Understanding the impact of missing multilevel attributes on choice

Alaa El-Gayar
Institute of Marketing, School of Business and Economics, Humboldt University Berlin
Daniel Guhl
HU Berlin

Cite as:
El-Gayar Alaa, Guhl Daniel (2021), Understanding the impact of missing multilevel attributes on choice. *Proceedings of the European Marketing Academy*, 50th, (94511)

Paper from the 50th Annual EMAC Conference, Madrid, May 25-28, 2021

Understanding the impact of missing multilevel attributes on choice

Abstract:
Numerous consumer decision contexts are characterized by a lack of transparency in terms of incomplete product information. In mystery product selling, firms intentionally withhold product information that remains hidden from consumers until after the transaction. Despite extensive research in this domain, with particular attention to price discrimination and revenue management, the impact on consumer behavior remains unclear. This study develops and compares several models to incorporate the impact of missing brand information in discrete choice models. In this regard, we draw on behavioral theories to construct a latent class model that directly accounts for inference-making strategies. Using choice-based conjoint, we elicit consumer preferences for a mystery product at the example of the market for jeans. The results indicate that consumers do not ignore and moderately evaluate missing brand information. Finally, we discuss practical implications for willingness-to-pay and profits.

Keywords: Choice Modeling, Mystery product, Inference

Track: Methods, Modelling & Marketing Analytics

1 Introduction

In recent years, mystery (also referred to as opaque, probabilistic or non-transparent) products have been adopted by many firms in practice, most prominently in the travel (e.g., *priceline.com* and *hotwire.com*) and retailing industries (e.g., *groupon.de* and *bananarepublic.gap.com*). Real-world examples of mystery product selling include combining existing products of one supplier or combining the products from different suppliers, where the supplier identity is masked, and consumers' receipt of their preferred brand is risky. The focus of this study lies in the latter form. As others have noted, this risky prospect might induce curiosity, fun, or an exciting experience (e.g., Japanese *fukubukuro* or lucky bags). On the other hand, behavioral theories suggest that risk is negatively associated with an alternative's attractiveness.

Mystery product selling has long been a topic of interest in strategic operations management. While most studies in this field employ theoretical models to analyze the decision problem from a supplier perspective, little effort has been made so far to examine the behavioral underpinnings of consumer demand. However, a salient feature of mystery products is that firms intentionally impose a situation that lacks transparency, and consumers are well aware that crucial product information is missing. As a result, consumers are prompt to form beliefs, i.e., make inferences about the missing information. This constitutes a unique decision context and raises the following research questions: How do consumers treat missing attribute information during choice in such a setting? Do consumers infer missing supplier information, and if so, how do these inferences affect their evaluation of alternatives? Though prior research provides valuable insights on individuals' inference-making in the context of missing attribute information, there have been few attempts for its inclusion in choice models (Wu, Swait, & Chen, 2019). To fill this gap, we propose a latent class model that incorporates inference-making strategies in a choice-based decision frame and further compare it to other models of differing behavioral assumptions. The paper is structured as follows: a short overview of related literature is presented in chapter 2. Next, chapter 3 outlines different modeling approaches and their underlying behavioral assumptions. In chapter 4, we present the results of an empirical application and briefly discuss implications for willingness-to-pay (WTP) and profits. We conclude with possible directions for future research.

2 Related Literature

The present work builds on two existing lines of research: In chapter 2.1 we show by means of selected examples in literature that strategic operations management focuses on the mechanisms and the market conditions under which mystery product selling may be profitable; however, it simplifies assumptions about consumer behavior. As illustrated in chapter 2.2, there is a rich stream of literature that focuses on consumer behavior under incomplete product information in general, but not necessarily in the context of mystery product selling. Moreover, there have been only few applications to choice data.

