form of the original certificate that still hangs on his office wall—Giuliano Stiglitz remains the sole owner of Osteria L'Intrepido.



## Hedonic price models

### *Zagat Survey* data

Our restaurant price and attribute data set of approximately 1,700 observations was collected from the 2008 New York City *Zagat Survey* (Figure 9.8), a popular restaurant guide that includes ratings aggregated from reader opinions. We chose to use these 10-year-old *Zagat* data so that it would correspond to the 2008 *Wine Spectator* experiment described above.

Figure 9.8. 2008 *Zagat* New York restaurant guide.



The *Zagat Survey* is a restaurant guide series, founded in New York in the 1970s, that publishes print and digital ratings and short reviews of thousands of restaurants in dozens of cities around the world. According to the *Zagat* editors, the ratings and judgments expressed in the *Zagat Survey* come from the opinions of several thousand voluntary annual survey participants, who participate by mail or online and are not paid for their

participation. Editors compile aggregate ratings. Restaurants do not have to pay to be included in *Zagat*.

Beyond the usual self-selection bias that is inherent to any voluntary, open-access survey, additional bias is likely to result from the almost certain presence of some restaurant-industry fluffers posing as voters; in its once-a-year pre-Google format, *Zagat* might have had less power to detect fluffing than other sources of crowdsourced reviews, such as Yelp! (see Luca and Zervas, 2015).<sup>44</sup>

Participation in the *Zagat Survey* is open to the public. Survey forms, which can be completed and mailed in or submitted online, elicit simple ratings for food, décor, and service on a four-point scale.

*Zagat* subsequently aggregates the data and converts to its 30-point scale using an opaque formula and unobservable editorial discretion. We do not know the extent to which data are filtered or checked for fluffing. Statistics about the composition of survey participants are not released by *Zagat*. Given the opacity of the *Zagat* formula, these ratings data cannot be viewed as representing the opinions of a cross-section of New York City restaurant-goers or any other readily describable consumer group. However, the *Zagat Survey* has been a respected source of consumer restaurant information for decades, especially in its home base of New York. New York City restaurants' *Zagat* ratings, like their *New York Times* or *Michelin* ratings, are considered to be important indicators (and, in some cases, maybe even influencers) of a restaurant's esteem.

In 2008, the *Zagat Survey* monetized its content mostly through print books,<sup>45</sup> often co-branded as corporate gifts, with some contribution also coming from paid online subscriptions. The company was subsequently acquired by Google, which now publishes *Zagat* ratings free online as part of its local map-based content. We do not know to what extent Google's version of the survey today corresponds to the process used in 2008.

*Zagat* publishes separate ratings for food, décor, and service for each restaurant. Each rating is an integer between 1 and 30. There is no aggregate rating. The approximate price of a meal at each restaurant, per person (as determined by the *Zagat* editors via research and reader feedback) is published in the *Zagat* guide, along with the restaurant's rating, a short review quoting the pithiest remarks of the survey participants, plus some practical information about each restaurant. Several students at the University of British Columbia assisted with data entry.

---

<sup>44</sup> “Fluffers” and the practice of “fluffing,” more commonly known in the tech industry as “astroturfers” and “astroturfing,” refers to the posting of glowing reviews by undercover insiders or paid contractors who are posing as impartial customers. Starbucks CEO Howard Schultz was one notable practitioner of this craft.

<sup>45</sup> Many of which, it bears mention, were sold to restaurants.

### Wine Spectator and Zagat combined data set

The full data set that we use in the regressions in this chapter combines data from *Wine Spectator* (on award-winning restaurants in New York City) with data from the New York City *Zagat Survey*. Our data set consists of one observation for every restaurant that was rated by the 2008 New York City *Zagat Survey*, along with the restaurant's three subjective *Zagat* ratings, the restaurant's neighborhood (within New York City, e.g. "Flatiron District"), the types of cuisine and number of different cuisines served at the restaurant, and the restaurant's average meal price per person as estimated by *Zagat*.

The data set initially included 1,712 ratings and other information for New York City restaurants. We dropped three outlying restaurants with prices above \$200 from the data set. No other restaurants in the remaining 1,709 had prices above \$155. The three outlying observations were not included in any of our regressions. We add variables to the data set that indicate the 141 restaurants in the data set that received *Wine Spectator* Awards of Excellence, with different indicator variables for different levels of award. We explain all of these variables in more detail below.

A summary and brief description of all the variables used in our models and analysis in the combined data set is shown in Table 9.9 (Table 9.9a lists objective restaurant attributes such as location, meal price, and cuisine type, whereas Table 9.9b lists subjective opinion attributes: ratings and awards). Table 9.10 summarizes the data set numerically. The variables shown in Tables 9.9 and 9.10 are explained in the paragraphs that follow. Tables 9.9 and 9.10 do not list or describe the 31 individual cuisine type variables; these are detailed in Tables 9.15 and 9.16.

Table 9.9. Combined data set, *Zagat + Wine Spectator*: List of variables with brief descriptions.

#### **9.9a. Objective attributes**

##### ***Location***

###### Raw

Manhattan {0, 1}, binary : indicator for restaurants located in Manhattan

##### ***Meal price***

###### Raw

Price [5, 155], quasi-continuous (integer values only) : average meal price in USD reported by *Zagat*

###### Constructed

ln(Price) [1.61, 5.04], continuous : natural logarithm of meal price

RankPct(Price) [0, 1], continuous : "rank percentile" of meal price (price rank converted to a scale from 0 to 1)

## Table 9.9 (continued)

### ***Cuisine types***

#### Raw

31 cuisine type variables {0, 1}, binary (see Table X) : indicators for individual cuisine types with >1% frequency

CuisineFreq [1, 236], quasi-continuous (integer values only) : frequency of cuisine type in data set

MeanCuisinePrice [9.25, 82], continuous : mean (raw) price of cuisine type in data set

#### Constructed

NumCuisines {1, 2, 3, 4}, ordinal : number of cuisines listed for each restaurant, from 1 to 4 cuisines

TopCuisines {0, 1}, binary : indicator for restaurants with popular cuisine types (CuisineFreq > 75)

### **9.9b. Opinion attributes**

#### ***Zagat ratings***

#### Raw

FoodRating [11, 28], quasi-continuous (integer values only) : *Zagat* food rating points on a 1-to-30 scale

DecorRating [3, 28], quasi-continuous (integer values only) : *Zagat* décor rating points on a 1-to-30 scale

ServiceRating [7, 28], quasi-continuous (integer values only) : *Zagat* service rating points on a 1-to-30 scale

#### Constructed

RankPct(FoodRating) [0, 1], continuous : “rank percentile” of price (price rank converted to a scale from 0 to 1)

RankPct(DecorRating) [0, 1], continuous : “rank percentile” of price (price rank converted to a scale from 0 to 1)

RankPct(ServiceRating) [0,1], continuous : “rank percentile” of price (price rank converted to a scale from 0 to 1)

#### ***Wine Spectator Awards of Excellence***

#### Raw

BasicAward {0, 1}, binary : indicator for restaurants with the “Basic” (lowest) level of *Wine Spectator* award

BestOfAward {0, 1}, binary : indicator for restaurants with the “Best Of” (middle) level of *Wine Spectator* award

GrandAward {0, 1}, binary : indicator for restaurants with the “Grand” (highest) level of *Wine Spectator* award

#### Constructed

HigherAward {0, 1}, binary : indicator for restaurants with either a *Wine Spectator* “Best Of” or “Grand” award

Table 9.10. Combined *Zagat* and *Wine Spectator* data set: Summary statistics of raw data.

Not including 31 individual cuisine type indicator, cuisine frequency, and mean cuisine price variables, which are summarized in Tables 9.11–9.14.

<b>Variables</b>	<b>Obs.</b>	<b>Freq<sup>1</sup></b>	<b>Mean</b>	<b>StdDev</b>	<b>Min</b>	<b>Max</b>
<b>Location</b>						
Manhattan	1709	1412	.826	.379	0	1
<b>Meal price</b>						
RawPrice	1709		40.204	17.811	5	155
<b>Cuisine type</b>						
Cuisines	1709		1.303	.546	1	4
PopCuisine	1709	838	.490	.500	0	1
<b>Zagat ratings</b>						
FoodRating	1709		20.837	2.629	11	28
DecorRating	1709		16.559	4.352	3	28
SvcRating	1709		18.309	3.024	7	28
<b>WS awards</b>						
BasicAward	1709	91	.053	.225	0	1
BestOfAward	1709	43	.025	.157	0	1
GrandAward	1709	7	.004	.064	0	1
HigherAward	1709	50	.029	.169	0	1
AnyAward	1709	141	.083	.275	0	1

<sup>1</sup> Frequency for binary variables only.

<sup>1</sup> See Table 9.12 and Section 9.5.1 for derivation.

### Price data

All 1,709 observations in our *Zagat* data set include estimated average meal prices, in U.S. dollars, as estimated by *Zagat*. These are denominated in integer values representing whole U.S. dollar amounts. We call these “raw prices.” The mean raw price is \$40.62 before dropping the three high-priced outliers, \$40.20 if the outliers are excluded. Median price, with or without the three outliers included, is \$38. The three restaurants with outlying prices (hundreds of dollars more than any others) were excluded from all regressions. The range of raw prices, after dropping the three outliers, is \$5 to \$155. Table 9.11 summarizes meal price data.

Table 9.11. Combined *Zagat* and *Wine Spectator* data set: summary statistics.

Summary of Raw Price (USD)					
Percentiles		Smallest			
1%	12		5		
5%	16		7		
10%	21		8	Obs	1709
25%	28		9	Sum of Wgt.	1709
50%	38			Mean	40.204
				Std. Dev.	17.811
		Largest			
75%	49		132		
90%	62		138	Variance	317.226
95%	71		146	Skewness	1.428
99%	98		155	Kurtosis	7.580

Price Variables	Obs.	Mean	Std. Dev.	Min	Max
<b>1a.</b> RawPrice	1709	40.203	17.811	5	155
<b>1b.</b> lnPrice	1709	3.600	.444	1.609	5.043
<b>1c.</b> PriceRankPct <sup>1</sup>	1709	.500	.289	.0006	1.000

<sup>1</sup> See Table 9.12 for derivation.

### Conversion to rank percentiles

We use three different price variables: “RawPrice” (in U.S. dollars, i.e. prices as they originally appear in the *Zagat* guide); “lnPrice” (the natural logarithm of RawPrice); and what we call “RankPct,” a price-percentile variable. The intuitions behind RankPct are, first, that human beings make judgments more on the basis of relative numbers than in terms of absolute numbers (Dehaene, 2007); and second, that percentile scaling, like log transformation, also serves to reduce the biasing effects on population means and regression results that can result from a few high-magnitude observations (i.e. the rank of the highest-magnitude observation is encoded without regard to its distance from the second-highest).

Beginning with the 1,709 raw price observations (with mean price of \$40.20, standard deviation of \$17.81, and range from \$5 to \$155), we inversely rank-order the observations from 1 to 1,709, such that #1 is assigned to the lowest price in the *Zagat* set (\$5) and #1,709 is assigned to the highest price (\$155). This inverse rank-ordering generates a variable we call “PriceRank.” Because *Zagat* prices are all integers, there are many “ties” in the rankings, i.e. restaurants with the same prices that share the same ranks, which results in far fewer than 1,709 different unique values. For example, there are 37 restaurants in the data set with an average meal price of \$35, all of which share a PriceRank of 719; there are 53 restaurants with an average meal price of \$36, all of which share a PriceRank of 764; and PriceRank does not take any values between 719 and 764). Finally, we linearly re-map PriceRank onto a more intuitive 0-to-1 scale (so that an increase of 0.01 in PriceRankPct represents climbing by one percentile point in the price rankings). We do this by simply dividing PriceRank by 1709 (such that the restaurant with the highest price, #1,709, is assigned to  $1709/1709 = 1$ ; and the restaurant with the lowest price, #1, is assigned to  $1/1709 = 0.059$ ). This generates the final price rank percentile (“PriceRankPct”) variable that we use in all regressions in this chapter (other than the robustness checks of the functional form of the price variable).

We also convert *Zagat* ratings, which in their raw form are on a 1-to-30 scale (see Section 9.5.5), to “rank percentiles” using an analogous process, to obtain “FoodRankPct,” “DecorRankPct,” and “ServiceRankPct.”

The end result of the four RankPct transformations described above, after converting raw ordinal data to rank variables and then to RankPct variables, is that all four of our final “RankPct” variables have 1,709 observations, mean values near 50, minimum values near 0 (.006 to .009), maximum values near 1 (.998 to 1), and standard deviations that vary within a narrow range (.287 to .289).<sup>46</sup>

---

<sup>46</sup> Given that multiple restaurants sometimes have the same values for price or *Zagat* ratings, and that such “ties” are resolved by assigning the same rank to such restaurants, the minimum value in the data set for some of these variables is greater 0.059 and the maximum value may be less than 1. For instance, there are

Table 9.12. Transformations on raw *Zagat* data to obtain RankPct variables.

Price and *Zagat* food, décor, and service ratings are converted from raw USD prices and *Zagat* ratings points into “rank percentiles” using the following procedure. First, all 1,709 observations are rank-ordered by the given variable, where 1 is lowest and 1,709 is highest. These rank values are then linearly scaled to percent values between 0 and 1, where 1 is the highest-ranked restaurant and 0 is the lowest-ranked restaurant.

Data are taken from the 2008 *Zagat Survey* for New York City. All 1,709 observations include prices and all three *Zagat* ratings. Raw prices and ratings take only integer values.

Step 1. Raw data	Obs.	Mean	Std. Dev.	Min	Max
RawPrice	1709	40.20363	17.81083	5	155
Zagat RawFoodPts	1709	20.83733	2.628521	11	28
RawDecorPts	1709	16.55939	4.351856	3	28
RawServicePts	1709	18.30895	3.024381	7	28
Step 2. Ranks	Obs.	Mean	Std. Dev.	Min	Max
PriceRank	1709	855	493.3829	1	1709
FoodRank	1709	855	490.0538	1	1705.5
DecorRank	1709	855	492.0701	1.5	1707
ServiceRank	1709	855	490.8968	1	1708.5
Step 3. RankPct	Obs.	Mean	Std. Dev.	Min	Max
PriceRankPct	1709	.500	.289	.0006	1.000
FoodRankPct	1709	.500	.287	.0006	.998
DecorRankPct	1709	.500	.288	.0009	.999
ServiceRankPct	1709	.500	.287	.0006	1.000

eight restaurants with the highest observed raw *Zagat* food rating of 28, so each of these eight restaurants is assigned a FoodRank of  $(1702 + 1703 + \dots + 1709) / 8 = 1705.5$  (rather than 1709), which is then converted to a FoodRankPct of .998 (rather than 1), as shown in the rightmost column of the “FoodRankPct” row of “Step 3” in Table 9.12. On the other hand, only one restaurant has the highest price of \$155, so the maximum PriceRankPct in the data set is exactly 1.



The process of deriving RankPct variables and a summary of the resulting variables are shown in Table 9.12. In the models we state in this chapter (Sections 9.6–9.10), we generally use RankPct variables for price and *Zagat* ratings as both dependent and independent variables. The one exception is Models 1.3.1a and 1.3.2a, which test the robustness of our models that use PriceRankPct as a dependent variable by comparing them to the same models using RawPrice and lnPrice instead. (See Sections 9.6–9.7 for the statements of these models and their results).

### Location data

In anticipation that Manhattan restaurants might differ systematically in price and other regards, we created a binary (dummy) variable for location called “Manhattan” that is assigned a value of 1 for restaurants in Manhattan (83% of the data set, thus a mean value of .83 for the Manhattan binary variable) and 0 for restaurants in the four other “outer boroughs” of New York City: Brooklyn, Queens, the Bronx, and Staten Island (cumulatively 17% of the data set).

Figure 9.13. Location variable: relative frequency of NYC meal prices listed in *Zagat Survey*, 2008: Manhattan vs. non-Manhattan restaurants.  
*Epanechnikov kernel density plots (for definition, see Chapter 3).*

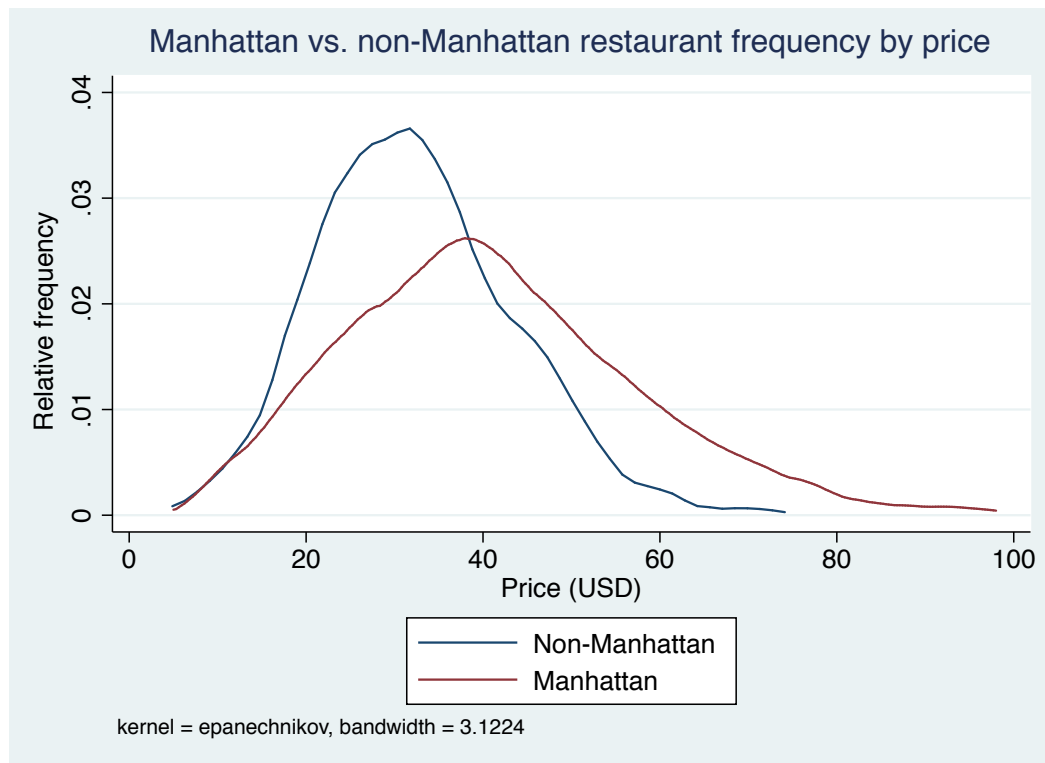


Figure 9.13 compares the price distributions of Manhattan and non-Manhattan restaurants. This plot and several subsequent frequency plots are smoothed using an Epanechnikov kernel density function (Epanechnikov, 1969), a common smoothing technique that is discussed in Chapter 3. In Figure 9.13 and subsequent plots, when we refer to the “relative frequency” of price  $x$ , we mean the number of observations with prices of  $x$  (or, in a smoothed kernel density plot or histogram, as discussed in Chapter 3, in the interval around the bin containing  $x$ ), divided by the total number of observations (or bins).

### Cuisine data

Each restaurant is assigned to one or more cuisine categories, which can represent either a regional culinary theme (e.g. “Italian”), a type of food (e.g. “Seafood”), another type of culinary modifier (e.g. “Traditional”), or some multi-word mix of the above. A total of 135 different cuisine types appear in the data set.

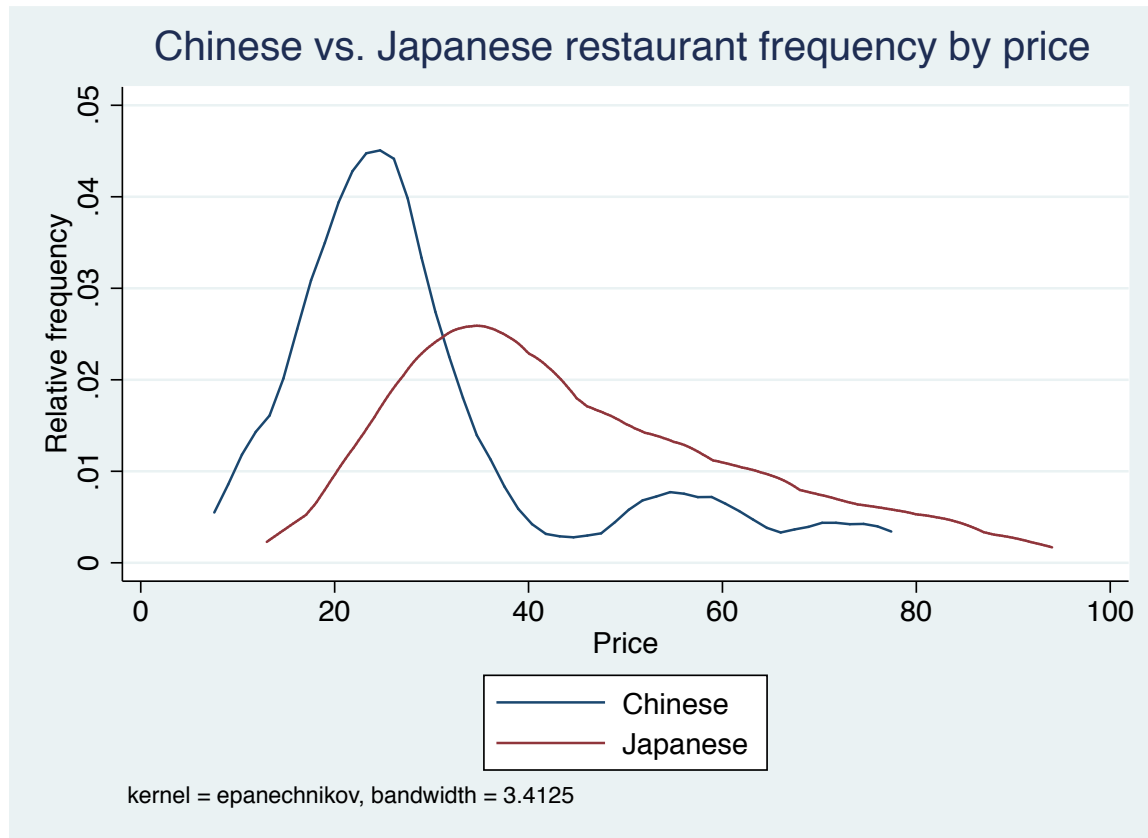
Up to four cuisine categories can be assigned to each restaurant. 1,206 restaurants (71% of all restaurants) list only one cuisine, 441 restaurants (24%) list two cuisines, 50 restaurants (3%) list three cuisines, and 12 (0.7%) list four cuisines. We construct an ordinal variable called “NumCuisines,” which simply represents the number of different cuisine types listed for a restaurant. This variable takes a positive integer from 1 (every restaurant listed in *Zagat* has at least one cuisine type) to 4 (the maximum number of cuisine types listed in the 2008 edition of the guide).

We hypothesize that for restaurant meal prices vary systematically, to some extent, by cuisine type. For instance, Japanese restaurants might intuitively be expected to be more expensive, on average, than Chinese restaurants. Figure 9.14 verifies this Japanese-vs.-Chinese-restaurant intuition graphically with Epanechnikov kernel density plots that compare the relative price frequency distributions of the two cuisine types in the *Zagat* data set.

