Credence Services: Content, credibility, and usefulness of online reviews

Research-in-Progress

Shannon Lantzy
R.H. Smith School of Business
University of Maryland
College Park, MD
slantzy@rhsmith.umd.edu

Katherine Stewart
R.H. Smith School of Business
University of Maryland
College Park, MD
kstewart@rhsmith.umd.edu

Rebecca Hamilton
Professor of Marketing
537 Hariri Building
McDonough School of Business
Georgetown University
Washington, D.C. 20057
rebecca.hamilton@georgetown.edu

Abstract

Credence products are those whose quality is difficult or impossible for consumers to assess, even after consuming the product (Darby & Karni, 1973). For example, it is difficult to assess the technical skill and knowledge of a physician even after a visit. This research is focused on the content, structure and consumer perceptions of online reviews for credence services. We start by examining how the content and structure of real online reviews of credence services systematically differs from those of experience services (Nelson, 1970). We find that online reviews of credence services are more likely to contain unsupported claims than reviews of experience services. We experimentally examine consumer perceptions of reviews, varying both their structure and content. Consumers rationally discount the credibility of credence claims when presented with short, simple reviews but we expect more complex argument structure and inclusion of experience attributes in the review to attenuate this effect.

Keywords: information asymmetry, credence services, online reviews

Introduction

Despite living in the information age, consumers in credence markets still face a significant information deficit. Markets for credence products and services are differentiated from markets for search and experience products by the difficulty of obtaining accurate information about quality. For example, although a consumer might be able to evaluate the price of an entree (a search attribute) prior to eating at a restaurant, it is very difficult to evaluate its tastiness (an experience attribute) without having consumed a meal, and even after eating the meal, a consumer cannot verify the claim that it was made with organic ingredients (a credence attribute). Economic theory predicts that consumers of credence services are vulnerable to a wide range of fraud such as being overcharged or over-treated (Dulleck & Kerschbamer, 2006; Asher Wolinsky, 1993). A number of studies establish that fraud occurs in real markets: foreign visitors are subjected to longer and higher-priced taxi rides than locals (Balafoutas, Beck, Kerschbamer, & Sutter, 2011), as many as half of recommended automotive repairs are actually unnecessary (Rasch & Waibel, 2012; Schneider, 2012), and typical consumers receive unnecessary surgeries at a higher rate than more informed patients (e.g. physicians; Dulleck & Kerschbamer, 2006; Gruber, Kim, & Mayzlin, 1999; Rehavi & Johnson, 2013). These studies demonstrate that information asymmetries lead to real overtreatment and overcharging.

Online reviews allow consumers to share information about their service experiences with others. For most services, sharing this information should reduce market information asymmetries between potential consumers and service providers (Huang, Lurie, & Mitra 2009). However, because by definition consumers cannot assess the quality of credence attributes, consumer reviews of credence services (e.g., doctors, auto mechanics) are of dubious credibility. Indeed, many doctors have strongly resisted the legitimacy of consumer reviews (Andrews, 2008; Jain, 2010), even requiring patients to sign documents promising never to review their doctors (ElBoghdady, 2012). Nonetheless consumers have numerous forums to review credence services (e.g., RateMDs.com, Angie's List, Yelp), and while there are a number of studies that focus on one credence service type (e.g. physician reviews), to the best of our knowledge there are no studies which compare credence services to search or experience services, or studies which include multiple service types in one sample (Huang, Lurie, & Mitra, 2009; Mudambi & Schuff, 2010; Park & Kim, 2008).

We investigate the broad question "Are online reviews a useful and credible source of information for consumers of credence products?" More specifically, we ask: what information is provided in credence service reviews, and how do consumers perceive them? We answer these questions in three related studies. The first study, a qualitative content analysis of real online reviews, addresses the question "what information is provided?" A controlled experiment – designed to measure consumer perceptions of credence versus experience claims – addresses the question, "how do consumers perceive credence service reviews?" A third study is proposed to test the perceived usefulness and credibility of credence claims alone and when combined with either experience data or credence data, as well as to replicate and generalize the findings from the first experiment.

Our analysis of real reviews demonstrates that online reviews of credence services significantly differ from those of experience services in their content and structure. Reviews of credence services include more information about credence attributes and less about experience attributes, and include more unsupported claims than reviews of experience services. Our first experimental study suggests that consumers find credence claims significantly less credible than experience claims, and credence service reviews significantly less credible than experience service reviews.

These systematic differences in the content, structure and consumer perceptions of credence vs. experience service reviews have several important implications. First, broad conclusions from existing online review literature may not extend to credence service reviews. For example, the established finding that higher ratings lead to higher sales may not apply if the high ratings are associated with reviews that are not perceived to be credible (Chevalier & Mayzlin, 2006). Furthermore, information platforms such as Yelp and Angie's List do not differentiate between credence and experience reviews in the design of their review systems. These platforms may be able to improve consumer perceptions of their reviews by designing review forms and templates tailored to credence vs. experience and search products and services.