2.1 *Strategic operations management of mystery products*

Jiang (2007), and Fay and Xie (2008) propose a monopolistic setting where two horizontally differentiated products can be offered in a regular or mystery version. Their analysis suggests that in certain cases, consumer preference heterogeneity for brands allows firms to segment the market and price discriminate among customers by simultaneously offering both product types. Moreover, by charging a lower price for the mystery product, firms do not only serve existing customers who are indifferent towards regular products (Gönsch, 2020), but also attract new customers who otherwise would not have purchased the firm's product. This could lead to market expansion and profit increase. In contrast, Shapiro and Shi (2008) and Jerath, Netessine, and Veeraraghavan (2010) among others investigate a competitive setting where multiple firms need an intermediary to sell a mystery version of their products. The latter elaborate a two-period dynamic model where firms may offer cheaper direct last-minute sales or mystery products through an intermediary depending not only on consumer brand preferences but also on demand in the first selling period. Despite some sparse attempts to depart from simple assumptions about consumer demand (Huang & Yu, 2014), there has been little research on conceptualizing the decision problem from a consumer's perspective. In one exception, Xie, Anderson, and Verma (2017) conduct a choice-based conjoint (CBC) study about regular and mystery hotel rooms. However, they apply a simple multinomial logit framework and do not account for consumers making inferences about missing product information.

2.2 *Consumer product evaluation under incomplete information*

Extant literature has, in general, outlined several behavioral strategies to explain consumers' product evaluations under missing attribute information. For example, consumers may follow a concreteness principle and evaluate alternatives by utilizing only observed attribute levels and ignoring missing ones. Other studies suggest that consumers use their memory, prior knowledge, or experience to fill in missing attribute levels (Bradlow, Hu, & Ho, 2004). Finally, researchers have investigated how consumers form inferences about missing information based on situationally available information that they observe. Most notably, Wu et al. (2019) point out two general decision rules: In within-attribute (other-brand) information processing strategies, Meyer (1981), among others, argues that consumers impute a discounted average value based on all known possible manifestations of the missing attribute. By comparison, researchers as Johnson and Levin (1985) demonstrate a within-alternative (same-brand) information processing strategy and look at the interactions between several attributes. Though we cannot completely rule out the possibility of inference-making based on this decision rule, in particular, missing brand inferences based on price. We argue that in the context of mystery product selling if products are similar enough and consumers believe that there is a random chance of receiving any brand, their inferences should be independent of the manifestation of other attributes in the same or other alternatives. Therefore, our focus lies on within-attribute inference-formation.

3 Methodology

In the proposed frameworks, we consider for practical considerations a multi-alternative choice scenario with regular products and one mystery product with undisclosed brand information. Based on random utility theory, each individual consumer $i = 1, \ldots, I$ chooses an alternative $j = 0, \ldots, J$ that maximizes her utility given the evaluation of the observed attribute levels and possibly a mystery feature presented in a given choice scenario $t = 1, \ldots, T$:

$$u_{ijt} = brand_{ijt} \cdot \delta_i^\top + x_{ijt} \cdot \beta_i^\top + \alpha_i \cdot price_{ijt} + \varepsilon_{ijt} \tag{1}$$

$$u_{i0t} = \beta_i^0 + \varepsilon_{i0t} \tag{2}$$

where x_{ijt} is a row vector of non-brand and non-price attribute values and β_i is a row vector of corresponding individual-level taste parameters. δ_i and α_i are preference parameters that capture the impact of the observed brand and price on utility. The random error term ε_{ijt} consists of all unobserved factors affecting choice and is modeled as an i.i.d. extreme value distribution. Equation (2) for the no-buy and Equations (1) for the regular alternatives without a mystery feature are straight-forward. In Equations (3)-(5) we implement and later compare different behavioral assumptions to model the impact of missing (brand) information on the utility of one mystery alternative in the choice scenario:

$$u_{ijt} = x_{ijt} \cdot \beta_i^\top + \alpha_i \cdot price_{ijt} + \varepsilon_{ijt} \tag{3}$$

