

A Preliminary Test of the Resonance Marketing Hypothesis: Guidance for Future Research Assessing Multi-Attribute Preferences in Horizontal Competition

Eric K. Clemons
The Wharton School

clemons@wharton.upenn.edu

Josh Wilson
The Wharton School

wilsonsj@wharton.upenn.edu

Abstract

The Resonance Marketing hypothesis suggests that in an era of nearly perfect online information consumers will be able to know everything about products that interest them, including their exact descriptions, the best available sellers, and prices throughout the market. Consumers will then purchase those products that offer the greatest consumer surplus, which may not be the products with lowest prices in areas where individual consumers have the strongest preferences. We divide the Resonance Marketing hypothesis into three testable subordinate hypotheses. We test one of the three, Verifiability: Consumers will not be deceived by false or manipulated content, and will be able to assess products and their reviews and be able to make rational purchase decisions. Preliminary data suggest that testing Verifiability will be difficult and will require complex experimental design. The paper ends with insights into how one would more completely test the Verifiability Hypothesis and suggestions for future research.

1. Introduction

This paper offers an experimental examination of one element of the Resonance Marketing hypothesis. The Resonance Marketing hypothesis was initially put forward by Clemons [7]. It states that in an era of nearly perfect online information consumers will be able to know everything about products that interest them, including their exact descriptions, the best available sellers, and prices throughout the market. Consumers will then purchase those products that offer the greatest consumer surplus, which may not be the products with lowest prices in areas where individual consumers have the strongest preferences. The hypothesis has been used to explain a range of phenomena.

The hypothesis seems plausible. More favorable reviews, if believed, should lead to more trial. Additionally, if consumers' experience is consistent with the reviews then trials should lead to adoption.

The hypothesis has been accepted by other authors, and has been defended theoretically; see, for example works by Markopoulos [20, 21] and Dye and Lahiri [13, 14]. The hypothesis has been subject to some empirical testing; see, for example, works by Clemons and his colleagues [7, 8, 9, 10] and Duan, Gu, and Whinston [15, 16]. But there has been limited exploration of the mechanisms that would enable the Resonance Marketing hypothesis and little or no

URI: <http://hdl.handle.net/10125/41833>
ISBN: 978-0-9981331-0-2
CC-BY-NC-ND

controlled experimentation.

In this paper we conduct a preliminary study of one mechanism that enables Resonance Marketing. In order to do so we subdivide the Resonance Marketing hypothesis into three subordinate hypotheses for the purpose of experimentation and exploration. We then explore one of them in more detail in this paper.

- The Trial Hypothesis — favorable reviews will lead to trial.
- The Verifiability Hypothesis — False, scammed, or astroturfed reviews will immediately be seen as false by consumers as soon as they try the product, and the reviews will be rejected immediately. Only reviews that are true and accurate will influence repeat purchase.
- The Rational Adoption Hypothesis — Consumers will be informed about those categories that matter to them because it is now so easy to access information online. Consumers will act in accordance with their informed preferences, unless otherwise constrained from doing so.

Preliminary findings were mixed. The design of simple experimental studies could not fully support the Verifiability Hypothesis. The results of our study of the Rational Adoption hypothesis will be reported in a companion paper.

The problems we encountered with the test of the Verifiability Hypothesis should not have been unanticipated. Indeed, they are inherent in the concept of Resonance Marketing. The Resonance Marketing hypothesis argues that perfect information will accelerate the launch of new products and increase horizontal differentiation. As horizontal differentiation increases, products will compete on their individual characteristics and their precise location in their products' multidimensional product attribute space [7]. However, consumers with very different innate preferences will respond very differently if asked to sample the same products, depending on whether they strongly like or strongly dislike the products. Similarly, consumers with very different innate preferences may respond very differently to reviews that enthusiastically endorse products in a category they like or that they dislike. Most significantly, consumers may respond differently to inaccurate and disparaging reviews of truly superior products, depending on whether or not the product is in a category that they like or that they dislike. This suggests that a more complex experimental design may be required to assess the Verifiability Hypothesis.

The structure of this paper is as follows. Section 2 reviews selected relevant literature. Section 3 provides a formal statement of our hypotheses. Section 4 provides an overview of our methodology. Section 5 reviews our experimental examination of the Verifiability Hypothesis. Finally, section 6 reviews limitations and suggestions for future research. It presents future research designs for exploring the Verifiability Hypothesis.

2. Review of Relevant Literature

2.1. Fit Cost and Consumer Behavior

Hotelling [18] and Salop [25] both contributed to the framework used to analyze consumers' choice of products to purchase. Both assume that consumers buy the product that creates the greatest surplus for them. Both start with the assumption that consumers have a maximum willingness to pay for their individual ideal product within a category. Both assume that a consumer's maximum willingness to pay for an instance of a product within a category is the consumer's valuation for their ideal product, minus the distance between the ideal product and the actual product being considered times a unit compromise cost, called the unit fit cost. The distance times the unit fit cost can be considered either the fit cost or the compromise discount. When consumers choose among competing products, they do not buy the cheapest one, but rather the one for which they receive the greatest surplus, which is their maximum valuation minus price and minus the compromise discount.

Hotelling places consumers on an abstract unit line, called the Hotelling Line. Salop acknowledges that the space of all possible products is not linear and uses a unit circle, the Salop Circle. The Salop Circle simplifies some of the arithmetic analysis by allowing us to ignore boundary conditions; rationally, firms would want to avoid locating their product too close to an edge, since there are no consumers to attract on the far side of the edge.