In order to use cuisine types as independent variables in regressions and thus explore variations in price by cuisine, we first narrow down the list of 135 cuisines to a smaller set of cuisines that occur with enough frequency that their price trends might be reasonably be analyzed in summary statistics or regressions. Of the 135 cuisine types, 31 cuisine types (23%) are assigned to 1% or more (more than 17 restaurants) of the 1,709 restaurants in the data set. The remaining cuisine types (104 of 135, or 77%) occur in fewer than 1% of restaurants (17 restaurants or fewer). Tables 9.15 and 9.16 list the 31 cuisine types that occur with a frequency of 1% or higher. Table 9.15 lists these 31 cuisine types by frequency, and Table 9.16 lists them by mean meal price.

Figure 9.14. Cuisine type variables: Relative frequency of NYC meal prices listed in *Zagat Survey*, 2008: Chinese vs. Japanese restaurants.

*Epanechnikov kernel density plots (for definition, see Chapter 3).*



Source: Zagat Survey, New York City, 2008.

Footnotes to Table 9.15 (table appears on following page 195)

- <sup>1</sup> The >1% frequency threshold corresponds to Total Frequency > Int(1709 x 1%) = >17 occurrences of the cuisine type.
- <sup>2</sup> Frequency rank, total frequency, and total percent frequency are calculated based on occurrences of a cuisine type as any one of a restaurant's cuisines (i.e. up to 4 cuisines per restaurant). Total frequencies are thus not mutually exclusive.
- <sup>3</sup> Frequency count of occurrences as first (Cuisine1), second (Cuisine2), third (Cuisine3), or fourth (Cuisine4) listed cuisines.
- <sup>4</sup> Total does not equal sum of column or row because frequencies are not mutually exclusive (i.e. some restaurants have multiple cuisines with >1% frequency, and thus are counted in more than one of the cuisine frequencies listed above).

Table 9.15. Cuisine types occurring in at least 1% of restaurants,<sup>1</sup> by frequency of occurrence.

<b>Cuisine type</b>	<b>Freq Rank<sup>2</sup></b>	<b>Total Freq<sup>2</sup></b>	<b>Total Pct Freq<sup>2</sup></b>	<b>Freq as Cuisine<sub>1</sub><sup>3</sup></b>	<b>Freq as Cuisine<sub>2</sub><sup>3</sup></b>	<b>Freq as Cuisine<sub>3</sub><sup>3</sup></b>	<b>Freq as Cuisine<sub>4</sub><sup>3</sup></b>
Italian	1	236	13.8%	231	4	1	
American (New)	2	174	10.2%	174			
French (Bistro)	3	93	5.4%	86	7		
Sushi	4	80	4.7%	70	10		
American (Traditional)	5	79	4.6%	79			
French	6	77	4.5%	52	12	5	8
Seafood	7	76	4.4%	35	39	2	
Steakhouse	8	75	4.4%	50	23	2	
Pizza	9	72	4.2%	25	45	1	1
Italian Northern	10	71	4.2%	68	3		
Mediterranean	11	58	3.4%	35	22	1	
Thai	12	50	2.9%	40	8	2	
Mexican	13	48	2.8%	45	3		
Sandwiches	14	41	2.4%	22	11	5	3
Indian	15	40	2.3%	40			
Eclectic / Int'l	16	35	2.0%	29	6		
Greek	17	34	2.0%	33		1	
Hamburgers	18	33	1.9%	21	12		
Japanese	19	31	1.8%	27	4		
Italian Southern	20	30	1.8%	25	5		
Small Plates	21	29	1.7%	2	25	2	
Chinese	22	28	1.6%	27	1		
Tuscan	23	23	1.3%	20	3		
Dessert	23	23	1.3%	10	12	1	
Bakery	23	23	1.3%	19	4		
Vietnamese	26	21	1.2%	11	9	1	
Turkish	27	19	1.1%	16	3		
Coffeeshop / Diner	27	19	1.1%	11	7	1	
Kosher	29	18	1.1%	6	8	4	
Dim Sum	29	18	1.1%	14	4		
Coffeeshouse	29	18	1.1%	6	2	10	
<b>Restaurants w/ TopCuisines</b>							
(8 shaded cuisines; see Sec. 9.10)		838 <sup>4</sup>	49.0%	777	95	10	8
<b>w/ cuisines in full model</b>							
(31 cuisines w/ >1% freq.)		1408 <sup>4</sup>	82.4%	1329	292	39	12
<b>All Zagat-rated restaurants</b>							
(135 cuisines appearing in <i>Zagat</i> NYC)		2212	100.0%	1709	441	50	12

Footnotes to Table 9.15 appear on previous page 194.

Table 9.16. Cuisine types occurring in at least 1% of restaurants,<sup>1</sup> by mean meal price.

*Sorted by mean raw price. Includes all occurrences of cuisine types (as 1st, 2nd, 3rd, or 4th listed cuisine).*

<b>Cuisine type</b>	<b>Price Rank<sup>2</sup></b>	<b>Mean P<sup>2</sup></b>	<b>StdDev<sub>2</sub></b>	<b>Min<sub>2</sub></b>	<b>Max<sup>2</sup></b>	<b>Freq<sup>2</sup></b>	<b>Pct Freq<sup>2</sup></b>
Steakhouse	1	57.2	16.5	26	146	75	4.4%
French	2	54.8	31.1	14	132	77	4.5%
Kosher	3	54.2	16.6	26	71	18	1.1%
Italian Northern	4	51.3	12.5	27	83	71	4.2%
American (New)	5	49.6	19.9	13	127	174	10.2%
Tuscan	6	49.0	8.6	34	62	23	1.3%
Japanese	7	48.7	28.7	13	155	31	1.8%
Seafood	8	47.7	17.7	24	155	76	4.4%
Sushi	9	46.7	19.3	23	138	80	4.7%
Small Plates	10	44.6	10.4	30	71	29	1.7%
Mediterranean	11	43.7	13.4	21	92	58	3.4%
Italian	12	42.4	12.2	12	88	236	13.8%
Italian Southern	13	41.9	12.2	25	73	30	1.8%
French (Bistro)	14	41.8	8.8	25	67	93	5.4%
Greek	15	39.6	13.1	20	73	34	2.0%
Eclectic / Int'l	16	38.2	16.1	16	83	35	2.0%
American (Traditional)	17	35.9	14.7	13	80	79	4.6%
Chinese	18	33.1	19.0	11	74	28	1.6%
Mexican	19	33.0	10.3	12	54	48	2.8%
Indian	20	32.6	9.1	12	57	40	2.3%
Turkish	21	32.3	7.4	12	40	19	1.1%
Vietnamese	22	27.6	14.5	9	54	21	1.2%
Thai	23	27.6	8.0	19	61	50	2.9%
Pizza	24	27.5	9.8	8	35	72	4.2%
Dim Sum	25	26.0	9.3	15	49	18	1.1%
Dessert	26	24.7	11.2	14	53	23	1.3%
Bakery	27	21.6	6.8	9	39	23	1.3%
Hamburgers	28	21.3	10.0	10	63	33	1.9%
Coffeehouse	29	20.7	3.6	16	27	18	1.1%
Coffeeshop / Diner	30	19.8	3.5	14	24	19	1.1%
Sandwiches	31	17.7	7.7	9	39	41	2.4%
<b>Restaurants w/TopCuisines</b> (8 shaded cuisines; see Sec. 9.10)		46.2	18.8	12	155	838 <sup>3</sup>	49.0%
<b>w/ cuisines in full model</b> (31 cuisine types w/ >1% freq.)		41.4	18.4	8	155	1408 <sup>3</sup>	82.4%
<b>w/o cuisines in full model</b> (104 cuisine types w/ <=1% freq.)		34.6	13.3	5	108	301	17.6%
<b>All Zagat-rated restaurants</b> (135 cuisines appearing in <i>Zagat</i> NYC)		40.2	17.8	5	155	1709	100.0%

<sup>1</sup>The >1% frequency threshold corresponds to Total Frequency > Int(1709 x 1%) = 17 occurrences of each cuisine type.

In all variants of our two full regression models (Models 1.1a–1.3a and 2.1a–2.3a), which are stated and tested in Sections 9.6–9.9, we include binary indicator variables for the 31 cuisines with frequency of 1% or higher.

### Zagat ratings data

The three *Zagat* rating variables are called “Food,” “Service,” and “Décor.” All three variables are denominated as integer values, representing *Zagat* points. *Zagat* ratings are limited to the interval between 1 and 30. The maximum observed value for all three ratings is 28. The minimum observed ratings are 11 (Food), 7 (Service), and 3 (Décor).

We convert *Zagat* ratings of food, décor, and service into “RankPct” form in our regressions. In addition to the natural advantages of rank percentile analogous to the ones described with respect to price variables, given that *Zagat* ratings for food, décor, and service have different ranges, RankPct scaling also serves to make regression coefficients of the different *Zagat* scores directly comparable to each other, and thus to be able to draw more intuitive inferences about the relative effect sizes of various restaurant attributes.

Figure 9.17. Raw *Zagat* rating variables by raw price: Food, Service, and Décor.

*Epanechnikov kernel density plots (for definition, see Chapter 3).*

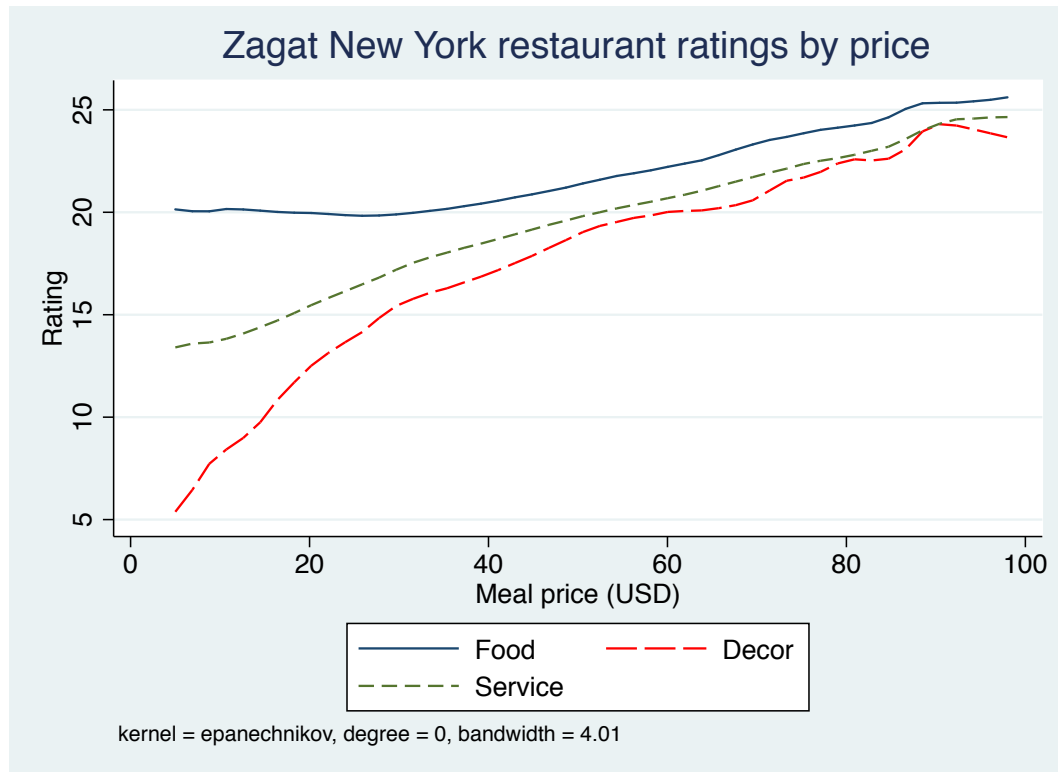


Figure 9.17 plots smoothed averages of three types of *Zagat* ratings (Food, Décor, and Service) by raw price (for ease of interpretation).

A comparison between the three *Zagat* ratings curves in Figure 9.17 gives the initial impression that restaurants in different price segments might be behaving differently. In particular, food ratings, in contrast to service and décor ratings, appear to be relatively flat (in fact, to be precise, slightly downward-sloping) at price points below about \$30. This phenomenon will be explored in much greater detail through the price tercile analysis presented in Models 2.3.1a–2.3.3a (and simplified Models 2.3.1b–2.3.3b) and Figures 9.22–9.23 below.

#### *Wine Spectator* awards in the combined data set

We construct three dummy variables to indicate whether each of the 1,709 restaurants in the *Zagat* data set received a *Wine Spectator* Award of Excellence: one variable for each of the three award levels (from low to high, “Basic Award,” “Best Of Award,” and “Grand Award”). These were simply compiled by matching the 2008 list of *Wine Spectator* Award of Excellence winners in New York City, published online, and matching it with the *Zagat Survey* data set. Like the compilation of the *Zagat* data set, this was done with the assistance of student researchers at the University of British Columbia. A total of 141 restaurants (8% of all restaurants) in the *Zagat* data set received some level of *Wine Spectator* Award of Excellence. Of these, 91 restaurants (5.3%) received the “Basic” award, 50 restaurants (2.9%) received the “Best Of” award, and 7 (0.4%) received the “Grand” award.

The relative price frequency distributions of *Wine Spectator* award-winning restaurants at the three levels is illustrated in Figure 9.18a, which plots the relative frequency by raw price of restaurants winning each of the three award levels (red for Basic, orange for Best Of, and yellow for Grand) and restaurants with no award (blue).

Table 9.18b shows comparative raw-price means by *Wine Spectator* award level. Restaurants with no award have a mean price of \$38.28 (slightly below the population mean price of \$40.20), restaurants with a Basic award have a mean price of \$54.15, restaurants with a Best Of award have a mean price of \$72.79, restaurants with a Grand award have a mean price of \$90.71, and restaurants with any level of *Wine Spectator* award have a mean price of \$61.65.

Figure and Table 9.18. Raw restaurant meal price (USD) by *Wine Spectator* award level.

Figure 9.18a. Frequency distribution of awards by raw price.

*Epanechnikov kernel density plots (for definition, see Chapter 3).*

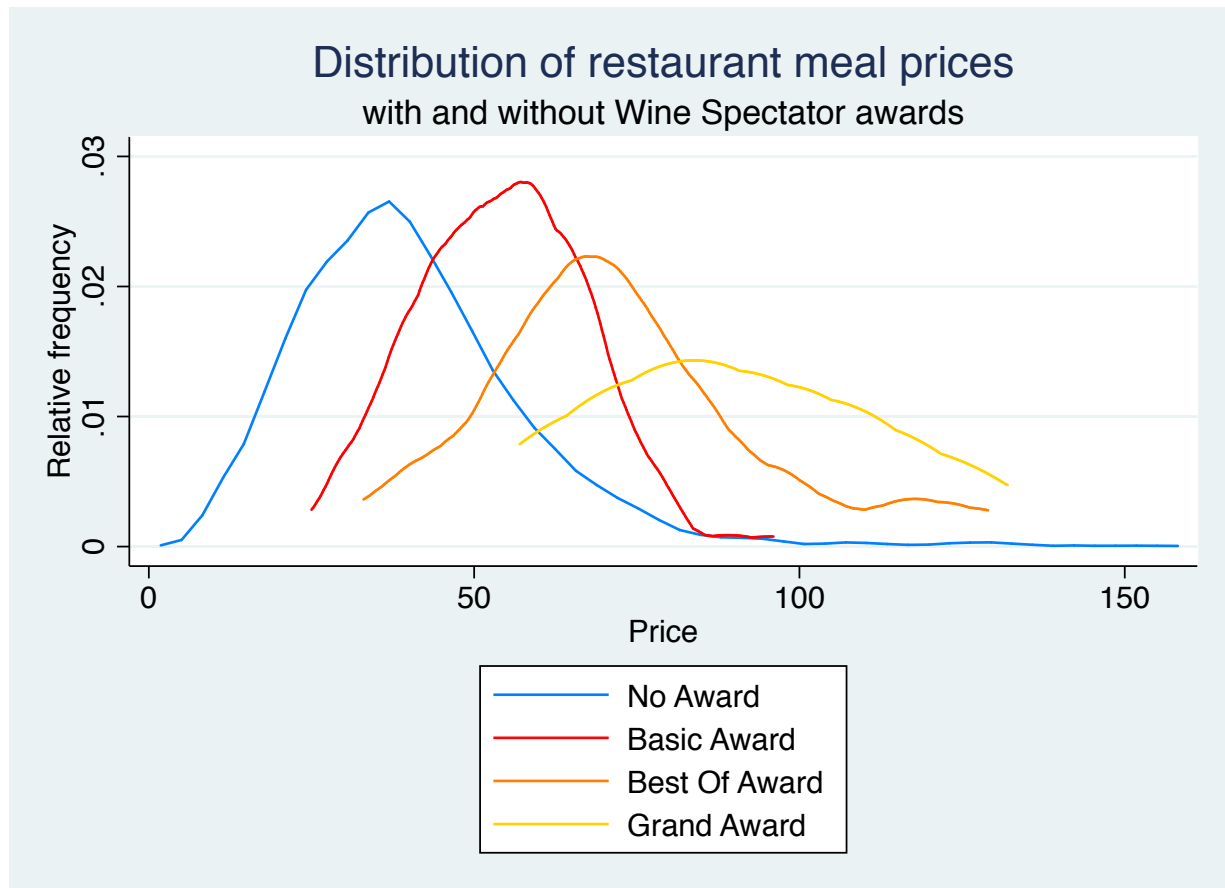


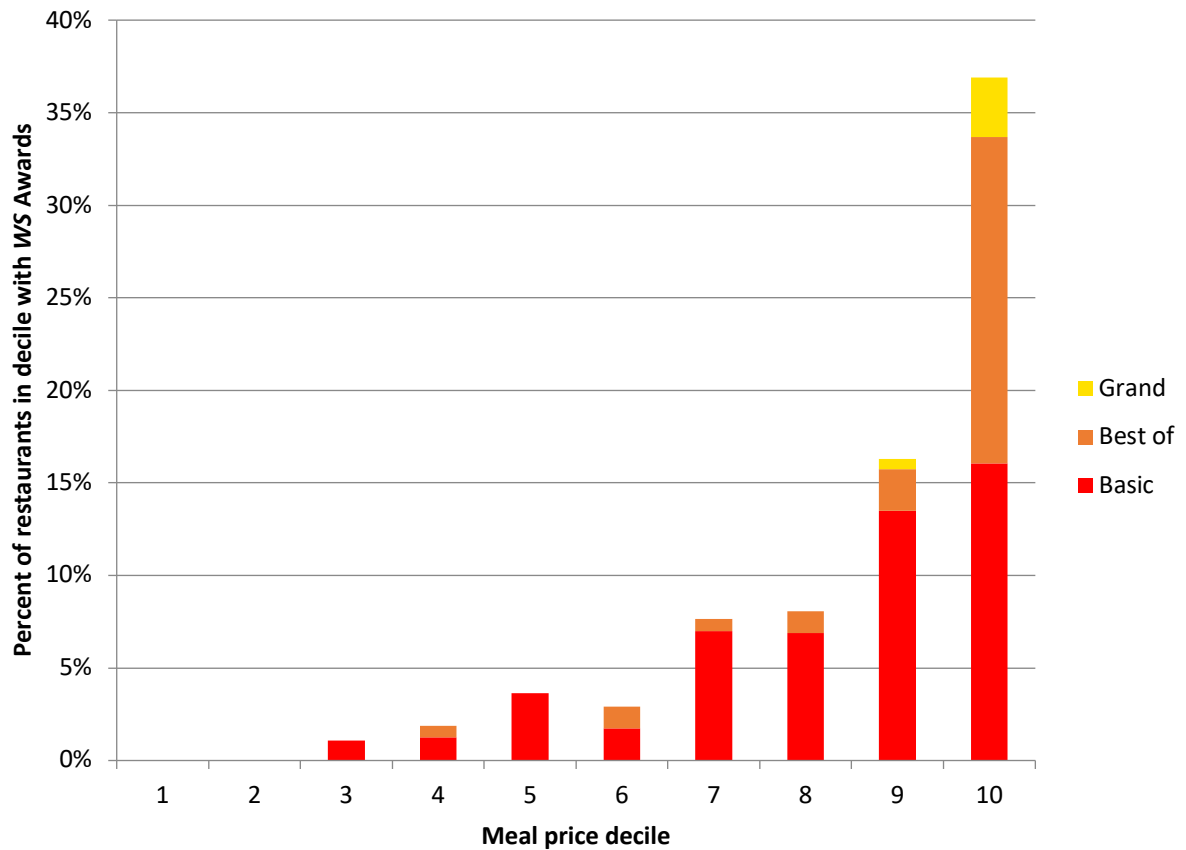
Table 9.18b. Comparative raw-price means by award level.

Award level	Obs	Mean	Std. Dev.	Min	Max
No award	1568	38.275	16.284	5	155
Basic	91	54.154	13.042	25	96
Best Of	43	72.791	21.942	33	129
Grand	7	90.714	24.777	57	132
Any award	141	61.652	19.924	25	132



Figure 9.19 shows a histogram of awards by (*Zagat*) price decile, which shows that award-winning restaurants (especially “Best Of” and “Grand” award winners) crowd into the highest two price deciles. The large number of winners in the highest price decile is driven by a spike in “Best Of” (Level 2) winners, which, like “Grand” award (Level 3) winners, are relatively scarce below the 80th price percentile.

Figure 9.19. *Wine Spectator* award variables: Cumulative plot of the percent of restaurants with three levels of awards by price frequency decile (10 bins).



Given the small incidence of Grand awards, we construct another binary variable called “HigherAward” that combines “Best Of” and “Grand” awards. 57 restaurants (3.3%)—50 “Best Of” award winners plus 7 “Grand” award winners—take values of 1 for this binary variable, and the remaining 1,652 restaurants (96.7%) take values of 0.

### Model 1: Regressions of price on restaurant attributes, ratings, and awards

We now proceed to estimating the first of our hedonic price models, which we call Model 1. In this model and its several variants that we explain below, the average meal price of a restaurant is the dependent variable, and the independent variables are restaurant attributes. In Models 1.1a–1.3a, the dependent variable is *PriceRankPct*, the rank percentile of meal price, as discussed in Section 9.5.2 and derived using the process shown in Table 9.12. The intuition behind our choice of *PriceRankPct* as the dependent variable in our main models is explained in Section 9.5.2.

In Model 1.1a, the independent variables are the restaurant’s location (“Manhattan”), number of different cuisines (“NumCuisines”), and the 31 binary variables representing cuisine types, with frequency of 1% or higher. Model 1.2a adds *Zagat* food, décor, and service ratings to the independent variables. Model 1.3a adds *Wine Spectator* awards. Models 1.3.1a and 1.3.2a use the same set of independent variables as Model 1.3a (the full price model, including *Zagat* ratings), but differ in their dependent price variables. In Model 1.3.1a, the dependent variable is the natural logarithm of meal price, denoted as “lnPrice.” In Model 1.3.2a, the dependent variable is raw meal price (in U.S. dollars), denoted as “RawPrice.” These variants of Model 1 serve to test the robustness of our *PriceRankPct* models by comparing the full *PriceRankPct* version of Model 1 (Model 1.3a) to variants that use the same set of independent variables but use different functional forms of price than *PriceRankPct*.