One modeling approach for the mystery alternative presented in Equation (3) (henceforth referred to as the Exclusion model) omits explicit representation of the mystery feature. This is motivated by the extant literature on exclusion as an information processing strategy, which suggests that consumers only use explicitly available information in making choices and neglect missing attribute information.

$$u_{ijt} = \zeta_i + x_{ijt} \cdot \beta_i^\top + \alpha_i \cdot price_{ijt} + \varepsilon_{ijt} \tag{4}$$

In comparison, Equation (4) departs from this assumption and incorporates the missing feature as a random intercept denoted by ζ_i (referred to as the Random model). Though this approach assumes that there may be a systematic and heterogeneous utility value associated with the mystery brand, it does not provide a behavioral explanation.

$$u_{ijt|l} = \delta_{i|l} + x_{ijt} \cdot \beta_i^\top + \alpha_i \cdot price_{ijt} + \varepsilon_{ijt} \tag{5}$$

In Equation (5), we develop a latent class model (referred to as the Inference model) with the underlying behavioral assumption that consumers do not ignore missing attribute information, but rather impute a distinct value $\delta_{i|l}$ for it, which is the lth element in the brand utility row vector δ_i. Consumers do not know from which brand the mystery product is coming but develop some heterogeneous beliefs based on the set of possible outcomes $l = 1, \ldots, L$. The number of inference strategies or possible brands behind the mystery alternative corresponds to the number of latent classes L. Once consumers form beliefs about the missing information, they make their choices as if there was no uncertainty and all information was available. To differentiate from the brand preference parameter in Equation (1), we represent the inferred value by using the subscript l that indicates the class membership. This modeling framework is built upon the work of Wu et al. (2019), where inference-formation and perception bias of a feature-based attribute is modeled under different competitive framing contexts. Since, in our case, the missing information is evident to consumers, we do not consider perception bias. Moreover, our inference model explicitly incorporates specific inference-making strategies for missing attributes with multiple (more than two) levels that are subject to the individual preference utilities (i.e., the brand order can differ across individuals).

As we do not directly observe consumers' inferences, we assume the probability of occurrence of all possible imputed values for the brand of the mystery alternative. Thus, the class probabilities ω_l for the latent classes are defined on the population level and are modeled as a standard multinomial logit format, where γ_l is a class-specific constant:

$$\omega_l = \exp(\gamma_l) / \sum_{l'=1}^{L} \exp(\gamma_{l'}) \tag{6}$$

It follows that the likelihood of observing a particular choice pattern is a weighted average of the conditional probabilities that an individual i belongs to a latent class l. Let $j_t^* = j_1^*, j_2^*, \ldots, j_T^*$ be the set of observed choices for individual i and $\theta_i = \{\delta_i, \beta_i, \alpha_i\}$ the individual-level parameters, the unconditional choice probability is given by:

$$L_i = \sum_{l=1}^{L} \omega_l \int_{-\infty}^{\infty} \prod_{t=1}^{T} P_{i j_t^* t | l}(\theta_i) f(\theta_i | \bar{\theta}, \Sigma) d\theta_i \tag{7}$$

Considering that analytical integration of Equation (7) is not possible, we apply maximum simulated likelihood by taking $R = 200$ Halton draws for each $\theta_i \sim MNV(\bar{\theta}, \Sigma)$. For identification purposes, we fix $\gamma_{l=1}$ at 0. Price enters linearly into the utility function, all categorical variables are effect-coded, ζ_i and the no-buy option are dummy variables. By exchanging specific model components, the other models are estimated accordingly. Note that we estimate γ in the inference model and the mean and standard deviation of ζ_i in the random model. Heterogeneous preferences across brands are necessary for differentiation between both models.