In reality, both Hotelling and Salop have created useful but restrictive simplifications. No product is located in a uni-dimensional space of alternative products, whether a straight line or a circle. Cars are large or small, fast or not, fuel efficient or not, nimble off road or comfortable for touring, in almost any combination of attributes. Beers are hoppy or not, light roast or not, all malted barley or based on a mix of ingredients, low alcohol or high alcohol, again in almost any combination of attributes. Products exist in a complex and multidimensional product attribute space (PAS). Not surprisingly, consumers have multi-attribute utility functions, and determining fit cost involves a weighted Euclidean distance function.

2.2. The Resonance Marketing Hypothesis and the Role of Uncertainty

The Resonance Marketing hypothesis argues that online content and the nearly free access to accurate

information will alter consumer behavior in predictable ways. Long before Salop and Hotelling, marketing has that known that consumers shop to maximize their expected utility. Consumers have always attempted to purchase products that left them the greatest surplus, that is, the products that were closest to their ideal, or that had the lowest expected fit cost.

However, in the presence of significant uncertainty, a consumer would not know where an unfamiliar product was actually located in its product attribute space. There was an expected value and a band of uncertainty around that expected value, and the product might be located anywhere within the band of uncertainty. This band of uncertainty represented a wide range of locations that a product could actually take in its product attribute space. Each location had its own corresponding *compromise discount*, based on the location of the consumer's ideal product, the possible location of the product in the band of uncertainty, the distance between the two locations, and the consumer's individual fit cost [7]. The numerical average of all these possible compromise discounts led to the expected compromise discount, also called the *uncertainty discount* [7]. This uncertainty discount is not a result of risk aversion, but is simply the result of the fact that the product could be placed other than where the consumer hoped it would be located in product attribute space. Figure 1 shows a product in a simplified single-dimension product attribute space. The consumer's ideal product is located at the asterisk P. The consumer's maximum willingness to pay is V for products at P. The further to the left or right the product is moved the less the consumer is willing to pay for it, as shown by the downward sloping lines on either side of P. X_L and X_R mark the consumer's willingness to pay at the extreme edges of the range of uncertainty, and the average of all WTP values within the range of uncertainty yields U, the uncertainty discount.

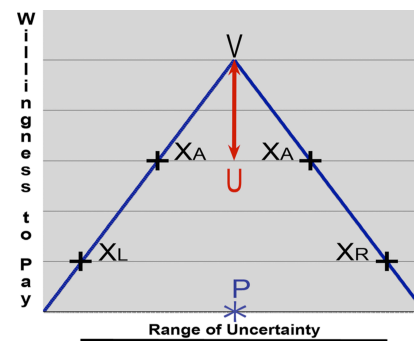


Figure 1.—Impact of uncertainty on a customer's willingness to pay.

Marketers have long understood that the uncertainty discount blocked new product introduction. Advertising alone could communicate that a product existed, but could not communicate its precise location in its PAS. Marketers have historically had at least four mechanisms for reducing uncertainty about

a product's attributes, in order to reduce the uncertainty discount. It differs from both The Long Tail [1, 4, 5] and Trading Up [27] in its emphasis on consumers' access to and use of information.

- More *content-based advertising*, to attempt to communicate what a product was in terms that would locate it precisely in its PAS.
- Providing *deep discounts* to encourage initial trial and eliminate the uncertainty discount by allowing the customer to become familiar with the product through use. Selling a product at or below cost is not a viable long-term strategy, but it may be very effective if a significant fraction of the initial trials lead to loyal long-term customer relationships.
- Providing *free samples*, which allows the customer to try the product without needing to acquire it or pay for it; this is the ultimate deep discount.
- Offering an *extended test* of the product, or selling the product with an option to return it within a specified time period without penalty. At the time of their launch, Lexus and Infiniti offered guest drives or multi-day test drives to acquaint customers with their vehicles. Early Palm and Apple smart phones allowed 30 day free returns. Free samples and deep discounts are inappropriate ways to introduce new consumer durables, since there will not be a repeat purchase for months or years. Here, extended tests were the most appropriate alternatives.

Each can be effective in an environment where products are targeted at a mass market. But none is appropriate when the expected total market is very small. The cost of ad impressions per purchase, or the cost of products sold at or below cost, or the cost of extended tests, will all be prohibitive if only a small fraction of consumers contacted actually adopt the product. None of the traditional mechanisms for reducing uncertainty is entirely appropriate for niche markets and resonance marketing [6].

The Resonance Marketing hypothesis argues that organic informedness, from various online sources, replaces advertising. There is now a vast literature on the impact of online content, word of mouth, and viral marketing on consumer purchasing behavior. The following are only a small and representative sample of the early literature [6, 11, 12, 24, 28]. This reduction in uncertainty “unblocks” innovation and has contributed to product proliferation, niche markets, and resonance marketing [7].

Resonance marketing argues that firms no longer need to invest in advertising to reach consumers and therefore no longer need to fight to defend huge but low margin fat spots like Budweiser, Miller General Draft, Coors Banquet, Bud Light, Miller Lite, and Coors Light. Resonance Marketing argues that firms can now innovate, stake out, and profit from a large portfolio of small but high margin sweet spots. In-

deed, Resonance Marketing can be viewed as the move from massive firms using advertising to defend their fat spots, to smaller and more agile firms competing on the basis of numerous sweet spots defended without advertising.

2.3. Using Advertising When It Can be Effective

Advertising can be effective if there is a small range of the product attribute space that has passionate consumers. In this case it makes sense to develop products to appeal to those consumers and to promote the product actively [21]. Here, the cost per ad impression divided by the number of consumers who actually adopt will allow advertising to be effective.

Alternatively, if it were possible to target ads cheaply and effectively, once again advertising would be relevant to the introduction of products, even resonance products with an expected small market niche.

This is consistent with the Resonance Marketing hypothesis because it supports the idea that reducing uncertainty is essential to successful launches of resonance products.

