Pricing for a credence good: an exploratory analysis

Matthew G. Nagler, Fredi Kronenberg, Edward J. Kennelly, and Bei Jiang

Abstract
Purpose – This paper explores the role of observable product characteristics and label wording in consumers’ valuations for credence goods, products for which key characteristics may not be fully evaluated even after purchase. The objective is to draw conclusions with relevance to pricing policy.
Design/methodology/approach – Hedonic price equations are estimated for a dietary supplement called black cohosh, taken by women for relief from menopausal symptoms.
Findings – Consumers respond in expected ways to label words that directly indicate product characteristics: for example, paying more for a product labeled as suitable for vegetarians. But surprising results occur for some nonspecific label words (e.g., “guaranteed” is associated with lower prices), suggesting that consumers view these words as indirect signals with respect to unobservable qualities. Additionally, consumers pay more for packages containing more units (e.g., tablets) even when the time supply of product is held constant; this outcome is consistent with the notion that sheer quantity reassures consumers about value and could indicate a reaction to uncertainty in the overall value proposition.
Research limitations/implications – The study relies on list prices in place of transacted prices, so consumers’ true valuations may not be reflected with complete accuracy. The study should be repeated in the future using scanner data on actual product transactions.
Practical implications – The indirect signals transmitted by label words and other observable attributes play a key role in consumers’ valuations for credence goods, and so are highly relevant to pricing strategy.
Originality/value – Previous studies have considered willingness-to-pay for label-indicated credence qualities, but have not looked at the role of indirect signals/indicators.

Keywords Pricing policy, Credence qualities, Information, Uncertainty, Actaea racemosa, Herbal medicine

Paper type Research paper
1. Introduction

Credence goods are products for which key characteristics may not be fully evaluated by the consumer even after purchase and consumption (Darby and Karni, 1973). They are distinguished from experience goods, which consumers may judge satisfactorily after the purchase; and search goods, which may be satisfactorily evaluated prior to purchase (Nelson, 1970). In assigning value to credence goods, consumers must accept a degree of uncertainty over the long-term – either taking the value proposition “on faith,” or else placing trust in someone else’s word that the product holds value, such as a professional expert or another consumer who gives a testimonial.

In situations of uncertainty, a range of observable signals may influence consumers’ perceptions of product quality (Rao and Monroe 1989, Kirmani and Rao 2000). Studies have shown that consumers rely on observable variables as surrogate indicators of various unobservable characteristics (Quester et al., 2000; West et al., 2002; Jo, 2005; Johnson and Folkes, 2007). But what role do observable characteristics play in product valuation, particularly in the highly uncertain situation posed by credence goods? For example, the appearance and feel of a sweater, as well as the phrases “Shetland,” “Made in Scotland,” and even “Dry Clean Only” on the sweater’s label, all reassure consumers that a sweater genuinely contains wool, something that the consumer cannot generally verify even after purchasing and wearing the sweater. How should indirect indicators such as these be accounted for in developing a pricing strategy?

This paper offers an exploratory study of the role of observable product characteristics and label wording in consumers’ valuations for credence goods. We posit that observables fall into two categories: those that consumers value intrinsically or as direct indicators of the qualities they name; and those that additionally provide value relating to the uncertainty of
credence goods, either as indirect signals of the presence of an unobservable quality, or as a source of general reassurance about the value of the product. We expect that valuations with respect to the second category of observable characteristics will exhibit both direct effects, incorporating the intrinsic valuation of these characteristics, and also indirect effects attributable to their indicator and assurance roles.

To examine evidence of the hypothesized effects, we estimate hedonic price equations for an over-the-counter dietary supplement called black cohosh, used by women for relief from menopausal symptoms. Among other explanatory variables, we consider the number of active ingredients in the product, whether the product is a retailer’s brand, whether the product has been rated online, and whether certain words of potential interest are used on the product label. In regard to this last set of variables, we follow Levitt and Syverson (2008), who examine the effect on house sale prices of the precise words used by agents to market houses through the Multiple Listing Service (MLS). Our results show consumers responding in expected ways to label words that directly indicate product characteristics. For example, consumers are willing to pay a premium for supplements that claim to be suitable for vegetarians or to contain a standardized set of ingredients. But surprising results enter in for certain nonspecific label words; for instance, the seemingly positive word “guaranteed” correlates with lower prices. These results suggest that consumers do indeed sometimes view label words as signaling indirectly the presence or absence of unobservable, valuation-relevant product qualities.

We also find that consumers are willing to pay more for packages that contain more units (i.e., tablets or capsules) even when the time supply of product is held constant, a result that runs counter to prior findings for prescription drugs (e.g., Suslow, 1996; Berndt et al., 2003) that consumers prefer achieving a given therapeutic effect via a lower dosing frequency. This result
appears to reflect a reassurance role that sheer quantity plays for the consumers of credence goods: that is, consumers may value quantity, all else being equal, because it reassures them about the value proposition of the purchase, something over which they may experience considerable doubt. Thus we find support for the notion that observable characteristics may manifest indirect effects related to general reassurance of the consumer.

The rest of the paper is structured as follows. Section 2 reviews the literature relating to credence goods and surrogate indicators. Section 3 outlines our conceptual framework and advances hypotheses. Section 4 provides background on the market for black cohosh and presents our empirical evaluation of the theory in the context of that market. Conclusions, inclusive of managerial implications, are discussed in Section 5.