4 Empirical Study

4.1 Survey design

We employed a web-based survey using Sawtooth Software and a convenience sample of mostly students from a major German university. We divided our questionnaire into three parts. The first part included questions related to past jeans purchase behavior, denim brand awareness and purchase intention, as well as brand credibility for *Levi's*, *T. Hilfiger* and *Guess*. The second part consisted of CBC tasks, where each respondent faced a choice among a fixed number of 4 alternatives in a sequence of 14 choice scenarios. After a pretest with students, we selected denim jeans for our application, as the jeans' brand had a relatively higher attribute importance than other product categories. We created the conjoint stimuli according to a randomized design and described them by the three most important attributes obtained from pretesting (fit: *skinny, straight, boot-cut, wide*; color: *blue, black, light blue*; and price: *\$69, \$79, \$89, \$99, \$109*), and also included brand to construct a mystery product with a hidden supplier identity. For the brand levels, we pretested several international denim brands and picked three that have a high level of discrimination in terms of purchase intention. Other selected attribute levels correspond to major jeans characteristics found in the online stores of the chosen brands. In each choice scenario, respondents encounter two regular and a mystery purchase choice through an intermediary called *hotdeal* with a no-refund and non-transferable purchase policy. In addition, respondents were able to opt-out of the market by selecting the no-buy option. The *hotdeal* offer involves an equal probability of getting a jeans from any of the mentioned brands (Fay & Xie, 2008) and is otherwise specified by all other attributes. The reasoning behind this design is that if the average consumer is not aware of the firm's selling strategy, her best guess would be random chance (i.e., equal probability for each brand. We will test this assumption in our analysis). Our experimental design closely mimics real online choice situations, in which consumers face a trade-off among mystery and regular products. The mystery version used is similar to the real online shopping experience found on one of the biggest mystery product sellers *hotwire.com*. In our study, we informed respondents that they should treat each choice scenario in the CBC task independently of all others and that there is no sequence learning possible. In the final part of the questionnaire, respondents gave information about their socio-demographic and socioeconomic background.

4.2 Data description

A total of 963 respondents completed all survey questions. We eliminate respondents who exhibited straight-lining behavior, and the slowest and fastest 5% for the CBC part as well as for the complete survey. To later assess the predictive performance of our models, we create an estimation sample by randomly sampling 90% of respondents each with 12 randomly selected choice scenarios. The final sub-sample consists of 672 respondents used for analysis.

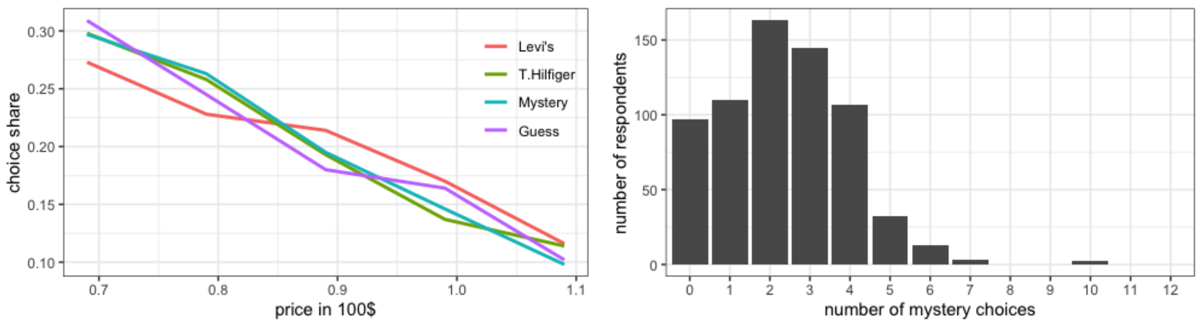

Figure 1: Average choice share per price (left) and Frequencies of mystery choices (right).