The five variants of Model 1 are written below as Equations 1.1a–1.3.2a. For simplicity, in our notation, *CuisineTypes* denotes a vector of the 31 cuisine type indicator variables in our model, and its coefficient,  $\beta_3$ , denotes a vector of coefficients corresponding to each of these cuisine type variables.  $\beta_0$  denotes the constant, and  $\varepsilon$  denotes an error term:

$$PriceRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \varepsilon \quad (1.1a)$$

$$PriceRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 FoodRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \varepsilon \quad (1.2a)$$

$$PriceRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 FoodRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \beta_7 BasicAward + \beta_8 HigherAward + \varepsilon \quad (1.3a)$$

$$\ln Price = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 FoodRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \beta_7 BasicAward + \beta_8 HigherAward + \varepsilon \quad (1.3.1a)$$

$$RawPrice = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 FoodRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \beta_7 BasicAward + \beta_8 HigherAward + \varepsilon \quad (1.3.2a)$$

### Model 1 results

The main results of our Model 1 regressions are shown in Table 9.20. Columns 1, 2, and 3 (proceeding from left to right) show results for Models 1.1a, 1.2a, and 1.3a (all using PriceRankPct as the dependent variable). Column 3 is shaded to indicate that Model 1.3a is the full model. Column 4 shows results for Model 1.3.1a (using log price as the dependent variable), and Column 5 shows results for Model 1.3.2a (using raw price as the dependent variable).

Our main finding, shown in the results of Model 1.3a (column 3) in Table 9.20, is that, controlling for location, number of cuisines, and the three *Zagat* ratings, *Wine Spectator* Awards of Excellence are associated with relatively higher meal prices. *BasicAward* and *HigherAward* coefficients are positive and highly statistically significant ( $p < 0.001$ ). Since the dependent variable in Model 1.3a is PriceRankPct, this means that the model predicts that a restaurant listed in *Zagat* that has a basic award or a higher award will be about 5 or 6.8 percentile points higher in average meal price than another *Zagat*-rated restaurant with equivalent location, cuisine, and ratings that does not have an award.

In Model 1.3a (the full model), as shown in Table 9.20, column 3, Manhattan location and *Zagat* “Décor” and “Service” ratings are associated with relatively higher prices ( $p < 0.001$ ), and the number of cuisines served by a restaurant is associated with lower prices ( $p < 0.05$ ). Interestingly, however, coefficients for food ratings are statistically indistinguishable from zero. (We explore this phenomenon in greater depth in Model 2; see Sections 9.8–9.9.)

We observe statistically significant cuisine effects in 19 of the 31 cuisine-type coefficients in Model 1.3a. Steakhouse, Northern Italian, Tuscan, Sushi, Italian, French Bistro, Seafood, French, New American, Greek, and Small Plates cuisine types are associated with higher prices; whereas Indian, Coffeeshop/Diner, Hamburgers, Sandwiches, Dessert, Pizza, Bakery, and Thai cuisines are associated with lower prices.

Table 9.20. Full Models 1.1a–1.3.2a: Results of regressions of meal price on location and cuisine attributes, Zagat ratings, and awards, using price rank percentile as the dependent variable for main results, and log-price and raw-price as a robustness check.

<u>Price regressions (full)</u>	(1) Model 1.1a	(2) Model 1.2a	(3) Model 1.3a	(4) Model 1.3.1a	(5) Model 1.3.2a
<i>Dependent variable</i>	PriceRankPct <sup>1</sup>	PriceRankPct <sup>1</sup>	PriceRankPct <sup>1</sup>	lnPrice	RawPrice
<i>Independent variables</i>	Location+cuis.	+Zagat ratings	+WS awards	Same as 1.3a	Same as 1.3a
Manhattan <sup>2</sup>	0.157*** (0.0128)	0.160*** (0.0101)	0.155*** (0.0102)	0.198*** (0.0157)	8.405*** (0.605)
NumCuisines <sup>3</sup>	-0.0497*** (0.0119)	-0.0184* (0.00823)	-0.0170* (0.00816)	-0.00882 (0.0128)	-1.203* (0.534)
Steakhouse	0.293*** (0.0250)	0.177*** (0.0174)	0.169*** (0.0177)	0.223*** (0.0298)	9.220*** (1.779)
ItalianNorthern	0.256*** (0.0239)	0.125*** (0.0203)	0.123*** (0.0205)	0.145*** (0.0284)	4.764*** (1.337)
Tuscan	0.242*** (0.0357)	0.0882** (0.0294)	0.0943** (0.0294)	0.0935* (0.0404)	2.005 (1.763)
Kosher	0.0621 (0.0687)	0.0933 (0.0562)	0.0937 (0.0562)	0.135 (0.0812)	6.327* (3.111)
Sushi	0.104*** (0.0297)	0.0840*** (0.0197)	0.0897*** (0.0199)	0.170*** (0.0304)	5.547*** (1.645)
Italian	0.137*** (0.0170)	0.0866*** (0.0122)	0.0867*** (0.0122)	0.121*** (0.0170)	2.904*** (0.687)
FrenchBistro	0.100*** (0.0231)	0.0828*** (0.0174)	0.0826*** (0.0175)	0.114*** (0.0227)	2.485** (0.922)
Seafood	0.136*** (0.0268)	0.0772*** (0.0179)	0.0738*** (0.0177)	0.0942*** (0.0259)	2.963 (1.580)
French	0.232*** (0.0304)	0.0758*** (0.0193)	0.0732*** (0.0195)	0.162*** (0.0339)	10.80*** (2.267)
NewAmerican	0.225*** (0.0199)	0.0755*** (0.0132)	0.0706*** (0.0131)	0.0982*** (0.0194)	4.401*** (1.034)
Greek	0.106* (0.0432)	0.0688* (0.0281)	0.0702* (0.0278)	0.112** (0.0406)	2.521 (1.557)
SmallPlates	0.109** (0.0333)	0.0588* (0.0276)	0.0603* (0.0271)	0.0595 (0.0358)	0.687 (1.583)
Japanese	0.152** (0.0536)	0.0551 (0.0294)	0.0578 (0.0295)	0.126* (0.0540)	7.730* (3.650)
ItalianSouthern	0.0970** (0.0375)	0.0511 (0.0316)	0.0522 (0.0311)	0.0878* (0.0379)	1.393 (1.599)
Mediterranean	0.110*** (0.0274)	0.0268 (0.0212)	0.0272 (0.0211)	0.0326 (0.0294)	-0.264 (1.154)
AmericanTrad	0.00445 (0.0315)	0.0154 (0.0227)	0.0150 (0.0226)	0.0279 (0.0319)	1.547 (1.208)
Eclectic	0.0238 (0.0460)	0.00213 (0.0283)	0.00303 (0.0286)	-0.00301 (0.0420)	0.0352 (1.587)

Table 9.20 continues on the following page. Footnotes on p. 205. Std. errors in parentheses; \* =  $p < 0.05$  \*\* =  $p < 0.01$  \*\*\* =  $p < 0.001$ .

Table 9.20, continued	(1) Model 1.1a	(2) Model 1.2a	(3) Model 1.3a	(4) Model 1.3.1a	(5) Model 1.3.2a
<i>Dependent variable</i>	PriceRankPct <sup>1</sup>	PriceRankPct <sup>1</sup>	PriceRankPct <sup>1</sup>	lnPrice	RawPrice
Mexican	-0.0642* (0.0311)	-0.00396 (0.0185)	-0.000982 (0.0186)	0.00579 (0.0344)	-0.715 (0.865)
DimSum	-0.228*** (0.0432)	-0.0131 (0.0218)	-0.0144 (0.0220)	-0.00258 (0.0435)	0.666 (1.045)
Turkish	-0.100* (0.0396)	-0.0434 (0.0256)	-0.0401 (0.0255)	-0.0241 (0.0463)	-1.997 (1.300)
Vietnamese	-0.115* (0.0550)	-0.0492 (0.0252)	-0.0487 (0.0251)	-0.174*** (0.0509)	-4.051*** (1.126)
Chinese	-0.133* (0.0579)	-0.0494 (0.0322)	-0.0505 (0.0321)	-0.109 (0.0570)	-2.319 (2.073)
Coffeehouse	-0.230*** (0.0487)	-0.0692 (0.0413)	-0.0764 (0.0398)	-0.104 (0.0807)	-2.008 (2.540)
Indian	-0.117*** (0.0329)	-0.0829*** (0.0218)	-0.0797*** (0.0217)	-0.0763* (0.0338)	-4.804*** (1.195)
CoffeeshopDiner	-0.259*** (0.0298)	-0.0800* (0.0336)	-0.0825* (0.0332)	-0.121 (0.0677)	-1.187 (1.907)
Hamburgers	-0.299*** (0.0355)	-0.0980*** (0.0187)	-0.0982*** (0.0186)	-0.231*** (0.0499)	-3.889*** (1.130)
Sandwiches	-0.307*** (0.0337)	-0.0978*** (0.0179)	-0.0999*** (0.0177)	-0.343*** (0.0459)	-5.225*** (1.027)
Dessert	-0.0905 (0.0512)	-0.120* (0.0514)	-0.117* (0.0507)	-0.206** (0.0784)	-6.086* (2.736)
Pizza	-0.224*** (0.0238)	-0.120*** (0.0198)	-0.118*** (0.0197)	-0.175*** (0.0309)	-5.699*** (0.985)
Bakery	-0.189*** (0.0500)	-0.138*** (0.0318)	-0.134*** (0.0310)	-0.274*** (0.0715)	-10.29*** (2.152)
Thai	-0.222*** (0.0257)	-0.172*** (0.0218)	-0.169*** (0.0216)	-0.193*** (0.0332)	-7.633*** (1.198)
FoodRankPct <sup>1</sup>		0.0195 (0.0176)	0.0140 (0.0176)	-0.0255 (0.0275)	4.395*** (1.076)
DecorRankPct <sup>1</sup>		0.355*** (0.0204)	0.347*** (0.0205)	0.531*** (0.0311)	19.19*** (1.338)
ServiceRankPct <sup>1</sup>		0.363*** (0.0246)	0.358*** (0.0247)	0.553*** (0.0381)	19.02*** (1.487)
BasicAward <sup>4</sup>			0.0503*** (0.0149)	0.0504* (0.0206)	2.511* (1.073)
HigherAward <sup>4</sup>			0.0681*** (0.0130)	0.180*** (0.0309)	15.26*** (2.541)
_cons	0.376*** (0.0205)	-0.00974 (0.0171)	-0.00185 (0.0171)	2.870*** (0.0305)	11.24*** (1.094)
N	1709	1709	1709	1709	1709
F	55.85	218.9	233.2	114.3	97.83
r2	0.418	0.736	0.738	0.728	0.671
df_m	33	36	38	38	38
df_r	1675	1672	1670	1670	1670

Table 9.20, continued: Footnotes

<sup>1</sup> Price and *Zagat* food, décor, and service ratings are converted from raw USD prices and *Zagat* ratings points into “rank percentile” (RankPct) variables using the procedure shown in Table 9.12 and described in section 9.5.

<sup>2</sup> “Manhattan” is a binary variable that takes 1 for restaurants located in Manhattan, and 0 for restaurants located outside of Manhattan.

<sup>3</sup> “NumCuisines” is an ordinal variable taking positive integer values from 1 to 4, representing the number of different cuisine types listed for each restaurant. Of the 1,709 restaurants in the data set, 1,206 (71%) list one cuisine, 441 (26%) list two cuisines, 50 (3%) list three cuisines, and 12 (0.7%) list four cuisines.

<sup>4</sup> Binary variables indicating restaurants winning *Wine Spectator* awards in 2008. “Basic Award” indicates restaurants winning the first (lowest) level of award ( $n = 91$ ). “Higher Award” indicates restaurants winning either the middle (Best Of) and highest (Grand) levels of awards ( $n = 50$ ).

Comparing full Model 1.3a (with ratings and awards) to the model with ratings but not awards (1.2a), the effects we observe appear generally robust to the inclusion or exclusion of *Wine Spectator* awards. The signs of all coefficients are generally consistent across all three models, even for statistically insignificant coefficients. The addition of ratings results in considerably more predictive power in the model, with R-squared increasing from 0.418 to 0.736; but the addition of awards adds relatively little predictive power to the Model 1.2a, with R-squared increasing to 0.738 in the full Model 1.3a.

Results from Models 1.3.1a (the log-price model, column 4 of Table 9.20) and 1.3.2a (the raw-price model, column 5 of Table 9.20) suggest that this award-price association is fairly robust to different functional forms of the dependent price variable. Both the *BasicAward* and *HigherAward* coefficients are positive and statistically significant in Models 1.3.1a and 1.3.2a, although the *BasicAward* coefficient has lower statistical significance ( $p < 0.05$ ) in Models 1.3.1a and 1.3.2a than it does in Model 1.3a. The biggest difference in award coefficients between Models 1.3a, 1.3.1a, and 1.3.2a lies in the difference between *BasicAward* and *HigherAward* coefficients. In the log-price model (1.3.1a), the *HigherAward* coefficient is more than three times higher than the *BasicAward* coefficient (0.180 vs. 0.0504), and in the raw-price model (1.3.1b) the *HigherAward* coefficient of 15.26 (representing a meal price that is higher by \$15.26, holding other factors constant) is about six times higher than the *BasicAward* coefficient of 2.51 (representing a meal price that is higher by \$2.51); whereas in the price-rank-percentile Model (1.3a), the *HigherAward* coefficient of 0.068 is only about 35% higher than the *BasicAward* coefficient of 0.050.

*Zagat* rating coefficients for décor and service are positive and highly statistically significant ( $p < 0.001$ ) in Models 1.3a, 1.3.1a, and 1.3.2a. There is a positive and statistically significant coefficient for *Zagat* food rating in the raw-price model, whereas there is no statistically significant association between food and awards in the log-price or price-rank-percentile models. Given that even Basic awards are given to higher-than-averaged-priced restaurants (\$54.15 mean price vs. \$40.20 for the full data set; see Table

9.18b), we attribute this effect to the proportionally larger influence exerted by high-priced restaurants when raw, unscaled prices are used.

The signs of cuisine-type coefficients are generally consistent across models 1.3a, 1.3.1a, and 1.3.2a. 18 of 19 cuisine-type coefficients that are statistically significant in Model 1.3a (price rank percentile) are also statistically significant in Model 1.3.1a (log price), all with the same sign. Vietnamese cuisine has a statistically significant and negative association with price in Models 1.3.1a and 1.3.2a, but no statistically significant effect in Model 1.3a. R-squared is 0.738 in the price-rank-percentile model, 0.728 in the log-price model, and 0.671 in the raw-price model. We infer that the choice of the functional form for dependent price variables does not substantially alter our main regression results, although there is some variation between the models. In the rest of this chapter (other than in some graphical illustrations), we refer only to price rank percentile, and not to raw price or log price, when reporting results.

## Model 2

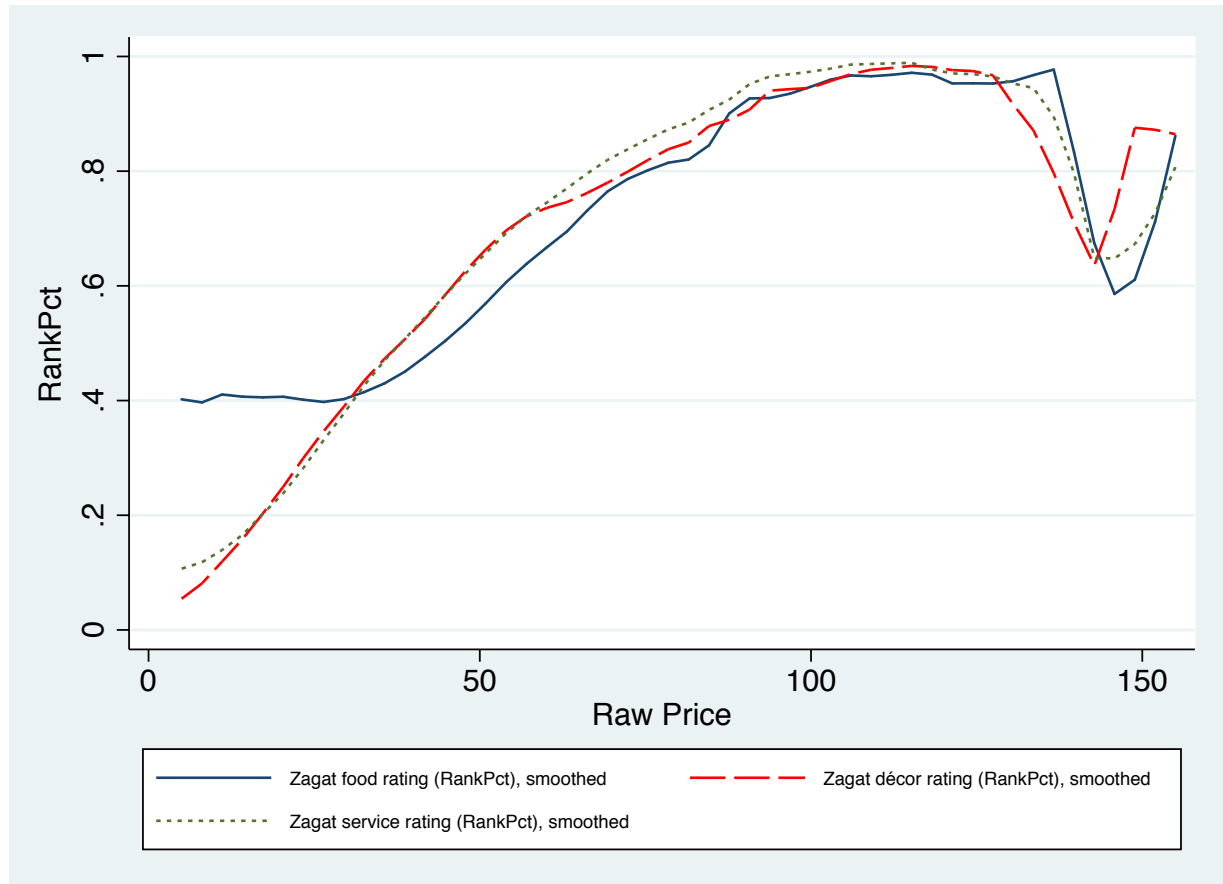
Model 2 moves *Zagat* food rating to the left side of the regression equation and meal price to the right side of the equation. We use food rating as the dependent variable in order to test the counter-intuitive hypothesis that higher-priced restaurants do not serve higher-rated food, if all other restaurant attributes in our model are held constant.

We initially developed this hypothesis from looking at the uncontrolled relationships illustrated in Figure 9.17 (see Section 9.5.5), but the initial hypothesis was limited to the low end of the price spectrum: visually, the uncontrolled food-ratings curve (with food ratings as the Y-axis and raw price as the X-axis) appears to be relatively flat—or even slightly downward-sloping—at meal prices below about \$30. Consistent with this hypothesis, the regression results from Model 1.3a show no significant effect of food ratings on price across the full data set when other restaurant attributes are held constant.

Figures 9.21–9.24 seek to refine this intuition by plotting *Zagat* food, décor, and service ratings (using kernel density smoothing) as a function of restaurant price. Figure 9.21 plots the frequency of *Zagat* ratings (in rank-percentile form) on the Y-axis against raw price on the X-axis.

Figure 9.21. *Zagat* Food, Service, and Décor ratings (RankPct) by raw price.

Frequency plots with Epanechnikov kernel density smoothing (see Chapter 3 for details on this smoothing method).



The curves for *Zagat* food, décor, and service ratings shown in Figure 9.21 paint an interesting picture of the uncontrolled relationship between price and food rating. We use raw price rather than price rank percentile as the X-axis in Figure 9.21 to examine this relationship in an intuitive way (i.e. with respect to actual restaurant meal prices rather than price percentiles). Consistent with Figure 9.17, the food ratings curve appears to be flat at prices below about \$30, whereas the décor and service ratings curves slope upward over that same range. Across the range from about \$30 to about \$125, in contrast, food, décor, and service rankings are all upward-sloping and appear to move together. Then, surprisingly, across the highest price range (from about \$125 to the maximum observed price of \$155), the positive relationship between prices and all three *Zagat* ratings seems to disappear.



Next, motivated by the visual relationships we observe in Figure 9.21, we split the data set into three “Price Terciles,” which correspond roughly to the bottom, middle, and top thirds of the meal-price distribution. Figures 9.22–9.24 plot the relationships between the three *Zagat* ratings (on the Y-axis, and again in rank-percentile form) and raw prices (on the X-axis). Figure 9.22 visually illustrates these relationships for the bottom Price Tercile of restaurants (from \$5 to \$32), Figure 9.23 shows them for the middle Price Tercile (\$33 to \$45), and Figure 9.24 shows them for the top Price Tercile (\$45 to \$155).