Literature Review

According to economics of information theory (EIC), products are distinguished by the time and cost required for consumers to evaluate the product's qualities. Nelson (1970) was the first to differentiate between product qualities that may be evaluated by the consumer before purchase (i.e. search qualities or attributes), and qualities that can only be evaluated after purchase (i.e. experience qualities). In this literature, products whose most important qualities are search qualities (or "attributes") are classified as search products, and products whose most important attributes are "experience" attributes were classified as experience products. Darby and Karni (1973) extended this framework to "credence qualities which are expensive to judge even after purchase" (p69, Darby & Karni, 1973).

Nelson, Darby and Karni initiated a rich stream of theoretical and empirical research on credence product markets. The economics literature has largely focused on theoretical outcomes of fraud (i.e. over-treatment, over-charging, and under-treatment) and other aspects of seller behavior that are unique to credence products (Balafoutas et al., 2011; Beck, Sutter, & Kerschbamer, 2010; Dulleck, Kerschbamer, & Sutter, 2011a; Dulleck & Kerschbamer, 2006; Emons, 1997; Kerschbamer, Dulleck, & Sutter, 2009; Kerschbamer, Sutter, & Dulleck, 2009; Liu, 2011; Mimra, Rasch, & Waibel, 2012; Asher Wolinsky, 1993). The marketing and information systems literature has focused on how companies can overcome consumers' lack of information about product quality, e.g., through branding or other marketing strategies (e.g., Bloom & Pailin, 1995; Galetzka, Verhoeven, & Pruyn, 2006; Lim & Chung, 2011; Srinivasan & Till, 2002).

It has been posited that online reviews reduce consumers' information asymmetry and may be more credible for credence and experience information than seller advertising. For example, Lim and Chung (2011) find that consumers use word of mouth to change their evaluations of credence attributes; this example demonstrates that consumers may use word of mouth information sources for credence quality evaluation. Indeed, surveys of consumers show that online reviews of doctors – a classic credence service – are increasing in number and use (Fox & Jones, 2009). Huang et al. (2009) suggest that the internet and online word of mouth (OWOM) are blurring the lines between search and experience attributes, though they do not evaluate OWOM effects on credence products.

Review Content

Most products are multi-attribute products with a mix of search, experience, and credence attributes (Darby & Karni, 1973; Ford, Smith, & Swasy, 1990; Lim & Chung, 2011; Srinivasan & Till, 2002). However, most online review research treats each review as a single entity, subsequently studying and reporting on the effects of review-level measures. There is a growing body of research that acknowledges the existence of information about distinct product attributes within reviews; this research is focused on automatically extracting attribute information from reviews using natural language processing and other techniques (see Decker & Trusov, 2010, for a review). Since we know that multiple attributes of a product are mentioned in reviews, and products are composed of a mix of search, experience, and credence attributes, we reason that online reviews are likely to contain mentions of a mix of search, experience, and credence attributes.

Online reviews are written by consumers at least in part in an effort to share information with other consumers. The motivations for sharing product information may include helping other consumers (altruism), affiliating with others, and building an online reputation (e.g., Bateman, Gray, & Butler, 2006; Kraut & Resnick, 2010; Moe & Trusov, 2010; Hamilton, Schlosser, & Chen, 2014; see Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004 for a review). Although altruism should motivate consumers to post accurate information, other goals may lead reviewers to post biased information. If a reviewer is motivated to affiliate with others, she may selectively write about attributes that have already been discussed in prior reviews (Hamilton, Schlosser, & Chen, 2014). A reviewer seeking to build her online reputation may be more critical in order to signal her expertise (Schlosser 2005).

In this work, we expose a conflict between the motivation to provide credible information and motivation to provide important information in credence reviews. For reviews of experience services, these two motivations are aligned: experience attributes are the most important for evaluating experience services and (by definition) reviewers can easily evaluate them after purchase, making their reviews of these

attributes credible. As such, we expect reviews of experience service providers to contain a large amount of experience information. In other words, the motivations to provide important and credible information are aligned in online reviews of experience providers. For credence services, however, there is a tradeoff. Credence attribute information is the most important for evaluating credence services. However it is difficult for reviewers to evaluate credence attributes (by definition), and therefore information about credence attributes in online reviews would be less credible than other information. If reviewers want to provide the most important information, reviews will contain information about credence attributes whereas if they want to provide the most credible information, they will supply information about experience attributes.