The mean completion time for the sub-sample is 3.3 for the CBC part and 12.8 minutes for the entire survey. The choice shares for the alternatives 25.6% (1), 24.6% (2), 24% (3) and 25.9% (no-buy) provide a sound basis for our estimation. The brand choice shares, conditional on their inclusion in a choice scenario, are 35% (*Levi's*), 27.6% (*T.Hilfiger*), 19.6% (mystery) and 19.3% (*Guess*), hinting at a particular brand preference order. This is likewise reflected by the purchase intention and brand credibility scales. We can descriptively observe that the average choice share for the mystery alternative depends on the price level, as it is also the case for the different brands, represented as a downward sloping demand curve in Figure 1 (left). Moreover, Figure 1 (right) illustrates that the fraction of mystery-choices varies between 0-83% on the individual level, and implies that this feature is not entirely unattractive to respondents (i.e., ~86% of respondents pick a mystery product at least once, with a mode of 2).

4.3 Estimation results

In total, we compared four models: Exclusion model (*EXC*), Random model (*RAN*), Inference model (*INF (2)*), and the Inference model (*INF (1)*) with fixed and equal prior beliefs; i.e., $\omega_l = 1/L$ or $\gamma_1 = \gamma_2 = \gamma_3 = 0$. All models in this comparison allow for preference heterogeneity. First, the reported results in Table 1 for the utility parameters are robust across all models and confirm the earlier expectation of the brand preference rank order. All being equal, the brand *Levi's* is preferred to *T.Hilfiger* and *Guess*, a skinny fit preferred to a straight fit, and a black jeans is preferred to a blue and light blue jeans. The price parameter measured in 100$ is negative for all models. Despite a systematic off-set, we do not observe an uncertainty effect for the mystery feature on average, as its evaluation is between the second and least preferred brand. Computing the class sizes for the *INF (2)* model yields an inference probability of 7% for *Levi's*, 23% for *T.Hilfiger*, and 70% for *Guess*. Second, we estimated all the models with homogeneous preferences and observed that incorporating preference heterogeneity enhances model fit by large. Moreover, the findings demonstrate that accounting for the mystery feature (by a behavioral explanation such as inference-making) further increases model performance, as evident by comparing the log-likelihood (LL) and Akaike information criterion (AIC) of the *EXC* model to the *RAN* and the two *INF* models. While we implemented a setting with equal assignment probabilities, we nevertheless observe surprising behavioral patterns in mystery product choice, which indicate that consumers form beliefs that deviate from the communicated probabilities. Consequently, when the class probabilities are freely estimated, the improvement in AIC suggests that the *INF (2)* model is the better model compared to all others. Turning to the in-sample hit rate and probability, we observe higher performance for the *INF* models, though the difference is relatively small and the *INF (1)* model with two fixed γ parameters moderately outperforms. To assess the predictive ability of our models, we create two holdout samples with our remaining observations in the original sample. First, we predict choice probabilities of the two left-out choice scenarios for each respondent in the sub-sample. To this end, we calculate the conditional individual-level estimates by allocating each respondent to the inference class with the highest ("posterior") class probability. Second, we do the prediction for the remaining 74 respondents in the original sample, each with 12 randomly selected choice scenarios by integrating over the population density. We observe for the out-of-sample hit rate and probability a slight edge for the *INF* models in terms of predicting new observations for the respondents in the estimation sub-sample. However, looking at the out-of-sample hit rate and probability for new respondents, we conclude that the models perform similarly. This is not likewise reflected by the LL after the integration over heterogeneity.

Table 1: Estimation results.