2.4. Using Sampling When It Can be Effective

Two recent papers by Dey and Lahiri present an interesting finding on the effective use of sampling for resonance products [13, 14]. They focus on computer games. They explore the use of two part-tariffs in the sale of new games. This sales strategy allows consumers to make an initial purchase to reduce uncertainty about the game and then to make a second purchase, if they desire to do so, that adds functionality and capabilities not present in the first purchase.

The essential idea is that the first purchase reduces uncertainty on fit by reducing uncertainty about the game's actual location in its complex and multi-attribute product attribute space. The second purchase allows those consumers who are interested to spend more heavily on the game. Without the reduction in uncertainty few consumers would make the full purchase.

The two-part purchase can reduce revenue slightly, if some consumers are satisfied with the first purchase and therefore do not make the second purchase. But more often the two-part purchase will increase revenue significantly, since many more consumers who are delighted with the first purchase have their uncertainty reduced, and then make the second purchase. This work shows that a modified form of sampling can work to promote resonance products. In this case the modified form of sampling was a base purchase and a second purchase that augments the first. A limited time license and a longer-term license could work as well. The seller can engage in providing samples even for resonance products in this case because the cost of the samples is paid in part or in full by the buyer.

This work is once again consistent with the Resonance Marketing hypothesis because it once again

supports the idea that reducing uncertainty is essential to successful launches of resonance products.

2.5. Response to Reviews and to Misleading Reviews

The resonance marketing hypothesis assumes that reviews are accurate, and that consumers can therefore trust them. For reviews to remain accurate, however, it is necessary that consumers be able to assess reviews, and thus reject those that are inaccurate or worse, are deliberately misleading or fraudulent. If consumers cannot be misled or influenced by false reviews then there will be little or no incentive to post false reviews, at least in categories that rely primarily upon repeat purchases. Thus, resonance marketing will work over the long term if and only if reviews remain reliable and trusted, which in turn requires that consumers be able to detect and reject false reviews.

We know that online reviews influence consumers' online choices [26]. We know that reviews influence what other reviewers write [23]. Indeed, this is true even when reviewers initially disagree with prior reviews [2]. But do reviews alter perception when they disagree with the consumer's experience?

It has long been accepted by entrepreneurs in the beer industry that sophisticated beer drinkers' perceptions of their beers cannot be influenced by reviews and thus cannot be manipulated by advertising, by cues such as branding, labels, or packaging, or by manipulation of online content.¹ This is an essential element of common strategy of the entire craft brewing industry, which has been based on making good beer and letting word of mouth promote the product.

And yet there is reason to believe that this may be false. In the wine industry we have long known that reviews, price, and packaging are no indication of how a wine will fare among consumers in a blind tasting (see, for example, [17]). Neural Science allows us to measure the sensation of pleasure as it arises in the brain of a subject, which has demonstrated that loyal Coke drinkers are truly happy to drink Pepsi, and vice versa, as long as each thinks he is drinking his preferred beverage. That is, not only will an avowed Pepsi drinker tell you he is happy drinking a Coke, as long as you fool him by putting it in a Pepsi can, his brain will indicate that he truly *is* happy as long as the deception is undetected [22].

Worse yet, there are experiments to suggest that beer drinkers are as easy to deceive as consumers of wine and soft drinks [3, 19]. If you tell a group of MIT students that they are drinking an industrial lager laced with vinegar they will reject it. If you tell a

group of students that they are drinking a craft brewer's new Belgian Sour Lambic they will embrace it, even if, in fact, it is the same industrial lager laced with vinegar.

Is there any evidence to support the idea that sophisticated consumers of craft beer are any different from wine drinkers and cola drinkers? Is there any evidence to support the views of the craft brewing industry that they are safe from the industrial major breweries, safe from their advertising, and safe from the possibility manipulation of online content, because their customers would never be influenced? This question is vitally important to the future of craft brewing, and indeed to the future of all new niche super-premium product categories. If consumers can be deceived then large companies with large promotional budgets will dominate.

3. The Hypotheses

The Resonance Marketing Hypothesis argues that trial and initial purchase will be encouraged by favorable information, which reduces the uncertainty discount and increases consumers' willingness to pay for products that are initially unknown and unfamiliar to them. For this hypothesis to remain valid users have to trust reviews; that is, reviews truly must reduce consumers' uncertainty. This will only be true if sellers and their competitors have no incentive to post false reviews (astroturfing), greatly reducing the value of community content reviews. We do not test The Trial Hypothesis in this paper. We begin instead by testing the Verifiability Hypothesis, which is essential to the Trial Hypothesis.

The Verifiability Hypothesis suggests that consumers can immediately determine if a review is accurate or not. This has obvious implications for consumer behavior. A consumer may be influenced to try a product based on a false positive review, but a consumer will not continue to purchase a product after initial purchase based on a false review. We further divide the Verifiability Hypothesis into two hypotheses, which we examine in our experiments.

False Praise Hypothesis H1A: *False positive reviews do not lead to consumer acceptance of a product.*

False Criticism Hypothesis H2B: *False negative reviews do not lead to consumer rejection of a product.*

If the verifiability hypothesis holds, false reviews do not alter consumer's long-term behavior. If false reviews do not influence consumers' behavior, firms will not invest in creating them. Consumers should trust online reviews. This is essential for the plausibility of The Trial Hypothesis, and indeed, for the Resonance Marketing hypothesis generally.