Since there is clear motivation to provide experience information for experience providers and a tension between providing experience and credence information for credence providers, we predict the following: *H1: There will be a larger proportion of credence attribute information in reviews of credence providers than reviews of experience providers.*

Extant research suggests that the credibility of an online review depends on the structure and content of the information in the review. Cheung et al. (2009) found that the argument strength of online word of mouth communications significantly influences consumers' perceptions of the credibility of a review. Nelson (1970) proposed and Ford et al. (1990) demonstrated that messages with more objective information rather than subjective information would be more credible. Bringing these two streams of research together, we propose that online reviews can be analyzed as arguments, broken down by their argument structure (Kim & Benbasat, 2006; Racherla, Mandviwalla, & Connolly, 2012). We use the classic Toulmin (1958) argument structure to break down reviews into strands of argument about individual product attributes. Toulmin's framework includes six types of components: claim (the message's conclusion), grounds or data (data to support the claim), warrant (logical link between grounds and claim), backing (supports the warrant), rebuttal (reasonable restrictions on the claim), and qualifier (words that modulate the degree of certainty of the claim). A claim is an assertion, such as "this doctor was prompt," whereas data is information that would support the claim, such as "this doctor arrived for my appointment three minutes before the scheduled time."

Crucial to our work, consumers are likely to have data about experience attributes after purchase (e.g., the number of minutes before or after an appointment time), whereas they are unlikely to have data about credence attributes (e.g., whether the doctor possessed critical and up-to-date knowledge of a particular ailment). Therefore, when consumers mention experience attributes in a review, they may offer data about that attribute, but when they mention credence attributes, they are not likely to offer data. We derive the following hypothesis: *H2: Data about credence attributes will be offered less frequently than data about experience attributes.*

Consumer Perception of Reviews

As discussed above, economics of information theory posits that products contain a mix of search, experience, and credence attributes. The theoretical *service* type is defined by the most important *attributes* for the focal product. Thus, credence attributes are the most important qualities to evaluate for credence services, and experience attributes are the most important qualities to evaluate for experience services (cf. Darby & Karni, 1973; Lim & Chung, 2011). Theoretical predictions of consumer and seller behavior in markets for credence goods hinge on the assumption that consumers are aware that they lack important information about credence products. For example, diagnosis of auto repair services is a credence attribute (Dulleck & Kerschbamer, 2006). Credence service theory assumes that consumers are aware that they cannot evaluate the diagnosis, and furthermore they know that evaluating whether the diagnosis is correct is one of the most important facets of evaluating the service (Darby & Karni, 1973; Dulleck & Kerschbamer, 2006). Yet this crucial assumption about consumer knowledge of credence attributes and credence providers seems to be untested in the empirical literature. One lab experiment examining the behavior of sellers and consumers in a credence market leads to mixed results. In some cases consumers behave as theory predicts (i.e. consumers distrust credence sellers at a level in line with equilibrium results) whereas other consumers choose to do business with sellers even when the rational action would be to forego treatment (in other words, they pay too much or are over-treated; Dulleck, Kerschbamer, & Sutter, 2011b). Because this experiment only measures transaction outcomes and not consumer information processing, we cannot be sure whether consumers consider the relative importance

of credence attributes. We propose a test of consumers' perception of the relative importance of credence and experience attributes in credence versus experience reviews. ***H3a(b): Consumers will perceive credence (experience) attributes as more important for evaluating credence (experience) providers than for experience (credence) providers.***

Ford, Smith, and Swasy (1990) set out to directly test the consumer-side propositions generated by Nelson (1970) and Darby and Karni (1973). Predicting that consumers would be more skeptical of experience claims than search claims and more skeptical of credence claims than experience claims, Ford, Smith, and Swasy conducted experiments measuring consumers' skepticism of the different types of claims in an advertising context. Their results partially confirmed the economics of information propositions. They showed that consumers were indeed more skeptical of the validity of experience claims over search claims, but their results failed to reach significance for the test between experience and credence claims. The authors believe that this was due to a flaw in the design of their credence claim manipulation, which led to a confound in the study. Post-hoc, the authors realized that most of their credence claims were "nonperformance-related" whereas all experience claims were performance-related. My study uses the same types of claims across service provider types to avoid this confound (e.g., we test claims about knowledge for both experience and credence service providers). Aside from their study, we are not aware of any explicit tests of consumer skepticism of the credibility of credence claims. Testing the original propositions of consumer's perceptions of credence quality claims is one of a number of intended contributions for this research.