Figure 9.22. *Zagat* Food, Service, and Décor ratings (RankPct) by raw price, for the bottom Price Tercile (\$5 to \$32)

*Frequency plots with Epanechnikov kernel density smoothing (see Chapter 3 for details on this method).*

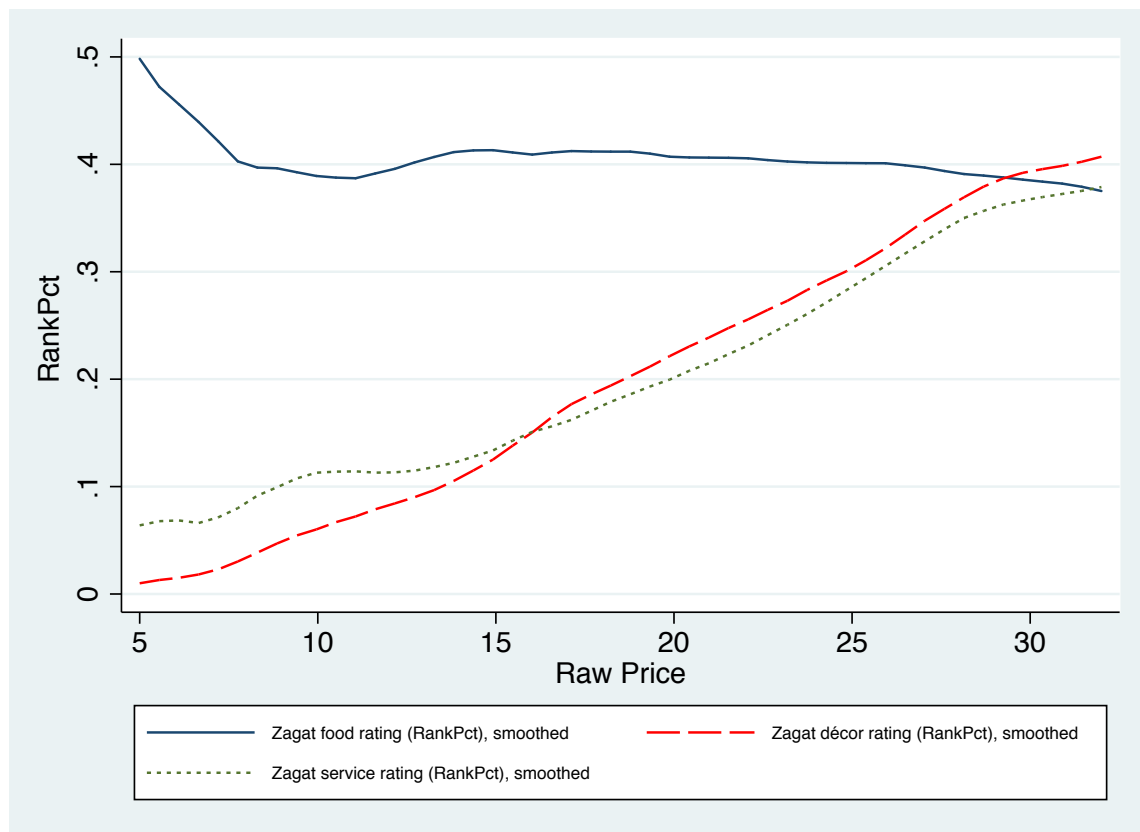


Figure 9.23. *Zagat* Food, Service, and Décor ratings (RankPct) by raw price, for the middle Price Tercile (\$33 to \$45)

*Frequency plots with Epanechnikov kernel density smoothing (see Chapter 3 for details on this method).*

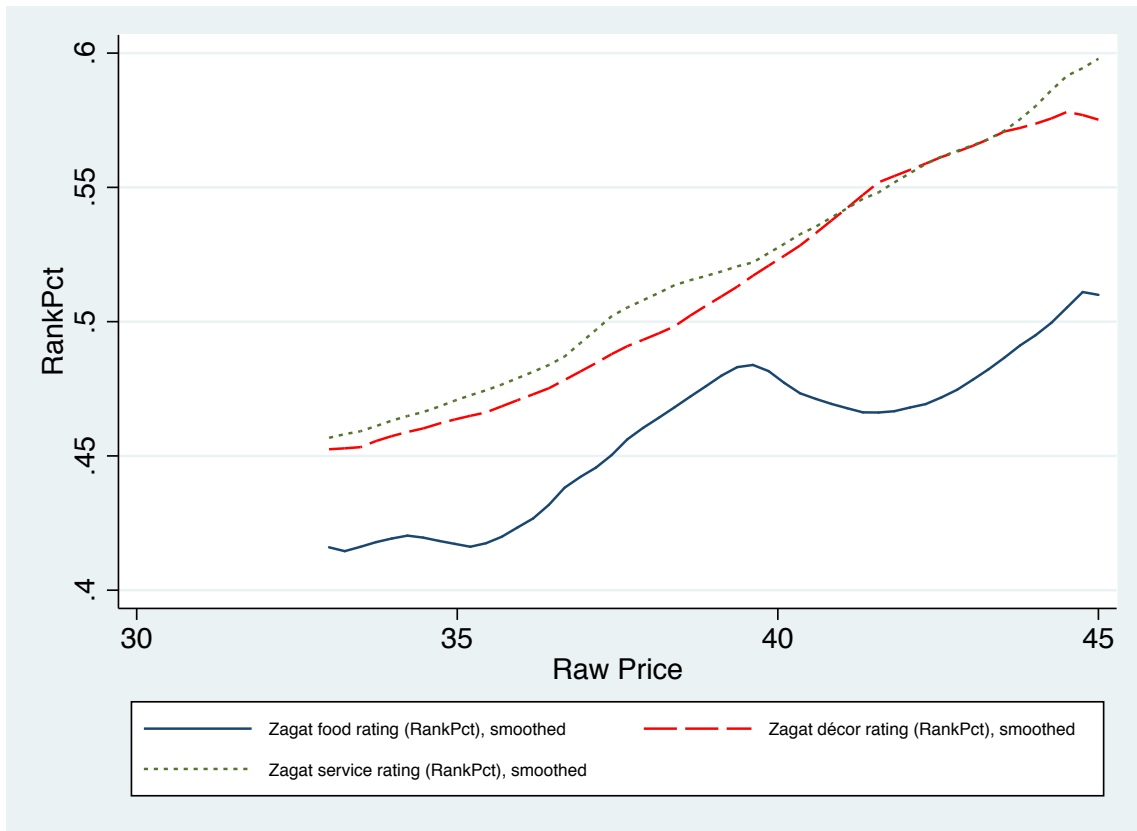


Figure 9.24. *Zagat* Food, Service, and Décor ratings (RankPct) by raw price, for the middle Price Tercile (\$46 to \$155)

*Frequency plots with Epanechnikov kernel density smoothing (see Chapter 3 for details on this method).*

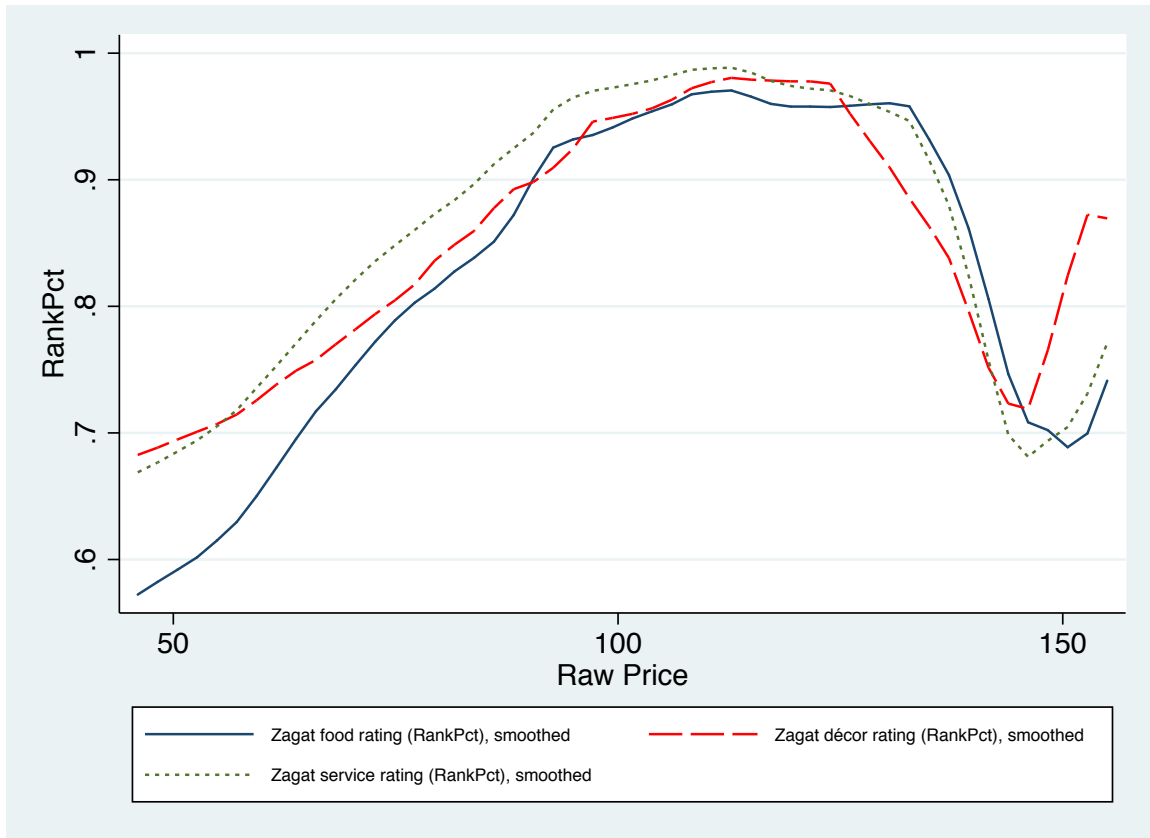


Figure 9.22, showing the bottom Price Tercile, supports the possibility—again, without controls—that there may be an inverse relationship between prices and food ratings in low-priced restaurants. Figure 9.23 suggests a positive relationship between prices and ratings in the middle tercile, and Figure 9.24 shows this relationship appearing to dissolve at prices above \$125. However, there are only nine observations in this highest price range, so we do not develop a testable hypothesis about high-price effects. There are more than 600 observations of restaurants with prices \$32 or lower, so the low-price effects may be testable. With Model 2 (specifically Models 2.3.1a–2.3.3a, which regress food ratings on price and restaurant attributes for each of the three Price Terciles individually), we seek to test the hypothesis that higher-priced restaurants do not serve higher-rated food, holding other attributes constant, within the cheapest third of restaurants. More generally, Models 2.3.1a–2.3.3a test the hypothesis that restaurants in different price categories may differ systematically in the relationships between their prices and various hedonic attributes.

Models 2.1a–2.3a are equivalent to Models 1.1a–1.3a, except that *FoodRankPct* is moved to the left side of the regression equation, becoming the independent variable, whereas (in Models 2.1a–2.3a, using the full data set) *PriceRankPct* is moved to the right side of the equation, becoming a dependent variable. Model 2.1a includes price, location, and cuisine attributes (but not ratings or awards) as independent variables. Model 2.2a adds *Zagat* décor and service ratings as independent variables. Model 2.3a, the full food-ratings model, adds *Wine Spectator* awards as independent variables. Models 2.3.1a–2.3.3a have the same set of independent variables as Model 2.3a, but partition restaurants into three groups simply by dividing the data set, by price, into three terciles (using Stata’s “xtile” function) that each have similar, but not exactly equal, numbers of observations. We call these three groups “Price Terciles.” Price Tercile 1 (the bottom tercile) has 603 restaurants with meal prices between \$5 and \$32. Price Tercile 2 has 567 restaurants with meal prices between \$33 and \$45. Price Tercile 3 has 539 restaurants with meal prices between \$46 and \$155.<sup>47</sup> We regress food ratings on price and restaurant attributes for Price Terciles 1, 2, and 3 separately in Models 2.3.1a, 2.3.2a, and 2.3.3a, respectively.

The six variants of Model 2 are written below as Equations 2.1a–2.3.3a. Again, *CuisineTypes* denotes a vector of the 31 cuisine type indicator variables in our model, and its coefficient,  $\beta_3$ , denotes a vector of coefficients corresponding to each of these cuisine type variables.  $\beta_0$  denotes the constant, and  $\varepsilon$  denotes an error term:

---

<sup>47</sup> The three Price Terciles do not have exactly the same number of members because *Zagat* prices are integer-valued, and the terciles were partitioned using the rule that all restaurants with the same price must be assigned to the same tercile. Thus any way of partitioning the terciles would result in such an imbalance. If, for instance, Tercile 1’s top price of \$32 were moved into Price Tercile 2, this would mean that 54 restaurants would move to Tercile 2, leaving 549 restaurants in Tercile 1 and 621 in Tercile 2.

$$FoodRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 PriceRankPct + \varepsilon \quad (2.1a)$$

$$FoodRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 PriceRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \varepsilon \quad (2.2a)$$

$$FoodRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 PriceRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \beta_7 BasicAward + \beta_8 HigherAward + \varepsilon \quad (2.3a)$$

$$FoodRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 PriceRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \beta_7 BasicAward + \beta_8 HigherAward + \varepsilon : \mathbf{PriceTercile = 1} \quad (2.3.1a)$$

$$FoodRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 PriceRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \beta_7 BasicAward + \beta_8 HigherAward + \varepsilon : \mathbf{PriceTercile = 2} \quad (2.3.2a)$$

$$FoodRankPct = \beta_0 + \beta_1 Manhattan + \beta_2 NumCuisines + \beta_3 CuisineTypes + \beta_4 PriceRankPct + \beta_5 DecorRankPct + \beta_6 ServiceRankPct + \beta_7 BasicAward + \beta_8 HigherAward + \varepsilon : \mathbf{PriceTercile = 3} \quad (2.3.3a)$$

### Model 2 results

The results of the six regressions corresponding to Models 2.1a–2.3.3a are shown together in Table 9.25 (columns 1 to 6, respectively). Cuisine types are sorted in the order of the magnitude of their coefficients in full Model 2.3a, from the most positive to the most negative.

The regression results of full Model 2.3a (column 3 of Table 9.25), like those of Model 1.3a, show no statistically significant relationship between price and food ratings. Model 2.2a (column 2 of Table 9.25) also shows no price-food ratings relationship when *Wine Spectator* awards are not included in the model.

Table 9.25. Full Models 2.0a-2.3a: Results of regressions of Zagat food rating rank percentile on location and full cuisine attributes; Zagat décor and service ratings, and Wine Spectator Awards of Excellence, for the full data set and three individual price tertiles.

<u>Food rating regressions (full)</u>	(1) Model 2.1a	(2) Model 2.2a	(3) Model 2.3a	(4) Model 2.3.1a	(5) Model 2.3.2a	(6) Model 2.3.3a
<i>Dependent var.</i>	FoodRnkPct	FoodRnkPct	FoodRnkPct	FoodRnkPct	FoodRnkPct	FoodRnkPct
<i>Independent vars.</i>	+Price	+Zagat rat.	+WS awards	Same as 2.3a	Same as 2.3a	Same as 2.3a
<i>Data set</i>	Full	Full	Full	Bottom price tertile (<=\$32)	Middle tertile (\$33 to \$45)	Top price tertile (>=\$46)
Manhattan <sup>2</sup>	-0.126*** (0.0170)	-0.0464** (0.0158)	-0.0507** (0.0159)	-0.0278 (0.0240)	-0.0902*** (0.0258)	-0.120*** (0.0284)
NumCuisines <sup>3</sup>	0.00884 (0.0130)	0.0209 (0.0116)	0.0217 (0.0116)	0.0371 (0.0198)	0.0197 (0.0227)	0.00137 (0.0172)
<i>31 cuisine variables omitted</i>						
PriceRankPct <sup>1</sup>	0.461*** (0.0280)	0.0398 (0.0357)	0.0284 (0.0359)	<b>-0.453***</b> (0.124)	0.138 (0.111)	<b>0.471***</b> (0.101)
DecorRankPct <sup>1</sup>		-0.101*** (0.0274)	-0.106*** (0.0274)	-0.145** (0.0512)	0.00118 (0.0455)	-0.0922* (0.0411)
ServiceRankPct <sup>1</sup>		0.670*** (0.0279)	0.663*** (0.0279)	0.521*** (0.0543)	0.634*** (0.0422)	0.889*** (0.0458)
BasicAward <sup>4</sup>			0.0171 (0.0202)	0.00232 (0.0650)	-0.00474 (0.0454)	-0.000265 (0.0189)
HigherAward <sup>4</sup>			0.132*** (0.0251)		0.171 (0.182)	0.0255 (0.0220)
_cons	0.397*** (0.0267)	0.235*** (0.0266)	0.245*** (0.0268)	0.387*** (0.0433)	0.160* (0.0631)	-0.208* (0.0841)
N	1709	1709	1709	603	567	539
F	25.13	52.34	53.95	13.87	.	.
r <sup>2</sup>	0.281	0.455	0.460	0.334	0.425	0.657
df_m	34	36	38	36	34	31
df_r	1674	1672	1670	566	530	504

Standard errors in parentheses; \* =  $p < 0.05$  \*\* =  $p < 0.01$  \*\*\* =  $p < 0.001$ .

(a) No observations of this variable in price tertile.

<sup>1</sup> Price and Zagat food, décor, and service ratings are converted from raw USD prices and Zagat ratings points into “rank percentile” (RankPct) variables using the procedure shown in Table 9.12 and described in section 9.5.

<sup>2</sup> “Manhattan” is a binary variable that takes 1 for restaurants located in Manhattan, and 0 for restaurants located outside of Manhattan.

<sup>3</sup> “NumCuisines” is an ordinal variable taking positive integer values from 1 to 4, representing the number of different cuisine types listed for each restaurant. Of the 1,709 restaurants in the data set, 1,206 (71%) list one cuisine, 441 (26%) list two cuisines, 50 (3%) list three cuisines, and 12 (0.7%) list four cuisines.

<sup>4</sup> Binary variables indicating restaurants winning Wine Spectator awards in 2008. “Basic Award” indicates restaurants winning the first (lowest) level of award ( $n = 91$ ). “Higher Award” indicates restaurants winning either the middle (Best Of) and highest (Grand) levels of awards ( $n = 50$ ).

The only significant positive price-food ratings relationship we observe is in Model 2.1a (column 1 of Table 9.25), which does not control for décor and service ratings. Across the full data set, then, a higher restaurant meal price, holding other attributes including décor and service ratings constant, does not generally predict a higher food rating.

In full Model 2.3a (column 3 of Table 9.25), we observe some statistically significant relationships between cuisine types and food ratings, holding price and other restaurant attributes constant. Five of the 31 cuisine types (Bakery, Dessert, Vietnamese, Sushi, and Pizza) have statistically significant positive coefficients, and thus predict higher food ratings, holding other factors constant. Nine cuisine types (Italian, New American, Small Plates, Turkish, French Bistro, Northern Italian, Coffeehouse, Traditional American, and Coffeeshop/Diner) have statistically significant negative coefficients, predicting lower food ratings. The *Zagat* décor rating coefficient is negative and statistically significant, and the *Zagat* service rating coefficient is positive and statistically significant. Higher (Best Of and Grand) *Wine Spectator* awards predict higher food ratings, whereas Basic awards are not correlated with food ratings in Model 2.3a.

Models 2.3.1a–2.3.3a (columns 4 to 6 of Table 9.25) provide support for our counter-intuitive hypothesis that the relationships between price and food ratings may be different in different Price Terciles, and that higher prices do not predict higher food ratings in the bottom third of restaurants. On the contrary, we observe we observe a statistically significant *negative* effect ( $p < 0.001$ ) of price on food ratings in the bottom Price Tercile. The PriceRankPct coefficient in Tercile 1 (Model 2.3.1a) is -0.453, meaning that within the bottom Price Tercile (column 4 of Table 9.25), the higher a restaurant’s average meal price, holding other factors constant (including décor and service ratings), the *lower* its food rating. In the top Price Tercile (column 6 of Table 9.25), on the other hand, we observe a positive and statistically significant effect ( $p < 0.001$ ) of price on food ratings, with a positive coefficient of 0.471 that is approximately equal in magnitude to, but with the opposite sign as, the equivalent coefficient in the bottom Price Tercile. The price-food rating association thus inverts between the bottom and top Price Terciles. In the middle tercile (column 5 of Table 9.25), we observe no statistically significant effect of price on food ratings.<sup>48</sup>

This surprising inversion between the price-food ratings relationships in Terciles 1 and 3 suggests that splitting the data set into three levels of restaurants by price level may offer additional insights about the relationship between prices and food ratings.

---

<sup>48</sup> Statistically significant negative décor effects and positive service effects, on the other hand, are consistent between the bottom and top Price Terciles. (The middle tercile, meanwhile, shows statistically significant positive effects of service, but no statistically significant effect of décor.)

### Simplified models and results

The last two sets of models we present in this chapter are aimed at providing a more practically useful way of predicting restaurant prices (as in Model 1.3a) and food ratings (as in Models 2.3a–2.3.3a) with many fewer independent variables. We begin with an intuition derived from observing the (uncontrolled) relationships between cuisine frequencies and mean prices by cuisine that are shown in Table 9.16. At first glance, the most popular cuisines appear to have relatively higher prices. We hypothesize that some of the predictive power of the 31 binary cuisine-type variables in Models 1.3a and 2.3a might be captured by a single, much simpler, variable that simply splits the data set approximately in half by the frequency of their cuisine types. At least one of the eight most frequent cuisine types (Italian, New American, French Bistro, Sushi, Traditional American, French, Seafood, and Steakhouse) is served in a total of 838 restaurants, or about half (49%) of all restaurants in the data set. We construct a binary variable called “TopCuisines” that takes a value of 1 if a restaurant offers any one of these eight cuisine types, and 0 if none of these eight cuisine types are offered.

To see how the TopCuisines variable is constructed, please refer back to Tables 9.15 and 9.16 above, where the top eight cuisines are shaded. Table 9.15 shows the selection of the top eight cuisines by frequency and reports summary frequency statistics for the TopCuisines group, and Table 9.16 reports summary raw-price statistics for the TopCuisines group. As is shown in Table 9.16, restaurants serving top cuisines have a mean price of about \$46, compared with about \$40 for the full data set.

Using this variable, we run regressions that are identical to the ones in Models 1.1a–1.3a (for the models with price as a dependent variable) and Models 2.1a–2.3.3a (for the models with food ratings as a dependent variable), except that all 31 binary variables for cuisine type are replaced with the single “TopCuisines” binary variable. We designate these as Models 1.1b–1.3b and 2.1b–2.3.3b, respectively. The results, reported in Tables 9.26 (for Models 1.1b–1.3b) and 9.27 (for 2.1b–2.3.3b), suggest that the TopCuisines variable can capture some, but not all, of predictive power of individual cuisine variables.

R-squared for Model 1.3b (Table 9.26, column 3)—which predicts PriceRankPct from the number of cuisines, top cuisines, *Zagat* ratings, and *Wine Spectator* awards—is 0.69, versus 0.74 for full model 1.3a. Model 1.3b is simpler, and its F-statistic is substantially higher than the F-statistic for Model 1.3a. Model 1.3b may also be more practically useful in terms of being adapted for real-world predictions, and likely is more robust to changes in the concentrations of different cuisine types that may occur naturally over time as a result of changing consumer trends, social fashions, restaurant-owner demographics, and so on. Similarly, Model 2.3b (Table 9.27, column 3), which replaces the 31 cuisine-type variables in Model 2.3a with the single TopCuisines variable,



sacrifices some predictive power but has a substantially higher F-statistic than its full-model equivalent.

We also view the simplified “b” variants of Models 1 and 2 as a type of robustness check against the possible overspecification of the full “a” variants of these models. Tables 9.26–9.27 show that all of the statistically significant main effects we report for Models 1.3a and 2.3a (other than individual-cuisine effects, obviously) are also statistically significant, with the same signs, in Models 1.3b and 2.3b. These include, in Model 1.3b, the positive effect of Manhattan location, the negative effect on price of number of cuisines, and the positive effects of décor ratings, service ratings, and *Wine Spectator* awards on price; and in Model 2.3b, the negative effect of Manhattan location, the negative effect of décor rating, the positive effect of service rating, and the positive effect of higher *Wine Spectator* awards (but not Basic awards) on food ratings. As expected, in TopCuisines, the one new variable in these models, has a positive and statistically significant effect ( $p < 0.001$ ) on price. TopCuisines also has a negative and statistically significant effect ( $p < 0.001$ ) on food rating.

As for the Price Tercile comparisons, the statistically significant main effects we report for Models 2.3.1a–2.3.3a (Table 9.25, columns 4 to 6) are also statistically significant, with the same signs, in Models 2.3.1b–2.3.3b (Table 9.27, columns 4 to 6), including the negative effect of décor and positive effect of service on food rating, and the inversion of the effect of price on food rating between the bottom tercile (where, in Model 2.3.1b, the effect is negative and statistically significant, with a coefficient of  $-0.386$ ,  $p < 0.01$ ) versus the top tercile (where, in Model 2.3.3b, the effect is positive and statistically significant, with a coefficient of  $0.483$ ).

## Discussion

The first aim of this chapter was to explore the effect of *Wine Spectator* awards on restaurant prices. In both the full and simplified versions of our models, we find evidence that restaurants with awards charge more for meals, all other things being equal, suggesting that restaurants receive some return on their purchase of an award. We also find evidence in both the full and simplified for a few unexpected effects in *Zagat*-rated restaurants in New York City—most notably, that (holding location and cuisine attributes constant): (1) higher *Zagat* service and décor ratings predict higher meal prices, but higher food ratings do not; (2) higher service ratings predict higher food ratings, but higher décor ratings predict lower food ratings; and (3) in the lowest third of restaurants by price (below \$33 per meal per person), higher prices predict lower food ratings, whereas in the highest

third of restaurants by price (above \$45 per meal per person), higher prices predict higher food ratings.

With respect to our main effects of *Wine Spectator* awards on meal price, it is worth noting that the “Award” signal may not be generated purely by the *Wine Spectator* award itself, but could occur in combination with other types of non-rationed restaurant awards and honorific designations that are also purchased by the same restaurants (or non-restaurants).<sup>49</sup> Anecdotally, the type of establishment that displays awards, certificates, and reviews on its walls (typically in the front of the house, near the host stand) often tends to display more than one such award. Each framed certificate on such an award wall may have its own partial cost and its own partial return on investment.