2. Literature Review

The theoretical literature on credence goods centers on the phenomenon of expert services, particularly situations where diagnosis is bundled with treatment, such as car repairs and appendectomies (Darby and Karni, 1973; Dulleck and Kerschbamer, 2006). Expert services are classified under credence goods as the consumer may not be completely clear on how well the practitioner performed them even after they are delivered, given the intricacies involved. The literature focuses on the potential for fraud, market mechanisms that may tend to prevent it, and the methods and institutions that deal with the resultant problems when these mechanisms fail.

A separate category of credence goods, consumer products with valued putative qualities that the consumer is unable to confirm after purchase and use, form the basis for the limited empirical literature with relevance to pricing policy. These include, for example, products that claim to have a certain origin (e.g., “Swiss,” “bottled at the winery”) or to have been made
through a certain production process (e.g., “hand-made,” “free range,” “cold-filtered”). As with expert services, one must take a key valued characteristic on faith: one does not find out, even after acquiring the product, whether one’s watch is truly Swiss or one’s beer truly cold-filtered. 

The relevant studies have focused for the most part on consumers’ willingness to pay for label-indicated credence qualities in this category. For example, Lusk et al. (2003) consider willingness to pay for beef labeled as not having been from cattle administered growth hormones or fed genetically modified corn. Loureiro and Lotade (2005) investigate consumers’ willingness to pay a premium for coffee labeled as fair trade, shade grown, or organic. Other analyses in this vein have been conducted by Morgan et al. (1979), Roe et al. (2001), Larson (2003), and Huang and Lin (2007). None of these studies look at the valuation of observable characteristics or the role of these in relation to the valuation of credence qualities.

Meanwhile, a rich literature on surrogate indicators shows that consumers use a range of observable variables to infer levels of unobservable product qualities (Quester et al., 2000; West et al., 2002; Jo, 2005; Johnson and Folkes, 2007). Valuation and pricing studies for surrogate indicators are scarce, however, particularly as relates to credence goods. In one such study of electricity prices, Roe et al. (2001) find that consumers are willing to pay a premium for green electricity offerings that bear a Green-e certification. While Green-e certifications are identified with a specific set of product characteristics, the authors speculate that consumers likely value Green-e labeling as conferring “name brand” status, indicating that “the quality and veracity of environmental benefits promised by the product have been guaranteed by a certification firm that consumers trust” (Roe et al., 2001, p. 924). The authors also find that consumers pay a premium, relative to standard green electricity offerings, for offerings branded by Pacific Gas &
Electric and Greenmountain.com. The article does not attempt to draw conclusions of general relevance to the class of credence goods.

3. Framework

Our analysis is based on Lancaster’s (1966) commodity characteristics theory of consumer demand, which proposes that the price a consumer is willing to pay for a product reflects the value she places on the individual characteristics that make up the product. This theory has been applied through the method of hedonic analysis to estimate implicit valuations for the characteristics. By regressing the market price on product attributes, one may separate out the degree of influence of each attribute on the price (Rosen, 1974). Early demonstrations of the approach are provided by Griliches (1971) and Ladd and Suvannunt (1976).

To capture the credence good context, let us represent the consumer’s utility from a product as a function of J observable characteristics, $X_j, j=1,\ldots,J$, and K unobservable characteristics, $Z_k, k=1,\ldots,K$. Thus the utility the consumer obtains from consuming the product is given by

$$U = U(X_1,\ldots,X_J; Z_1,\ldots,Z_K)$$

(1)

Because the $Z_k$ are not observable, they do not enter the hedonic model, which expresses the price for each brand $i$ only as a function of the observable $X_j$, to wit,

$$\log P_i = \alpha + \beta_1 X_{1i} + \ldots + \beta_J X_{Ji} + \xi_i$$

(2)

where $X_{ji}$ is the value of the $j$th characteristic for brand $i$, and $\xi_i$ is a stochastic disturbance term.

It is reasonable to assume that among the observable characteristics there will be characteristics that consumers value intrinsically, or, in the case of label wording, as direct
indicators of the qualities they name; and characteristics that they value both intrinsically/directly and because they indicate indirectly some subset of the unobservable characteristics (for example, as “Made in Scotland” indicates that a sweater is real wool). We may represent this by partitioning the $X_j$ and rewriting (2) accordingly,

$$\log P_i = \alpha + \beta_1 X_{i1} + \ldots + \beta_L X_{iL} + \beta_{L+1} X_{L+1,j} + \ldots + \beta_J X_{j,j} + \xi_j$$ (3)

where component “A” consists of the $L<J$ characteristics that consumers value purely directly, while “B” consists of characteristics valued both directly and as indirect surrogate indicators. The former set will generally take expected signs in the hedonic estimation, as consumers “take them at face value.” The latter set may offer more surprising results. To the extent that they have value both in reflecting directly indicated characteristics and indirectly indicated characteristics, these variables have a “split personality.” That is, coefficient estimates are the sum of two effects: the effect on price of the observed (or nominal) characteristic taken at face value, and the effect on price of the unobserved characteristic for which the variable is a surrogate indicator. Viewed through another lens, the fact that the variable indicates an unobserved characteristic means that it is in effect correlated with an unincluded variable, and this introduces bias into the estimation of the coefficient for the “face-value” characteristic. But the bias is functional: in empirically observing it, we are able to detect the characteristic’s second role.