	EXC		*RAN*		*INF (1)*		*INF (2)*	
	mean	sd.	mean	sd.	mean	sd.	mean	sd.
Parameter								
No-buy	-3.95*	1.38*	-4.19*	1.32*	-4.22*	1.24*	-4.25*	1.25*
T. Hilfiger	0.10*	0.40*	0.09*	0.32*	0.09*	0.41*	0.09*	0.36*
Guess	-0.69*	0.77*	-0.66*	0.71*	-0.70*	0.76*	-0.69*	0.68*
Straight	0.58*	1.14*	0.59*	1.15*	0.61*	1.18*	0.61*	1.17*
Boot-cut	-0.56*	0.96*	-0.63*	1.09*	-0.61*	1.07*	-0.62*	1.07*
Wide	-0.96*	1.29*	-0.92*	1.26*	-0.96*	1.33*	-0.95*	1.32*
Black	0.18*	0.36*	0.19*	0.39*	0.19*	0.40*	0.19*	0.41*
Light Blue	-0.20*	0.05	-0.21*	0.12	-0.22*	0.17*	-0.21*	0.16*
Price	-4.72*	1.43*	-4.87*	1.43*	-5.06*	1.65*	-5.00*	1.53*
ζ			-0.48*	0.39*				
γ_2					fixed at 0		1.20*	
γ_3					fixed at 0		2.30*	
In-Sample Fit								
LL	-8327.95		-8264.54		-8277.29		**-8238.45**	
AIC	16691.89		16569.07		16590.59		**16516.90**	
Hit rate	0.75		0.75		**0.77**		0.76	
Hit probability	0.60		0.60		**0.62**		0.61	
Out-of-Sample Fit								
Hit rate[a]	0.60		0.60		0.60		**0.62**	
Hit probability[a]	0.50		0.51		**0.52**		**0.52**	
LL per obs.[a]	-1.194		**-1.187**		-1.195		-1.188	
Hit rate[b]	0.45		**0.46**		**0.46**		0.45	
Hit probability[b]	0.33		0.33		0.33		0.33	
LL per obs.[b]	-1.021		-1.018		-1.018		**-1.015**	

Note: $^*p < 0.05$; [a]New obs. same individuals; [b]New obs. new individuals

4.4 Managerial implications

Next, we investigate whether the differences between the *RAN* and *INF (2)* models translate into different managerial implications as measured by WTP. Therefore, we elicit hypothetical WTP indirectly through multiple sequential CBC questions (Miller, Hofstetter, Krohmer, & Zhang, 2011). We treat WTP as the reservation price or the specific price at which consumers are indifferent between buying and not buying a given product:

$$WTP_{rj} = \left(\beta_r^0 - v_{rj}\right)/\alpha_r \tag{8}$$

For each of the $R = 10000$ draws from the population distribution of the *RAN* and *INF (2)* models, we define β_r^0 as the utility of the no-buy option, and v_{rj} as the value of a product's non-price attributes (Miller et al., 2011). We compute the WTP average over draws for a blue mystery jeans with a wide fit, and obtain 73.3$ in case of the *RAN* and 65.5$ in case of the *INF (2)* model. This demonstrates that not accounting for inference-making in mystery product selling may result in an overestimation of WTP, consequently inefficient pricing decisions.

Finally, we simulate a specific market to explore whether the two models yield different profit implications, and whether offering products through a mystery channel in addition to a regular channel increases profits. Therefore, we compare two different scenarios: In a first step, we look at a market where each brand offers only regular products, specifically blue jeans with a skinny, straight, boot-cut and wide fit. In a second step, each brand in the market also introduces a blue mystery jeans with a wide fit through an intermediary. For simplification purposes, we assume prices $price_j$ and variable costs c^v_j for regular products are fixed and do not vary between scenarios. Prices start at 95$ (*Levi's*), 85$ (*T. Hilfiger*), and 75$ (*Guess*) for a wide blue jeans and increase for more desirable fits. Price elasticities in our simulation are -3 and -4; hence, reasonable markups for profit-maximizing firms imply costs between 54$ and 86$. The intermediary has variable costs equal to the brands' average variable costs for a wide blue jeans and sets the mystery product price to maximize her profit. To determine profits π_j, we normalize the market volume mv to a unit, set fixed costs c^f_j to zero and calculate the market share ms of each product j in the market by considering the choice probabilities $P_{rj}(\hat{\theta}_r)$ over draws from the population distribution:

$$\pi_j = mv \cdot (price_j - c^v_j) \cdot ms_j - c^f_j, \text{ where } ms_j = 1/R \sum_{r=1}^{R} P_{rj}(\hat{\theta}_r) \tag{9}$$