In sum, we find that *Wine Spectator* Awards of Excellence deliver a measurable return on a restaurant’s revenue per customer. Thus, for some restaurants, the choice to purchase Excellence from *Wine Spectator* may well be a prudent business decision.

---

<sup>49</sup> Anecdotally, there are many other types of non-rationed awards and other forms of Excellence available to U.S. businesses on the open market. For instance, Enron, shortly before its collapse, paid millions of dollars to the Arthur Andersen accounting firm in exchange for being certified as financially sound, and the Better Business Bureau (BBB) gave an A+ rating to an imaginary business whose owner was listed as Adolf Hitler.

## Chapter 10: Opportunities in immortality

*On strategic planning in the Omega-3 fish oil industry*

Fresh from the trenches of their headquarters in Salt Lake City and Denver and Oslo they came, scarred by their latest battles of wits with the world's leading muckrakers. From battleship-sized ice trawlers harvesting krill off the shoulders of Antarctica, they stepped onto dry land and boarded planes bound for Tenerife, just west of Africa, where volcanic mountaintops peek out into the pink and blue skies of the southeast Atlantic.

There may be no better place for sunsets in February than the Canary Islands and no place better to watch them from than the Ritz-Carlton Abama, the site of the 2016 annual meeting of the [Global Organization for EPA and DHA Omega-3s](#) (GOED), a 501(c)(6) not-for-profit trade association based in Salt Lake City.

Together, GOED's 200-plus members represent the vast majority of the world's major producers of Omega-3 fish oil supplements. The meeting was officially called the "GOED Exchange." It was the time for the competition to join forces, for the bloc to strategize as one.

In the high-flying year of 2016, there was a lot to discuss: the Omega-3 industry was under attack in recent media reports, but it had also just been named (by the *Nutrition Business Journal*) as the most gangbusters niche in the young history of the booming supplement industry, with nearly a billion and a half dollars in annual revenue.

Streaming into the sprawling orange stucco Ritz-Carlton Abama estate as I arrived were mostly executives from European companies: DSM, which controls 25% of the global Omega-3 market; Aker BioMarine, a Norwegian firm with the world's biggest Antarctic krill-harvesting operation; and others such as Golden Omega, Bioprocess Algae, Nycomed, Herbalife, plus some of the top brass from [Nutraceuticals World](#) (*NW*) and the [Nutrition Business Journal](#) (*NBJ*), the industry's two main advertorials of record.

I was on assignment for *Men's Health* magazine. My task was to update readers on the pros and cons of Omega-3 supplements. Although *Men's Health* ultimately paid me for the job—and well by industry standards (\$2 per word)—it might have been the worst four thousand dollars spent all year by Rodale, the publisher of *Men's Health*. The final revision of my piece bore so little resemblance to the long-form nutraceutical-industry hatchet job I originally submitted that it was almost unrecognizable as my own work. Gone were all my framing quotations from the Biblical Apocrypha, medical philosophers, and heterodox physicians; my stories of animal torture by dead scientists; and my unsubstantiated claims that a news media conglomerate called Advance Publications had passed the Turing Test and replaced most of its human writers with sockpuppets that had possibly turned scores of newsrooms across the American Midwest into ghost offices.



Figure 10.1. Ritz-Carlton Abama, Tenerife

I'm almost certain that it took my *Men's Health* editor more time to revise my piece than it would have taken him to write it from scratch. He even took it upon himself to conduct additional interviews (we will return to this point later). My editor was a consummate magazine professional and a gentleman, and he was very apologetic about the whole thing. He gave me the option to publish the piece under a pen name if I was no longer comfortable with the content. I thought the pen-name idea sounded a bit grandiose, figured that nobody probably looked much at *Men's Health* bylines in any case, and told him to go ahead and use my name. We parted ways amicably.

Full disclosure: *Men's Health* didn't actually send me to Tenerife. I went there by myself, on my own dime, partly out of bloodthirsty curiosity and partly because the hotel

rate was phenomenal: 150 euros per night for a room whose normal price was about 500 euros in low season. It was the best conference group discount I had ever seen. How did GOED possibly pull this off? The prospect that the organization might be run by unbelievable ballers drew me in. Once I'd been granted media credentials for the event and booked one of the Ritz's incredible rooms at the conference rate, I blew the rest of the *Men's Health* fee on a round-trip ticket to Tenerife via Amsterdam.

From the opening bell, the GOED meeting was a class act. The Abama had its own little complex of private beaches, hot pockets of volcanic rock protecting crystalline coves from the warm desert winds that blew in from Western Sahara. The bountiful lunch buffet, with its fresh seafood, was possibly the best hotel food I had ever eaten. All of it was included in my media rate of zero.



*Figure 10.2. Swimming pool, Ritz-Carlton Abama Resort, Tenerife.*

On the first night, after closing my veranda doors and twisting open their slats to expose the streams of moonlight but before climbing onto my expansive king mattress, I looked in the mirror and asked myself if I was the kind of man who would take this kind of gift from this organization under false pretenses, and then use it as his sustenance while attempting to execute a secret plot to subvert the organization.

After some deliberation, I decided that I was that kind of man. But with the deliberation came a compromise: I decided not to use the Thai snake oil that I had ordered on Taobao in China and brought to the Canary Islands for the purpose of performing some kind of outrageous, still-to-be-scripted Borat-type stunt to shame the Omega-3 people, in the middle of their conference, on video. Snake oil, it turned out, was rich in Omega-3s, and I'd smelled an opportunity for a guerrilla YouTube short where the industry experts make asses of themselves by taking snake oil seriously—which, in fact, on their own terms, they should.

I had brought a DSLR and tripod to capture the snake-oil brouhaha. But in the end it wasn't a very hard decision to ditch the whole idea: the concept of the stunt was still vague, and without a wingman I probably didn't have the balls to do it anyway, especially now that I knew a few of the other conference participants by name and could see that they were mostly just slightly alcoholic normies, and that most of them appeared to be physically stronger than me.

For my own protection, in case any of my adversaries ended up passing through my room, I zipped the two bottles of snake-oil capsules into a compartment in the darkest interior reaches of my backpack, where not even a burglar could find them.

The organized program started the next morning in a big auditorium whose audio-visuals were best-in-class. The plenary sessions featured inspirational storytellers and well-spoken scientists highlighting the most compelling available evidence for many different categories of health benefits of fish oil. Among the possible indications for Omega-3 fatty acid supplements mentioned in the presentations were high triglycerides, coronary heart disease, cancer, stroke, obesity, diabetes, metabolic syndrome, pre-menstrual cramps, hypertension, pregnancy complications, sexual dysfunction, and childhood autism.

One of the first things I learned in Tenerife was that the main competition for the non-prescription Omega-3 supplements business, and for the entire nutraceutical industry (this term had recently replaced the stigmatized “natural medicine” in the lingo), was the prescription pharmaceutical industry. It was explained in one particularly insightful plenary presentation that obtaining FDA approval for a prescription medicine costs hundreds of millions of dollars, whereas FDA approval to label an over-the-counter medicine as having similar types of health benefits could be achieved with an orders-of-magnitude-lower amount of capital.

Although prescription pills would always command significant price multiples over their nutraceutical cousins during the period of patent protection even if the marginal costs of manufacturing the two were comparable, the nutraceutical segment was lower-hanging-and-still-plump fruit for smaller investors that seek to capture some of America's health-care boom.





Figure 10.3. Thai snake oil capsules.

Marketing an over-the-counter health remedy presents a different kind of business challenge than getting doctors to write prescriptions. Direct-to-consumer sales and media messaging are central to the strategy. To this end, another plenary presenter profiled the typical Omega-3 consumer.

The presenter was a British, and therefore charming and articulate-sounding, market research expert who consulted to the Omega-3 industry. She was an upper-middle-class white woman in her 40s or 50s. She described the typical Omega-3 consumer as an upper-middle-class woman in her 40s or 50s who had developed a deep fear of her own mortality. She did not specifically describe the Omega-3 consumer as white.

Although I first associated the Brit with the Grim Reaper, her storytelling had a warm, unpretentious feel that was not entirely neutralized by its unsettling message that the fear of death was a business opportunity to be mined.



*Figure 10.4. The Grim Reaper.*

At the magnificent, fresh-fish-themed outdoor gala dinner on an old sugar plantation, a friend I'd made—a bright, up-and-coming young fish-oil-supplement producer from the Midwest in his early thirties who knew all about John Harvey Kellogg and had even better conspiracy theories than I did—explained to me that Senator Orrin Hatch had won the right for the supplements to include FDA-approved health claims with the [Dietary Supplement Health and Education Act of 1994](#).

From my friend I also learned that much of Hatch's Utah constituency had made its money by building multi-level marketing organizations (MLMs), which are basically Avon-style companies with direct-to-consumer sales reps who earn commissions by cold-



calling and selling products to their friends and neighbors.

Utah's friendly Mormons, many of whose cold-calling skills have been galvanized on their missions, turned out to be particularly good at this type of sales. Hatch's bill was a victory for the Mormon-run MLMs against pharmaceutical giants whose rule had previously been protected by an FDA ban on supplement label claims. He has been the hero of the industry for decades, with Herbalife one of his biggest corporate backers.



*Figure 10.5. Sen. Orrin Hatch (R-UT) looks on in consternation as then-Sen. Joe Biden (D-DE) discusses an FDA crackdown on the supplement nutraceutical Andro, whose presence on the market threatened prescription-drug competitors.*

In 2016, Hatch introduced new [legislation](#) to make supplements tax-deductible and eligible for inclusion in Health Savings Account (HSA) and Flexible Spending Account (FSA) plans. The *New York Times* [reports](#) that Hatch “credits a daily regimen of nutritional supplements for his vigor at 77,” and that his son Scott is a lobbyist for the supplement industry. Hatch's vigor is no joke, and he continues to fight hard to bring more bacon

home to his compadres in Utah.

The first recorded use of fish innards as medicine came from the Biblical Apocrypha in the fourth century A.D., when Tobias, protagonist of the Book of Tobit (11:13–15), “taking of the gall of the fish, anointed his father’s eyes...and recovered his sight.” The good son was rewarded with cattle, camels, money, and one or more giant feasts.

For most of the first millennium and a half A.D., fish-oil treatments were only [in vogue](#) amongst the Icelandic Vikings—until the late 1700s, when Nordic immigrants brought fish oil to the U.K. and marketed it as a treatment for rheumatism and tuberculosis.<sup>50</sup> In the 1800s, it caught on as a cure-all remedy across the British Empire, and by the early 20th century, it was a fixture in the American medicine cabinet. My grandmother fed my mom a spoonful every day.

The business of fish oil was suddenly threatened, however, by the prescription-medicine boom of the 1960s, when pharmaceuticals began stealing share from traditional remedies. But in 1971 came a stroke of good fortune for fish oil: a well-publicized [Danish study](#) in the *Lancet* observing the fine cardiopulmonary health of a cohort of Eskimos in Greenland who ate almost exclusively whale blubber and oily fish for breakfast, lunch, and dinner. The Danish researchers attributed the Eskimos’ low levels of heart disease and diabetes to two polyunsaturated omega-3 ( $\eta$ -3) fatty acids, known as EPA and DHA, that were abundant in the fish they hunted.

Fish oil was really taken to the big leagues in 1994, when [Dr. Alexander Leaf](#), former chief of medicine at the Massachusetts General Hospital, published [a paper](#) in the prestigious *Proceedings of the National Academy of Sciences* finding that “acute intravenous administration of an emulsion of either fish oil or purified omega-3 fatty acids could prevent ischemically-induced ventricular fibrillation.”<sup>51</sup>

Translated into English and clarified, the *PNAS* finding was that shooting up Omega-3s might help prevent heart problems in dogs whose coronary arteries had been sliced open by Dr. Leaf and his colleagues and subsequently sutured with hydraulic cuffs to stop the bleeding. The research team observed that their maimed dogs had the tendency to live the rest of their short, excruciating lives with slightly less irregular heartbeats than if they had been injected with fish oil before their arteries were sliced. For understandable reasons, media reports of Dr. Leaf’s result often left out details about the study’s gruesome limitations, e.g. that the scope of the finding was limited to maimed dogs. The observation in the health media was that taking Omega-3 fatty acid supplements would save human beings from heart disease and extend human life.

---

<sup>50</sup> Mark Kurlansky, *Cod: A Biography of the Fish that Changed the World* (Penguin, 2010).

<sup>51</sup> Billman, G.E., Hallaq, H., and Leaf, A. (1994). Prevention of ischemia-induced ventricular fibrillation by omega 3 fatty acids. *Proc Natl Acad Sci* 91(10):4427–4430.



*Figure 10.6. [The late Dr. Alexander Leaf](#), a pioneer in the study of the health benefits of injecting dogs with fish oil after slicing open and re-cuffing their arteries.*

One of the most vocal cheerleaders for Dr. Leaf's findings was the American Heart Association, one of his major supporters. In 1994, the same year as his landmark paper was published, the AHA [organized a conference](#) in Houston to “focus on the mechanisms by which Omega-3 fatty acids elicit their therapeutic effects.” The conference proceedings were later turned into an [official AHA advisory](#) encouraging Omega-3 consumption, thus elevating the theory straight into the medical mainstream.<sup>52</sup>

[Slicing dogs' arteries](#) and then [shooting the dogs up with fish oil](#) would turn into a productive second career for Dr. Leaf after his mandatory retirement at the age of 70. In his magnum opus—published when he was 83—Dr. Leaf turned his attention from dogs to rats. Having discovered that “one can quickly remove the hearts from several one-to-two-day-old rat pups and separate the individual myocytes enzymatically,” Dr. Leaf and his team set the rat hearts artificially beating on their own, put them into an EPA or DHA bath under a microscope, and recorded the dismembered hearts beating with a video camera.

The team of scholars observed that the cells in the dismembered hearts that were mixed with fish oil started to beat a bit more slowly than the cells in the dismembered hearts that had not been mixed with fish oil. The conclusion was that human beings could be saved from cardiac arrhythmia, and thus from sudden heart attacks, by consuming fish-oil supplements every day. The AHA [published](#) the result in its flagship scientific journal, *Circulation*.<sup>53</sup>

By the turn of the 21st century, the buzz about fish oil was about to turn into a feeding frenzy. Many entrepreneurial nutritional-supplement peddlers, some of them backed by big private equity, had laid the groundwork for getting in on the action. In a milestone for the industry, four years of lobbying by a core group of nutraceutical executives finally led to a 2004 ruling by the F.D.A. that Omega-3 supplements could be legally labeled as reducing the risk of coronary heart disease. Capsule factories were springing up everywhere. New multi-national bio-marine corporations were being formed, some of which trawled Antarctica to harvest krill, a newly discovered source of Omega-3s.

In 2013, with annual sales of Omega-3 fish oil supplements up to \$1.3 billion,<sup>54</sup> the industry was stung by a *New England Journal of Medicine* [paper](#) reporting that a randomized, controlled Omega-3 trial studying 12,536 subjects for six years had found “no significant

---

<sup>52</sup> Stone NJ (1996). Fish consumption, fish oil, lipids, and coronary heart disease. *Circulation* 94: 2337–2340. Revised as Kris-Etherton PM, Harris WS, and Appel LJ (2003). Omega-3 Fatty Acids and Cardiovascular Disease: New Recommendations From the American Heart Association. *Arteriosclerosis, Thrombosis, and Vascular Biology* 23: 151-152

<sup>53</sup> Leaf A, Kang JX, Xiao YF, and Billman GE. Clinical Prevention of Sudden Cardiac Death by n-3 Polyunsaturated Fatty Acids and Mechanism of Prevention of Arrhythmias by n-3 Fish Oils. *Circulation* 107: 2646-2652

<sup>54</sup> Sales of Omega-3 supplements were \$1.2 billion in 2013, according to the *Nutrition Business Journal*.

benefit of  $\eta$ -3 fatty acids in reducing the risk of death from cardiovascular causes or hospital admission for cardiovascular causes.”<sup>55</sup> The next year, an even bigger blow came when the Greenland research project that had jump-started the whole industry was [essentially discredited](#) by George Fodor and his colleagues at the University of Ottawa Heart Institute, who produced a [meta-study](#) in the *Canadian Journal of Cardiology* rejecting the initial premise that the Greenland Eskimos had less heart disease than other populations.

If there is one core talent common to the leadership of GOED, it is experience creating and distributing nutritional propaganda across media platforms with extraordinarily broad reach and minimizing the damage caused by scientific skeptics. Many of the organization’s executives and members have ties to *Nutraceutical World* or its main competitor, *The Nutrition Business Journal (NBj)*, which is a division of a Boulder company called New Hope Natural Media (NHNM), “[the premiere \[sic\] digital marketplace that connects the healthy lifestyle industry from supply to shelf.](#)” NHNM runs the enormous Natural Products Expo West, the increasingly enormous Natural Products Expo East, and a new trade-show venture in China. The long-term impact of the 2010–2014 attacks on GOED interests appears to have been minimal.

GOED also publishes and sells *Delicious Living*, *Functional Ingredients*, *Natural Foods Merchandiser*, and white-paper reports such as the disarmingly candid \$1,495-per-copy *NBj Special Diets Report*, which [presents](#) the industry’s inside view of its customers in an “Analysis of the Five Dominant Food Tribes: Gluten-free, Paleo, Vegan/Vegetarian, Raw and Biohacking.” To its credit, the *NBj* website and other marketing materials are totally transparent about its positioning as a marketing tool for nutraceutical companies rather than an objective news source. The *Special Diets Report* abstract begins:

“We have become a nation of food tribes...a whopping 44% of adults now say food restrictions, food allergies, or avoidance of certain ingredients dictate what they eat...The special-diets category will reach \$144 billion by 2018, a scale that has massive ramifications from farm to shelf...The gluten-free tribe represents a roughly \$22 billion market segment, Paleo has ridden CrossFit into the mainstream. ... Entire brands live and die by...terms like ‘non-GMO,’ ‘plant-based,’ and ‘local.’”

At the time of the 2016 GOED meeting, the most popular Omega-3 supplement on Amazon.com was Dr. Tobias’ Optimum Omega 3 Fish Oil capsules, priced at \$29.97 per bottle of 180 softgels.

---

<sup>55</sup> Upon further inspection, the former study looks solid, the latter [not so much](#): it’s an association study rather than a randomized experiment, and therefore much less reliable (some would say worthless) as a source of causal inference. Secondly, a huge majority of men above a certain age have prostate cancer, so their deaths get over-reported as prostate cancer deaths. When a cancer patient dies, he almost always dies of cancer.

In line with new FDA rules, Dr. Tobias' supplements advertised that they would "support heart, brain, joints, immune system" and "promote weight loss" while having "NO Fishy Aftertaste."



Figure 10.7. Dr. Tobias' Optimum Omega 3 Fish Oil capsules.

Customers (or sockpuppets, at least) loved the product: at the time of the 2016 GOED conference, it was rated 4.6 stars in more than 6,000 reviews on Amazon.com. By comparison, Michael Pollan's *The Omnivore's Dilemma* had about 1,400 reviews, and the bestselling King James version of the *Apocrypha*, which contains the Book of Tobias, had about 400 reviews.



As far as I could tell, Dr. Tobias was not represented on Tenerife, which, at least in my mind, only pumped up his mystique. Who was this doctor, and why was his product so well liked?

The brand turned out not to be named after the Apocryphal fish healer but rather after its founder and C.E.O., Tobias Ihde, 44, who [explained](#) on the company's website that his background is in "sports science" and that "Dr. Tobias" is "the nickname that my American friends had given me." Although Ihde did have a doctoral degree from a good school, it wasn't a medical or sports science degree but rather a [Ph.D. in economics](#) from the University of Kiel.

During and briefly after graduate school, Dr. Tobias consulted for DaimlerChrysler and the [Droege Group](#), a "hybrid family-equity-business model" consultancy that "restructures and invests aiming at value enhancement for its own portfolio and for external clients." Dr. Tobias then moved to [TriSolutions](#) in Hamburg, where he specialized in risk and liquidity management for a decade.

[Finally, in 2013](#)—almost too late, but not quite—Dr. Tobias launched his supplements outfit, [DTI UG](#), as well as his own risk-management-consulting shop, Dr. Ihde und Partner, whose current [website](#) advertises recent projects such as "US secondary market life insurance: How recoverable?" (Apparently, the problem is that life-insurance derivatives are underperforming due to not enough insurance customers dying.) But supplements saved Dr. Tobias from a life of corporate drudgery. After a few well-timed months and a good showing in [LabDoor's quality rankings](#) of Omega-3 supplements, Dr. Tobias was killing it, and his day job didn't matter anymore.

Doctor or not, his success was well earned. His intentions seemed genuine. He worked his way up in the corporate world, and somewhere along the way, his curiosity led him elsewhere: he traveled the world, met and married a [Canadian woman](#), and started teaching [Anusara Hatha yoga](#) with her in the suburbs of Toronto and Hamburg. He discovered his inner hippie. One day, he put up a shingle in the rapidly growing Omega-3 industry and hit the jackpot. Dr. Tobias had followed his heart in a number of different directions, and one of them paid dividends.

But what, exactly, is the biomedical expertise of a risk-management-economist-turned-life-insurance-actuary-turned-yoga-instructor who started moonlighting as a drug-capsule entrepreneur a couple of years ago? Should we trust this man as an expert on the biochemistry of the human circulatory system, trust him with the delicate microbial balance of our inner bodies?

Not every fish-oil doctor is an amateur. There were also many brilliantly qualified scientists on the sunset cruise to Tenerife. Through their innovations, fish oil efficiency continues to march ahead. Every year, new technology allows us to more easily and efficiently deconstruct fish into smaller and smaller fish components—livers, spoonfuls of

liver oil, capsules—as we do to so many of the other foods that we once ate whole.

Could it be that it's fish, not Omega-3 fatty acids, that our bodies really want? What if the nutrition is actually packed into the flesh and bones and cartilage that we yank out prior to encapsulation? Could there be an evolved sensory reason that good fish tastes delicious whereas krill-oil capsules taste tolerable at best? Or is challenging pills at large is a fool's errand, as sacred to our culture as the sun was to the ancients?

My last hope of finding a duplicitous villain to shame in Tenerife came in my meeting with Adam Ismail, GOED's executive director. He was the closest thing the conference had to a heartthrob: a gentle, intelligent man in his late thirties whose dark brown hair, thick eyebrows, and slim figure made him look a decade younger. The day after the banquet, as the conference was winding down, I met Ismail for a beer at a terrace table overlooking the volcanic coastline and a blazing fuchsia-and-turquoise sky.



*Figure 10.8. Lobby bar, Ritz-Carlton Abama Resort, Tenerife.*

By the second sip from my beautifully tapered glass of fresh Estrella beer, my plan had unravelled further. From the start, every ludicrous claim made by my adversary seemed to be both true and compelling. In the frenzied world of un-vetted health journalism, my eloquent subject pointed out, there are always plenty of voices to re-broadcast and plenty



of others to fight against, and Ismail expressed rightful annoyance at the recent media reports that Omega-3s were actually *harmful* to human health. He made me wonder: what kind of insane vigilante would actually devote his professional life to *shaming* Omega-3 supplements? What were his motivations? What were mine?