Component B may also include characteristics whose second role is to reassure the consumer about the value of the product in a less specific sense. As consumers cannot evaluate key characteristics of credence goods prior to purchase, they will continue to experience uncertainty about the value of these goods as they approach the decision point. This uncertainty is unpleasant. Research has shown that consumers attempt to reduce cognitive dissonance, the
psychological discomfort related to uncertain decisions, by seeking exposure to information that reinforces the decision (Engel, 1963; Sweeney et al., 2000). Such processes are engaged prior to decisions as well as after them (Russo et al., 1998). While surrogate indicators of the relevant unobserved characteristics would serve this purpose, so potentially could any observable characteristics that reassure the consumer about value, even if unrelated to the specific source of lingering uncertainty about value.

Among the “component B” variables that might serve as indicators of unobserved qualities or as sources of general reassurance are a full range of consumer-perceivable pieces of information – information provided on the product label, over the Internet, or by friends and other personal sources; as well as aspects of the product that may be viewed directly. In this context, the wording of a label conveys useful information about the sophistication of a product’s marketing strategy which might, in turn, tip off the presence of underlying quality. While any brand could use certain words, some brands do while others do not. This is because some brands (and not all) have made the investment in identifying which words work best. Marketing sophistication signals product quality in that marketing strategy development consists of a product-specific sunk investment that has value to the producer only to the extent that quality is ultimately delivered and repeat purchases generated (Nelson, 1974). This line of reasoning explains why “mere wording” on a label can be an effective signal to consumers, and why we might expect label words to have significant and meaningful effects in hedonic estimation.

In summary, we hypothesize as follows:

\[ H1. \] Observable characteristics that consumers value purely intrinsically or as direct indicators of qualities they name will generally take expected signs in hedonic regression. The coefficients will reflect each characteristic’s intrinsic value.
Observable characteristics that take an unexpected sign or show unexpected significance in hedonic regression may be playing a role as an indirect indicator of an unobservable characteristic or as a source of general reassurance about product value. The coefficients will reflect the sum of each characteristic’s intrinsic value and its surrogate indicator or reassurance value.

4. Empirical analysis

4.1 The market for black cohosh

We now operationalize our conceptual framework to analyze the statistical relationship of observable characteristics to prices for the dietary supplement black cohosh. A plant native to North America, black cohosh (Latin name *Actaea racemosa*, formerly *Cimicifuga racemosa*), has been used historically by Native Americans for a number of medicinal purposes. Over the past 50 years, it has gained popularity in Europe, and more recently North America, as an herbal supplement for treating menopausal hot flashes. The randomized clinical trials that have been conducted have produced mixed results, some showing benefit, and others not (Kronenberg and Fugh-Berman, 2002; Fabricant and Farnsworth, 2005; Borrelli and Ernst, 2008). Yet, despite the lack of certainty provided by the trial results, black cohosh remains very popular, ranked recently as one of the top 10 dietary supplements based on retail sales in the United States (Cavaliere et al., 2009). Continued strong interest in this supplement suggests that consumers generally believe it to be effective in treating menopause.

With respect to assessing the role of observable characteristics in the valuation of a credence good, black cohosh provides an ideal object for study. First, it exemplifies well the key characteristics of a credence good. As with other medications, the consumer may not with
complete certainty distinguish an authentic black cohosh product from an adulterated product or one with no black cohosh (placebo) even after using it. Sometimes symptoms disappear when one takes a placebo; other times, symptoms fail to disappear when one takes an authentic medicine. False label claims – such as that the product contains black cohosh when it does not – might go undetected. Moreover, information on observable characteristics of different black cohosh brands is readily available from a range of retail outlets and supplemental sources, as we discuss below. Black cohosh authenticity, while not observable by consumers, can readily be studied in a lab, opening the possibility for follow-on research relating observable characteristics with unobservable product quality. Finally, the market prominence of black cohosh means that results relating to this product are likely to have far-reaching significance.

Dietary supplements such as black cohosh are subject to regulation in the United States under the Dietary Supplement Health and Education Act (DSHEA). Following DSHEA’s main premise, the rules that cover labeling and good manufacturing practices for dietary supplements more closely resemble those that govern foods than the FDA’s rigorous drug regulations. Accordingly, supplements require no premarket clinical testing or approval. Moreover, since DSHEA was enacted in 1994, the number of products to which it has applied has grown from 4,000 to approximately 30,000, while FDA funding for supplement oversight has declined. Consequently, as noted by Wechsler (2007), the agency has recently faced severe constraints in its efforts to enforce its rules against supplement mislabeling and contamination. This problem may continue to grow as economic trends encourage consumers to substitute herbal remedies for prescription drugs. (See Tanner, 2009.)

Indeed, mislabeling of black cohosh has been recently documented. In a laboratory study of 11 products labeled as pure black cohosh, Jiang et al. (2006) detected adulteration in four.
Three of these products contained marker compounds for a lower-cost Asian *Actaea* species, but not those expected for black cohosh. The fourth contained both Asian *Actaea* and black cohosh. The medicinal uses of Asian *Actaea* differ from those of North American black cohosh, and, as noted in National Pharmacopoeia Committee (2005, p. 50), the health consequences of substitution are not known.³