4. Experiments Explore the Hypotheses

The experiment is designed to test both parts of

¹ We had numerous conversations with Ron Barchet and Bill Covalesski, founding CEO and President of Victory Brewing, and their guests, over the period 2004 through the present. We had similar conversations with Devin Kimble and Scott Robertson, founder and brew master of Brewwerkz in Singapore, and their guests, with Carl Setzer and LIU Fang, founders of Great Leap Beer in Beijing, and with numerous other brewers and their customers around the world.

the Verifiability Hypothesis, The False Praise and the False Criticism hypotheses. The experimental design is straightforward. Four categories of beers were selected, Pilsner, India Pale Ale (IPA), Wheat, and Imperial Stout.² In each category we selected a *better* and a *worse* beer. While quality is subjective, within a category quality is more objective. Fans of IPAs will generally agree on the characteristics of a superior IPA, and fans of an imperial stout will likewise be in agreement on what constitutes a superior imperial stout. For each of the eight beers selected we used reviews taken from Ratebeer.com, which is a reliable site with millions of reviews.³ For each of the eight beers we also prepared accurate reviews, describing the strengths and limitations of each beer. Not surprisingly, the reviews of the superior beers were more favorable than the reviews of the inferior beers.

The eight beers are listed below in table 1, at the end of the paper. As should be clear from table 1, our assessments of the superior and the inferior beer in each pairing enjoyed significant support in the reviewing community.

Subjects were given one beer in each of two categories; that is, in each of two categories a subject was given either a superior or an inferior beer to taste. Moreover, each beer was accompanied by a review. Some subjects received reviews that were correctly paired with the beer that the review assessed. Some subjects received reviews that incorrectly paired with a beer in the same category; that is some subjects might receive a high quality Pilsner paired with the review of a low quality Pilsner. Finally, as a limited control, some subjects received a beer to sample without receiving any review.

None of the beers was identified by name during the experiment. We did this to ensure that impressions and expectations were created by the review and not by associations with the name of an individual beer. We also did this to ensure that we did not either inflate or harm the reputation of a beer unfairly by pairing it with a mismatched review. All beers were presented in an identical glass designed for tasting, with no cues to the beers' true identity other than the review paired with it. As noted above, reviews were accurate in the description of the beers' types, but may not have accurately reflected the beers' quality within its type. All eight reviews are included in appendix A

Our expectation was that false positive reviews

² Pilsner is the most popular beer category in the world, a pale, crystal clear, relatively low alcohol beer that originated in Pilsen and Buda. Over time Pilsners have morphed into the industrial beer popular around the world. India Pale Ale is an amber ale, higher alcohol than a Pilsner and much more intensively flavored with bitter hops. There are numerous sub-styles of wheat beer, ranging from low alcohol and inexpensive German Hefeweizen and Belgian Wit to stronger and more complex Belgian Triples.

³ Ratebeer.com has worked with us to limit Astroturfing. Additionally, with millions of reviews from tens of thousands of reviewers, it would not be easy to create a successful astroturfing campaign on this site.

would not improve consumers' assessment of an inferior product, and the false negative reviews would not damage a consumer's assessment of a superior product.

The experiment was conducted at Victory Brewing Company. All subjects had expressed an interest in craft brewing, and all subjects had just completed the tour of the Victory brewing facility in Downingtown, Pennsylvania. We have 606 data points, where a data point represents a single individual sampling a single beer. Since on average a participant sampled 2 beers, this represents data from 303 distinct individuals. Data are summarized in table 2 below.⁴

	Number of Observations	
	Correct Review	Deceptive Review
Shock Top	79	29
Golden Monkey	41	32
Storm King	48	60
Black Ops	16	52

	Number of Observations		
	Correct Review	Deceptive Review	No Review
Bass Ale	12	21	4
Hop Devil	42	21	24
Prima Pils	19	25	19
Miller	22	29	11

Table 2.—Number of observations

5. Findings from the Beer Experiment

Findings from the beer experiment provided only mixed support for either part of the Verifiability Hypothesis. There were clear superior and inferior beers in each experimental pairing. Subjects received correctly paired reviews with the beers they were given, or incorrectly paired reviews. (Remember, subjects did not receive both beers in a pair, but only received one or the other. The intent of the pairing was to allow us to check for inversions in the subjects' collective assessments, not to test whether individual subjects inverted their own preferences.)

Our findings were mixed. All eight summary observations are listed in table 3 below. Note that in two of the eight cases, false criticism did hurt superior beers. Pairing the superior beers Black Ops and HopDevil with false and critical reviews did lower consumers' assessment of the beer, compared to assessments with correct reviews. Moreover, false negative reviews lowered subjects' assessment of both of these beers enough to result in both Black Ops and HopDevil being ranked below the inferior beer to which they were being compared.

In hindsight this mixed support should not have been surprising. It has been noted that with the most extreme resonance products average assessments are

⁴ Readers will note that we added the treatment "no review" after starting the experiment. Also, we stopped the experiment early, when we realized that there was no simple explanation for our findings.

not uniformly good indicators of a product's popularity or its assessment. Some resonance products are extreme, and extreme products often produce extreme reactions. Indeed, with resonance products extreme reactions are often essential for a product's success. "It is better to have some people love you and others hate you, than to have everyone like you" [7], because only people who like a product will pay its higher price.

Additionally, there is actually greater variation in beers than in many other beverages, such as wine. Rieslings are different from Chardonnays, and Cabs are different from Zins, but none would be considered shocking. Beers use different grains, different degrees of light or dark roasting, different hops and different degrees of hopping, and different yeasts. Some, like a Lambic, could indeed be considered shocking, since they represent all the sanitation problems that Louis Pasteur sought to eliminate with Pasteurization.