When necessary, Ismail willfully accepted the role of the industry's whipping boy and stood up for its interests against the attack from Big Pharma, whose main message was that over-the-counter nutraceuticals were unsafe. It's understandable why Big Pharma would want to devote significant resources to discrediting the OTC supplement industry: Big Pharma's central consumer-facing marketing challenge is to convince consumers, and by proxy insurance companies, to pay enormous premiums for prescription drugs. If there's a competing non-prescription version of Vascepa selling on Amazon.com for one-tenth the price, then Vascepa's job is to convince the public that the cheap substitute is ineffective or dangerous. What can Ismail do but defend his product against the muckrakers with pockets deeper than the seas?

Vascepa, one of the three main competitors in the prescription Omega-3 space, is, like the non-prescription version, a bottle of gelcaps, but it is differentiated from the OTC product in certain ways.



Figure 10.9. Vascepa branding, packaging, and pill design.

The most concrete difference between prescription and non-prescription fish oil is that the prescription version contains only EPA and not DHA. Non-prescription supplement producers suggest that DHA also has beneficial properties, and thus argue that their product may offer more therapeutic benefits than the prescription version does. Prescription Omega-3 capsule advocates, on the other hand, argue that their product is purer and more consistent in quality, and that the therapeutic benefits of DHA are negligible. In this case, the most reliable route to turning a successful over-the-counter drug into a successful prescription drug may be to discredit the safety of its OTC version, and scaring existing OTC users into spending ten times more, ideally via a cost-opaque insurance agency barrier.

Prescription Vascepa capsules typically contain 1000 mg EPA and are meant to be taken four times per day, whereas the non-prescription supplements typically contain 800 mg EPA. Taking five of Dr. Tobias' capsules per day would thus give you the same EPA dose as four capsules of Vascepa. A 30-day supply (120 g) of Dr. Tobias' capsules would cost you about \$25 on Amazon.com. A 30-day supply (120 g) of Vascepa capsules would cost you or your insurance company about \$275 at a pharmacy.

Vascepa is the only product currently made by the Irish pharmaceutical company Amarin (NASDAQ:AMRN), which in 2012 received F.D.A. approval for the drug as a treatment for a medical condition known as "hypertriglyceridemia" (high blood triglyceride levels). Hypertriglyceridemia has no known symptoms, and the only basis for its classification as a disease is its association with other cardiovascular conditions and measurements. In 2011, after clinical trials showed that the drug reduced triglyceride levels and was poised for F.D.A. approval, AMRN stock traded as high as \$19.50 per share, giving the company an erstwhile market capitalization of more than \$5 billion. That heroic valuation may be justifiable if enough consumers and/or insurance companies are in fact willing to pay \$275 per patient per month for a month's supply of the same active ingredient that can be extracted and sold profitably in the non-prescription market for \$25.

But Amarin ran into a few bumps in the road to riches. In 2012, the F.D.A. approved Vascepa as a treatment for patients with blood triglyceride levels between 200 mg/dL and 500 mg/dL, but to the company's dismay, the agency rejected its application for approval as a supplemental treatment for general cardiovascular disease. This had been the real jackpot Amarin's shareholders were hoping for, as there's a much bigger market for heart disease than there is for hypertriglyceridemia (the medical condition of having high triglycerides).

The F.D.A.'s rejection of Vascepa's cardiovascular indication was made on the basis that there was no scientific evidence that lowering blood triglycerides, or any of the other effects observed against placebo controls in Amarin's trials, would reduce the risk or

symptoms of cardiovascular disease. A high triglyceride measurement may be a marker of other health problems, but simply bringing down the level of a measurement may not affect the underlying condition it marks. A scratchy voice may be sign that you've got a sore throat, but strapping a voice box over your mouth so that your voice sounds normal again won't likely cure your condition.

This problem, which is known as “treating the numbers,” is thought by many observers to be rampant in the pharmaceutical industry, where a huge jackpot can be won by carving out any bit of F.D.A.-approved space, especially if you have your own disease to go with it. In a crowded marketplace with fierce competition, it can be easier for a relatively small pharmaceutical company to come to market with a treatment for an obscure or novel condition than it would be to compete for share with the pharmaceutical giants to treat a common diagnosis.

Hypertriglyceridemia, in spite of being asymptomatic, has been discussed in the medical literature since the 1960s. This may be why the F.D.A. was willing to accept it as a legitimate medical condition for which Vascepa could be indicated. But this was a small victory for Amarin, a poor substitute for the cardiovascular prize the company wanted. The F.D.A.'s refusal to approve EPA as a general cardiovascular treatment severely limited Vascepa's upside, sending AMRN shares plunging to less than two dollars and leading the company to lay off half its sales staff.

Amarin's 2016 Vascepa sales were \$125 million, which would be a dream come true for most Omega-3 producers but is just a blip on the radar screen of big Western pharma. The company has still not given up running trials, lobbying, and even suing for the right to use broader indications for the prescription of the product, even after two of its subsequent F.D.A. appeals were rejected. At press time, AMRN shares were back up to \$3.13, giving the company and its one-and-only EPA pill a valuation of \$845 million—still more than half the size of the entire non-prescription Omega-3 supplement industry.

Our monopolistic and uniquely American systems of patent protections and medical licenses, which in turn supports the government's practice of paying (or licensing health insurance companies that charge to consumers) health-insurance premiums that support the outrageous prices of prescription drugs like Vascepa, for which virtually identical over-the-counter alternatives are available at one-tenth the cost, is what keeps doctor's visits, hospital visits, blood tests, X-rays, medical devices, prescriptions, prescription medications, and pretty much any other medically certified product you can imagine at prices that are orders of magnitude above prices for the same products in any other country on Earth.

The vertical cartel of insurance companies, health-care providers, pharmaceutical companies, and academic licensing institutions has been so successful in cooperatively exploiting the scientific method and F.D.A. approval process, and in maintaining a

complex network of licensing barriers that gate the medical industry, that it has collectively managed to create and maintain a scheme under which the American consumer pays \$275 for a bottle of pills whose marginal cost of production is the same as a bottle of Tic-Tacs, and another \$275 for the five minutes of a doctor's time that it takes to prescribe them.

One of the most surreal things about the age of pharma is the way it has shifted the career choices of massive numbers of talented young people, especially on the East Coast, in recent years. Right now there's more money to be made running pills than doing almost anything else. There is a lost class of creators, makers, and doers in America who would rather be designing fuel-efficient jetpacks, self-chilling beer bottles, shoes that don't wear out, or bread that bakes itself. Instead they are being deployed to spend their days filling out mad libs on Word and completing the paperwork necessary convince the FDA to approve a placebo called Omega-3 ethyl ester to treat a disease called hypertriglyceridemia, which (a) is asymptomatic; (b) can be diagnosed only by comparing a numerical measurement to an arbitrary numerical threshold; and (c) has a highly suspect ontological status in human pathology.

People choose this path because people respond to economic incentives, even when they are perverse. A career in almost any segment the of pill-paperwork business is more likely to be lucrative than a job as an industrial designer, a professor, a small business owner, a musician, or a governor.

Pill paperwork is how you become wealthy in America today. The sudden diversion of talent to the medical bureaucracy is a labor market failure for the ages. It is the legacy of our time. It is what we have done with our city on a hill. How will our work be judged, a thousand years from now?



Figure 10.10. Adam Ismail on PBS' Frontline.

When Ismail and I first started talking, he mentioned that it had been a tough week. Just before coming to Tenerife, he'd appeared as the nutraceutical industry whipping boy in a PBS *Frontline* hatchet job called "Supplements and Safety," which was produced in collaboration with CBC and the *New York Times*. Ismail's nemesis in the *Frontline* piece, the investigative protagonist of the exposé of the Omega-3 industry, was one "R. Preston Mason, Ph.D., Harvard Medical School." The sequence begins with Mason handing two beakers of oil to Gillian Findlay, the *Frontline* host, so that she can smell the difference between the prescription and OTC fish-oil pills:

MASON: This is an FDA-approved product—

FINDLAY [voice-over]: Dr. Preston Mason is a Harvard University researcher. Here he's comparing prescription-quality fish oil to the oil found in over-the-counter supplements.

MASON [handing the F.D.A.-approved prescription oil to Findlay]: And give it a smell.

FINDLAY [smiling]: Smells a little bit fishy but not—not bad.

MASON: Right. Smells—you're going to have always some smell.

FINDLAY [voice-over]: One of the issues with fish oil is it's delicate. It's extracted as a byproduct from oily fish like anchovies. As the fish get crushed, the oil is exposed to oxygen. And it doesn't take much oxygen to turn the oil rancid.

MASON [handing the OTC oil to Findlay]: This is a common supplement for fish oil [*sic*]. See what that smells like.

FINDLAY [smelling then wincing, totally grossed out]: Oh!

MASON: [breaking out into a wide smile]: What?

FINDLAY: That doesn't smell good. That's—that smells like it's going bad.

MASON [beaming, cackling]: Yeah. Right. Yeah. It's a very strong, fishy smell.

FINDLAY [voice-over]: If it was simply an odor issue, that would be one thing. But oxidized oil contains oxidized lipids, one of the building blocks of cells. We've long known that lipids, when oxidized, can be harmful.

MASON [serious again]: So oxidized lipid triggers inflammatory responses within our body, particularly in our cells. And if we ingest oxidized lipid, we can trigger these inflammatory changes that can lead to things like cardiovascular disease.



*Figure 10.11. Frontline sequence with Mason and Findlay.*



Figure 10.12. Findlay's conclusion.

In researching Mason, the first thing I was surprised to find out was that his name is not listed anywhere either in Harvard Medical School's [staff and faculty directory](#). When you search for Mason plus Harvard, the first thing that comes up is the *Frontline* appearance of "Preston Mason, Ph.D., Harvard Medical School." Beyond that are some author affiliations in several published medical articles, alternatively as "professor," "faculty," and "lecturer" at Harvard Medical School.

Mason, who has the manicured-WASP look of a daytime soap-opera star, graduated in 1985 from Gordon College, a small Christian institution of 1,700 undergraduates with an acceptance rate of 93%. He was subsequently accepted to the University of Connecticut M.D.-Ph.D. program and dropped out of the M.D. program, but did complete a Ph.D. in cell biology and biophysics from the University of Connecticut in 1989.

The current [Cardio-Metabolic Health Congress page](#) on Mason, as well as a 2004 [journal article](#) published by Mason, also lists him as having an M.B.A., but that degree is not mentioned in any of his other bios. In 2013, he did almost surely receive some kind of honorary degree from Gordon, his alma mater.

As of April 2018, Mason's LinkedIn page, which does not mention the M.B.A., claimed that after completing the Ph.D., he was an assistant professor at the University of Connecticut, and was then an associate professor at Drexel University for the decade following that. It also stated that he had joined Harvard Medical School in October 2002 and had been working there for the past 15 years and 7 months. His position was listed as



Lecturer on Medicine.

Mason's LinkedIn page also stated that he had joined the faculty of the Brigham & Women's Hospital the month before joining Harvard, in September 2002, and had also been working there ever since: 15 years and 8 months. For those same 15 years, plus one—starting in 2001—he has also been President and Founder of Elucida Research LLC. Of Mason's several supposed longtime employers, PBS selected the most prestigious for his TV byline: Harvard.



Figure 10.13. R. Preston Mason on PBS' Frontline.

<https://www.pbs.org/wgbh/frontline/video/how-beneficial-is-fish-oil/>

Calls to the Harvard Medical School and Harvard University registrar were unsuccessful in verifying whether R. Preston Mason is currently employed for pay by Harvard University or ever was employed for pay by Harvard University. To be clear, I am not alleging that Mason faked his Harvard credential: I am merely stating that the credential and what it means, like so many of the credentials out there, are unverifiable by a member of the general public.

If some aspects of Mason's credentials and work history are still murky, what is perfectly clear is that Mason's main job is as a pharmaceutical executive for Elucida Research LLC of Newton, Mass., a pharmaceutical company that lobbied the F.D.A. on behalf of Vascepa ([vascepa.com](http://vascepa.com)), the prescription fish-oil supplement; that he is an



advisory board member of Cardax Pharmaceuticals, the manufacturer of Vascepa; and that he is founder and principal of Elucida Research LLC, which is paid by Cardax to publish research about the benefits of Vascepa and the ills of its competitors. Amongst recent work sponsored in part by Cardax includes a publication that would make Dr. Alexander Leaf proud: “Eicosapentaenoic Acid and Atorvastatin Active Metabolite, Alone or in Combination, Reversed Glucose- and Oxidized LDL-Induced Endothelial Dysfunction Measured Ex Vivo in Rats.” The P.I. is “R. Preston Mason, Ph.D., Department of Medicine, Division of Cardiology, Brigham & Women’s Hospital, Harvard Medical School, and president and founder, Elucida Research LLC.”

Credentials are an elitist business to the core. But they do serve a gating function in science. The Harvard name is supposed to mean something; even in a white elitist world, it’s still supposed to help sort out the geeks from the quacks. That’s also the role *Frontline* and the taxpayer-funded Public Broadcasting Service that produces it are supposed to play. Their only job is to separate the wheat from the chaff. What is most disturbing about the negligence of PBS is not that they didn’t bother to verify Mason’s place of employment—a reasonable person could have been fooled by all his fraudulent bylines in fraudulent academic journals—but that they either didn’t know or didn’t disclose that his livelihood was as a full-time mercenary scientist for Vascepa. His only real paying job, then as now, was to sign their praises and crush their foes.

When I watched the *Frontline* piece later that night, I felt genuinely sorry for the way the editors had impaled Ismail. I wondered, though not aloud, whether he would have fared better if he had focused his airtime on questioning Mason’s qualifications and vested interests. But the idea that a perfect gentleman like Ismail might stoop to his opponents’ muckraking level seemed inconceivable.

Who was to blame for \$1.3 billion in possibly zero-sum spending? My Tenerife exposé was a train wreck: Ismail was clean; the Grim Reaper was just a kind middle-aged lady who was scared to die, and who might have even been helping other death-fearing individuals by positioning fish oil as having maximum psychotherapeutic utility; and R. Preston Mason, the main opponent of Omega-3s, was at least an enemy of science and possibly even a professional con man. I thought about the snake oil, how I had so flippantly buried it in my bag, and I knew then that I was a coward, trembling before my receding prey.

I did meet one true asshole in Tenerife, but she was too irrelevant to be a scapegoat. She was a Ph.D. scientist who had taken a job as Media Spokesperson for a big nutraceutical company. The Media Spokesperson was a thirty-something woman who dressed in stiff suits, shuffled papers, and periodically scanned the room with eyes that stirred with a disquieting mix of opportunism and fear. I asked her politely if she would like to chat about her company’s research, but when I disclosed that I was a *Men’s Health*

reporter, she turned into a monster and declined to speak further without the presence of corporate counsel. The job of “Spokesperson” was narrower than I had imagined.

The last time I saw the Spokesperson was in the airport when we were leaving Tenerife. She was checking in ahead of me in line for her overnight flight back to America via Germany. As a traveler on a corporate expense account, the Spokesperson was booked in business class. However, she had booked her fiancé into economy for the long journey home. The gate agent asked, just to be sure, whether this was what she had intended, and the Spokesperson gave a confident nod. I felt sad for her fiancé, who stood behind her as she obtained their different-colored boarding passes.

Whatever class they were in, the fiancé had probably been hoping that he could spend the flight with his beloved on this special occasion: he had just proposed to her in Tenerife during the conference, which had won the Spokesperson an extra round of applause after her plenary session.

Since time immemorial, consumer demand has given birth to markets. It was not GOED’s idea to scare middle-aged, upper-middle-class white women in urban areas about death, nor to coerce such women into seeking and believing in a new treatment for the terminal human condition. The Omega-3 industry merely showed up to participate in a market that arose from the eternal cruelties of nature, human emotion, and aging. Nutraceuticals showed up to compete with billion-dollar companies that had paid skyscrapers’ worth of credentialed clinicians to set up and control trials so as to maximize their chances of showing positive effects while complying with the FDA’s minimum standards of scientific evidence. The challenge is fierce, but the bounties are worth the bloodshed.

We, consumers of content, are making small decisions, a hundred times every day, not to challenge the validity of the science we consume through the media. Yes, there are the cost-opaque intermediaries to blame, the agency problems and moral hazards, and the fact that we just don’t have enough time on our hands to be skeptical consumers in most situations. But whether directly or indirectly, whether through our wallets or our employers or our elected leaders, we are still willfully choosing to demand these pharmaceutical products at these prices.

When *Men’s Health* had first approached me and asked me to do the Omega-3 article, I was flattered that the magazine would think of me given my own lack of medical or health credentials. They massaged my ego further by saying that they had recruited me because they were skeptical of the truth behind the supplements’ health claims, and they needed a mischievous investigator like me to do the story. I saw it as a chance to write for a new publication with a big readership, and maybe even to expose some bullshit in the process.

In the draft of the article that I first submitted to *Men’s Health*, my main independent

source of medical opinion was Rishi Goyal, M.D., Ph.D., who is an emergency-room physician and professor of medicine at Columbia University with a joint appointment in Columbia's comparative literature department. Rishi is one of the smartest medical thinkers I have ever met. In spite of all this, he does not refer to himself as "Dr. Rishi."

For our purposes here, I have dug up a selection from the deleted Goyal material. "Dietary supplements," Goyal told me, "especially vitamins and minerals, probably had a role in our lives up to a few generations ago when deficiencies really caused recognizable disease: pellagra from vitamin B3 (niacin) deficiency, or scurvy from Vitamin C deficiency." These days, Goyal suggests that many common foods "are so fortified or over-fortified that we just aren't going to get these conditions now." But that doesn't stop us from trying to stop the conditions, and make money doing so. The underlying problem, according to Goyal, is that "the medicalization of everyday life has made it seem that biomedical science has solutions to all the daily bodily insults, infirmities and humiliations that plague us—in the words of *Hamlet*, the 'slings and arrows of outrageous fortune.'"

At this point Dr. Rishi was really getting going. "Like magic beans," he sang, "the dietary supplement industry claims that their products can improve everything from male sexual vigor to workplace productivity to cardiovascular health. But as is often the case when technology promises everything, it delivers nothing. Repeated trials and studies show that this shell game is nothing more than a high-profit scam.... The best that can be said of these sham tonics and remedies is that they're placebos, that at least they don't hurt anybody. But that's not the truth, is it: contaminated herbal supplements have been implicated in causes of liver failure; cases of sudden cardiac death and muscle necrosis have been attributed to amphetamine-like diet pills; and all fat-soluble vitamins (A, D, E, K) have significant toxicity in high doses because of their storage in the body."

"If there's any benefit to be found in food," Goyal concluded, "it probably comes from eating foods whole. The reasons for this are likely multiple: there are probably beneficial compounds we haven't isolated yet; we probably fail to realize the benefits of some compounds we have isolated; certain compounds need to be consumed in conjunction with others to be beneficial; and the process of removal or production of supplements might alter their positive attributes."

During the editorial process, *Men's Health* cut all of my quotes from, paraphrases of, and references to Goyal. To replace the Goyal quotes, *Men's Health* supplied me with a list of pre-approved sources from whom I was to solicit quotes, including three university professors and one weight-loss guru named Alan Aragon, M.S., author of *The Lean Muscle Diet*. One thing that particularly distinguished Goyal from the experts that *Men's Health* instructed me to quote was his frequent use of the word "probably."

Sometimes poetic justice comes back to bite you in the ass. A year after the *Men's*

*Health* article was published, I was doing some research for this chapter when one of my Google searches landed me on an article entitled “[Why You Don’t Need to Take Fish Oil Supplements](#).” It turned out that the article was from *Men’s Health*, published on the auspicious day of April 20, 2016, and that I was the author. The last revision I’d seen contained so few remaining fragments of my own writing that I hadn’t bothered to read it all the way through, either before or after it was published. This made me a bit nervous, although it was comforting to know that I had never been contacted by anyone about it, and that there were almost no comments at the end.

I had been prepared for the general unrecognizability of the writing style, but the first thing that jolted me a bit was the article’s boastful subtitle. “Here’s your new prescription,” it said, as if I were a doctor. I figured that *Men’s Health* readers were in the habit of forgiving this kind of self-aggrandizement. I remained calm and kept going.

Next came some of the quotations from the obligatory *Men’s Health* experts that I had emailed on the magazine’s behalf. Nothing too remarkable there. I tried not to obsess about all the phrases I would never have composed in a million years, like “here’s the bottom line for the average guy” and “the good news is that you don’t have to subsist entirely on anchovies and sardines.”

But then something stopped my heart and poisoned my soul: there was a new paragraph-long quote, inserted roughly at the point in the article where the Goyal material had once been, from an additional source that I did not remember contacting or even being asked to contact. The quote read as follows:

“These labels can be confusing,” says R. Preston Mason, Ph.D., who researches Omega-3s at Brigham and Women’s Hospital. “They make it sound as though dietary fish oil supplements are some sort of approved Omega-3 fatty acid medication, but they’re not. They’re not intended to prevent or treat disease.”<sup>56</sup>

Upon reflection, I thought about how Mason’s point wasn’t actually that different from Goyal’s. Was I only faithful to Goyal because of my higher degree of confidence in the validity of his elitist credentials vs. Mason’s? Or because Goyal was a friend and an all-around bang-up guy? Was Mason? I’d never met the man. How could I know?

In the age of bullshit, it can feel impossible, however hard we try, to immunize ourselves against pseudo-science. It is a delusion—it was my delusion—to think that I could steer clear from the massive pile of shit that we’re all wading in, or even to shovel most of it out of my own house. Yet I obviously could have avoided this rotten outcome by exercising even the least bit of intellectual responsibility. It was my choice not to insist that *Men’s Health* let me have a final read before I went to press with the story, my choice

---

<sup>56</sup> Goldstein, R. “Why You Don’t Need to Take Fish-Oil Supplements.” *Men’s Health*, April 20, 2016.

to let them publish it under my real name.

It was also my decision not to conduct basic research on my publisher, Rodale, before signing a contract with them. If I had done even cursory research on Rodale, I would have discovered that the company was founded by J.I. Rodale, one of the greatest nutrition quacks in American history. He popularized the term “organic,” founded the nasty tabloids *Organic Gardening* and *Prevention*, and died of a heart attack in front of a live ABC studio audience while he was taping an appearance on the Dick Cavett show. Minutes before dying, Rodale, then 72, had told Cavett in his interview: “I’ve decided to live to be 100...I’ve never felt better in my life...I’m going to live to be 100, unless I’m run down by some sugar-crazed taxi driver.” Then he keeled over and croaked right there in his chair. Needless to say, the footage never aired, but it is still the stuff of legend amongst ABC old-timers.

So now, with the cycle completed, I march to the scaffold with J. I. Rodale, R.I.P., and R. Preston Mason, M.B.A.?, and I enter my own byline into the growing list of health and nutrition journalists that have discredited themselves by misleading the public into trusting the expertise of wretched charlatans. Mason’s not just a *Frontline* expert anymore: he’s my expert now. May God have mercy on our souls.

## Afterword

“[The] growth of punctilious discrimination as to qualitative excellence in eating, drinking, etc. presently affects not only the manner of life, but also the training and intellectual activity of the gentleman of leisure. He is no longer simply the successful, aggressive male—the man of strength, resource, and intrepidity. In order to avoid stultification he must also cultivate his tastes, for it now becomes incumbent on him to discriminate with some nicety between the noble and the ignoble in consumable goods. He becomes a connoisseur in creditable viands of various degrees of merit, in manly beverages and trinkets, in seemly apparel and architecture, in weapons, games, dancers, and the narcotics. This cultivation of aesthetic faculty requires time and application, and the demands made upon the gentleman in this direction therefore tend to change his life of leisure into a more or less arduous application to the business of learning how to live a life of ostensible leisure in a becoming way.”