How the potential for mislabeling affects consumers’ valuations of black cohosh products is an open empirical question. We are aware of no previous work estimating hedonic price equations for dietary supplements. Berndt *et al.* (1996), Suslow (1996), Cockburn and Anis (2001), and Berndt *et al.* (2003) have estimated such equations for prescription drugs. But pharmaceuticals differ from dietary supplements – and, indeed, most other consumer products – in three important ways. First, more stringent regulation and enforcement reduces the probability that pharmaceuticals fail to contain what their labels indicate. Second, pharmaceuticals are usually selected or recommended for the end user by a prescribing physician. This critically alters the structure of the decision process such that factors affecting valuations, hence prices, are likely to be quite different for products such as dietary supplements for which no prescriber is involved. Third, many consumers of pharmaceuticals have insurance plans that pay some portion of the price of the drug. Insurance co-payment may well have differential effects on consumers’ willingness to pay for different characteristics, causing the findings of hedonic analysis for prescription drugs to diverge from those for dietary supplements and most typical consumer products. In short, the earlier studies (of pharmaceuticals) are inadequate for the purpose of understanding the process by which consumers value the broader array of credence goods.
4.2 Data

Our exploratory hedonic study of black cohosh is based on the analysis of list (i.e., non-promotional) prices; in this, we follow Gandal (1994) and Roe et al. (2001). We visited 20 stores in New York City and the lower Hudson River Valley and collected retail list prices for pure black cohosh products, as well as multi-component products combining black cohosh with vitamins, minerals, and other active (usually botanical) ingredients. The 20 stores sampled represented a range of outlets from single-shop apothecaries to large national drug and vitamin store chains. Canvassing multiple stores allowed the collection of multiple price points per product, while also ensuring that a diversity of black cohosh brands would be represented. The resulting sample consisted of a total of 118 price observations, of which 54 observations represented 37 distinct pure black cohosh products, and 64 observations represented 44 distinct multi-component products. (A “distinct” product consists of a specific brand with a certain number of units per package, a certain formulation – e.g., tablet, liquid-filled capsule, etc. – and a distinct set of ingredients.)

The price data were supplemented by publicly observable product information. For each distinct product, we collected all the information appearing on the label. Using standard Internet search techniques (e.g., Google), we also found and collected online consumer ratings of the brands represented. All data were collected during the summer and fall of 2007.

In all, in addition to the retail price, the following variables were populated for each observation:

1. Ingredients, the number of commonly-perceived “active” ingredients listed on the label;
2. Veggie, a dummy variable indicating whether the product is suitable for vegetarians, based on the ingredients or an explicit label affirmation;

3. Standardized, a dummy variable indicating whether the label claims the product contains a standardized component or set of components, as determined by industry-recognized methods (a quality assurance measure);

4. Side Effects, a dummy variable indicating whether the label warns of side effects;

5. “Safe”, a dummy variable indicating whether the label contains the word “safe” or some derivative (e.g., “safely”);

6. “Effective”, a dummy variable indicating whether the label contains the word “effective” or some derivative;

7. “Guaranteed”, a dummy variable indicating whether the label contains the word “guarantee” or some derivative;

8. “Proven”, a dummy variable indicating whether the label contains the word “proven” or some derivative;

9. “Certified”, a dummy variable indicating whether the label contains the word “certified” or some derivative;

10. “Studied”, a dummy variable indicating whether the label contains the word “studied” or some derivative;

11. Rated, a dummy variable indicating whether an online consumer rating could be located for the particular brand of black cohosh;

12. Sum of Ratings, the total of the ratings located for the brand (with individual ratings normalized to a 1-point scale);
13. Units Per Day, the number of units per day in the recommended dosage (or the maximum number of units per day in those cases where a range was given);

14. Units Per Package, the number of units (e.g., tablets) per package;

15. Time Supply, the number of days of supplement supplied per package (calculated as variable #14 divided by variable #13).

16. NYC, a dummy variable indicating whether the store where the price was observed was located in New York City; and

17. Retailer Brand, a dummy variable indicating whether the product is a retailer’s brand (e.g., Whole Foods, Vitamin Shoppe).

The product’s formulation was also noted: tablet (including caplets), liquid-filled capsule (including gelcaps and softgels), or solid-filled capsule. Table I displays descriptive statistics for our data.

<TAKE IN TABLE I>

The variables above provide the basis for the regressors in our hedonic model. Variables #1 through #3 are hypothesized to indicate concrete product characteristics of intrinsic value to consumers, manifesting the sort of “face value” (i.e., direct) effects represented in (3) by component A. Consistent with H1, these variables are expected to enter with a positive sign; for example, a product with a greater number of active ingredients (variable #1) would presumably be more highly valued by consumers and would accordingly command a higher price.

While variables #4 and #6 constitute label words that may directly indicate product characteristics that consumers care about (side effects, safety, and effectiveness), we anticipate that they may also be used by consumers to indicate indirectly the presence or absence of other
unobserved qualities (e.g., adulteration). Thus these variables manifest the dual effects – direct and indirect – represented in (3) by component B. Whereas their intrinsic value to consumers suggests a negative coefficient for #4 and positive coefficients for #5 and #6, significant coefficients with the opposite signs would tend to support the notion of a dual role, consistent with $H2$.

Variables #7 through #10 constitute nonspecific label words that are likely unrelated to concrete product characteristics; these are hypothesized to manifest “component B” effects as indirect surrogate indicators of unobserved qualities that consumers care about. Significant coefficients on these variables of either sign would be consistent with $H2$.