Shock Top Correct > Shock Top Deceptive (False praise did not help)
Golden Monkey Correct < Golden Monkey Deceptive (False criticism did not hurt)
Golden Monkey Correct > Shock Top Deceptive (False praise did not produce inversion)
Golden Monkey Deceptive > Shock Top Correct (False criticism did not produce inversion)
Storm King > Storm King Deceptive (False praise did not help)
Black Ops Correct > Black Ops Deceptive (False criticism did hurt)
Black Ops Correct > Storm King Deceptive (False praise did not produce inversion)
Black Ops Deceptive < Storm King Correct (False criticism did produce inversion)
Bass Correct > Bass Deceptive (False praise did not help)
Hop Devil Correct > Hop Devil Deceptive (False criticism did hurt)
Hop Devil Correct > Bass Deceptive (False praise did not produce inversion)
Hop Devil Deceptive < Bass Correct (False criticism did produce inversion)
Miller Correct > Miller Deceptive (False praise did not help)
Prima Correct < Prima Deceptive (False criticism did not hurt)
Prima Correct > Miller Deceptive (False praise did not produce inversion)
Prima Deceptive > Miller Correct (False criticism did not produce inversion)

Table 3.—Summary of Experimental Findings

Both Black Ops and HopDevil are extreme products and there is no indication that all drinkers will rank them the way they would be ranked by their fans. Black Ops has strong notes of bourbon from its aging in bourbon barrels. HopDevil is extremely hoppy, and at one point was both Victory's best selling beer and its most hated beer simultaneously. Perhaps the fact that pairing these beers with false negative reviews

produced a reduction in subjects' assessments can be attributed to subjects having been influenced by positive reviews to rate the beers higher than they normally would have done. Perhaps the more negative reviews actually resonated with some consumers' true reactions to the beers, allowing them to express more negative assessments. Our sample sizes are not large enough for confidence intervals to be meaningful. More importantly, the high variance usually experienced when reviewing Black Ops and HopDevil can be attributed to different characteristics among beer drinkers and their preferences; our sample sizes are not large enough to support correlations between subjects' demographics and subjects' assessments.

In particular, we asked subjects to assess beers that they had not set out to purchase. There was thus no reason for us to assume we had asked them to taste beers in a category they enjoy. For a variety of reasons we could not therefore assume that they would rank the beers the way we would. Two examples might be useful here. Hyper-Hoppy triple IPAs are hopped to produce a level of bitterness never before experienced in a beer or any other beverage. If someone likes this category of beer, Devil Dancer, Old Crustacean, or Old Fog Horn might receive very high reviews from informed consumers who sought the beers out, bought them with full knowledge of what they were buying, and liked them. They might receive very low reviews from others who hated the style, even from experienced drinkers of craft beers. Wild Fermentation Ales, especially Lambics, are produced with whatever wild yeast and wild bacteria fall into the broad flat fermentation tanks. They are most notable for their cloudy appearance, lack of head, and profound barnyard stench. Again, if someone likes this category of beer, the best Belgian Lambics might receive very high reviews from informed consumers who sought the beers out, bought them with full knowledge of what they were buying, and liked them. But once again the best Lambics might receive very low reviews from others who hated the style, even from experienced drinkers of craft beers. There is no reason that subjects would agree with us on which beers were better and which beers were worse; indeed, if they hated a style, the best examples of it might indeed be the ones they liked least in our tastings.

Interestingly, in some instances false and misleading negative reviews actually increased subjects' ratings for several beers that we thought were superior. Figure 2 (at end) is a box plot diagram of subjects' assessments of two beers in a category where the best beers are extreme. The plots show assessments under conditions of *no review*, *accurate review*, and *deceptive review*. Since Bass Ale is the less intensely hopped of the two IPAs, the deceptive review praised it as a HopDevil while the accurate review correctly described a Bass. In the case of the Prima Pils, the Prima was the superior beer in the Pilsner pairing. Here the accurate review correctly described a Prima

Pils, while the deceptive review described a Miller.

In figure 2 it is clear that when presented with an accurate negative review that indicated that Bass Ale was a moderately hoppy British IPA subjects' assessments went down relative to the assessments of subjects that had no indication that this was a hoppy beer. With the inaccurate reviews, that described Bass as an extremely hoppy American IPA reviews went down even further. Is it possible that subjects believed the reviews, knew that they hated hops, and lowered their assessment precisely in accordance with the degree of hoppiness indicated in the reviews? This is consistent with the cola experiments [22], but we certainly do not have enough information to conclude that this is the correct explanation.

We can perform a similar analysis of the assessments for Prima Pils summarized in figure 2. Without any reviews, subjects assessed Prima Pils as roughly as enjoyable as Bass Ale. When they learned, correctly, that it was an extremely hoppy Pilsner, often considered the best Pilsner in the US, their assessments went down. When subjects were informed, incorrectly, that it was an industrial Pilsner like a Miller their assessments also went down, but not as much. An interpretation consistent with the data would be that subjects *were* influenced by reviews. They did not like anything they were told was very bitter. They did not like anything they were told was a cheap industrial lager. Once again, this is consistent with the cola experiments and with Dan Ariely's experiments [3, 19]. Subjects were influenced by our reviews, but not in the manner that would be expected solely by the strength of the reviews. Once again we do not have enough information to determine the cause of the anomalous impact of the reviews.

Two examples will help illustrate the problems of confounding horizontal (differentiation) and vertical (quality) assessments. Both reviews were taken from Ratebeer.com. TheBeerLover, a reviewer with over 1,000 reviews, gave Victory HopDevil the highest possible ratings (AROMA 10/10 APPEARANCE 5/5 TASTE 10/10 PALATE 5/5 OVERALL 20/20). His comments were unrestrained:

"This beer must be born of original sin, as it pretty much seduces anyone and everyone who has ever had the pleasure of tasting it. I have had this beer hundreds of times now, but each time is almost identical to my first. This is one hell of a good beer. HopDevil pours to a beautiful deep amber/burnt orange color with a nice white head, and a lively carbonation. The nose is the first thing that draws you in. Very piney/citric waves of zesty and spicy hop aromatics flood the nose. There is a nice underpinning of malt in the nose, but hops take center stage. The palate is very firm, and this is what makes HopDevil a little different than other American IPA's, and why it is so loved. There is a sea of smooth maltiness on the palate, good biscuity and caramel malt flavors coat the tongue. The finish is what takes you straight to

hell. A blistering, long, dry, resinous, piney and citric hop bite tattoos the tongue, and lingers for an eternity. What more do I need to say? This is an incredible beer from one of America's best breweries. HopDevil goes very well with spicy fare. I love to match this beer with volcano hot buffalo wings, or with spicy Thai or Indian dishes."