Credence attributes were originally defined as those which "cannot be evaluated in normal use" (Darby & Karni, 1973, p68). Following Darby and Karni's logic, we expect consumers to perceive credence claims as less credible than experience claims because the consumers making the claims do not have the knowledge or expertise required to make an accurate evaluation. Darby and Karni's proposition leads directly to our next hypothesis: ***H4: Reviews that contain claims about credence attributes will be perceived to be less credible than reviews that contain claims about experience attributes.***

In the economics of information literature borne of Darby and Karni's original paper the distinction between credence attributes and credence products is only implicit. Both theoretical and experimental work focus on products whose performance is completely opaque, rather than partially opaque, to the consumer (Dulleck et al., 2011b; Dulleck & Kerschbamer, 2006; Emons, 2000, 1997; Wolinsky, 1984, 1993). In their research, a key assumption is that consumers cannot assess the overall quality of a credence product, or at least that it is very difficult to do so in comparison to experience or search products, whereas in reality most products, even credence products, have search and experience attributes that consumers can assess. A subsequent assumption is that consumers are aware of their inability to adequately assess the overall quality of a credence product. From this assumption arises an untested proposition: consumers are aware that it is more difficult or impossible to assess the overall quality of credence products compared to experience products. This proposition is suggested by our reasoning for hypotheses 3 and 4. Together they suggest that information in a review of a credence service provider will contain important but not credible information, or credible but not important information, while reviews for experience providers can contain credible and important information. That being so, then consumer reviews would be perceived as less helpful for evaluating credence services than for evaluating experience services, controlling for the information content within. We hypothesize: ***H5: Consumers will perceive reviews for credence service providers as less helpful than reviews of experience service providers, controlling for the information within the review.***

We will test the hypotheses in three studies. Study 1 will examine the content of real online reviews of credence and experience providers as well as consumer perception of credibility and helpfulness of the reviews. Study 2 will test consumers' perceptions of experience and credence attributes and service provider reviews in a controlled experiment. In Study 3 we plan to further test consumers' perceptions of credence and experience attributes for credence providers, as well as their perceptions of reviews with data versus claims. In **Table 1** each hypothesis is listed along with the studies that will test it.

Hypotheses	Relevant Studies
H1: There will be a larger proportion of mentions of credence attributes in reviews of credence providers than reviews of experience providers.	Study 1
H2: Data about credence attributes will be offered less frequently than data about experience attributes.	Study 1
H3a: Consumers will perceive credence attributes as more important for evaluating credence providers than for experience providers.	Study 2
H3b: Consumers will perceive experience attributes as more important for evaluating experience providers than for credence providers.	Study 2
H4: Reviews that contain claims about credence attributes will be perceived to be less credible than reviews that contain claims about experience attributes.	Studies 1 and 2
H5: Consumers will perceive reviews for credence service providers as less helpful than reviews of experience service providers, controlling for the information within the review.	Studies 1 and 2

Table 1: Table of hypotheses and corresponding studies designed to test each.

Study 1

Content Analysis: We analyzed the content of 158 online reviews from Yelp.com. We randomly sampled ~30 reviews each from credence service providers (doctors and mechanics), and experience service providers (hair stylists, masseuses, and house cleaners) from the population of service providers in the Washington, DC metro area. We first selected service provider types from the literature (e.g. physicians are a classical credence service), and then randomly sampled from within the types. It is possible to introduce bias through selection of a specific provider type. We try to alleviate this concern by including multiple provider types (i.e. "replicates") in the analysis, to test whether it is one provider type or another that skews the results.

To determine the information content of the reviews, we divided each review into small snippets of text then systematically coded all the individual attribute mentions within each review as well as the argument structure of the mentions (i.e. whether the mention was in the form of data or a claim; Toulmin, 1958). A subset (40%) of the coding was done by two independent coders; PRL scores for reliability were 0.79 for structure and 0.83 for attribute type (Rust & Cooil, 1994); the remaining 60% were coded by Coder 1. The attribute types (e.g. knowledge, promptness) were derived from free-elicitation surveys of service provider attributes (Lehmann, 1979). Three judges used the definitions of search, experience, and credence from prior work (Darby & Karni, 1973; Ford, Smith, & Swasy, 1988) to independently categorize each attribute. Disagreements were resolved via discussion.

Initial Results: Our analysis (see Figure 1) shows that credence attributes were more likely to be mentioned in reviews for credence providers ($M_{credprov}$ = 2.70) than for experience providers ($M_{expprov}$ = 0.29; t=-9.92, p<0.001), supporting H1. The reverse was true for experience attributes ($M_{expprov}$ =5.28 vs. $M_{credprov}$ =4.28; t= 1.60, p=0.056). Experience attributes were more likely to appear in the form of data (57%) than credence attributes (33%; X^2 (1, 970) = 37. 49, p<0.001), supporting H2. Thus, both the content and structure of reviews differs across credence and experience services.