Variables #11 and #12, which together represent consumers’ valuation of Internet-based ratings of black cohosh products, pose a more complex manifestation of “component B” effects. We anticipate that consumers look to ratings of black cohosh products as an indicator of quality, on the hope that other consumers’ prior experiences provide relevant information. In particular, consumers likely consider both the number of ratings a product received and the average level of rating as indirect indicators of product quality: a high number of ratings signals a product that has been market-tested and is likely more reliable, whereas a high average rating indicates that the product is on average well-liked. Thus we include Sum of Ratings, calculated as a straight sum of the normalized ratings observed for the product, which we expect to enter with a positive sign. Meanwhile, Rated is included as a control for whether the product has been rated at all. With the sum of ratings is held constant as a separately modeled component, it is unclear a priori what sort of signal the binary indicator of whether the product is rated or not would represent. A significant negative sign on Rated, for instance, could show that a product’s being rated, all else equal, conveys a negative, indirect indicator effect, consistent with $H2$. 
Variables #13 through #15 are incorporated in two alternative specifications that represent alternative theories on consumers’ valuation of quantities. In the first specification, Units Per Day and Units Per Package are included to measure consumers’ valuation of the quantity of product included in a package, controlling for the daily dosage level. Here, a positive coefficient is expected on Units Per Package, on the expectation that consumers value additional doses, consistent with $H1$. An alternate specification uses Time Supply in place of Units Per Package to ascertain whether consumers distinguish the sheer quantity of tablets or capsules from the quantity of therapeutic doses when valuing the products. We anticipate that consumers might value sheer quantity of itself: a large bottle with lots of pills may provide visual reassurance that one is obtaining substantial value for the money spent, even if the apparent value is illusory. Thus significant positive coefficients on Time Supply and Units Per Day in the second specification would suggest that quantity conveys indirect, reassurance value consistent with $H2$.

Finally, variables #16 and #17 represent geographic and brand effects. NYC is included because it is expected that retail prices might differ for stores in New York City versus the Hudson River Valley. Retailer Brand is expected to take a negative sign, reflecting the tendency of store private labels to command lower prices than brand-name products.

4.3 Results

Table II presents our regression results estimating the standard hedonic model in (2), using Units Per Package and including explanatory variables #1-#13, #16, and #17. The left panel displays regression runs using the full sample of products, while the right panel displays runs on the
subsample of pure black cohosh products. The basic regression is displayed in columns I and IV. Columns II and V display a variant of the basic regression with formulation fixed effects included. Columns III and VI add store fixed effects and exclude the variable NYC. Table III repeats the regressions from Table II, but substitutes Time Supply for Units Per Package.

The results are roughly consistent across the different specifications. The R-squared measures, ranging from 0.64 and 0.68 for runs on the full sample without fixed effects to near 0.93 for runs on the pure black cohosh subsample including all fixed effects, indicate a very good fit for our model. F-values range between 6.40 and 13.51, in all cases well above the 1% threshold for significance of the regression.

Variables #1 through #3, representing direct effects, all came in with the positive signs predicted. Consumers seem willing to pay a premium for supplements that claim to be suitable for vegetarians, or that claim to contain a standardized set of ingredients. They also seem to value additional active ingredients in the multi-component products. Beyond vitamins and minerals, the multi-component products in our sample contain perceived active ingredients ranging from botanical components such as valerian (a sleep aid) and soy isoflavones (believed by some to be estrogenic and, thus, helpful with menopausal symptoms) to, in the case of one product, glandular extracts (including bovine pituitary, ovary and uterus). As the data in Table I indicates, many multi-component products have more than just a few such ingredients.

Several of the proposed surrogate indicator variables were significant. Product prices are generally higher when the label mentions possible side effects, higher when it uses the word “safe” or some derivative of “safe,” and lower when it uses “guarantee” or some derivative.
Based on the coefficients from column IV in Table II, the mention of side effects on the label equates to a 20% price premium, the presence of “safe” a 51% premium, and the presence of “guarantee” a discount of 13%. The effects for Side Effects and “Guaranteed,” in particular, run contrary to the signs expected based on the connotations of the terms. Consonant with $H2$, these effects seem to show that consumers are attempting to infer unobservable product qualities based on signals provided by various sets of label words. While the mention of side effects and the use of “safe” appear to reveal a level of investment in marketing that signals superior product quality, the use of “guarantee” seems to suggest the opposite – a lack of marketing sophistication that reflects poorly on the underlying product.

The significant positive coefficients on Sum of Ratings reflect the positive influence on price we had predicted for both the number of ratings and their level, all else equal. However, the significant negative coefficients for Rated in our regressions indicate lower prices, $ceteris paribus$, for rated products relative to unrated products. Taking these outcomes together, consumers seem to view online-rated products as slightly more valuable than unrated products if they have many positive ratings; however, they are viewed as substantially less valuable than unrated products otherwise. The results seem to demonstrate the indirect indicator effect for Rated we had anticipated, consonant with $H2$. It may be that consumers suspect online ratings are staged if there are only a few of them, so that their presence indicates a disingenuous manufacturer whose product cannot be trusted. In fact, we cannot be certain that any of the products in the sample have a sufficient number of high ratings to convince consumers that the ratings reflect positive value overall. For 3 out of the 5 regressions in which both Rated and Sum of Ratings were significant, the estimated coefficients show a breakeven point greater than the maximum value of Sum of Ratings for our sample. Across all 5 regressions of relevance, the
lowest breakeven point for Sum of Ratings was approximately 14.8; Sum of Ratings takes a value greater than 14.8 for 18 out of the 40 observations of online-rated products in our sample. Figure 1 shows that the distribution of Sum of Ratings is strongly bimodal, with products either taking on a value of less than 4 (i.e., clearly below the breakeven point), or else more than 15 (potentially above it).