In contrast, Fluffy, with over 2,000 reviews was unimpressed. (AROMA 1/10 APPEARANCE 3/5 TASTE 2/10 PALATE 1/5 OVERALL 5/20). His comments are helpful and provide clear insight into confounding horizontal and vertical assessments.

"good for what it was, but im not a IPA fan, sorry. my wife bought this for me, not realizing i dont like the style. so im not a good judge for stuff like this."

Horizontal preferences clearly can dominate vertical quality differences. Our findings therefore suggest that more complex and more robust tests are needed to assess the Verifiability Hypothesis.

6. Limitations and Directions for Future Research

Our work has three principal limitations. First, because we were constrained by the amount of beer that IRB would permit us to serve and the conditions under which we were allowed to serve it, and by our limited access to subjects over time, we observed a subject under the most artificial conditions, standing in front of observers, rating a beer after one sip after reading a review of it. How different would consumers' behavior be if they were drinking the beer over the course of days or weeks, under more natural conditions? Would the effect of false reviews rapidly erode with more exposure to a beer?

Second, because our samples were so small, we were observing subjects without a large enough number to attempt to correlate the impact of reviews on subjects' expertise and experience with a category of beer or preference for that category of beer. We cannot determine the causes of the subjects' behavior.

Finally, our subjects sampled beers we gave them, rather than types of beer for which they had revealed enough interest to make a purchase. It's possible that these categories of beer might have been of little interest to the subjects because they did not like them. It's possible that they might never have bought them after seeing a review or a description of the category, so their preferring the less extreme example of the category might not be inconsistent with the Resonance Marketing hypothesis. An ideal beer in a category a subject does not enjoy is likely to receive a weaker assessment than an insipid version of the same category; this is especially true of beers that are extremely hoppy, extremely sour, or have strong flavors of bourbon or coffee. It is therefore possible that subjects' personal rankings of superior or inferior beers might differ from the generally accepted rankings. A subject who disliked hoppy beers might actually prefer a

pseudo-Pilsner like Miller to a traditional hoppy Pilsner like Prima Pils. A subject who disliked hoppy beers might likewise actually prefer a less authentic India Pale Ale like a modern Bass Ale to a beer like HopDevil, which retains a more authentic level of hoppy bitterness. Subjects' assessments of some beers may have been inverted from our expectations because subjects truly enjoyed the less extreme, less authentic examples the tasted. Moreover, strong positive reviews might have caused some subjects to expect to dislike beers in styles that they did not like

These subjects probably would not have bought HopDevil or Prima Pils no matter how strong the accurate reviews were. They would not have tried the beers and therefore would not have been influenced by favorable or unfavorable reviews. They would not have adopted these beers anyway, with or without deceptive reviews. The subjects' behavior may not be inconsistent with either the False Praise or the False Criticism hypotheses. Importantly, false reviews may not alter these consumers' long-term consumption behavior.

Each of the limitations suggests directions for future research.

- With a larger budget researchers could create consumer panels that assessed beers after consuming six beers or even a case at home. Consumers would be asked to review the beer after a month, rather than after a sip. This is consistent with the firms have historically conduct their assessment of consumer reactions to a product.
- With a much larger subject population we could assess whether subjects were more easily influenced by reviews that were consistent with their own preferences, or were more easily influenced by reviews for products that were outside their categories of interest. This is essential to assessing the limitations of the Verifiability hypothesis.
- With a much larger set of beers from which to choose we might ask subjects to select the categories that they were most likely to purchase. We could then offer subjects beers only from categories that were important to them. Perhaps subjects who enjoyed specific categories of beer would immediately reject inaccurate reviews in those categories. That would provide a much stronger test of the Verifiability Hypothesis and of its applicability to the Resonance Marketing Hypothesis. It would let us assess whether accurate or inaccurate reviews were likely to affect selection and adoption in categories in which adoption was likely.

The principal contribution of the paper is highlighting the complexity of accurately assessing the Verifiability Hypothesis and suggesting directions for future research exploring it. Since resonance products exhibit strong horizontal differentiation, consumers cannot be expected to have preferences consistent

with vertical comparisons within a category.