—Thorstein Veblen, *Theory of the Leisure Class* (1899)

The empirical parts of this dissertation have examined consumer demand for beer, wine, and food. The first two of these three goods, at least, would perhaps fall under Thorstein Veblen’s class of “manly beverages.” According to most accounts, Veblen wasn’t much of a manly drinker himself. The only alcohol he mentions in *TLC* is wine, for the purpose of submitting wine knowledge as an example of a badge of conspicuous proof that one has spent long periods of time devoted to unproductive endeavors.

Veblen was generally coy on the question of whether he approved, morally speaking, of alcohol consumption in general. Would he have classified all alcoholic beverages as forms of conspicuous leisure, or could some of them be considered “useful” goods and thus an example of “productive” industry? To answer this question would require a definition more precise than any provided by Veblen of just what it is that makes an industry (or a human want) “productive.” Certainly consuming alcohol, to a considerable extent, counteracts physical productivity on the part of the alcohol consumer; and certainly its role in society is both social and honorific.

Yet alcohol is also a staple good that is demanded and consumed on a daily basis by billions of consumers, and is produced in staggering quantities by some of the world’s most efficient factories. Alcohol’s effects are not purely social: alcohol impacts the human body and brain in a number of fundamental ways, some of which can be physically or mentally beneficial and others of which can be physically or mentally detrimental. The same might be said of looms, cars, or cancer treatments. To classify entire categories of such goods as “productive” or “unproductive,” perhaps, is to deny the complexity that surrounds almost every kind of human consumption.

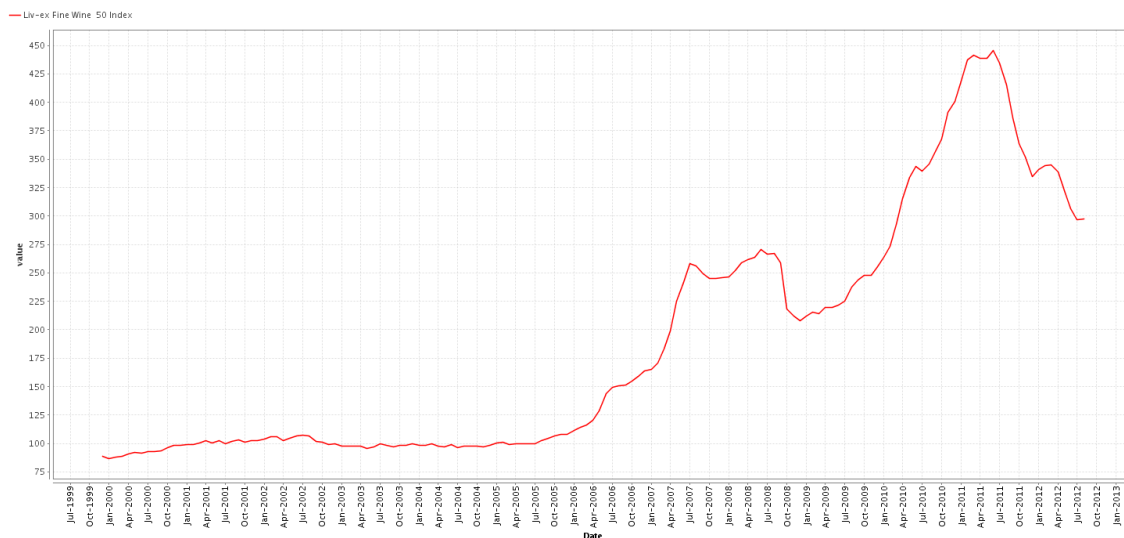
For an atheist without any stated first principles, Veblen displayed a strong metaphysics. Although he was a declared atheist, at times he seemed to advocate for

living like the Luddites or Transcendentalists, whose first principles derived from God. Yet there is little evidence that Veblen himself lived according to any first principles. On the contrary, he was often one of the main antagonists of his own book. In his own life he displayed little of the material industriousness that he admired in his writing. If there is a self-hating undercurrent in *TLC*, though, there are other ways in which the unstated metaphysics of Veblen’s framework are self-validating. Whether due to his physical constitution or intellectual devotion, Veblen was not much for sports, and *TLC* classifies sports as one of the four main unproductive leisure activities of the upper classes (warfare, government, religious observances, and sports). Sports was more of a white, upper-class activity than it would be a century later.

Today, luxury shopping is a leisure sport—especially luxury wine shopping. In the major cities of the United States, Europe, and Asia, consumers are willing to pay more than \$3,000 for a single bottle of newly released Château Pétrus red wine. At supermarkets in these same cities, a basic good that serves the same useful function—dry, unspoiled red wine of 12 to 13 percent alcohol, packaged in a sturdy 750ml bottle with a reliable resealable enclosure—can be bought for about three dollars, or 1,000 times less.

In 1986, according to the *New York Times*, the release price for a bottle of Château Pétrus—the same bottle that now sells for \$3,000 at release—was \$49.50 (Fabricant, 1986). As the sixty-fold increase in the price of Pétrus and the explosion of the LivEx index since 2005 over the subsequent 32 years suggests (Fig. i.1), fine wines have delivered good returns to some investors over the past few decades.

Figure a.1. Liv-Ex 50 prices, 1999 to 2012.



Source: [Information Gain](#)

Even in 1986, though, there were some antique wines selling at auction for more than \$150,000 per bottle. What could be sustaining these thousand-fold spreads, over time, between ordinary prices and super-premium prices? As luxury products have grown into a bigger segment of many consumer-products markets than they once were, if differences between products are not (or are not mainly) useful, then what attributes are consumers actually paying for? In the preceding chapters, this dissertation has made a few attempts to answer this question, including the discussion of “Veblen attributes” introduced in Chapter 3.

Let me be clear that although a bottle of Château Pétrus is canonically full of bullshit attributes or Veblen attributes, my goal has not been to try to explain any prices as high as the ones for Pétrus. Such outlying prices—which sometimes seem to rise and fall as unpredictably as the volatile aromatic molecules falling in and out of solution, giving wine its bouquet—would be unlikely to be described well to any generalizable model; and analyzing them might yield little insight on the market as a whole. Where I observe consumer prices, my approach is generally not to try to explain the prices of like Petrús, Screaming Eagle, Pliny the Elder, Beluga caviar, or other types of “cult” goods that often seem to fall in and out of fashion unpredictably, like volatile aromatic compounds in and out of solution.

Rather than focusing on cult goods, I have used a variety of different experimental techniques (and a few non-empirical techniques as well) to observe bullshit in some of the mainstream segments of wine, beer, and food—attributes and prices that I hope may be relevant to the day-to-day experience of a non-billionaire.



## References

*Occasionally annotated. Not all references are cited in the dissertation; some were simply sources of general wisdom and contributed to my thinking.*

- AB InBev (2019). Press release reporting FY2018 financial results. <https://www.ab-inbev.com/content/dam/universaltemplate/ab-inbev/investors/reports-and-filings/quarterly-reports/2019/FY18-Press-Release-Final-EN.pdf>
- Adams, W. J., and Yellen, J. L. (1976). Commodity bundling and the burden of monopoly. *The Quarterly Journal of Economics*, 1976: 475–498.
- Addressi, E., Mancini, A., Crescimbene, L., Padoa-Schioppa, C., and Visalberghi, E. (2008). Preference transitivity and symbolic representation in capuchin monkeys (*Cebus apella*). *PLoS ONE*, 3:e2414.
- Akerlof, G. A. (1970). The market for ‘lemons’: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84:3, 488–500.
- Akerlof, G. A., and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3), 715–753.
- Alcott, B. (2004). John Rae and Thorstein Veblen. *Journal of Economic Issues*, 38(3), 765–786.
- Alessandri, J., Darcheville, J.-C., Delevoeye-Turrell, Y., and Zentall, T. R. (2008a). Preference for rewards that follow greater effort and greater delay. *Learn. Behav.*, 36, 352–358.
- Alessandri, J., Darcheville, J.-C., and Zentall, T. R. (2008b). Cognitive dissonance in children: justification of effort or contrast? *Psychon. Bull. Rev.*, 15, 673–677.
- Alison, R. and Uhl, K. (1964). Influence of beer brand identification on taste perception. *Journal of Marketing Research*, 1, 36-39. When subjects taste beers with the labels and brands exposed, they rate their favorite beers higher than other brands. When labels are concealed, they show virtually no preferences for certain beers over others, and often score other beers significantly higher than their supposed favorites.

- Almenberg, J., and Dreber, A. (2011). When does the price affect the taste? Results from a wine experiment. *Journal of Wine Economics*, 6:1, 111–121. With a setup similar to Lee et al. (2006), finds that price information matters only when it is given before tasting, and not after, and also finds that positive price signal placebo effects are stronger in females than males.
- Alston, J., Norton, G. W., and Pardey, P. G. (1995). *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*. Ithaca, NY: Cornell University Press.
- Alston, J., Fuller, K., Lapsley, J., Soleas, G., and Tumbler, K. P. (2015). Splendide Mendax: False label claims about high and rising alcohol content of wine. *Journal of Wine Economics*, 10:3, 275–313.
- Amerine, M.A., Panghorn, R.M., and Roessler, E.B. (1965). *Principles of Sensory Evaluation of Food*. New York: Academic Press. Introduces the triangle test.
- Amerine, M.A., and Roessler, E.B. (1976/1983). *Wines: Their Sensory Evaluation*. New York: W.H. Freeman. 1st Ed., 1976. Enlarged and Revised, 1983.
- Anderson, K., and Nelgen, S. (2011). *Global Wine Markets, 1961 to 2009: A Statistical Compendium*. Adelaide: University of Adelaide Press.
- Anderson, K., and Wittwer, G. (2015). Modeling global wine markets to 2018: exchange rates, taste changes, and China’s import growth. *Journal of Wine Economics*, 8:2.
- Anderson, K., and Wittwer, G. (2017). The UK and global wine markets by 2025, and implications of Brexit. Mimeo, Wine Economics Research Centre, University of Adelaide.
- Arantes, J., and Grace, R. C. (2008). Failure to obtain value enhancement by within-trial contrast in simultaneous and successive discriminations. *Learn. Beh.*, 36, 1–11.
- Aronson, E., and Mills, J. (1959). The effect of severity of initiation on liking for a group. *J. Abnorm. Soc. Psychol.*, 59, 177.
- Ashenfelter, O., Ashmore, D., and Lalonde, R. (1995). Bordeaux wine vintage quality and the weather. *Chance*, 8, 7–14.

- Ashenfelter, O. (2008). Predicting the quality and prices of Bordeaux wine. *The Economic Journal*, 118:529, F174–F184. Finds that an econometric model taking climate into account can better predict the future prices of Bordeaux wines than critics' ratings.
- Ashenfelter, O. (2010). Predicting the quality and prices of Bordeaux wine. *Journal of Wine Economics*, 5, 40–52.
- Ashenfelter, O., and Jones, G.V. (2013). The demand for expert opinion: Bordeaux wine. *Journal of Wine Economics*, 8:3, 285–293.
- Ashton, R. (2012). Reliability and consensus of experienced wine judges: expertise within and between? *Journal of Wine Economics*, 7:1, 70-87.
- Ashton, R. (2013). Is there consensus among wine quality ratings of prominent critics? An empirical analysis of red Bordeaux, 2004–2010. *Journal of Wine Economics*, 8:2, 225–234.
- Ashton, R. (2014a). Wine as an experience good: Price versus enjoyment in blind tastings of expensive and inexpensive wines. *Journal of Wine Economics*, 9:2, 171–182. Replicates the main result in Goldstein et al. (2008), finding that in a blind tasting, consumers prefer cheaper wines to more expensive wines.
- Ashton, R. (2014b). “Nothing good ever came from New Jersey”: Expectations and the sensory perception of wines. *Journal of Wine Economics*, 9:3, 304–319.
- Ashton, R. (2017). Dimensions of expertise in wine evaluation. *Journal of Wine Economics*, 12:1, 59–83.
- Ayres, C. E. (1944). *Theory of Economic Progress*. Chapel Hill: University Of North Carolina Press. Available at:  
[https://afee.net/downloads/AFEEbooks/AYRES\\_TEP/TEPHome.htm](https://afee.net/downloads/AFEEbooks/AYRES_TEP/TEPHome.htm)
- Bagwell, K., and Riordan, M. H. (1991). High and declining prices signal product quality. *The American Economic Review*, 224-239.
- Bagwell, L., and Bernheim, B. (1996). Veblen effects in a theory of conspicuous consumption. *American Economic Review*, 86:3, 349–372.

- Baker, J. (2019). Brewers Association launches supporter seal and website. Craftbeer.com, Editor's Picks, 29 January 2019. Available at: <https://www.craftbeer.com/editors-picks/brewers-association-launches-new-independent-craft-brewer-supporter-seal-and-website>
- Basmann, R. L., Molina, D. J., and Slottje, D.J. (1988). A note on measuring Veblen's theory of conspicuous consumption. *The Review of Economics and Statistics*, 70:3, 531–535.
- Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70:5, Part 2, 9–49.
- Becker, G. S. (1991). A note on restaurant pricing and other examples of social influences on price. *Journal of Political Economy*, 99:5, 1109–1116.
- Bello Acebrón, L., and Calvo Dopico, D. (2000). The importance of intrinsic and extrinsic cues to expected and experienced quality: an empirical application for beef. *Food Qual. Prefer.*, 11, 229–238.
- Benjamin, B.A. and Podolny, J.M. (1999). Status, quality, and social order in the California wine industry. *Administrative Science Quarterly*, 44, 563–589.
- Berns, G. S., Laibson, D., and Loewenstein, G. (2007). Intertemporal choice—toward an integrative framework. *Trends Cogn. Sci.*, 11, 482–488.
- Biddle, J. (1991). A Bandwagon Effect in Personalized License Plates? *Economic Inquiry*, 29:2, 375–388. Finds that “sales [of license plates] are related negatively to both current and lagged prices.”
- Bohannon, J., Goldstein, R., and Herschkowitsch, A. (2010). Can People Distinguish Pâté from Dog Food? *Chance*, 23:2, 43–46. In a double-blind tasting, only 17% of consumers are able to pick out the dog food from five blended organ meat products, four of which are intended for human consumption: liverwurst, Spam, pork liver pâté, duck liver mousse, and Newman's Own dog food.
- Bombrun, H., and D. A. Sumner (2003). What determines the price of wine? The value of grape characteristics and wine quality assessment. *AIC Issues Brief*, 18.
- Bowles, S., and Park, Y. (2005). Emulation, Inequality, and Work Hours: Was Thorstein Veblen Right? *The Economic Journal*, 115:507, F397–F412.

- Bradshaw, J. W. S. (1991). Sensory and experiential factors in the design of foods for domestic dogs and cats. *Proceedings of the Nutrition Society*, 50, 99–106.
- Bradshaw, J. W. S. (2006). The evolutionary basis for the feeding behavior of domestic dogs (*Canis familiaris*) and cats (*Felis catus*). *Journal of Nutrition*, 136, 1927–1931.
- Brantlinger, P., and Higgins, R. (2006). Waste and value: Thorstein Veblen and H.G. Wells. *Criticism*, 48:4, 453–475.
- Brochet, F. (2001). Chemical object representation in the field of consciousness. Working Paper, General Oenology Laboratory, Talence, France. Application presented for the Grand Prix of the Académie Amorum following work carried out toward a doctorate from the Faculty of Oenology, University of Bordeaux. Finds that experts not only rate wines higher but also use different, more positive adjectives when describing wine labeled as cru than wine labeled as village-level, even though the two bottles contain the same wine.
- Bureau of Indian Affairs (2019), in the U.S. Department of the Interior. Indian Affairs website. <https://www.bia.gov/>
- Cardebat, J.M. and Figuet, J.M. (2004). What explains Bordeaux wine prices? *Applied Economics Letters*, 11:5, 293–296. Reputation is an important determinant of price, but sensory characteristics are showing to have increasing power vs. previous analysis, which the authors attribute to more competition and fewer information asymmetries in the marketplace.
- Cardebat, J.M. and Figuet, J.M. (2009). Estimation of a hedonic price equation for Alsace, Beaujolais and Provence wines. *Applied Economics Letters*, 16:9, 921–927.
- Cardebat, J.M., Figuet, J.M., and Paroissien, E. (2013). Expert opinion and Bordeaux wine prices: an attempt to correct biases in subjective judgments. *Journal of Wine Economics*, 8:1, 1–22. Constructs a model incorporating meteorological data and finds high sensitivity of prices to expert opinion in Bordeaux, California, and Spain; also finds no Parker effect in Bordeaux.
- Cardebat, J., Faye, B., Le Fur, E., & Storchmann, K. (2017). The Law of One Price? Price Dispersion on the Auction Market for Fine Wine. *Journal of Wine Economics*, 12:3, 302–331.

- Catapano, R., Buttrick, N., Widness, J., Goldstein, R., and Santos, L. R. (2014). Capuchin Monkeys Do Not Show Human-Like Pricing Effects. *Frontiers in Decision Neuroscience*, 5:1330. Using a capuchin population previously trained in a token market at Yale's Comparative Cognition Laboratory, monkeys do not show increased valuation for foods (Jell-O pieces) that have their prices arbitrarily raised, suggesting that price signal effects, biases, and irrationalities may be uniquely human.
- Chalfant, J. A., & Alston, J. M. (1988). Accounting for changes in tastes. *Journal of Political Economy*, 96(2), 391-410.
- Chen, M. K., Lakshminarayanan, V., and Santos, L. R. (2006). How basic are behavioral biases? Evidence from capuchin monkey trading behavior. *J. Polit. Econ.*, 114, 517–537.
- Christensen, Z.T., Ogden, L.V., Dunn, M.L., and Eggett D.L. (2006). Multiple comparison procedures for analysis of ranked data. *Journal of Food Science*, 71:2, S132–S143.
- Cialdini, R.B. (1998). *Influence: The Psychology of Persuasion*. New York: Collins.
- Clement, T. S., Feltus, J. R., Kaiser, D. H., and Zentall, T. R. (2000). “Work Ethic” in pigeons: reward value is directly related to the effort or time required to obtain the reward. *Psychon. Bull. Rev.*, 7, 100–106.
- Cohen, G.A. (2002). Deeper into bullshit. In Buss, S., and Overton, L., eds., *Contours of Agency: Themes from the Philosophy of Harry Frankfurt Finding Oneself in the Other*.
- Cohen, G.A. (2012). Complete bullshit. Chapter 5, *Finding Oneself in the Other*. Princeton, NJ: Princeton University Press.
- Costanigro, M., McCluskey, J., and Mittelhammer, R. (2007). Segmenting the wine market based on price: Hedonic regression when different prices mean different products. *Journal of Agricultural Economics*, 58:3, 454–466.
- Cumbers, A., Davis, J., and McMaster, R. (2015). Theorizing the Social Provisioning Process Under Capitalism: Developing a Veblenian Theory of Care for the Twenty-First Century. *Journal of Economic Issues*, 49:2, 583-590.
- Davis, D. D., and Holt, C. A. (1993). *Experimental Economics*. Princeton: Princeton University Press.

- Dehaene, S. (1997). *The Number Sense: How the mind creates mathematics*. Oxford: Oxford University Press.
- Deliza, R., and MacFie, H. J. H., 2007: The generation of sensory expectation by external cues and its effect on sensory perception and hedonic ratings: A review. *Journal of Sensory Studies*, 11:2, 103–128.
- DiGian, K. A., Friedrich, A. M., and Zentall, T. R. (2004). Discriminative stimuli that follow a delay have added value for pigeons. *Psychon. Bull. Rev.*, 11, 889–895.
- Dodds, W. B., and Monroe, K. B. (1985). The effect of brand and price information on subjective product evaluations. *Adv. Consum. Res.* 12, 85–90.
- Dorfman, J. (1934). *Thorstein Veblen and His America*. New York: Viking Press.
- Duesenberry, J. (1949). *Income, Saving, and the Theory of Consumer Behavior*. Cambridge: Harvard University Press. Presents the “relative income hypothesis.”
- Egan, L. C., Bloom, P., and Santos, L. R. (2010). Choice-induced preferences in the absence of choice: evidence from a blind two choice paradigm with young children and capuchin monkeys. *J. Exp. Soc. Psychol.*, 46, 204–207.
- Egan, L. C., Santos, L. R., and Bloom, P. (2007). The origins of cognitive dissonance evidence from children and monkeys. *Psychol. Sci.*, 18, 978–983.
- Epanechnikov, V. A. (1969). Non-Parametric Estimation of a Multivariate Probability Density. In *Theory of Probability and its Applications*, S. 156.
- Fabricant, F. (1986). Dining out; a French star loses one. *New York Times*, April 27, 1986. Available at: <https://www.nytimes.com/1986/04/27/nyregion/dining-out-a-french-star-loses-one.html>
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Evanston, IL: Row Peterson.
- Festinger, L., and Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *J. Abnorm. Soc. Psychol.*, 58, 203.
- Frank, R. H. (2005). Positional externalities cause large and preventable welfare losses. *American Economic Review*, 95:2, 137-141. Robert Frank’s take on the Veblen effect, cited by Bagwell and Bernheim (1996).

- Frankfurt, H. (2005). *On Bullshit*. Princeton, NJ: Princeton University Press.
- Friedman, D., and Sunder, S. (1994). *Experimental Methods: A primer for economists*. Cambridge: Cambridge University Press.
- Friedman, M. (1962/1976). *Price Theory*. Chicago: Aldine Publishing Company.
- Friedrich, A. M., and Zentall, T. R. (2004). Pigeons shift their preference toward locations of food that take more effort to obtain. *Behav. Processes*, 67, 405–415.
- Fuller, K. B., and Alston, J. M. (2012). The demand for California wine grapes. *Journal of Wine Economics*, 7:2, 192-212.
- Galbraith, J. K. (1973). A New Theory of Thorstein Veblen. *American Heritage Magazine*, 24:3.
- Galbraith, J. K. (1977). *The Age of Uncertainty*. BBC television series. Dir. Andre Deutsch.
- Gipson, C. D., Miller, H. C., Alessandri, J. J. D., and Zentall, T. R. (2009). Within-trial contrast: the effect of probability of reinforcement in training. *Behav. Processes*, 82, 126–132.
- Goldstein, R. (2008a). *The Wine Trials*. New York: Workman Publishing.
- Goldstein, R. (2008b). What does it take to get a *Wine Spectator* award of excellence? “Blind Taste” blog article.
- Goldstein, R. (2010). Book review, “Robert M. Parker’s Wine Bargains: The World’s Greatest Wine Values Under \$25.” *Journal of Wine Economics* 5:1, 209-216. Available at: <https://www.wine-economics.org/journal/details-content/volume-5-2010-no-1/book-review-robert-m-parker-parkers-wine-bargains-the-worlds-greatest-wine-values-under-25/>
- Goldstein, R., and Campbell, S. (2010). *The Beer Trials*. New York: Workman Publishing.
- Goldstein, R., Almenberg, J., Dreber, A., Emerson, J. W., Herschkowitsch, A., and Katz, J. et al. (2008). Do more expensive wines taste better? Evidence from a large sample of blind tastings. *Journal of Wine Economics* 3(2), 1–9. In a large-scale blind tasting of more than 500 consumers and 6,000 glasses of wine, consumers as a whole show a small but significant preference for cheaper wines over more expensive wines. Amongst the



subset of wine professionals and experts, there is no significant correlation between price and preference.