<TAKE IN FIGURE 1>

The results with respect to consumer preferences over quantity of product per package reveal both the direct and indirect effects we had anticipated. In the regressions presented in Table II, Units Per Package is shown to have a significant positive influence on price, while the Units Per Day coefficient is negative (and significant in 5 out of 6 runs). However, when Time Supply is substituted for Units Per Package, as in the regressions presented in Table III, both Time Supply and Units Per Day take on strongly significant positive coefficients. While the Table II results suggest that consumers recognize that value resides in the number of days of supplement provided by a product, Table III indicates that they also experience a significant preference for more tablets or capsules even when the number of days of supplement therapy is held constant.

The latter finding indicates that consumers experience “money illusion” with respect to black cohosh. That is, they derive value from the nominal quantity of black cohosh tablets or capsules. This, we suspect, has to do with the nature of credence goods. As discussed in the introduction, because consumers are unable to observe key quality-relevant characteristics for credence goods, they must accept some uncertainty about the value proposition. As consumers grapple with this uncertainty, it is possible that they obtain reassurance by assigning value to measures of nominal quantity – number of units, the size of the package, and so forth – which are
easily and directly observable prior to purchase. Their salience and verifiability makes nominal quantities in some sense more satisfying and comforting, when one is deciding whether one’s money is well-spent, than attributes of the credence good that are subject to being taken on faith. Thus, the result seems to represent the indirect, reassurance effect we anticipated, consistent with H2.

The result runs counter to prior findings regarding prescription drugs: Suslow (1996) and Berndt et al. (2003) found consumers willing to pay more for drugs with lower required dosing frequencies. A possible explanation for the difference is that preference for a greater number of pills per day is not necessarily inconsistent with preference for a lower dosing frequency: one could take more pills in a day, but take them all at once rather than at several points during the day. We have not controlled for dosing frequency in our regressions, and, in fact, it was not possible to do so: many of the products we tested specify a number of pills per day, but do not specify when the pills must be taken. Alternatively, if prescription drugs do not similarly propel consumers to money illusion, it may be because their value proposition is affirmed by a trusted prescriber. This vacates some of the need for reassurance that consumers must otherwise obtain from other sources. Moreover, the prescriber certifies the dosing of the prescription drug, providing confidence that time supply is the key quantity-related measure of value, rather than the sheer quantity of pills.

Finally, findings for the variable NYC indicate a significant difference in price between the two geographic areas incorporated in our study. Prices are lower for the same products when purchased in New York City, perhaps because the city has a greater concentration of stores selling botanical products then the suburban Hudson River Valley, implying greater competition. Meanwhile, Retailer Brand shows the significant negative price effect we had predicted for
private labels. Using the coefficients from column IV in Table II, products sell in New York City at an 18% discount, while retailer brands sell at a 48% discount relative to non-retailer brands.

5. Conclusions

We have used hedonic regression in the market for black cohosh dietary supplements to explore the basis for consumers’ valuations of a credence good. We proposed that consumers value observable characteristics relating to such goods in two ways: intrinsically, or, in the case of label words on product packaging, as direct indicators for unobservable qualities that they name specifically; or more indirectly, either as indicators with respect to unnamed unobserved qualities or as a source of general reassurance concerning product value. We hypothesized that characteristics valued solely in the first way would take expected signs in the hedonic regression, reflecting intrinsic value. Meanwhile, an unexpected sign or unexpected significance might identify characteristics that also carried indirect value of the consumer. Our findings with respect to black cohosh were in line with these hypotheses.

5.1 Managerial implications

To develop successful pricing strategies for credence goods, it is not enough to know consumers’ valuations of directly-indicated product attributes: our results indicate that managers must also understand the role of indirect signals. One must take account specifically of the role that label words and other surrogate indicators play in helping consumers infer unobserved attributes of interest. For example, it is not enough to know how much consumers value the “organic” label on a bag of coffee beans; one must also understand the cues transmitted by the other words on
the label, and by other product attributes consumers are able to observe, as regards the truth of the “organic” label.

The use of cues may lead to surprising results. Indicators are sometimes not what they appear: a label phrase that might appear positive for the product could send a negative signal, and vice versa. For example, a seeming positive phrase like “guaranteed” might connote sleaziness to consumers, indicating a lower quality product. On the other hand, apparently negative label messaging, such as a warning that a product could have side effects, may reassure the consumer that its manufacturer is trustworthy. Managers who gain a good understanding of the nature of consumers’ inferences might effectively use label words and other observable product attributes strategically to foster a perception of value. The product may then be priced accordingly.

But managers should also bear in mind that consumers might be sensitive to signal manipulation. Our results show, for example, that consumers reduce their valuations for black cohosh brands that are Internet-rated but that have only a few ratings, suggesting that they may suspect, and consequently discount, ratings manipulation. If consumers believe that product signals are being concocted or manipulated by the supplier, not only might this neutralize the effects of the manipulation, but it also could compromise consumer trust in the brand and its supplier. For this reason, managers should use caution in making strategic use of signals.

Another important managerial insight arising from this research is that, with respect to products with key unobservable qualities, reassuring the consumer is paramount. Our finding that consumers value nominal quantities of black cohosh tablets and capsules might have to do, as we have suggested, with the need to be reassured about the value proposition. The broad implication of this is that re-enforcing language on the product package and in advertising, the
use of trusted spokespersons and testimonials, and other marketing mix devices focused on reassurance, may be particularly helpful for credence goods. Moreover, highly visible, salient, and verifiable product attributes, because they are the most reassuring, might carry substantial weight in pricing for such goods. As a rule of thumb, the importance of an attribute to consumers’ valuations for a credence good may be thought of as the product of the intrinsic value of the attribute to the consumer multiplied by the degree of its verifiability and/or salience.