In addition to Equation (9), brands may either sell their products to the intermediary with a constant markup of 10% on costs, retain 90% of the intermediary's profit, or both. For the two models, profit-sharing would render brand profits higher than in the initial market, but only for mystery product prices higher than the optimal price. On the other hand, setting a markup would lead to a brand profit curve that converges but never quite reaches profit levels in the initial market due to substitution effects. As illustrated in Figure 2, the combination of both may yield higher brand profits at the optimal mystery product price of 93$ (for the two models). Moreover, the *RAN* model underestimates profits for *Levi's* (7.29$ vs 7.33$) and *T. Hilfiger* (6.32$ vs. 6.43$), and overestimates profits for *Guess* (5.97$ vs. 5.82$) compared to the *INF(2)* model in the initial market. Also profit increase is slightly underestimated by the *RAN*

Figure 2: Anticipated profits for brands before and after introducing a mystery product.

model for *Levi's* (0.4% vs 0.7%) and *T. Hilfiger* (0.6% vs. 0.7%), and overestimated for *Guess* (0.7% vs. 0.6%). This setup is one example showing that there may be differences in profit implications resulting from the two models and that selling mystery products may profit firms.

5 Conclusion and Directions for Future Research

In summary, our proposed *INF* model provides a structural framework to understand consumer choice behavior under incomplete product information. Running a CBC analysis, we find empirical evidence for inference, and demonstrate the model's superiority in fit. Drawing on the elicited preference structures, we compute WTP and market shares, and observe market expansion effects after introducing a mystery product. Depending on the mystery product price, we observe different magnitudes of the cannibalization effect. Possible extensions of our model may include capturing: consumer's tendency to discount the inferred value; higher variance of the logit error for mystery alternatives due to increased risk; or probabilistic reasoning in terms of computing the mystery feature utility as an expected utility value from the regular brands. Further applications on different product categories are worth exploring.

References

Bradlow, E. T., Hu, Y., & Ho, T.-H. (2004). A learning-based model for imputing missing levels in partial conjoint profiles. *Journal of Marketing Research*, *41*(4), 369–381.

Fay, S., & Xie, J. (2008). Probabilistic goods: A creative way of selling products and services. *Management Science*, *27*(4), 674–690.

Gönsch, J. (2020). How much to tell your customer? - a survey of three perspectives on selling strategies with incompletely specified products. *European Journal of Operational Research*, *280*(3), 793–817.

Huang, T., & Yu, Y. (2014). Sell probabilistic goods? a behavioral explanation for opaque selling. *Marketing Science*, *33*(5), 743–759.

Jerath, K., Netessine, S., & Veeraraghavan, S. (2010). Revenue management with strategic customers: Last-minute selling and opaque selling. *Management Science*, *56*(3), 430–448.

Jiang, Y. (2007). Price discrimination with opaque products. *Journal of Revenue and Pricing Management*, *6*(2), 118–134.

Johnson, R. D., & Levin, I. P. (1985). More than meets the eye: The effect of missing information on purchase evaluations. *Journal of Consumer Research*, *12*(2), 169–177.

Meyer, R. J. (1981). A model of multiattribute judgements under attribute uncertainty and informational constraint. *Journal of Marketing Research*, *18*(4), 428–441.

Miller, K. M., Hofstetter, R., Krohmer, H., & Zhang, Z. J. (2011). How should consumers' willingness to pay be measured: An empirical comparison of state of the art approaches. *Journal of Marketing Research*, *48*(1), 171–184.

Shapiro, D., & Shi, X. (2008). Market segmentation: The role of opaque travel agencies. *Journal of Economics and Management Strategy*, *17*(4), 803–837.

Wu, F., Swait, J., & Chen, Y. (2019). Feature-based attributes and the roles of consumers' perception bias and inference in choice. *International Journal of Research in Marketing*, *36*(2), 325–340.

Xie, X., Anderson, C. K., & Verma, R. (2017). Customer preferences and opaque intermediaries. *Cornell Hospitality Quarterly*, *58*(4), 342–353.