Next Steps: We recruited consumers using Amazon's mTurk platform (http://mturk.com) and asked each to read 12 of the 158 reviews, then respond to a survey measuring their perceptions of the usefulness and credibility of each review. We will analyze how the content and structure of these real reviews affects their perceived usefulness and credibility to test H4 and H5.
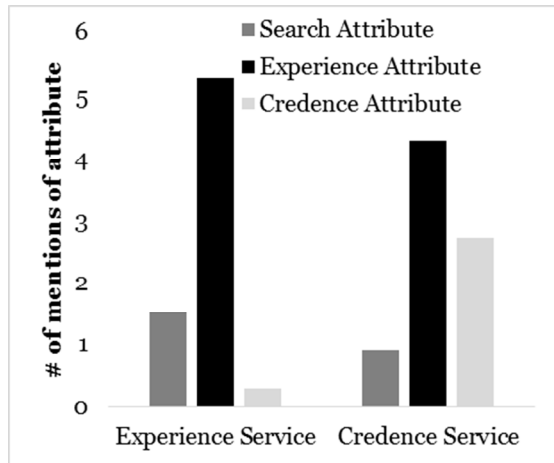
Figure 1: Number of attribute mentions in Experience and Credence reviews, by attribute type.
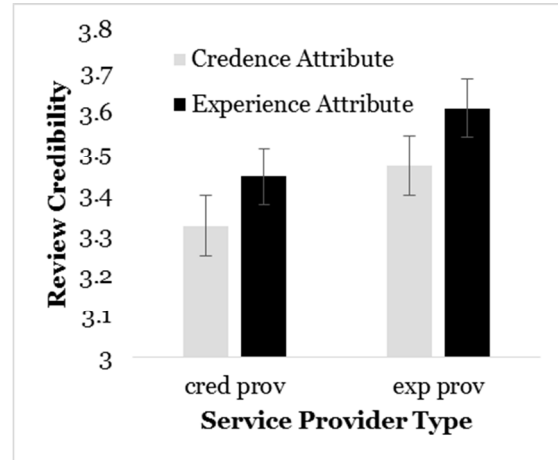

Figure 2: Perceived credibility of reviews by service and attribute type

Study 2

In Study 2, we examine consumer perceptions of reviews for credence and experience services.

Design: MTurk workers (N=356, ~60 per provider replicate) participated in the study in exchange for $0.50. Participants were asked to imagine a scenario in which they had recently moved to a new city, and needed to find a service provider. For each service provider (hair stylist, house painter, masseuse were experience providers and doctor, mechanic and financial advisor were credence providers), participants read an online review and then answered dependent measures. Each participant read and evaluated six different reviews. Reviews were randomly presented using a 2 (service provider type: credence or experience) x 3 (provider replicates) x 2 (attribute type: credence or experience) x 3 (attribute replicates) mixed experimental design in which service provider type and its replicates were manipulated between subjects and attribute type and its replicates were manipulated within subjects. Experience attributes included communication skill, personality, and ease of scheduling an appointment. Credence attributes included knowledge, trustworthiness, and intelligence. First, we asked participants to indicate how helpful and useful the review would be in selecting a service provider (Kempf & Smith, 1998). Participants rated the credibility of each review (dependable, honest, reliable, sincere, and trustworthy) and each reviewer (an expert, experienced, knowledgeable, and qualified; Ohanian, 1990). After evaluating each review, participants were asked to rate the importance of each type of attribute for the focal service provider, then answer control questions. Controls included overall propensity to trust (Mayer & Davis, 1999), familiarity with provider type (Hamilton & Thompson, 2007), online review use, and demographics. Scales were bi-polar with five levels unless otherwise indicated.

Results[1]: A 2x3x2x3 mixed model regression on review credibility showed significant main effects of service provider type (F(1, 2042)= 32.07, p<0.001) and attribute type (F(1,1927)=6.83, p=.009). Consumers found credence attribute claims to be less credible than experience attribute claims, and reviews of credence services less credible overall than reviews of experience services, supporting H4 (see Figure 2).

We analyzed the differences in importance of attributes to test whether credence attributes were perceived to be more important than experience attributes for credence providers (H3A) and whether experience attributes are more important for experience providers. We ran a similar 2 x 3 x 2 x 3 linear mixed model as above replacing importance as the dependent variable. We found main effects for service provider type

---

[1] We report the results of consumer perceptions of credibility. Results for perceptions of helpfulness are similar, and omitted for brevity. Analysis of replicates and covariates are also omitted for brevity.

(F(1, 19430=40.31, p<0.001) and attribute type (F(1,1939)=94.47, p<0.001). We also found a significant interaction between provider type and attribute type (F(1, 1934)=34.87, p<0.001). For credence providers, credence attributes are perceived as significantly more important than experience attributes (F(1,994)=88.05, p<0.001), supporting H3A. We expected to find that experience attributes would be perceived to be more important for experience providers than for credence providers (H3B). We did not find support for this hypothesis (F(1,1021)=.166, p=0.68). These results are shown in Figure 3.

Thus, as predicted by rational models, consumers seem to discount the credibility of credence attribute claims and credence service reviews.