7. References

- [1] Anderson, C. *The Long Tail: Why the Future of Business Is Selling Less of More*, New York: Hyperion Press, (2006).
- [2] Aral, S. "The Problem With Online Ratings", *Sloan Management Review*, December 19, 2013, <http://sloanreview.mit.edu/article/the-problem-with-online-ratings-2/>.
- [3] Ariely, D., *Predictably Irrational, Revised and Expanded Edition: The Hidden Forces That Shape Our Decisions*, Harper Perennial (2010).
- [4] Brynjolfsson, E., Hu, J. Smith, M. D. From Niches to Riches: Anatomy of the Long Tail, *Sloan Management Review*, (Summer 2006) 64-71.
- [5] Brynjolfsson, E., Smith, M.D. Hu, Y. Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers, *Management Science*, Vol. 49, No. 11, (2003). 1580-1596.
- [6] Chevalier, J.A., Mayzlin, D. The Effect of Word of Mouth on Sales: Online Book Reviews, *Journal of Marketing Research* Vol. 43 No. 3, (2006) 345-354.
- [7] Clemons, E.K. How Information Changes Consumer Behavior And How Consumer Behavior Determines Corporate Strategy, *Journal of Management Information Systems*, Vol. 25, No. 2 (Fall 2008) 13-40.
- [8] Clemons, E. K, Gao, G. Consumer Informedness and Diverse Consumer Purchasing Behaviors: Traditional Mass-market, Trading Down, and Trading Out into the Long Tail, *eCommerce Research and Applications*, Vol. 48, No. 1 (Dec. 2009) 46-56.
- [9] Clemons, E. K., Gao, G., Hitt, L. M. When online reviews meet hyperdifferentiation: A study of the craft beer industry, *Journal of Management Information Systems*, Vol. 23, No. 2, (Winter 2006-2007) 149-171.
- [10] Clemons, E. K., Spitler, Gu, B. Markopoulos, P. Broadband and Hyperdifferentiation: Creating Value by Being *Really* Different, *The Broadband Explosion*, Cambridge: Harvard Business School Press, (2005) pp. 137-164.
- [11] Cordain, Loren, S Boyd Eaton, Anthony Sebastian, Neil Mann, Staffan Lindenberg, Bruce A. Watkins, James H. O'Keefe, and Janette Brand-Miller. "Origins and Evolution of the Western Diet: Health Implications for the 21st Century." *The American Journal of Clinical Nutrition* 81 (February 2005).
- [12] Dellarocas, C. The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Systems. *Management Science*, Vol. 49, No. 10 (2003) 1407-1424.

- [13] Dellarocas, C., Zhang, M. Awad, N. Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures. *Journal of Interactive Marketing*, Vol. 21, 4 (2007) 2-20.
- [14] Dey, P. and Lahiri, A., “Versioning of Video Games: Go Vertical in a Horizontal Market?”, *Proceedings, 49th International Conference On System Sciences*, Koloa, HI, 2016, pp: 5259-5268.
- [15] Dey, P. and Lahiri, A., “Versioning: Go Vertical in a Horizontal Market?” *Journal of Management Information Systems*, 2016 (To Appear).
- [16] Duan, W., Gu, B., and Whinston A. B., “Do Online Reviews Matter? — An Empirical Investigation of Panel Data”, *Decision Support Systems*, Vol. 45., No. 4, pp. 1007-1016.
- [17] Duan, W., Gu., B., and Whinston A. B., “The Dynamics of Online Word-of-Mouth and Product Sales — An Empirical Investigation of the Movie Industry,” *Journal of Retailing*, Vol. 84, No 2, pp. 233- 242.
- [18] Goldstein, R., Herschkowitsch , A., and Walters, T. (eds.), *The Wine Trials 2011*, Fearless Critic Media (2011)
- [19] Hotelling, H. Stability in Competition, *The Economic Journal*, Vol. 39, (1929) 41-57.
- [20] Lee, L., Shane, F., and Ariely, D. “Try It, You’ll Like It — The Influence of Expectation, Consumption, and Revelation on Preferences for Beer”, *Psychological Science* Vol. 17, No, 12 (2006) pp. 1054-1058.
- [21] Markopoulos, P. M., and Clemons, E. K., “Reducing Buyers’ Uncertainty about Taste-Related Product Attributes”, *Journal of Management Information Systems*, Vol. 30, No. 2. (2013) pp. 269-299.
- [22] Markopoulos, P. M., Aron, R., and Ungar, L. H., "Information Market for Product Attributes: A Game Theoretic, Dual Pricing Mechanism”, *Decision Support Systems* (2010 49:2).
- [23] McClure, S. M., Jian, L., Tomlin, D., Cypert, K. S., Montague, L. M., Read, P., “Neural Correlates of Behavioral Preference for Culturally Familiar Drinks”, *Neuron*, Vol. 44, No. 2, pp.379–387 (2004)
- [24] Moe, W. W., and D. A. Schweidel. “Online Product Opinions: Incidence, Evaluation, and Evolution”, *Marketing Science* Vol. 31 No. 33, (2012) pp. 372-386.
- [25] Riedl, J., Konstan, J., Vrooman, E. *Word of Mouse: The Marketing Power of Collaborative Filtering*, Warner Business Books, (2002).
- [26] Salop, S. C., Monopolistic Competition with Outside Goods, *The Bell Journal of Economics*, Vol. 10, No. 1 (1979) 141-156.
- [27] Senecal, S., and Nantel, J. “The Influence of Online Product Recommendations on Consumers' Online Choices”, *Journal of Retailing*, Vol. 80, No. 2, (2004) 159-69.
- [28] Silverstein, M. and Fiske, N. *Trading Up: The New American Luxury*, New York: Penguin Group, (2003).
- [29] Ye, Q., Law, R., and Gu, B. “The Impact of Online User Reviews on Hotel Room Sales,” *International Journal of Hospitality Management*, Vol. 28 No. 1, pp. 180-182.

8. Large Tables and Figures

Category	Superior Beer	Inferior Beer	Comments
Pilsner	Victory Prima Pils	Miller	Miller and Budweiser started as Pilsners decades ago but are now more accurately classified as Industrial Pale Lagers. Prima Pils remains a true Pilsner.
	Ratings 2384	Ratings 2169	
	Wtd Avg 3.58/5	Wtd Avg 1.60/5	
	Overall 93% Style 99%	Overall 0% Style (NA)	
IPA	Victory Hop Devil	Bass Ale	Bass Ale was the first IPA and originally was the standard by which all IPAs were measured. Over time Bass reduced its hoppiness, and other IPAs became more popular. Victory Hop Devil remains an extremely hoppy IPA, but in the American rather than the English style.
	Ratings 3168	Ratings 2182	
	Wtd Avg 3.78/5	Wtd Avg 3.11/5	
	Overall 98% Style 98%	Overall 45% Style 57%	
Wheat	Victory Golden Monkey	Shock Top	Golden Monkey is made in the style of a Belgian Tripel Wheat. Shock Top is an American Industrial version of Belgian Wit, also a Wheat beer but in a different Belgian style. To an expert they are not truly comparable. But they are both wheat beers, of extreme and obvious differences in quality.
	Ratings 2501	Ratings 1301	
	Wtd Avg 3.64/5	Wtd Avg 2.56/5	
	Overall 95% Style 97%	Overall 13% Style (NA)	
Imperial Stout	Brooklyn Black Ops	Victory Storm King	Black Ops is bourbon barrel aged, the most expensive and the most highly regarded sub-category of Imperial Stout. Its flavor is distinctive, and consumers who dislike bourbon may strongly dislike Black Ops. Storm King is more traditional and not entirely comparable. Note that reviewers here have inverted the beers' rankings, reflecting the range of responses to bourbon barrel aging.
	Ratings 933	Ratings 3022	
	Wtd Avg 3.96/5	Wtd Avg 4.00/5	
	Overall 99% Style 94%	Overall 100% Style 96%	