- Gronow, J. (1997). *The Sociology of Taste*. London: Routledge.
- Gustafson, C. R., Lybbert, T. J, and Sumner, D. A. (2016). *Agricultural Economics* 47:1, 91–103.
- Hadj Ali, H., and Nauges, C. (2007). The pricing of experience goods: the example of en primeur wine. *American Journal of Agricultural Economics*, 89:1, 91-103.
- Hadj Ali, H., Lecocq, S., and Visser, M. (2008). The impact of gurus: Parker grades and en primeur wine prices. *Economic Journal*, 118, 158–173.
- Hart, J. (2018). Drink beer for science: an experiment on consumer preferences for local craft beer. *Journal of Wine Economics* 13:4, 429-441.
- Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The Economics of New Goods*. Chicago: University of Chicago Press, 207–248. A thought-provoking challenge to the CPI's way of dealing with technological progress.
- Heineken.com (2019). Found in 192 Countries. From Heineken U.S. website. Available at: <https://www.heineken.com/my/found-in-192-countries>
- Heslinga, J., Reinders, R.J., Schilte, M., and Sekuur, M. (2010). De globalisatie van Heineken. University of Groningen. Available at: <http://www.frw.rug.nl/persons/groote/cursus/global%20village/websites/0910/heineken/index.html>
- Hobsbawm, E. J. (1952). Economic fluctuations and some social movements since 1800. *The Economic History Review*, 5:1, 1-25.
- Hodgson, R. T. (2008). An examination of judge reliability at a major US wine competition. *Journal of Wine Economics*, 3:2, 105–113. When expert judges at the California State Wine Fair are given an identical wine three times during the course of the competition, there is little correspondence between the three scores awarded to the wines. Judges who perform well on this consistency test one year are no more likely than chance to perform well on the same test the next year.

- Hodgson, R. T. (2009). How expert are “expert” wine judges? *Journal of Wine Economics*, 4:2, 233–241. Finds that the chances of winning a medal at wine fairs in California mimic the binomial distribution, and thus can best be explained by chance alone.
- Holt, D. B. (2004). *How Brands Become Icons: The Principles of Cultural Branding*. Cambridge: Harvard University Press.
- Hughey, M. W., and A. J. Vidich (1993). Veblen, Weber and Marx on Political Economy. *International Journal of Politics, Culture and Society*, 6:4.
- International Wine and Spirit Research (IWSR). Vinexpo report (2017): Global study 2015–2020. January 2017.
- Isakson H. R., and Sproles, S. (2008). A brief history of Native American land ownership. In: Simons R. A., Malmgren, R., and Small, G. (Eds.). *Indigenous Peoples and Real Estate Valuation*. Research Issues in Real Estate 10. Boston: Springer.
- Jacoby, J., Olson, J. C., and Haddock, R. A. (1971). Price, Brand Name, and Product Composition Characteristics as Determinants of Perceived Quality. *Journal of Applied Psychology*, 55, 570–579.
- Jaeger, D. A., and Storchmann, K. (2011). Wine retail price dispersion in the United States: searching for expensive wines? *American Economic Review*, 101:3, 136–141. Finds considerable price dispersion in the U.S. wine market.
- 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.
- Jones, G. V., and Storchmann, K. (2001). Wine market prices and investment under uncertainty: An econometric model for Bordeaux crus classés. *Agricultural Economics*, 26, 115–133.
- Just, D. (2015). Is the ban on deception necessary or even desirable? *Food Policy* 83, 5–6.
- Kacelnik, A., and Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour*, 63:2, 245–250.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

- Kahneman, D., and Tversky, A. (1977). *Prospect Theory: An Analysis of Decision Making Under Risk*. McLean, VA: Decisions and Designs.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47, 278.
- Kahneman, D., and Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, 99–127.
- Kalman, P. (1968). Theory of Consumer Behavior When Prices Enter the Utility Function. *Econometrica*, 36, 496–510.
- Knauth, O. (1949). Considerations in the Setting of Retail Prices. *Journal of Marketing*, 14:7, 1–12.
- Kőszegi, B., and Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review* 97:4, 1047–1073.
- Kraft Heinz (2019). Press release announcing FY2018 financial results. <http://ir.kraftheinzcompany.com/news-releases/news-release-details/kraft-heinz-reports-fourth-quarter-and-full-year-2018-results>
- Lakshminarayanan, V. R., Chen, M. K., and Santos, L. R. (2011). The evolution of decision-making under risk: framing effects in monkey risk preferences. *J. Exp. Soc. Psychol.*, 47, 689–693.
- Lakshminarayanan, V., Chen, M. K., and Santos, L. R. (2008). Endowment effect in capuchin monkeys. *Philos. Trans. R. Soc. B.*, 363, 3837–3844.
- Landon, S., and Smith, C.E. (1997). The use of quality and reputation indicators by consumers: the case of Bordeaux wine. *Journal of Consumer Policy*, 20, 289–323.
- Lecocq, S., and Visser, M. (2006). What Determines Wine Prices? Objective vs. Sensory Characteristics. *Journal of Wine Economics*, 1:1, 42–56. Uses a hedonic price equation using sensory and objective product attributes to estimate prices of Bordeaux and Burgundy wines; finds an effect of critics' ratings and high heterogeneity in consumer preferences.
- Lee H., and Sumner D.A. (2013). The economic value of wine names that reference place in the U.S. market: Analysis of 'champagne' and sparkling wine. In Giraud-Héraud

E. and Pichery MC. (Eds.), *Wine Economics: Applied Econometrics Association Series*. London: Palgrave Macmillan.

Lee, L., Amir, O., and Ariely, D. (2009). In search of homo economicus: cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research*, 36:2, 173–187. Finds that consumers commit fewer transitivity violations when rating products that elicit stronger, more emotional responses.

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. When vinegar is added to beer, people like the beer less if they know there's vinegar in it—but only if this extrinsic information cue is presented *before* tasting, not after, suggesting that the information affects the sensory experience directly, rather than the cognitive ratings process.

Leibenstein, H. (1950). Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand. *Quarterly Journal of Economics*, 64:2, 183–207.

Leibowitz, G. (2017). The 1 thing you need to do to write a 'New York Times' best-selling book. Inc.com, 17 July 2017. Available at: <http://inc.com/glenn-leibowitz/want-to-publish-a-new-york-times-best-selling-book.html>

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), 15374.

Levy, D. J., and Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.*, 22, 1027–1038.

Lewis, G. and Zalan, T. (2014). Strategic Implications of the Relationship Between Price and Willingness to Pay: Evidence from a Wine-Tasting Experiment. *Journal of Wine Economics*, 9:2, 115–134.

Library of Congress (2013). The Gibson Girl's America: Drawings by Charles Dana Gibson. Exhibition with commentary. Available at: <https://www.loc.gov/exhibits/gibson-girls-america/the-gibson-girl-as-the-new-woman.html>

- Lindsay, J.A. (2018). Academic grievance studies and the corruption of scholarship. *Areo*, 2 October 2018. Available at: <https://areomagazine.com/2018/10/02/academic-grievance-studies-and-the-corruption-of-scholarship/>
- Lindsay, J.A., and Boyle, P. (2017). The conceptual penis as a social construct. *Cogent Social Sciences* 3:1330439. Available at: <https://www.skeptic.com/downloads/conceptual-penis/23311886.2017.1330439.pdf>
- Luca, M. and Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62:12, 3412–3427. Available at: <http://people.hbs.edu/mluca/fakeitillyoumakeit.pdf>
- 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.
- Makens, J. C. (1964). The pluses and minuses of branding agricultural products. *Journal of Marketing*, 28, 10–16.
- Marglin, S. A. and J. Schor (Eds.) (1990). *The Golden Age of Capitalism: Reinterpreting the Postwar Experience*. Oxford: Oxford University Press.
- Marshall, A. (1890). *The Principles of Economics*. London: Macmillan.
- Matthews, T. (2008). *Wine Spectator* has been scammed. Wine Spectator Forums. Available at: <https://forums.winespectator.com/topic/wine-spectator-has-been-scammed>
- McAdams, R. (1992). Relative preferences. *Yale Law Journal*, 102:1, 101–104.
- McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M. and Montague, P. R. (2005). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, 44, 379–387. Shows, with the help of functional magnetic resonance imaging (fMRI), that having a subject’s favorite soft drink brand’s name (Coca-Cola) on a drink makes it taste better than if the drink is unlabeled.

- McCormick, K. (1986). A note on Leibenstein's Veblen Effect. *Journal of the History of Economic Thought* 8:1, 38–39.
- Mencken, H. L. (1919). Professor Veblen. In *Prejudices: First Series*. New York: Alfred A. Knopf. Ch. 5, 59–83. Available at: <https://play.google.com/store/books/details?id=hg7bAAAAAA7>
- Merriam-Webster.com (2018). “honor.” Retrieved 27 November 2018. Available at: <https://www.merriam-webster.com>.
- Merriam-Webster.com (2018). “leisure.” Retrieved 18 October 2018. Available at: <https://www.merriam-webster.com>.
- Merriam-Webster.com (2018). “treasure.” Retrieved 14 December 2018. Available at: <https://www.merriam-webster.com>.
- Micallef, J. V. Is it whisky or whiskey and why it matters. *Forbes*. Available at: <https://www.forbes.com/sites/joemicallef/2018/05/17/is-it-whisky-or-whiskey-and-why-it-matters/#3e739ac07561>
- Michel, K. E. (2006). Unconventional diets for dogs and cats. *Veterinary Clinics of North America: Small Animal Practice*, 36:6, 1269–1281.
- Milgrom, P., and Roberts, J. (1986). Price and advertising signals of product quality. *Journal of Political Economy*, 94, 796–821.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.
- Miller, J. R., Genc, I., and Driscoll, A. (2007). Wine price and quality: In search of a signaling equilibrium in 2001 California cabernet sauvignon. *Journal of Wine Research*, 18, 35–46.
- Miu, C. (2001). Relationships among wine prices, ratings, advertising, and production: examining a Giffen good. *University Avenue Undergraduate Journal of Economics* 5:1.
- Mohanty, G. F. Experimentation in textile technology, 1788-1790, and its impact on handloom weaving and weavers in Rhode Island. *Technology and Culture*, 29:1, 1–31.
- Morganstern, O. (1948). Demand theory reconsidered. *Quarterly Journal of Economics* 62:1, 164–201.

- Morss, E. (2019). Tasting reports from the Lenox Wine Club, 2012–2019. Morss Global Finance. Available at: <https://www.morssglobalfinance.com>.
- Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78, 311–329.
- Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, 81, 729–754.
- Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: the case of Swedish wine consumers. *European Economic Review*, 39, 1697–1716.
- Ng, Y. (1987). Diamonds are a government's best friend: burden-free taxes on goods valued for their values. *American Economic Review* 77:1, 186–191.
- O'Rourke, K. H., Rahman, A. S. & Taylor, A. M. (2013). Luddites, the industrial revolution, and the demographic transition. *J. Econ. Growth* 18:373. Available at: <https://www.usna.edu/Users/econ/rahman/Luddites12June2011.pdf>
- Oczkowski, E. (1994). A hedonic price function for Australian premium wine. *Australian Journal of Agricultural Economics*, 38, 93–110.
- Parker, R. M. (2009). *Parker's Wine Bargains: The World's Greatest Wine Values Under \$25*. New York: Simon & Schuster.
- Pennisi, L. (2002). Canine evolution: A shaggy dog story. *Science*, 298:5598, 1540–1542.
- Peynaud, E. (1987). *The Taste of Wine: The Art and Science of Wine Appreciation*. (M. Schuster, Trans.). San Francisco: Wine Appreciation Guild.
- Pickering, G. J. (2008). Optimizing the sensory characteristics and acceptance of canned cat food: use of a human taste panel. *Journal of Animal Physiology and Animal Nutrition*, 93:1, 52–60.
- Piggott, N. E., Chalfant, J. A., Alston, J. M., and Griffith, G. R. (1996). Demand response to advertising in the Australian meat industry. *American Journal of Agricultural Economics*, 78(2), 268-279.
- Pigou, A. C. (1913). The interdependence of different sources of demand and supply in a market. *The Economic Journal*, 23:89, 19–24.

- 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. Finds increased activity in the medial orbitofrontal cortex, a brain area implicated in pleasure and reward systems, upon the presentation of the information that wines are expensive, regardless of whether the wines are actually more expensive. In a fully blind tasting of the same wines without any price cues, subjects significantly prefer the cheaper wines to the more expensive wines.
- Pompilio, L., and Kacelnik, A. (2005). State-dependent learning and suboptimal choice: when starlings prefer long over short delays to food. *Anim. Behav.* 70, 571–578. doi: 10.1016/j.anbehav.2004.12.009
- Quandt, R. E. (2007). On wine bullshit: some new software. *Journal of Wine Economics*, 2(2), 129–135.
- R. A. Poldrack (London: Elsevier Inc.), 81–93.
- Rao, A. (2005). The Quality of Price as a Quality Cue. *Journal of Marketing Research* 42(4), 401-405.
- Rao, A.R. and Monroe, K.B. (1989). The effect of price, brand name, and store name on buyers' perceptions of product quality: an integrative review. *Journal of Marketing Research*, 36, 351–357.
- Riordan, M.H. (1986). Monopolistic competition with experience goods. *Quarterly Journal of Economics*, 101, 265–279.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82:1, 34-55.
- Round Creative (2018). Rebrand...and Reconnect with your Customers. Blog article, 12 March 2018. Available at: <https://round.co.uk/blog/rebrand-and-reconnect-with-your-customers/>
- Santos, L. R., and Chen, M. K. (2009). The evolution of rational and irrational economic behavior: evidence and insight from a non-human primate species. In *Neuroeconomics: Decision Making and the Brain* (P. W. Glimcher, C. Camerer, and E. Fehr, Eds.). Academic Press.



- Schaffer, M. (2009). *One Nation Under Dog: Adventures in the New World of Prozac- Popping Puppies, Dog-Park Politics, and Organic Pet Food*. New York: Henry Holt and Company.
- Schamel, G. and Anderson, K. (2003). Wine quality and varietal, regional and winery reputations: Hedonic prices for Australia and New Zealand. *The Economic Record* 79(246): 357–369.
- Schmalensee, R. (1978). A model of advertising and product quality. *Journal of Political Economy*, 86, 485–503.
- Schnabel, H., and Storchmann, K. (2010). Prices as quality signals: Evidence from the wine market. *Journal of Agricultural and Food Industrial Organization*, 8, 1–21. In the case of wine, the information effect dominates the quality effect; price signals respond positively to wine quality and negatively to increasing information.
- Schor, J. B. (1992). *The Overworked American: The Unexpected Decline of Leisure*. New York: Basic Books.
- Schuessler, J. (2018). Hoaxers slip breastaurants and dog-park sex into journals. *New York Times*, 4 October 2018. Available at: <https://www.nytimes.com/2018/10/04/arts/academic-journals-hoax.html>
- Scitovsky, T. (1945). Some consequences of the habit of judging quality by price. *Review of Economic Studies*, 12:32, 100–105.
- Scott, D. (2010). What Would Veblen Say? *Leisure Sciences*, 32:3, 288–294.
- Shapiro, C. (1983). Premiums for high quality products as returns to reputations. *Quarterly Journal of Economics*, 98, 659–679.
- Shibasaki, M., and Kawai, N. (2008). The effects of response cost and time on choosing a stimulus. *Shinrigaku Kenkyu* 79, 241–249. doi: 10.4992/jjpsy.79.241
- Shibasaki, M., and Kawai, N. (2011). The reversed work-ethic effect: monkeys avoid stimuli associated with high effort. *Jpn. Psychol. Res.* 53, 77–85. doi: 10.1111/j.1468-5884.2010.00449.x
- Shiv, B., Carmon, Z., and Ariely, D. (2005). Placebo effects of marketing actions: consumers may get what they pay for. *J. Mark. Res.*, 42, 383–393.

- Smith, A. (1759). *The Theory of Moral Sentiments*. London: A. Millar; Edinburgh: A. Kincaid and J. Bell. Available at:  
<https://books.google.com/books?hl=en&lr=&id=WOc7AQAAAMAJ&oi=fnd&pg=PA1&ots=4ud6QmjrvL&sig=vkWYj0ejfbWfepo78VP1Q0YCNfY#v=onepage&q&f=false>
- Smith, V.L. (1982). Microeconomic systems as an experimental science. *American Economic Review* 72:5, 923-955.
- Solomon, G. E. A. (1990). Psychology of novice and expert wine talk. *American Journal of Psychology*, 103, 495–517.
- Stata (2019). Stata 13 Manual: Kernel density function. Available at:  
<https://www.stata.com/manuals13/rkdensity.pdf>
- Starr, A., Tomlinson, R. B., and Brannon, E. M. (2018). The acuity and manipulability of the ANS have separable influences on preschoolers' symbolic math achievement. *Frontiers in Psychology* 10.3389. Available at:  
<https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02554/full#B10>
- Stevens, J. R. (2010). Intertemporal choice. In *Encyclopedia of Animal Behavior*, Vol. 2. (M. Breed and J. Moore, Eds.). Oxford: Academic Press, 203–208.
- Storchmann, K. (2012). Wine economics. *Journal of Wine Economics*, 7:1, 1–33.
- Storchmann, K., Mitterling, A., and Lee., A. (2012). The detrimental effect of expert opinion on price-quality dispersion: Evidence from the wine market. AAWE Working Paper No. 118.
- Sumner, W. G. (1911). *What Social Classes Owe To Each Other*. New York: Harper and Brothers.
- Swinnen, J. (2017). *Beeronomics: How Beer Explains The World*. Oxford: Oxford University Press.
- Tellis, G. J., and Wernerfelt, B. (1987). Competitive price and quality under asymmetric information. *Marketing Science*, 6:3, 240–253.
- Thrane, C. (2004). In defense of the price hedonic model in wine research. *Journal of Wine Research*, 15:2, 123–134.

- Tilman, R. (2006). Colin Campbell on Thorstein Veblen on Conspicuous Consumption. *Journal of Economic Issues* 40:1, 97–112.
- Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies*, 63, 1–22.
- Total Wine (2019). Online retail store. Available at: <http://totalwine.com>
- Toynbee, A. (1884). *The Industrial Revolution*. Boston: Beacon Press. Available at: <https://archive.org/details/industrialrevol00toyngoog/page/n6>
- Valenzi, E., and Eldridge, L. (1973). Effect of Price Information, Composition Differences, Expertise, and Rating Scales on Product- Quality Rating. *Proceedings of the 81st Annual Convention of the American Psychological Association*, 829–830.
- vanMarle, K., Aw, J., McCrink, K., and Santos, L. R. (2006). How monkeys (*Cebus apella*) quantify objects and substances. *J. Comp. Psychol.*, 120, 416–426.
- Vasconcelos, M., and Urcuioli, P. J. (2009). Extensive training is insufficient to produce the “work ethic” effect in pigeons. *J. Exp. Anal. Behav.*, 91, 143–152.
- Vasconcelos, M., Urcuioli, P. J., and Lionello-DeNolf, K. M. (2007). Failure to replicate the “work ethic” effect in pigeons. *J. Exp. Anal. Behav.*, 87, 383–399.
- Veblen, T. (1892). The price of wheat since 1867. *Journal of Political Economy* 1:1.
- Veblen, T. (1898). The instinct of workmanship and the irksomeness of labor. *American Journal of Sociology* 4:2, 187–201.
- Veblen, T. (1899). *Theory of the Leisure Class*. London: Unwin Books.
- Veblen, T. (1904). *Absentee Ownership: Business Enterprise in Recent Times, The Case of America*. Chicago: Decennial Publications, University of Chicago.
- Veblen, T. (1918). *The Higher Learning in America: A memorandum on the conduct of universities by business men*. New York: B. W. Huebsch.
- Veblen, T. (1919). *An Inquiry Into the Nature of Peace and the Terms of its Perpetuation*. New York: B. W. Huebsch.

- Veblen, T. (1921). *The Engineers and the Price System*. New York: B. W. Huebsch.
- Veblen, T. (Trans., Ed.) (1925). *The Laxdela Saga. Translated from the Icelandic, with an introduction by Thorstein Veblen*. New York: B. W. Huebsch.
- Waber, R. L., Shiv, B., Carmon, Z., and Ariely, D. (2008). Commercial features of placebo and therapeutic. *Jama*, 299:9, 1016–1017.
- Wallace, D. F. (2003). *Everything, and More: A compact history of infinity*. New York: W. W. Norton.
- Wanat, M. J., Kuhnen, C. M., and Phillips, P. E. M. (2010). Delays conferred by escalating costs modulate dopamine release to rewards but not their predictors. *J. Neurosci.*, 30, 12020–12027.
- Weil, R. L. (2005). Analysis of reserve and regular bottlings: Why pay for a difference only the critics claim to notice? *Chance*, 18, 9–15. In a triangle test, consumers perform little better than chance (40%) in distinguishing a reserve bottling from a regular bottling from the same winery.
- Weil, R.L. (2001). Parker vs. Prial: The death of the vintage chart. *Chance*, 14:4, 27–31.
- Weil, R.L. (2005). Analysis of reserve and regular bottlings: why pay for a difference only the critics claim to notice? *Chance*, 18:3, 9–15.
- Weil, R.L. (2007). Debunking critics' wine words: can amateurs distinguish the smell of asphalt from the taste of cherries? *Journal of Wine Economics*, 2:2, 136–144.
- Weinberg, J. (2008). Taste how expensive this is: A problem of wine and rationality. In F. Allhoff (Ed.), *Wine & Philosophy: A Symposium on Thinking and Drinking*. Malden, MA: Blackwell, 257–274.
- Wilson, D. (2014). Arnold Toynbee and the industrial revolution: the science of history, political economy and the machine past. *History and Memory*, 26:2, 133-161.
- Wine-searcher.com (2019). Database of wine prices. Available at: <http://wine-searcher.com>.
- Wine Spectator* (2008). “Awards of Excellence” issue, print edition, August 2008. New York: M. Shanken Communications.

*Wine Spectator* (2019). Website. Available at: <https://www.winespectator.com/>

Young, S. (2017). Is showering every day bad for you? New research says yes. *The Independent* (UK), 26 January 2017. Available at: <https://www.independent.co.uk/life-style/health-and-families/is-shower-everyday-good-or-bad-cleaning-hygiene-university-utah-amazon-village-microbes-skin-a7546216.html>

Zentall, T. R. (2010). Justification of effort by humans and pigeons: cognitive dissonance or contrast? *Curr. Dir. Psychol. Sci.*, 19, 296–300.

Zentall, T. R., and Singer, R. A. (2007). Within-trial contrast: pigeons prefer conditioned reinforcers that follow a relatively more rather than less aversive event. *J. Exp. Anal. Behav.*, 88, 131–149.

Ziliak, S., & McCloskey, D. N. (2008). *The Cult of Statistical Significance: How the standard error costs us jobs, justice, and lives*. University of Michigan Press.

Ziliak, S. (2008). Retrospectives: Guinnessometrics: the economic Foundation of “Student’s” t. *Journal of Economic Perspectives*, 22:4, 199–216.