	Review Credibility				Review Helpfulness			
Source	F	Sig.	Beta	Sig.	F	Sig.	Beta	Sig.
Intercept	185.52	.000	1.832793	0	4.32	.038	.420685	.056
Credence Attribute	6.83	.009	-0.233284	0.02	14.50	.000	.135499	.322
Credence Provider	32.07	.000	-0.195875	0.025	43.09	.000	.118902	.311

Table 2: Tests of fixed effects for covariates on review credibility and review helpfulness. Variables included in the model but omitted from table: replicate, gender (***), propensity to trust (***), familiarity with service provider (*), frequency of use of service provider, age, income (***), attribute importance (***), use of online reviews (***), negative review valence (***). Interactions between the main variables were not significant. *** p<0.01, ** p<0.05, * p<0.1
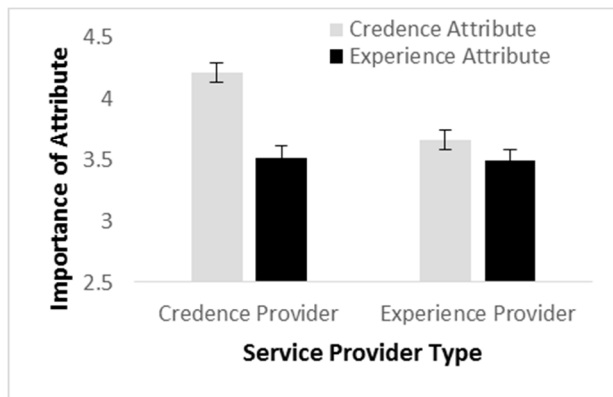

Figure 3: Interaction between attribute type and provider type on the importance of attributes

Discussion

In Study 1, we analyzed the content of online reviews for credence and experience providers to gain an understanding of how these reviews may systematically differ, and whether they are formed according to theory's predictions (e.g., that credence provider reviews have a higher proportion of credence attribute mentions than experience provider reviews). We found that when credence attributes are mentioned, they are more often mentioned as claims rather than data. Together, these findings demonstrate that consumers want to offer important information when writing reviews, although that information may be less credible than less important information. This finding may raise alarm for firms that offer credence services: consumers offer information which they may not be qualified to provide , as noted in the press with respect to physician reviews (ElBoghdady, 2012). In order to understand whether less-than-credible information affects consumer' perceptions of review credibility and helpfulness, we conducted a controlled experiment.

Study 2 confirmed our hypotheses that consumers would rationally discount the credibility of credence versus experience attribute claims. When a consumer sees a claim about service provider's knowledge, for example, that claim is perceived as less credible than a claim about how prompt the service provider is. Furthermore, consumers perceive that credence claims are more important for credence service providers. Thus we confirm that there is a tension between important and credible information for

reviews of credence service providers. Fortunately consumers perceive this tension, so the reviews are not necessarily harmful. However, it calls into question whether online reviews for credence providers can supply helpful information to consumers. This is an important finding for consumer welfare. In the press, service providers such as physicians have claimed that online reviews of credence are harmful to consumers, because they lack expertise and consumers may be too naïve to recognize this (ElBoghdady, 2012). In our experiment, consumers were not naïve. Rather, they were savvy to credence versus experience claims.

Limitations and Future Work

We acknowledge there are limitations to each of our studies. The simple design of our experimental study (Study 2) makes generalization difficult. We designed the reviews to be very short, with only one claim per review, and no data. Real reviews are usually longer with more than one attribute discussed within one review. To account for this limitation, we plan a third study to extend and replicate Study 2, increasing the complexity and length of reviews to increase external validity. Study 3 will be designed to test consumers' sensitivity to argument quality; whether information was presented as data, claims, or claims supported with data.

In addition to Study 3, we have collected data on consumer perceptions (through mTurk) of the real Yelp reviews from our Study 1 sample. We will analyze this data to determine whether it corroborates or refutes our findings from the controlled experiment in Study 2.

In Study 1, we use reviews only within one website (Yelp), and because we relied on human coders the sample size is limited. We attempted to use several tools to automate the text analysis, however they were not sophisticated enough to elucidate attribute type mentions within the reviews. In the future, when the technology becomes sophisticated enough, we hope to replicate this coding with a larger sample.

Conclusion

To the best of our knowledge, this research is the first to examine reviews of credence service providers in comparison to experience service reviews. While there are studies that focus on credence services alone (e.g. Gao, McCullough, Agarwal, & Jha, 2012; Lagu, Hannon, Rothberg, & Lindenauer, 2010), our unique sample of multiple provider types allows us to demonstrate how credence reviews are distinct from experience reviews. We found that credence service reviews are systematically different from experience service reviews in both information content and consumers' perceptions. Results imply the need for caution in generalizing findings from studies of one type of review to other types of reviews.