Table 1.—The eight beers used in the experiment.

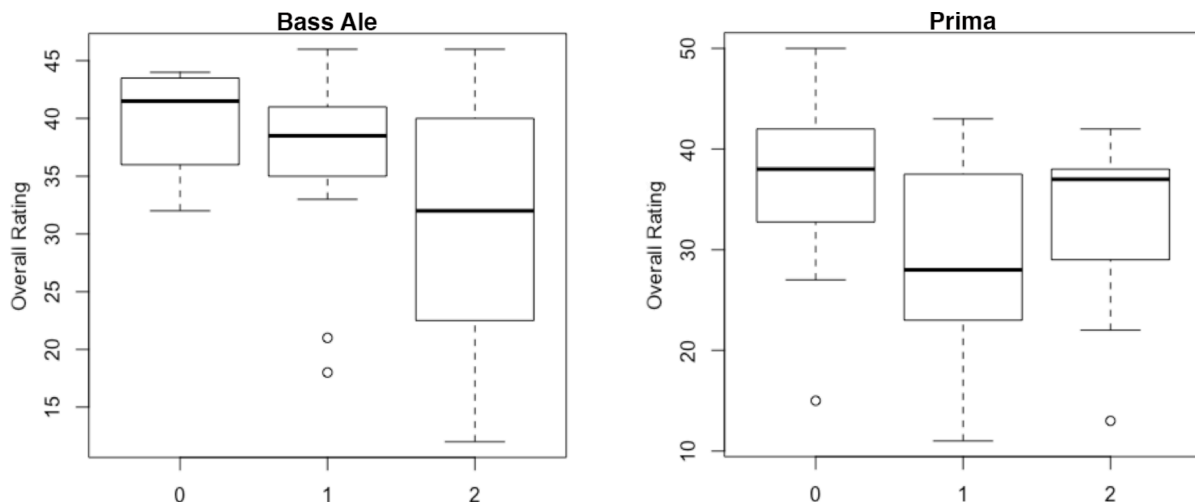


Figure 2.—Perverse Effects of Reviews on Ratings of Two Beers. 0 indicates no review, 1 indicates correctly paired review, 2 indicates incorrectly paired review.

Appendix A

Shock Top	A clean, well-made North American industrial wheat beer from one of the world’s largest brewers. Wheat, barley, and corn flavors predominate in the grain mix. Lightly hopped. Orange and hazy, with a small white head. Strong flavors from orange peel and secret spices add complexity to the nose and to the finish. Moderate to light in the mouth, for its type.
Golden Monkey	A rich, intensely flavorful American take on a Belgian Tripel. Pours with a golden body and huge frothy white head. Nose of lightly roasted malted barley, freshly baked bread, yeast, orange peel, citrus fruits, and spices. A huge beer, in all aspects, from its feel in the mouth to its long, sweet finish.
Storm King	A great American craft Imperial Stout, with huge complexity hiding its high alcohol content. Pours almost black, with a small tan to brown head. Both aroma and flavor exhibit the sweetness of dark roasted malts, married to notes of chocolate and bitter espresso, held together with a great slam of citrusy bitter west coast American hops. Huge feeling in the mouth, long sweet and bitter finish.
Black Ops	A great American craft Imperial Stout, finished with months of aging in bourbon barrels for additional complexity. Pours black and opaque in the glass, with a small tan to brown head. Both aroma and flavor exhibit the sweetness of dark roasted malts, married to notes of chocolate and bitter espresso, held together with the rich alcohol complexity of good bourbon. Huge feeling in the mouth, long sweet finish of chocolate, espresso, and alcohol.
Miller	A clean, well-made North American industrial beer, leaning more towards Czech Pilsner than towards a Bavarian Lager. Light flavor of light roasted malt, and sweet notes from adjuncts like corn and rice. Lightly hopped with American Idaho hops. Perfume and hay in the nose. Very little finish. Good value.
Prima Pils	A clean, crisp, refreshing craft Pilsner assertively hopped with whole flower Nobel Czech Saaz and German Tettnang hops. Lightly roasted all malt barley provides a solid backbone. Sweet malt, floral Noble hops, and grass in the nose. Long, dry, bitter finish. A great beer, and a great example of a very popular style.
Bass Ale	The original India Pale Ale, modified to suit modern palates. A rich Amber-roasted malt provides a solid backbone. Aroma of sweet malts and traditional English hops. The traditional English hops add a significant hoppy kick to the taste. Everything has been turned back a notch to suit modern drinkers, their food preferences, and their life-styles, making this an easy beer to love.
HopDevil	The classic American IPA. Heavy aromas of pine wood from the whole, American hop flowers. The malts are German imports, which add to the earthy aroma. A light body with huge flavors of citrus from the American hops. Very hoppy, with bitterness accompanying the citrus flavors long into the finish, which is nicely dry.