5.2 Limitations and future research

The results of this study are limited by our reliance on list prices in place of transacted prices. For us to know definitively that consumers have valued a product at a certain price, we would need to witness actual transactions at that price. Undoubtedly, a significant share of transactions at the stores covered by our study occurred at list prices, but we simply do not know what share, nor do we know the discounted prices that applied to the remaining transactions. In using list prices, we effectively assume that discounting has been applied equally across all SKUs in the study, a characteristic that our study shares with other hedonic studies that have used list prices (e.g., Gandal, 1994; Roe et al., 2001). One objective for future research would be to repeat the study using scanner data obtained over a period of time. This would enable correlation of dated transaction prices with contemporaneous consumer observable information from labels, the Internet, and other sources.

Our study is also limited by its small sample size, a consequence of our focus on a single narrowly-defined product. The number of total observations – just over 100 – is rather small for predicting the effects of 16 explanatory variables. This may account for the low statistical power inherent in our results and the consequent insignificance of some of our predictors. Studies of
additional consumer products and some broader categories could contribute a further test of our findings.

We must stress that, as an exploratory study, the purpose of the present work has been to gain preliminary insights, rather than draw firm conclusions, concerning the role of observable characteristics in the valuation and pricing of credence goods. While we have advanced hypotheses about these roles and found evidence consistent with our hypotheses, we cannot, based on the present study, rule out alternative explanations for some of our results. For example, the surprising signs on some of our label wording variables (e.g., “guaranteed”) might be rooted in something other than that they indicate the level of unobservable qualities of interest to consumers. Further research on credence goods incorporating a mix of methodologies is needed. One next step might be to measure and incorporate the value of consumer-unobservable characteristics in the analysis and so measure the possible surrogate indicator relationships directly. Additionally, survey evidence on what consumers say they are valuing would provide a broader perspective.
References


1 Products may obtain a Green-e certification if certain specific requirements are met with respect to the electricity supply sourcing; these include, among other things, that at least half of the electricity supply be generated from renewable sources, and that no specific contracts be designated with nuclear generation sources. See Roe et al. (2001, p. 924) for full details.

2 Recall, as discussed in the introduction, that consumers will tend to learn about credence qualities over time. Thus investments in quality for credence goods will eventually be rewarded, albeit to a lesser extent, and with a lag, relative to experience goods.


4 The limitations of using list prices instead of actual transacted prices, which we were unable to observe for this study, are discussed in Section 5.2.

5 Online retailers of dietary supplements, such as Amazon and drugstore.com, and dedicated review sites, such as Buzzillions.com, invite consumers to write product reviews. Typically any visitor to the site can write one. The review process is structured to allow the consumer both to offer a written comment and a numeric rating, usually on a 5-point scale. The comment and rating are then posted to the website for other visitors to view.

6 Variables #7-#10 may be thought of as constituting a special case of “component B” effects in which the direct indicator effects are zero.

7 Regressions for the pure black cohosh subsample exclude log of Ingredients as an explanatory variable. Pure black cohosh products, by definition, contain only one active ingredient.

8 “Money illusion” is an economic term that refers to the individual’s tendency to equate additional nominal quantities of money with additional spending power, e.g., ignoring or partially disregarding inflation.
**Table I  Descriptive Statistics**

<table>
<thead>
<tr>
<th></th>
<th>Full sample (N = 118)</th>
<th>Pure black cohosh subsample (N=54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>St. Dev.</td>
</tr>
<tr>
<td>Ingredients</td>
<td>10.093</td>
<td>17.535</td>
</tr>
<tr>
<td>Veggie</td>
<td>0.534</td>
<td>0.501</td>
</tr>
<tr>
<td>Standardized</td>
<td>0.797</td>
<td>0.404</td>
</tr>
<tr>
<td>Side Effects</td>
<td>0.195</td>
<td>0.398</td>
</tr>
<tr>
<td>&quot;Safe&quot;</td>
<td>0.169</td>
<td>0.377</td>
</tr>
<tr>
<td>&quot;Effective&quot;</td>
<td>0.186</td>
<td>0.391</td>
</tr>
<tr>
<td>&quot;Guaranteed&quot;</td>
<td>0.305</td>
<td>0.462</td>
</tr>
<tr>
<td>&quot;Proven&quot;</td>
<td>0.127</td>
<td>0.335</td>
</tr>
<tr>
<td>&quot;Certified&quot;</td>
<td>0.051</td>
<td>0.221</td>
</tr>
<tr>
<td>&quot;Studied&quot;</td>
<td>0.186</td>
<td>0.391</td>
</tr>
<tr>
<td>Rated</td>
<td>0.339</td>
<td>0.475</td>
</tr>
<tr>
<td>Sum of Ratings*</td>
<td>8.061</td>
<td>7.025</td>
</tr>
<tr>
<td>Units Per Day</td>
<td>2.475</td>
<td>1.431</td>
</tr>
<tr>
<td>Units Per Package</td>
<td>81.975</td>
<td>41.485</td>
</tr>
<tr>
<td>Time Supply</td>
<td>39.730</td>
<td>29.005</td>
</tr>
<tr>
<td>NYC</td>
<td>0.703</td>
<td>0.459</td>
</tr>
<tr>
<td>Retailer Brand</td>
<td>0.110</td>
<td>0.314</td>
</tr>
</tbody>
</table>