Given the opportunity for fraud in provision of credence services and the increasing prevalence of online reviews of credence providers, it is important to understand whether and how such reviews may be used by consumers to attempt to avoid fraudulent providers. We find that credence claims are made in online reviews despite their low credibility. However, we find evidence that consumers do rationally discount credence claims, as economic theory predicts. Credence service markets are characterized by severe information asymmetries; online reviews may supply useless information or they may be a valuable source of information to ease the consumers' information deficit. Our work will show whether and how online reviews in these markets differ from more traditional markets and how consumers perceive these differences.

Our findings will not only fill a gap in the literature on credence services and online reviews, but it will inform the design of product information platforms such as Yelp.com and Angie's List by providing guidelines for the types of review content and structure that may be most useful for different types of product categories.

References

Andrews, M. (2008, February). Rating Doctors : A Rank Practice ? *U.S. News and Wolrd Report*.

Balafoutas, L., Beck, A., Kerschbamer, R., & Sutter, M. (2011). *What drives taxi drivers? A field experiment on fraud in a market for credence goods.* Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1842097

Beck, A., Sutter, M., & Kerschbamer, R. (2010). for Expert Services : Theory and Experiment Guilt from Promise-Breaking and Trust in Markets for Expert Services : Theory and Experiment, (4827).

Bloom, P. N., & Pailin, J. E. (1995). Using information situations to guide marketing strategy. *Journal of Consumer Marketing*, *12*(2), 19–27. doi:10.1108/07363769510084876

Cheung, M. Y., Luo, C., Sia, C. L., & Chen, H. (2009). Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-line Consumer Recommendations. *International Journal of Electronic Commerce*, *13*(4), 9–38. doi:10.2753/JEC1086-4415130402

Chevalier, J., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, *41*(3), C01–354. doi:10.1509/jmkr.43.3.345

Darby, M. R., & Karni, E. (1973). Free competition and the optimal amount of fraud. *Journal of Law and Economics*, *16*(1), 67–88. Retrieved from http://www.jstor.org/stable/724826

Doctors Work Out Their Responses to Online Reviews - The Informed Reader - WSJ. (n.d.). Retrieved May 15, 2011, from http://blogs.wsj.com/informedreader/2007/07/30/doctors-work-out-their-responses-to-online-reviews/

Dulleck, U., & Kerschbamer, R. (2006). On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods. *Journal of Economic Literature*, *44*(1), 5–42. doi:10.1257/002205106776162717

Dulleck, U., Kerschbamer, R., & Sutter, M. (2011a). The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition. *American Economic Review*, *101*(2), 526–555. doi:10.1257/aer.101.2.526

Dulleck, U., Kerschbamer, R., & Sutter, M. (2011b). The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition. *The American Economic Review*, *101*(2), 526–555. Retrieved from http://www.ingentaconnect.com/content/aea/aer/2011/00000101/00000002/art00004

ElBoghdady, D. (2012, January 28). Some doctors try to squelch online reviews. *Washington Post*. Wshington, DC. Retrieved from http://www.washingtonpost.com/business/economy/some-doctors-try-to-squelch-online-reviews/2011/11/29/gIQA2KQhYQ_story.html

Emons. (2000). Credence Good Monopolists.

Emons, W. (1997). Credence Goods and Fraudulent Experts. *The RAND Journal of Economics*, *28*(1), 107. doi:10.2307/2555942

Ford, G., Smith, D. B., & Swasy, J. L. (1990). Consumer skepticism of advertising claims: Testing hypotheses from economics of information. *Journal of Consumer Research*, *16*(4), 433–441. Retrieved from http://www.jstor.org/stable/2489454

Ford, G., Smith, D., & Swasy, J. L. (1988). An Empirical Test of the Search , Experience and Credence Attributes Framework. *Advances in Consumer Research*, *15.* Retrieved from http://www.acrwebsite.org/volumes/display.asp?id=6817

Fox, S., & Jones, S. (2009). *The social life of health information. Pew Internet & Americal Life Project* (pp. 1–72). Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+Social+Life+of+Health+Information#1

Galetzka, M., Verhoeven, J. W. M., & Pruyn, A. T. H. (2006). Service validity and service reliability of search, experience and credence services: A scenario study. *International Journal of Service Industry Management*, *17*(3), 271–283. doi:10.1108/09564230610667113

Gao, G. G., McCullough, J., Agarwal, R., & Jha, A. (2012). Online physician ratings by patients: A changing landscape. *Jounral of Medical Internet Research*, *14*(1), e38.

Gruber, J., Kim, J., & Mayzlin, D. (1999). Physician fees and procedure intensity: the case of cesarean delivery. *Journal of Health Economics*, *18*(4), 473–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10539618

Hamilton, R. W., & Thompson, D. V. (2007). Is There a Substitute for Direct Experience ? Comparing Consumers ' Preferences after Direct and Indirect Product Experiences. *Direct*, *34*(December), 546–555.

Huang, P., Lurie, N. H., & Mitra, S. (2009). Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods. *Journal of Marketing*, *73*(2), 55–69. doi:10.1509/jmkg.73.2.55

Jain, S. (2010). Googling ourselves--what physicians can learn from online rating sites. *The New England Journal of Medicine*, *362*(1), 6–7. doi:10.1056/NEJMp0903473

Kempf, D. S., & Smith, R. E. (1998). Consumer Processing of Product Trial and the Influence of Prior Advertising: A Structural Modeling Approach. *Journal of Marketing Research*, *35*(3), 325. doi:10.2307/3152031

Kerschbamer, R., Dulleck, U., & Sutter, M. (2009). *The Impact of Distributional Preferences on The Impact of Distributional Preferences on (Experimental) Markets for Expert Services.*

Kerschbamer, R., Sutter, M., & Dulleck, U. (2009). *The Impact of Distributional Preferences on Experimental Markets for Expert Services*. IZA. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1526077

Kim, D., & Benbasat, I. (2006). The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores: Application of Toulmin's Model of Argumentation. Retrieved from http://pubsonline.informs.org/doi/abs/10.1287/isre.1060.0093

Lagu, T., Hannon, N. S., Rothberg, M. B., & Lindenauer, P. K. (2010). Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. *Journal of General Internal Medicine*, *25*(9), 942–6. doi:10.1007/s11606-010-1383-0

Lehmann, D. R. (1979). SEARCHING FOR IMPORTANT ATTRIBUTES AND APPROPRIATE LEVELS. *Advances in Consumer Research*.

Lim, B. C., & Chung, C. M. Y. (2011). The impact of word-of-mouth communication on attribute evaluation. *Journal of Business Research*, *64*(1), 18–23. doi:10.1016/j.jbusres.2009.09.014

Liu, T. (2011). CREDENCE GOODS MARKETS WITH CONSCIENTIOUS AND SELFISH EXPERTS. *International Economic Review*, *52*(1), 227–244. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1468-2354.2010.00626.x/full

Mayer, R. C., & Davis, J. H. (1999). The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of Applied Psychology*, *84*(1), 123–136. doi:10.1037//0021-9010.84.1.123

Mimra, W., Rasch, A., & Waibel, C. (2012). *Reputation in credence goods markets: Experimental evidence* (Vol. 49). Retrieved from http://www.mric.uni-muenchen.de/research/res_seminars_mric/seminar_mm1213/reputation_12-10-03.pdf

Mudambi, S. M., & Schuff, D. (2010). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185–200. Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=2898&context=misq

Nelson, P. (1970). Information and consumer behavior. *The Journal of Political Economy*, *78*(2), 311–329. Retrieved from http://www.jstor.org/stable/1830691

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising*, *19*(3), 39–52.

Park, D.-H., & Kim, S. (2008). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*, *7*(4), 399–410. doi:10.1016/j.elerap.2007.12.001

Racherla, P., Mandviwalla, M., & Connolly, D. (2012). Factors affecting consumers' trust in online product reviews. *Jounral of Consumer Behavior*, *11*, 94–104. doi:10.1002/cb

Rasch, A., & Waibel, C. (2012). What Drives Fraud in a Credence Goods Market?—Evidence from a Quasi Field Experiment, *49*(0). Retrieved from http://www.webmeets.com/files/papers/earie/2012/363/adac_2012_03_13.pdf

Rehavi, M. M., & Johnson, E. M. (2013). Physicians Treating Physicians: Information and Incentives in Childbirth. Retrieved from http://papers.ssrn.com/abstract=2295856

Rust, R., & Cooil, B. (1994). Reliability measures for qualitative data: Theory and implications. *Journal of Marketing Research*. Retrieved from http://www.jstor.org/stable/10.2307/3151942

Schneider, H. (2012). Agency problems and reputation in expert services: Evidence from auto repair. *Journal of Industrial Economics*, *LX*(3), 406–433. Retrieved from http://works.bepress.com/henry_schneider/3/

Srinivasan, S. S., & Till, B. D. (2002). Evaluation of search, experience and credence attributes: role of brand name and product trial. *Journal of Product & Brand Management*, *11*(7), 417–431. doi:10.1108/10610420210451616

Toulmin, S. (1958). *The uses of argument*. Cambridge, Eng: University Press.

Wolinsky, A. (1984). Product Differentiation with Imperfect Information. *The Review of Economic Studies*, *51*(1), 53. Retrieved from http://restud.oxfordjournals.org/content/51/1/53.short

Wolinsky, A. (1993). Competition in a Market for Informed Experts' Services. *The RAND Journal of Economics*, *24*(3), 380. doi:10.2307/2555964