*Statistics for Sum of Ratings are for the subsample of online-rated products only - N=40 and N=24 for subsets of the full sample and pure black cohosh subsample, respectively.
Table II  Regression Results - using Units Per Package

Dependent variable: log of Price

<table>
<thead>
<tr>
<th></th>
<th>Full sample (N = 118)</th>
<th>Pure black cohosh subsample (N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Formulation fixed effects</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Store fixed effects</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>log of Ingredients</td>
<td>0.1266***</td>
<td>0.1525***</td>
</tr>
<tr>
<td>Veggie</td>
<td>0.1845**</td>
<td>0.2055***</td>
</tr>
<tr>
<td>Standardized</td>
<td>0.3278***</td>
<td>0.3549***</td>
</tr>
<tr>
<td>Side Effects</td>
<td>0.1573</td>
<td>0.1882**</td>
</tr>
<tr>
<td>&quot;Safe&quot;</td>
<td>0.3619***</td>
<td>0.3749***</td>
</tr>
<tr>
<td>&quot;Effective&quot;</td>
<td>0.0573</td>
<td>0.0492</td>
</tr>
<tr>
<td>&quot;Guaranteed&quot;</td>
<td>-0.1734***</td>
<td>-0.1751***</td>
</tr>
<tr>
<td>&quot;Proven&quot;</td>
<td>-0.1412</td>
<td>-0.1019</td>
</tr>
<tr>
<td>&quot;Certified&quot;</td>
<td>-0.0234</td>
<td>-0.0384</td>
</tr>
<tr>
<td>&quot;Studied&quot;</td>
<td>-0.0874</td>
<td>-0.0920</td>
</tr>
<tr>
<td>Rated</td>
<td>-0.4115***</td>
<td>-0.3876***</td>
</tr>
<tr>
<td>Sum of Ratings</td>
<td>0.0235**</td>
<td>0.0249**</td>
</tr>
<tr>
<td>Units Per Day</td>
<td>-0.0559**</td>
<td>-0.0551**</td>
</tr>
<tr>
<td>Units Per Package</td>
<td>0.0056***</td>
<td>0.0055***</td>
</tr>
<tr>
<td>NYC</td>
<td>-0.1732***</td>
<td>-0.1874***</td>
</tr>
<tr>
<td>Retailer Brand</td>
<td>-0.4792***</td>
<td>-0.4438***</td>
</tr>
<tr>
<td>Constant</td>
<td>2.0952***</td>
<td>1.9406***</td>
</tr>
<tr>
<td></td>
<td>(0.1175)</td>
<td>(0.1358)</td>
</tr>
<tr>
<td>F-value</td>
<td>13.51</td>
<td>12.88</td>
</tr>
<tr>
<td>R²</td>
<td>0.6815</td>
<td>0.7008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively.
<table>
<thead>
<tr>
<th></th>
<th>Full sample (N = 118)</th>
<th>Pure black cohosh subsample (N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Formulation fixed effects</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Store fixed effects</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>log of Ingredients</td>
<td>0.1193***</td>
<td>0.1451***</td>
</tr>
<tr>
<td>Veggie</td>
<td>0.1991**</td>
<td>0.2159***</td>
</tr>
<tr>
<td>Standardized</td>
<td>0.3687***</td>
<td>0.3950***</td>
</tr>
<tr>
<td>Side Effects</td>
<td>0.2115**</td>
<td>0.2386**</td>
</tr>
<tr>
<td>&quot;Safe&quot;</td>
<td>0.4015***</td>
<td>0.4105***</td>
</tr>
<tr>
<td>&quot;Effective&quot;</td>
<td>0.0221</td>
<td>0.0106</td>
</tr>
<tr>
<td>&quot;Guaranteed&quot;</td>
<td>-0.1710**</td>
<td>-0.1681***</td>
</tr>
<tr>
<td>&quot;Proven&quot;</td>
<td>-0.1918</td>
<td>-0.1535</td>
</tr>
<tr>
<td>&quot;Certified&quot;</td>
<td>0.0737</td>
<td>0.0595</td>
</tr>
<tr>
<td>&quot;Studied&quot;</td>
<td>-0.0787</td>
<td>-0.0810</td>
</tr>
<tr>
<td>Rated</td>
<td>-0.4915***</td>
<td>-0.4690***</td>
</tr>
<tr>
<td>Sum of Ratings</td>
<td>0.0302**</td>
<td>0.0317***</td>
</tr>
<tr>
<td>Units Per Day</td>
<td>0.0643***</td>
<td>0.0827***</td>
</tr>
<tr>
<td>Time Supply</td>
<td>0.0066***</td>
<td>0.0065***</td>
</tr>
<tr>
<td>NYC</td>
<td>-0.1842***</td>
<td>-0.1975***</td>
</tr>
<tr>
<td>Retailer Brand</td>
<td>-0.5393***</td>
<td>-0.5500***</td>
</tr>
<tr>
<td>Constant</td>
<td>1.9253***</td>
<td>1.7804***</td>
</tr>
<tr>
<td>F-value</td>
<td>11.39</td>
<td>10.85</td>
</tr>
<tr>
<td>R²</td>
<td>0.6435</td>
<td>0.6637</td>
</tr>
</tbody>
</table>

Notes: Standard errors are in parentheses. ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively.