Value-based pricing in digital platforms

Christen, Tatjana; Hess, Manuel; Grichnik, Dietmar; Wincent, Joakim

Published in:
Journal of Business Research

DOI:
10.1016/j.jbusres.2022.07.042

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of Record

Link to publication

Citation for published version (APA):
Christen, T., Hess, M., Grichnik, D., & Wincent, J. (2022). Value-based pricing in digital platforms: A machine learning approach to signaling beyond core product attributes in cross-platform settings. *Journal of Business Research*, 82-92. https://doi.org/10.1016/j.jbusres.2022.07.042

Journal of Business Research 152 (2022) 82–92

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

Value-based pricing in digital platforms: A machine learning approach to signaling beyond core product attributes in cross-platform settings

Tatjana Christen [a], Manuel Hess [a,*], Dietmar Grichnik [a], Joakim Wincent [a,b]

[a] University of St.Gallen, Dufourstr. 40a, 9000 St.Gallen, Switzerland
[b] Hanken School of Economics, Arkadiankatu 22, 00100 Helsinki, Finland

ARTICLE INFO

Keywords:
Value-based pricing
Cross-platform
Signaling theory
Accommodation industry

ABSTRACT

Value-based pricing is known to be challenging, especially on online platforms, but is considered a superior pricing strategy. We investigate cross-platform pricing and other factors that influence perceived customer value in the context of the accommodation industry. This industry is characterized by powerful platforms (e.g., Booking.com) as well as small and medium-sized enterprises (SMEs) selling across platforms. We compare the importance of platform choice and seller history as underlying signals conveying value and thus defining pricing beyond core product attributes. Such actor-signaling-actions for value are neglected in previous research. We pay particular attention to how time-based price discrimination affects the importance of these non-core product signals. As cross-platform efforts increase the complexity of value-based pricing, we apply machine learning methods to model how SMEs can successfully predict pricing across platforms. We discuss our methodological and theoretical contributions to value-based pricing and signaling theory.

1. Introduction

Value-based pricing is known to be challenging for businesses but is recognized as a superior pricing strategy (Hinterhuber, 2008; Hinterhuber & Liozu, 2014; Ingenbleek et al., 2003). Setting prices based on value requires a strong focus on value creation for the customer and an understanding of which product attributes influence the total amount that a customer is willing to pay (Christopher & Gattorna, 2005; Forbis & Mehta, 1981; Yip, 2012). Prior research has highlighted the fact that pricing in the accommodation industry is complex due to the seasonality of demand and the inflexibility of the product supply, which is limited by the rooms available for a given booking time (Hung et al., 2010). Additional complexity arises as the nature of the product allows for time-based price discrimination—that is, setting different prices not just by seasons but for different days of the week—as well as dynamic pricing, that is, pricing based on the time difference between booking and the actual event (for example, last-minute discounts) (Abrate et al., 2012, 2022). Such platform and market complexity makes it challenging to set value-based prices and requires businesses to account for factors that go beyond-core product attributes.

Value-based pricing is challenging in general and especially on digital platforms. While prior studies have improved our understanding of pricing in the accommodation industry (Teubner et al., 2017; Wang & Nicolau, 2017), they do not inform value-based pricing theory on how to navigate pricing complexity in digital cross-platform settings (Abrate et al., 2022). Although accommodation platforms list different accommodation types to serve different customer segments, customers search across platforms and SMEs provide their offerings on several platforms (Stangl et al., 2016). Research has so far failed to reach an understanding of the patterns driving cross-platform pricing and the importance of platform proprietary factors that allow the prediction of optimal prices for seller profit (Abrate et al., 2022). Therefore, value-based price setting and optimization for profit remain challenging for businesses exposed to different platform characteristics, competitors' offerings, and time-based and dynamic price discrimination. In this study, we acknowledge this complexity when we investigated and developed an approach to predicting optimal pricing for sellers who list their offerings on several platforms. We began by identifying the importance of the channel brand (the platform's brand), seller history (the seller's experience, measured by the number of days spent operating the business), and timing. The interrelated effects of these three non-product factors on price have been insufficiently researched in prior literature, which

* Corresponding author.
E-mail addresses: tatjana.christen@unisg.ch (T. Christen), manuel.hess@unisg.ch (M. Hess), a.dietmar.grichnik@unisg.ch (D. Grichnik), joakim.vincent@unisg.ch (J. Wincent).

https://doi.org/10.1016/j.jbusres.2022.07.042
Received 9 August 2021; Received in revised form 14 July 2022; Accepted 20 July 2022
Available online 28 July 2022

focused merely on signaling value through product attributes (Cui et al., 2020; Magno et al., 2018; Reuer et al., 2012; Teubner et al., 2017; Wang & Nicolau, 2017). Our approach investigates signaling actions of actors and is summarized in the following research question: How important are channel brand and seller history (i.e., actors' signaling actions) in estimating value-based pricing under time-to-travel constraints and in a cross-platform setting?

To address this research question, we built on Spence's (1973) signaling theory to develop a conceptual framework of the impact of signals on price in the customer purchase decision. We identified the importance of both core and non-core product attributes in signaling value, where non-core product attributes refer to actors' signaling actions. We applied a machine learning approach to develop a predictive model that generated pricing insights for sellers in a cross-platform setting (Delen & Zolbanin, 2018). Our analysis used a sample of unique data from accommodation apartments in the U.K. and Austria that were advertised on Airbnb and Booking.com over a period of two years. By using SMEs' historical transactions in our sample, we mirrored the prevalent online platform conditions of limited access to data for sellers. We evaluated the predictive performance of a random forest model versus a linear regression model based on predictive accuracy—a measure that is commonly used in prediction models (Shmueli & Koppius, 2011)—to show the importance of different value-conveying signals on prices.

Our study makes several contributions to the literature. First, prior research has identified two major obstacles in value-based pricing: value quantification (identifying the value of a product for the customer, particularly the value drivers and customer or market segmentation (Liozu, 2017)) and value communication—helping the customer to understand the value of the product (Hinterhuber, 2008). The majority of papers targeting value quantification and communication have focused on core product attributes (Codini et al., 2012). Non-product-related core attributes, such as actions taken to signal value to customers, have not been integrated into prior research on value quantification. We contribute to the theory on value-based pricing (Kortge & Okonkwo, 1993; Raja et al., 2020) by integrating non-core product attributes, such as channel brand and seller histories on platforms. These attributes function as distinct signals to customers and are highly relevant to price prediction in cross-platform offerings (e.g., Gibbs et al., 2018; Magno et al., 2018; Teubner et al., 2017; Wang & Nicolau, 2017).

Second, prior research on value quantification has, to a large extent, relied on conjoint analysis (Backhaus et al., 2010; Codini et al., 2012; J. S. Kim, 2018). Although an established method in the marketing literature, conjoint analysis does not allow for integration of the higher complexity inherent in value-based pricing (Liozu, 2017). We contribute to value-based pricing theory by applying a more suitable machine-learning methodology for price prediction, which identifies the relative importance of core and non-core product attributes that signal value (Delen & Zolbanin, 2018; Shmueli & Koppius, 2011). Our model therefore better identifies the importance of certain product attribute factors in predicting the prices customers will pay in cross-platform selling.

Third, signaling (Bergh et al., 2014; Spence, 1973), as an underlying theory describing how core and non-core product attributes convey value signals to customers, has been used in digital settings to explain how sellers build buyers' trust and confidence in the quality of their products (e.g., Jean & Kim, 2021; Mavlanova, Benbunan-Fich, & Lang, 2016). Most previous research has focused on the frequency of signals, the growth or decay of signals over time, and the efficacy of signals on signal receivers (Connelly et al., 2011). However, the scholarship in the field has not considered the role of time from a receiver's (e.g., customer's) perspective. That is, although time-based price discrimination is common in the accommodation industry, especially with decreasing time-to-travel periods, we lack knowledge about the importance of time as a discriminating factor on core and non-core product attribute signals for value-based pricing. Our results show that the time factor is important for evaluating the efficacy of signals because the time pressure on the customer's purchase decision increases the importance of non-core product attributes in value-based pricing. Such findings have broader relevance for signaling theory, as they describe signaling actions of actors (the non-core product attributes) as value glue to create receiver interactions in a digital global context (Kromidha & Robson, 2021).

Finally, we believe our research has strong practical implications, because it shows how valuable insights can be gleaned from machine learning models based on limited historic data (such as the data available for many SMEs) and how they can be used to inform price decision-making.

2. Literature review and research background

2.1. Perspectives on pricing strategies

Prior literature indicates that pricing on digital online platforms is a complex topic, especially for new products. The application of a sophisticated pricing strategy requires core knowledge relating to multiple product dimensions, including product attributes, market demand, competitors' prices and, not least, the customer (Ingenbleek et al., 2003). Customer-oriented pricing, also referred to as value-based pricing or value-informed pricing (Ingenbleek et al., 2003, 2010; Raja et al., 2020; Simon, 1996), is often recognized as superior to cost- and competition-based pricing but is not employed as frequently (Hinterhuber, 2008; Hinterhuber & Liozu, 2014; Ingenbleek et al., 2003). Setting pricing based on a company's cost structure helps the firm to understand the price floor but does not reflect how much a potential customer is willing to pay (Ingenbleek et al., 2003). Therefore, cost-based pricing can result in prices that are too high and do not generate the optimal level of demand or prices that are too low and do not reflect the value perceived by the customer. In contrast, when pricing is based on market or competition levels, a generally accepted price level can be reached but can entail a loss in profit margin when the market price does not reflect individual customers' willingness to pay. Thus, potential positive effects of product advantage could be diminished by competition-based pricing practices (Ingenbleek et al., 2003, 2013; Pohland & Kesgin, 2018).

Value-based pricing strategies promise to resolve such tensions and inefficiencies but require sellers to consider a larger set of factors that convey value to customers than cost- or competition-based pricing (Ingenbleek et al., 2003). Following previous research on the advantages of value-based pricing, we link customer-perceived value and signaling theory in our framework and methodology, using the example of the accommodation industry.

2.2. The accommodation industry: Booking platform specifics and pricing factors

The accommodation industry is a setting in which SMEs compete fiercely on digital online platforms that host players such as Airbnb, Homelike, or HomeAway. Scholars have shown that price is of major relevance for long-term success in this kind of business (e.g., Hung et al., 2010). It is a major factor in product selection by the customer and can thus be considered a key differentiator on these platforms (e.g., Gibbs et al., 2018; Lockyer, 2005). Furthermore, the price paid per night directly influences the profits of the accommodation provider (Gibbs et al., 2018). This is due to the nature of the product, which has high fixed costs—such as utilities and rent. Each price increase beyond the fixed costs provides a relative increase in profit after deducting variable costs. Thus, price is of particular interest in the accommodation industry, especially for new businesses that lack the historical data or experience to track and analyze market insights that would allow them to leverage revenue or profit optimization mechanisms.

Researchers have highlighted the fact that pricing in the accommodation industry is complex due to the seasonality of demand and inflexibility of the product supply (Hung et al., 2010) and the possibility for time-based price discrimination and dynamic pricing (e.g. offering early-bird or last-minute discounts) (Abrate et al., 2012). Cross-platform selling increases the complexity for sellers because platforms have different characteristics and convey different branding signals towards customers. Often, such signals and platform characteristics (e.g., usability) affect the perceived value of products by customers.

2.3. The influence of product attribute signals on pricing

Signaling theory is essential to assessing pricing on online platforms. It explains how signals increase the perceived value of a product and thus affect the customer's purchasing behavior. Prior research has shown that signaling theory explains how individuals overcome the challenges of incomplete or asymmetric information in a selection process by focusing on and evaluating certain signals (Bergh et al., 2014; Spence, 1973). Signaling theory is already commonly used in entrepreneurship literature—for example, in identifying the effect of signals of firm value in venture capital funding decisions (Busenitz et al., 2005) or initial public offerings (IPOs) (e.g., Certo et al., 2001; Reuer et al., 2012). Likewise, signaling theory has been applied to small firms' digital presence in digital networks and its efficacy for internationalization (Kromidha & Robson, 2021).

Inherent in signaling theory is the concept of establishing trust between senders and receivers so that customers (i.e., receivers) decide to enter into the transaction—that is, make purchasing decisions (Gupta et al., 2009; Kirmani & Rao, 2000). Using and building signals to create trust is particularly relevant in an online context where face-to-face interactions between sellers and customers are usually absent (D. J. Kim et al., 2008; Kromidha & Robson, 2021; Li et al., 2014; Pappas, 2016), the seller tends to be unknown to the customer and information asymmetry is high (e.g., Jean & Kim, 2021; Mavlanova, Benbunan-Fich, & Lang, 2016). Such situations are frequent in accommodation bookings because many guests travel across countries and stay in accommodations that are new to them.

Focusing on the identified research gaps in the value-based pricing literature relating to value quantification and value communication (Hinterhuber, (2008), signaling theory makes it possible to theorize about the signals inherent in core and non-core related product attributes that are received by buyers and explains value drivers affecting different customers or market segments (Liozu, 2017). The majority of papers that have targeted value quantification identified product attributes as the major determinant for value-based pricing strategies but considered only core product attributes (e.g., Codini et al., 2012; J. S. Kim, 2018). They do not consider non-core product attributes such as the seller and his platform history as potential sources of signaling value nor differences in platform functioning as customer value attributes (Lee et al., 2005; Magno et al., 2018). As signaling theory would suggest, website signals are particularly important for online purchase situations with high levels of information asymmetry, such as when the seller is unknown to the customer (Mavlanova et al., 2016).

Previous literature has highlighted the challenge of trust on online platforms, which is due to information asymmetry from the customer's perspective (Lee et al., 2005). We apply signaling theory to explain how value signals can mitigate such problems in the online sales process. For example, scholars who have investigated the importance of reputation have found that it can be used to achieve price premiums as a direct signal to customers or an indirect signal via prominent partners (Henard & Dacin, 2010; Reuer et al., 2012). A study on Airbnb listings by Wang and Nicolau (2017) suggested that seller attributes are relevant price determinants on Airbnb and that they function as value signals for the customer. For example, the seller's number of listings positively influences prices. Similarly, other signals relating to the seller's business experience could be perceived—for example, professionalism and competency—and thus function as quality signals for the customer; the influence of perceived competency or abilities on trust has been found to be positive in prior research (e.g., Mayer et al., 1995). For instance, in a study on Airbnb listing prices, Magno et al. (2018) identified a positive relationship between the seller's business experience, measured by duration of membership on the platform, and the listing price. The authors explained this finding with reference to the more sophisticated pricing methods of more experienced hosts. However, given that booking platform contingencies limit sellers' abilities to employ sophisticated pricing methods, we argue that the seller's business experience is instead a seller history signal that can be observed by customers across platforms in several ways—for instance, by checking the duration of the seller's membership (Magno et al., 2018; Teubner et al., 2017; Wang & Nicolau, 2017). Therefore, we hypothesize the following:

Hypothesis 1. *Beyond core product attributes, seller history is an important factor in predicting price on digital online booking platforms.*

In a cross-platform selling environment, the channel brand (i.e., the platform) is a source of value that influences customer purchasing decisions, alongside inherent product characteristics. The relevance of platform choices is suggested by prior research that has found a positive relationship between a provider's reputation and customer's trust and also between provider reputation and customer price (D. J. Kim et al., 2008; Reuer et al., 2012; Teubner et al., 2017). The platform itself may, thus, function as an additional value signal for the customer in its role as the facilitator of the transaction and the customer's primary service address (Lee et al., 2005). Likewise, platforms invest in branding and value-signaling elements to ensure their legitimacy and credibility in the booking and payment transaction process, creating trust and thus increasing the customer's perception of their quality and value. Therefore, channel brand may increase the customer's perception of the value of a product and thus can be expected to be a relevant factor in predicting prices (Gregg & Walczak, 2010; Mayer et al., 1995; Reuer et al., 2012). Therefore, we hypothesize the following:

Hypothesis 2. *Beyond core product attributes, channel brand is an important factor in predicting price on digital online booking platforms.*

As signaling theory suggests that customers observe value-conveying signals before product purchases (e.g., Jean & Kim, 2021), we explore an additional signal, introducing the factor of time from the customer's perspective. That is, we consider the fact that customers may enter into transactions with different time-to-departure windows. We suggest that when the customer is under time pressure to make a decision, the time spent comparing offerings is likely to be reduced and the customer will be more open to receiving value signals that support their decision-making. Previous research has found that time constraints influence decision-making—for example, by accelerating the processing or filtering of information used for decision-making (Ben Zur & Breznitz, 1981; Cui et al., 2020; M. I. Hwang, 1994). Accordingly, we expect that purchases made under time pressure could make the customer more susceptible to easily observable quality signals that convey the value of certainty (instead of uncertainty). At shorter time-to-departure windows, seller history may convey signals of certainty by conveying a perception of credibility and reliability. Likewise, the platform brand may ensure certainty and credibility for the transaction process when customers have to decide quickly. Three main stages of timing from the customer's perspective are evident and correspond to platform-inherent selection modes for certain discount settings (i.e., platform contingencies): early-bird bookings with long time-to-departure windows, standard bookings, and last-minute bookings with the shortest time-to-departure windows. Based on signaling theory, we argue that customers pay more attention to signals conveying the value of certainty in shorter time-to-departure situations and hypothesize the following:

Hypothesis 3. *Beyond core product attributes, channel brand and seller history are more important predictors for customers booking with little time to departure than for those with a moderate or long amount of time before* *departure*.

3. Data and methods

3.1. Data collection

To test the hypotheses, we chose the accommodation industry where digital online platforms and the use of multiple channels are prevalent. One of the major challenges, which is also an opportunity, for SMEs selling on accommodation platforms is to understand and learn how to optimally set prices—by applying dynamic pricing, for instance (Abrate et al., 2022). We had unique access to data from SMEs selling their offerings on the booking platforms Airbnb and Booking.com through a research collaboration. SMEs have a variety of motivations to provide their offerings across platforms. Leveraging different channels, they aim to reach a broad customer group to reduce vacancies and optimize profits. The research setting and direct access to the SMEs and their actual profit and loss statements based on proprietary platform transaction data enabled us to address two research challenges of data collection identified by Abrate et al. (2022): first, including the effect of multichannel pricing and second, including pricing effects from "internal" platform data (e.g., discounts which are given during the booking process or extra fees for additional guests), which is otherwise inaccessible. The SMEs whose data we had access to offer apartments in Vienna, Austria, and Cardiff, Basingstoke, and Sheffield, United Kingdom. The data contained actual booking transactions and profit and loss accounts for each booking between January 2018 and December 2019, amounting to a total of 1,972 observations. An observation was made at the time a customer booked a certain apartment. The information gathered included, among other details, the travel period (check-in and check-out date), group size (number of guests staying), and total price paid by the customer for the booking.

3.2. Variables

Dependent variable. Based on prior research on value-based pricing and a customer's willingness to pay, we used observations of actual purchases by customers and defined the average price per night of a booking that the customer paid as our dependent variable (e.g., Yip, 2012). On and across booking platforms, the prices of accommodations per night are highlighted (Airbnb) or provided as filter criteria (Airbnb and Booking.com) and present an intuitive way for customers to compare prices and decide on purchases. We followed prior studies on hotel pricing factors by using price per night rather than total price (e.g., Wang & Nicolau, 2017). Our cross-platform setting was particularly helpful in allowing us to capture those aspects of value that prior research identified as influential on customers. For instance, these are the advantages of the offering compared to substitutes, competition affecting the perceived value, and the balance between the advantages and prices across different offerings. In online purchase situations in general, such information can be directly observed by customers using price comparison sites such as prevalent in the electronic market (Kocas, 2002) or otherwise compare offerings across platforms to increase perceived value for the price paid (Kortge & Okonkwo, 1993). Because the price that the seller sets versus the price that the customer paid differs in our setting due to platform fees set by Airbnb and Booking.com, it was important to use the final customer price to capture customer value. Therefore, we calculated for each observation the average price per night paid by the customer, including all fees (platform fee, cleaning fee, and local taxes).

Core product attributes. To capture the core product attributes essential for customer value, we used several baseline variables in our model. First, in line with prior research, we included the categorical variable "apartment" representing the booked apartment in each observation not only as a proxy for price-influencing core product attributes, such as the number of bedrooms, apartment size, apartment type, distance to the city center, etc. (Gibbs et al., 2018; Teubner et al., 2017), but also to control for individual seller effects as the apartment is a unique identifier for a seller. The apartment variable has one distinct value for each apartment.

In addition, we included *location factors* to capture potential local price differences in our data set (Teubner et al., 2017; Wang & Nicolau, 2017), using the categorical variables "country" and "city.".

We also accounted for *seasonal factors* that cause price differences in the accommodation industry, such as whether rooms are booked for holidays versus non-holidays or weekdays versus weekends (Abrate et al., 2012). Thus, we controlled for the "month of check-in" and for a numeric variable, "weekend factor," which calculated the share of weekend nights in a booking.

Finally, we included *customer factors* defined at the time of booking that describe the extent of usage of the product. These were the group size and the length of stay, both of which potentially influence the customer value and thus the customer price per night, and which are in addition needed as control variables for long-stay discounts and fees for additional guests. Thus, we included the numeric variables "number of guests" and "nights" of a booking, both being factors that directly affect customer value in terms of the extent of product usage.

Beyond-core product attributes. To measure non-core product attributes, we defined two variables, channel brand and seller history, to test their influence as a source of customer value beyond core product factors. Following prior research (D. J. Kim et al., 2008; Reuer et al., 2012; Teubner et al., 2017), we used a categorical variable for channel brand, in our case employing the categories "Airbnb" versus "Booking.com." Because the platform, as a facilitator customers rely on when making a transaction, is a value signal for the customer (Lee et al., 2005), it is therefore identified as a non-core product attribute. Likewise, for seller history we created a numeric variable representing the development of the seller's business experience over time, counting the number of days the seller had spent operating the business at the date when a booking occurred (i.e., when the customer decision was made). We measured a sellers' history on the respective platform in days, setting it to 1 at the beginning of the data observation period, January 1, 2018, so that the impact of increasing history on the development of prices could be observed over time. For example, when a booking was made on January 23, 2018, the seller's business history equaled 23. The date when each apartment was listed was unavailable in the data. By using a fixed starting point, we were able to measure the effect comparably among apartments and sellers and the effect on price of a booking over time. Prior research suggests that a seller's business history can be observed by customers across platforms in several ways—for instance, via the (increasing) number of reviews over time and the date when the seller joined the platform, which is visible on both the Airbnb and Booking.com platforms (Magno et al., 2018; Teubner et al., 2017; Wang & Nicolau, 2017). Thus, our created variable functioned as a proxy for such observable seller history signals, which function as non-core product-related attributes.

3.3. Model definition

To test the hypotheses, we defined five models. First, we defined Model 1 that served as the baseline model. It did not include the beyond-core product attributes channel brand and seller history. Model 2 represented the main model and included all factors. The comparison of Model 1 and Model 2 was used to analyze the importance of the beyond-core product attributes for price predictions and to test Hypothesis 1 and 2.

Second, we defined three sub models, Models 3, 4 and 5, to test Hypothesis 3. To analyze the importance of the beyond-core product attributes under different time constraints (the rate factor), we grouped the data into the rate groups "early-bird" (Model 3), "standard" (Model 4), and "last-minute" (Model 5) that corresponded with the lead time before booking. "Early-bird" bookings were those made at least 30 days before departure, "standard" bookings were made between 30 and eight days before departure, and "last-minute" bookings were made seven or fewer days before departure. This classification is based on practical use cases across platforms and corresponding discount settings (i.e., platform contingencies) that we were able to analyze and benchmark across platforms due to our collaboration with the SMEs.

3.4. Analytical approach

To find the best model to analyze the importance of the factors for price predictions, we evaluated two distinct approaches. We applied a random forest algorithm in R, comparing it to linear regression. The comparison of several models is common in evaluating the performance of a predictive model (Chatterjee et al., 2020). Particularly the random forest model can be better suited to tasks with unknown and complex underlying relationships (Delen & Zolbanin, 2018). Therefore, it represents a promising approach for determining the importance of individual customer value factors for predicting prices. Assumptions about data structure are not required, nor is an upfront model specification (Lantz, 2019). This allows our approach to be used as a blueprint for other contexts. We are aware that predictive models based on decision trees, such as random forests, are more difficult to interpret and do not disclose the direction of a factor. But they are suitable for assessing a factor's importance in explaining variance in a predicted variable and thus the factor's importance for the prediction. Many studies of pricing factors in the accommodation industry use regression models (e.g., Hung et al., 2010; Wang & Nicolau, 2017). Some scholars have acknowledged the superiority of the random forest model in prediction power, comparing it to linear regression, but have used linear regression due to its advantage as an explanatory model (Chatterjee et al., 2020). In contrast, our major criterion was predictive accuracy and the suitability of the model for exploratory analyses. We expected random forest to be a more accurate model, particularly well suited to small to medium-large data sets, and accepted the disadvantage of its inability to interpret the direction of factors. Our approach followed ideas in Lantz (2019) and was developed to provide a blueprint to those interested in analyzing small to medium-large data sets to derive actionable explorative insights.

4. Model results

4.1. Descriptive results

The data set contains 1,972 observations, from which 818 (41 %) are Airbnb bookings and 1,154 booking.com (59 %) bookings. Looking at the rate groups the data set contains 617 (31 %) early-bird bookings, 606 (31 %) standard bookings, and 749 (38 %) last-minute bookings. Descriptive statistics across booking characteristics are shown in Table 1.

Table 1
Distribution of bookings.

Guests (# of guests)		Length of stay (# of nights)		Platform		Country	
1	309 (16 %)	**1**	172 (9 %)	**Airbnb**	818 (41 %)	**Austria**	469 (24 %)
2	718 (36 %)	**2**	715 (36 %)	**Booking.com**	1,154 (59 %)	**UK**	1,503 (76 %)
3	326 (17 %)	**3**	437 (22 %)				
4	380 (19 %)	**4**	288 (15 %)				
5	161 (8 %)	**5**	147 (7 %)				
6	78 (4 %)	**> 5**	213 (11 %)				

4.2. Random forest in R

First, we set up and ran the random forest model on the full sample of 1,972 observations. Before running the model, the data were randomly split into a training data set (1,588 observations, approx. 80 %) and a test data set (384 observations, approx. 20 %). A separate test data set is needed as a hold-out sample to assess the predictive accuracy of the model (Delen & Zolbanin, 2018; Shmueli & Koppius, 2011). The random forest was run with the "randomForest" function in R on the training data with defined independent variables and a dependent variable. This procedure was done for the baseline model (without beyond-core product attributes) and the main model (with all factors). Similarly, the random forest was set up and run for the three sub models. For these smaller data sets, the splitting of the data into training and test data was repeated using the previous ratio (80:20). To ensure replicability of the random forest analysis, we always set the seed of the randomization to a fixed starting point before running the random forest algorithm (Lantz, 2019).

4.3. Linear regression versus random forest

4.3.1. Model performance: Linear regression versus random forest

To evaluate model performance, we assessed four performance criteria for prediction models. First, we examined the variance explained by the model as an indicator of overall model performance, although this was not appropriate for assessing predictive accuracy (Shmueli & Koppius, 2011). Second, therefore, we assessed three measures of predictive accuracy, namely, the root mean squared error (RMSE), the correlation between predicted values and actual values, and the mean absolute error (MAE) of predicted values and actual values (Lantz, 2019; Shmueli & Koppius, 2011). In the random forest models, the predictive accuracy evaluations were calculated on the test data set, as suggested by previous research (Delen & Zolbanin, 2018; Shmueli & Koppius, 2011). Table 2 details the comparison of the linear regression and random forest models. All four performance criteria were assessed in both models. The variance explained in the linear regression baseline (main) model was 60.72 % (60.75 %). In the random forest baseline (main) model, the variance that was explained increased to 67.99 % (70.99 %). To our knowledge, in predictive analytics, common thresholds have not been similarly established in literature than in traditional models such as linear regression. But, for regression models of real-world data who often show very low values, our results can be considered already quite good (Lantz, 2019). We then obtained the predictive accuracy of the two

Table 2
Comparison of the linear regression and random forest models.

Model	**Model 1 Baseline**	**Model 1 Baseline**	**Model 2 Main model**	**Model 2 Main model**
Method	LR	RF	LR	RF
Group (rate)	Full sample			
Dependent variable	Customer price per night			
Model evaluation				
Variance explained (in %)	60.72	67.99	60.75	70.99
RMSE	33.33	30.23	33.31	28.78
Correlation between predicted and actual values	0.78	0.84	0.78	0.88
MAE between predicted and actual values	22.35	20.36	22.34	17.77
MAE improvement vs LR model (rel.)		8.9 %		20.45 %

LR = linear regression; RF = random forest; RMSE = root mean squared error; MAE = mean absolute error.

model specifications. The random forest models were more accurate for all three criteria than the linear regression models. First, in the random forest baseline model, RMSE decreased to 30.23 (from 33.33 in the linear regression baseline model), the correlation of predicted and actual values increased to 0.84 (from 0.78 in the linear regression baseline model), and the MAE of predicted values decreased by 1.99 to 20.36 (from 22.35 in the linear regression baseline model), for an overall relative improvement of 8.90 %.

Second, in the random forest main model, RMSE decreased to 28.78 (from 33.31 in the linear regression main model), correlation of predicted and actual values increased to 0.88 (from 0.78 in the linear regression main model), and the MAE of predicted values decreased by 4.57 to 17.77 (from 22.34 in the linear regression main model), for an overall relative improvement of 20.45 %. The random forest models outperformed the linear regression models in all evaluation criteria.

4.3.2. Welch two-sample t-test: Linear regression versus random forest

To test the significance of model improvement, we conducted a significance test of the improvement of predictive accuracy between the random forest and linear regression models. To do this, we defined the mean of absolute errors of the predictions in each model–method combination $\mu_{model,method}$. First, we compared the baseline models $\mu_{base,LR}$ and $\mu_{base,RF}$. Then, we compared the main models $\mu_{main,LR}$ and $\mu_{main,RF}$. Comparing their means with a one-sided Welch two-sample *t*-test with the null hypothesis that $\mu_{model,LR} \leq \mu_{model,RF}$ and the alternative hypothesis that $\mu_{model,LR} > \mu_{model,RF}$ resulted in rejection of the null hypothesis at a *p*-value of 0.04332 for the baseline models and a *p*-value of 1.56E-05 for the main models. The MAE of the random forest models was thus significantly smaller than the MAE of the linear regression models in both cases—that is, the predictive accuracy improvement of using a random forest algorithm versus linear regression was significant.

4.4. Random forest

4.4.1. Model performance: Model 1 (baseline model) versus model 2 (main model)

The random forest baseline model and the four main models (main model and three sub models) are compared in Tables 2 and 3. First, we compared the predictive accuracy of the random forest main and baseline models. The variance explained in the baseline model was 67.99 %. In the main model, the variance that was explained increased to 70.99 %. Additionally, we obtained the predictive accuracy of the models. Predictive accuracy, based on the three criteria, was improved in the main model compared to the baseline model. In the main model, RMSE decreased to 28.78 (from 30.23 in the baseline model), the correlation of the predicted and actual values increased to 0.88 (from 0.84 in the baseline model), and the MAE of the predicted values decreased by 2.58 to 17.77 (from 20.36 in the baseline model), resulting in an overall relative improvement of 12.70 %.

4.4.2. Welch two-sample t-test: Baseline versus main model

To test Hypotheses 1 and 2 and determine whether beyond-core product attributes affected perceived customer value and price prediction, a significance test of the improvement of predictive accuracy between the random forest main model and the random forest baseline model was conducted. To do this, we calculated the mean of absolute errors of the predictions in the random forest baseline model $\mu_{base,RF}$ and the random forest main model $\mu_{main,RF}$. Comparing their means with a one-sided Welch two-sample *t*-test with the null hypothesis that $\mu_{base,RF} \leq \mu_{main,RF}$ and the alternative hypothesis that $\mu_{base,RF} > \mu_{main,RF}$ resulted in rejection of the null hypothesis at a *p*-value of 0.03095. The MAE of the main model was thus significantly smaller than the MAE of the baseline model, so the improvement of predictive accuracy obtained by including the beyond-core product attributes was significant.

Table 3
Random forest model accuracy evaluation.

Model	Model 1 Baseline model	Model 2 Main model	Model 3 Early-bird	Model 4 Standard	Model 5 Last-minute
Method	Random forest				
Group (rate)	Full sample	Full sample	Early-bird	Standard	Last-minute
Dependent variable	Customer price per night				
Model evaluation					
Variance explained (in %)	67.99	70.99	72.61	59.57	58.67
RMSE	30.23	28.78	33.69	31.36	23.37
Correlation between predicted and actual values	0.84	0.88	0.89	0.84	0.80
MAE between predicted and actual values	20.36	17.77	17.83	20.12	16.76
MAE of average-predicting model	40.37	40.37	48.33	40.41	29.50
MAE improvement (abs.)	20.01	22.60	30.50	20.29	12.73
MAE improvement (rel.)	50 %	56 %	63 %	50 %	43 %

RMSE = root mean squared error; MAE = mean absolute error.

4.4.3. ANOVA and rank test: Difference in rate groups for sub-models

To assess whether grouping the factor rate into the sub models “early-bird,” “standard,” and “last-minute” was meaningful, we conducted an ANOVA and a rank test. We assessed the beyond-core product attributes channel brand and seller history separately. ANOVA and rank tests show whether a factor differs between groups by testing the significance of the mean differences in this factor. Both the ANOVA and rank test confirmed that the beyond-core product attributes differed in the rate groups. Results are shown in Table 4.

4.4.4. Beyond-product attributes: Ranking, factor importance and significance of factors

To additionally support Hypothesis 1 and 2, and to test Hypothesis 3, we assessed the influence of the beyond-core product attributes channel brand and seller history on the prediction of prices under different time

Table 4
Results of ANOVA and rank test.

	Customer price per night		Seller history		Channel brand	
Rate group	Mean	sd	mean	sd	mean	sd
Early-bird (n = 617)	141	64.7	385	211	1.72	0.448
Standard (n = 606)	116	49.2	434	218	1.50	0.500
Last-minute (n = 749)	102	36.7	424	212	1.54	0.499
ANOVA						
Df	2		2		2	
F value	103.2		8.943		37.34	
Pr (>F)	< 2e-16 ***		0.000136 ***		< 2e-16 ***	
Kruskal-Wallis rank-sum test						
Df	2		2		2	
Chi-sq	161.72		20.41		72.03	
p-value	< 2.2e-16 ***		3.699e-05 ***		2.286e-16 ***	

*Significance levels: *** = 0.001; ** = 0.01; * = 0.05.*

constraints. Therefore, we obtained the factor importance of all of the variables, which, in random forest models, reflects the strength of the effect of individual factors. Further, to determine the significance of factor importance, we used the function RFpermute in R. This function extracts one sample tree from the random forest, which allows assessment of the increase in the mean standard error (incMSE). incMSE can be used to assess the importance of the factor: the higher the value, the more the mean standard error of the prediction would increase when applying a permutation (i.e., random allocation of values) to that variable; this means that the higher the value, the more important the factor to prediction accuracy. The significance of factor importance in a random forest model implies that the variable has a significant influence on decreasing the mean standard error of the prediction of the dependent variable. However, it must be noted that the random forest model does not allow interpretation of the direction of the effect, only the strength.

Findings of the random forest models are shown in Fig. 1 and Table 5. In the random forest baseline model and main model, all factors are significant in incMSE at a p-value of < 0.05. Both beyond-core product attributes (seller history and channel brand) are significant in incMSE.

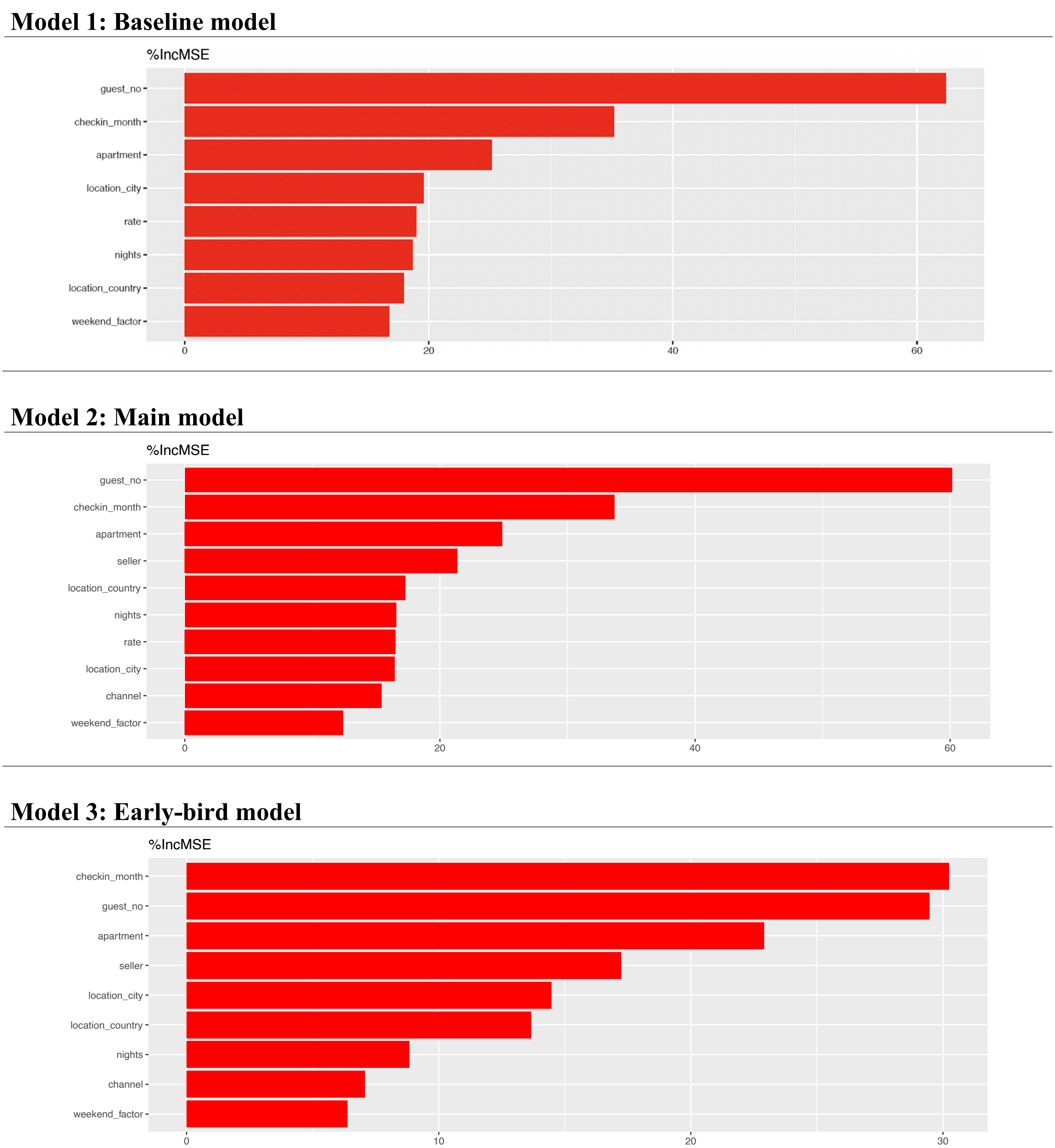
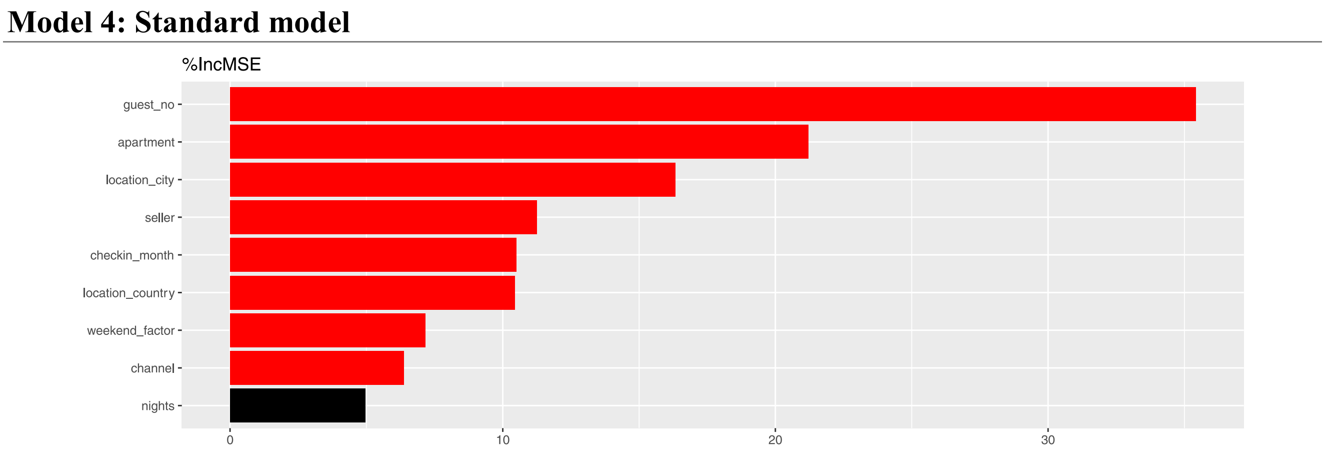


Fig. 1. Rank and significance of factors. Red indicates that the p-value is < 0.05. “Seller” = “Seller history.” “Channel” = “Channel brand.” IncMSE = increase in the mean standard error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



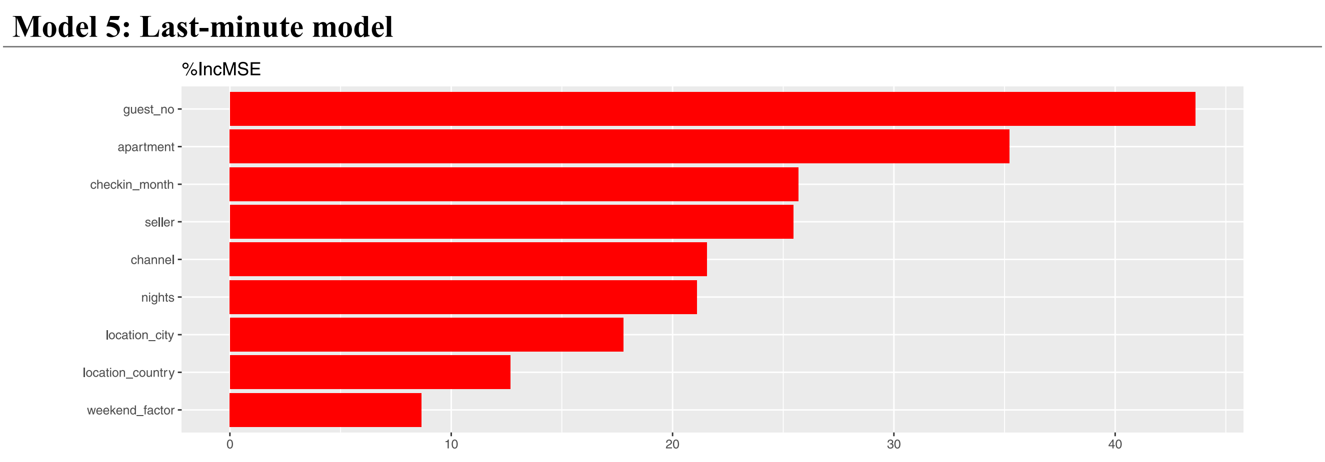


Fig. 1. (*continued*).

This supports Hypotheses 1 and 2, namely, that seller history and channel brand are relevant factors for the price prediction. Comparing the ranking of the beyond-core product attributes to the ranking of the core attributes, we found that seller history ranks fourth and channel ranks ninth out of ten variables in the main model. This indicates that seller history is more important than channel for the improvement of the price prediction and that it is relatively more important than most of the core attributes in the model. From the core product attributes, only number of guests, the check-in month and the apartment rank higher than seller history, followed by location factors “city” and “country”, length of stay and the booking lead time ranking before beyond-core product attribute channel. Weekend factor was found least important.

However, these rankings are found in main model and do not account for the time constraint, the “rate”, which is reflected in Hypothesis 3. In the main model, the factor rate is also significant in incMSE. This confirms, along with the ANOVA and rank test, that grouping the data based on the variable rate could generate further insights.

In all of the sub models (Models 3, 4, and 5), the factors seller history and channel brand were found to be significant in incMSE. Seller history ranked fourth in all three sub models, keeping the same rank as in the main model. Channel brand ranked highest in the “last-minute” model (ranking fifth, versus eighth in the “early-bird” and “standard” models and ninth in the main model). Both factors (seller history, channel brand) increased their relative share of factor importance versus the other factors in “last-minute” (12 %, 10 %) versus “early bird” (11 %, 5 %) or “standard” (9 %, 5 %) models. Both factors had their strongest effect on price predictions in Model 5, the “last-minute” model. This supports Hypothesis 3—that beyond-core product attributes are more important for price predictions when decision-making time is short.

5. Discussion

This study set out to investigate the importance of beyond-core product attributes such as channel brand and seller history as signals in estimating value-based pricing on digital platforms. We modeled the effects under time-to-travel constraints. In theorizing and confirming the importance of two beyond-core product attributes, we contributed to the research on value-based pricing, which tends to focus on core product attributes (e.g., Codini et al., 2012). We showed that both channel brand and seller history function as customer value-impacting price attributes and are impacted by the purchaser’s time constraints. Our findings contribute to the broad literature on price factors affecting purchases on digital platforms and, specifically, to the literature on the accommodation industry. Previous studies have investigated price determinants in this industry on a single platform, most often Airbnb, and assessed external factors for which data was publicly available (e.g., Gibbs et al., 2018; Magno et al., 2018; Teubner et al., 2017; Wang & Nicolau, 2017). To the best of our knowledge, our study is unique in enabling a cross-platform perspective within an SME context, and our results demonstrate the importance of platform choice—that is, channel brand—in predicting offering prices. This is a generalizable contribution to the understanding of dynamic pricing on digital online platforms (Abrate et al., 2012; Fassnacht & Unterhuber, 2016).

An increasing research focus on the interface of alternative marketing and entrepreneurial, small firms can be traced back to the 1980 s (e.g., Morris & Paul, 1987). Hansen et al. (2020) summarized the

Table 5
(Relative) strength and significance of factor importance (incMSE).

		Model 1: Baseline model			Model 2: Main model			Model 3: Early bird *Early bird*			Model 4: Standard *Standard*			Model 5: Last-minute *Last Minute*		
	Factor	% incMSE	Rank	Rel.	% incMSE	Rank	Rel.	% incMSE	Rank	Rel.	% incMSE	Rank	Rel.	% incMSE	Rank	Rel.
Dependent variable	**Customer price per night**															
Beyond-core product attributes	**Seller history**	–			21.38**	4	9 %	17.24**	4	11 %	11.24**	4	9 %	25.45**	4	12 %
	Channel brand	–			15.41**	9	7 %	7.07**	8	5 %	6.38**	8	5 %	21.56**	5	10 %
Time factor	**Rate**	18.98**	5	9 %	16.52**	7	7 %	–			–			–		
Core product customer value factors	**Guest No.**	62.37**	1	29 %	60.15**	1	26 %	29.46**	2	20 %	35.41**	1	29 %	43.61**	1	21 %
	Nights	18.67**	6	9 %	16.58**	6	7 %	8.84*	7	6 %	4.97(n.s.)	9	4 %	21.11**	6	10 %
	Weekend factor	16.76**	8	8 %	12.41**	10	5 %	6.38*	9	4 %	7.15**	7	6 %	8.66*	9	4 %
	Month of check-in	35.15**	2	16 %	33.68**	2	14 %	30.25**	1	20 %	10.50**	5	8 %	25.69**	3	12 %
	Apartment	25.15**	3	12 %	24.89**	3	11 %	22.90**	3	15 %	21.21**	2	17 %	35.23**	2	17 %
	City	19.57**	4	9 %	16.44**	8	7 %	14.47**	5	10 %	16.33**	3	13 %	17.78**	7	8 %
	Country	17.96**	7	8 %	17.29**	5	7 %	13.66**	6	9 %	10.43**	6	8 %	12.67**	8	6 %

Significance levels: *** = 0.001; ** = 0.01; * = 0.05; n.s. = not significant. Rel. = Relative share of importance versus other factors within model.

developments of the last 30 to 40 years and called for further research on this interface. Our study builds on such research through the cross-application of concepts from both literature on small firms and marketing competition on digital platforms. We apply signaling theory by linking beyond-product value signals to pricing in customer purchase decisions on digital platforms. We thus contribute to the establishment of signaling theory in the digitization processes of SMEs (e.g., Busenitz et al., 2005; Certo et al., 2001) and show the effect of value signals in an entrepreneurial context. In particular, we apply signaling theory to identify price optimization opportunities for small businesses across accommodation-booking platforms. By combining an entrepreneurial perspective and marketing concepts, our study furthers research on the marketing–digital platform interface. Our findings have relevance for signaling theory as they describe signaling actions of actors (the non-core product attributes) as value glue to create receiver interactions in a digital global context (Kromidha & Robson, 2021).

In our theorizing, we relied on underlying theoretical arguments from signaling theory (Bergh et al., 2014; Spence, 1973) that core and non-core product factors send signals to receivers, i.e., customers, and that customers observe and use such signals in their decision-making. Prior researchers have observed that sellers send signals about core product attributes (e.g., Jean & Kim, 2021; Mavlanova, Benbunan-Fich, & Lang, 2016), whereas we identified two non-manipulable value-conveying signals (channel brand and seller history) that we also describe as signaling actions of actors. We found support for our hypothesis that the importance of value signals on price prediction depends on timing. Previous studies examined the frequency of signals, their growth or decay over time and their effects on signal receivers (Connelly, Certo, Ireland, & Reutzel, 2011). In contrast, we considered the role of time from the receiver's (e.g., customer's) perspective. Both seller history and channel brand were found to be most important when customers were facing tight deadlines (i.e., when they were making last-minute bookings). In such situations, the relevance of value signals and signaling actions of actors increased. This effect could be explained by the situational time pressure affecting both seller and customer: the customer needs accommodation and the seller wants to avoid vacancies.

Our analysis provides several interesting methodological insights. Our results strongly support the application of random forest algorithms for price predictions and analysis of the effects of value signals on prices on digital online platforms. The random forest models outperformed the linear regression models in all four criteria. In particular, the mean absolute error of the prediction was reduced by 20.45 % (8.9 %) in the main (baseline) random forest model, compared to the linear regression model, and the improvement was statistically significant. The superiority of the random forest model in predictive accuracy, which parallels other scholars' methodological advances using new analytical models such as predictive analytics or machine learning in customer- or marketing-related business contexts (e.g., Giglio et al., 2019; Hwang et al., 2020; Salminen et al., 2019; Singh et al., 2017), should encourage researchers to be open to alternative analytical strategies that may better target their research questions (Delen & Zolbanin, 2018).

Our analytical approach addressed the existing challenge of value quantification in the value-based pricing literature (Hinterhuber, 2008; Liozu, 2017). We showed random forest to be a powerful method for analyzing customer value factors in online settings where value creation is complex, especially for those selling their offers across platforms. Like earlier scholars, we quantified the importance of value factors (Johnston & Mora Cortez, 2018; J. S. Kim, 2018; Wardell et al., 2008; Windisch, 2019) using the example of the accommodation industry (Abrate et al., 2022). Our predecessors identified core product attributes as important price factors for accommodation bookings (Hung et al., 2010; Magno et al., 2018; Wang & Nicolau, 2017), and research on digital channels has suggested how digital sales channels may variously impact pricing (Fassnacht & Unterhuber, 2016; Lee et al., 2005; Pappas, 2016; Stangl et al., 2016). In turn, our study makes it possible to rank the importance of such factors in a single model while showing their predictive accuracy and relevance for estimating pricing.

5.1. Practical contributions

Sellers using digital online platforms need to predict future prices accurately to avoid lost profits from prices that are too low and vacancies from prices that are too high. Our study showed that certain online activities, such as channel management, should be considered in pricing. In particular, we derived pricing insights and recommendations about two value signals. The findings on the effect of core and non-core product attribute signals are relevant for new businesses that operate on digital online booking platforms, but they can also be transferred to similar online contexts. Our results should, therefore, encourage new businesses to shape their online presence and, when doing so, acknowledge non-core product attributes.

Furthermore, our study highlights the strengths of the random forest model even when it is built on limited data. In our study, the random forest model outperformed linear regression in predictive accuracy. Our analytical approach can be used as a blueprint for comparing the performance of models and devising the most suitable predictive models for the creation of business-specific insights. We showed that a prediction model based on a random forest algorithm is suitable for price predictions on small data sets, mitigating the limitations of the data. Our approach further leverages the flexible, fast-learning nature of random forest models, which do not require model specification and data assumptions upfront (Delen & Zolbanin, 2018; Lantz, 2019). Our blueprint for setting up a predictive model can be used by practitioners to glean data-based insights for their businesses—something that is particularly relevant for new businesses that have only a short transaction history to employ in price optimization models. Overall, our findings highlight the strengths of the random forest model even when built on limited data. We encourage the model's application and adaptation to individual business contexts.

5.2. Limitations and future research

Besides its contributions, our study has limitations and offers opportunities for future research. We developed a framework to investigate the effect of value signals that go beyond core product attributes on prices on digital platforms. We controlled for important product features in our research, but future studies could expand our framework by investigating how a finer-grained display of product attributes—preconditioned by platform contingencies—could affect price prediction accuracy.

Besides optimizing value signals, new businesses operating in online markets, such as accommodation-booking platforms, must consider other issues. For example, the decision to operate through a specific channel involves not only pricing potential but also resources associated with operating on that channel—for instance, the commission paid to the channel for a booking and the time and effort required for product and price management. These factors can differ considerably. For example, the platform automatically transferred Airbnb payments from customer to seller for our sample cases in Austria, whereas payments for bookings on Booking.com had to be charged to the customer's credit card by the seller. Qualitative criteria also play a role in decision-making. For instance, selling on multiple channels could diversify risk and increase customer reach. Nevertheless, our insights into the effect of channels as value-signaling factors should be reflected in small businesses' decision-making.

Overall, our findings contribute to the discussion on the influence of value signals beyond core product attributes in customer purchase decisions. Future research should expand the investigation to additional sources of value signals and translate our analytical approach to other business contexts. Our results suggest that beyond-core product attributes are strongest when time is short. For this reason, beyond-core product attributes and the timing of purchase decisions should be given further attention in the study of price optimization.

6. Conclusion

The goal of this study was to understand the importance of channel and seller history as beyond-core product attributes and determinants of value-based pricing on digital online platforms. We drew from signaling theory to develop a theoretical framework of the effect of non-core product attributes as value signals on the prediction of prices and found evidence for their influencing role. We tested two value signals prevalent in the online customer purchase situation beyond the product itself: a seller history signal (the seller's experience, measured by number of days spent operating the business) and a channel brand signal (the platform's brand). Using a sample of non-public, cross-platform booking data from new businesses in the UK and Austria over a period of two years, we showed that a random forest model provided higher predictive accuracy than linear regression for analyzing the importance of such beyond-core product attributes for predicting prices. We also demonstrated that this model could produce valuable insights from limited data. The findings of our research are relevant both for future studies of signaling in value-based pricing strategies and for practitioners. The latter can draw on our findings about the influence of value signals on customer value and use our analytical approach to analyze pricing factors or other business questions.

CRediT authorship contribution statement

Tatjana Christen: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation. **Manuel Hess:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Conceptualization. **Dietmar Grichnik:** Supervision, Conceptualization. **Joakim Wincent:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abrate, G., Fraquelli, G., & Viglia, G. (2012). Dynamic pricing strategies: Evidence from European hotels. *International Journal of Hospitality Management, 31*(1), 160–168. https://doi.org/10.1016/j.ijhm.2011.06.003

Abrate, G., Sainaghi, R., & Maurib, A. G. (2022). Dynamic pricing in Airbnb: Individual versus professional hosts. *Journal of Business Research, 141*(March 2022), 191–199.

Backhaus, K., Becker, J., Beverungen, D., Frohs, M., Müller, O., & Weddeling, M. (2010). Incorporating willingness-to-pay data into online recommendations for value-added services. *18th European Conference on Information Systems, ECIS 2010*.

Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica, 47*(2), 89–104. https://doi.org/10.1016/0001-6918(81)90001-9

Bergh, D. D., Connelly, B. L., Ketchen, D. J., & Shannon, L. M. (2014). Signalling theory and equilibrium in strategic management research: An assessment and a research agenda. *Journal of Management Studies, 51*(8), 1334–1360. https://doi.org/10.1111/joms.12097

Busenitz, L. W., Fiet, J. O., & Moesel, D. D. (2005). Signaling in venture capitalist - New venture team funding decisions: Does it indicate long-term venture outcomes? *Entrepreneurship: Theory and Practice, 29*(1), 1–12. https://doi.org/10.1111/j.1540-6520.2005.00066.x

Certo, S. T., Daily, C. M., & Dalton, D. R. (2001). Signaling firm value through board structure: An investigation of initial public offerings. *Entrepreneurship: Theory and Practice, 26*(2), 33–50. https://doi.org/10.1177/104225870102600202

Chatterjee, S., Goyal, D., Prakash, A., & Sharma, J. (2020). Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application. *Journal of Business Research, October.*. https://doi.org/10.1016/j.jbusres.2020.10.043

Christopher, M., & Gattorna, J. (2005). Supply chain cost management and value-based pricing. *Industrial Marketing Management, 34*, 115–121. https://doi.org/10.1016/j.indmarman.2004.07.016

Codini, A., Saccani, N., & Sicco, A. (2012). The relationship between customer value and pricing strategies: An empirical test. *Journal of Product & Brand Management, 21*(7), 538–546. https://doi.org/10.1108/10610421211276321

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management, 37*(1), 39–67. https://doi.org/10.1177/0149206310388419

Cui, Y. (Gina), Kim, S. (Sam), & Kim, J. (2020). Impact of preciseness of price presentation on the magnitude of compromise and decoy effects. *Journal of Business Research, October*. https://doi.org/10.1016/j.jbusres.2020.10.017.

Delen, D., & Zolbanin, H. M. (2018). The analytics paradigm in business research. *Journal of Business Research, 90*(April), 186–195. https://doi.org/10.1016/j.jbusres.2018.05.013

Fassnacht, M., & Unterhuber, S. (2016). Consumer response to online/offline price differentiation. *Journal of Retailing and Consumer Services, 28*, 137–148. https://doi.org/10.1016/j.jretconser.2015.09.005

Forbis, J. L., & Mehta, N. T. (1981). Value-based strategies for industrial products. *Business Horizons, 24*(3), 32–42. https://doi.org/10.1016/0007-6813(81)90125-7

Gibbs, C., Guttentag, D., Gretzel, U., Mortin, J., & Goodwill, A. (2018). Pricing in the sharing economy: A hedonic pricing model applied to Airbnb listings. *Journal of Travel & Tourism Marketing, 35*(1), 46–56. https://doi.org/10.1080/10548408.2017.1308292

Giglio, S., Pantano, E., Bilotta, E., & Melewar, T. C. (2019). Branding luxury hotels: Evidence from the analysis of consumers' "big" visual data on TripAdvisor. *Journal of Business Research*, *119*(October 2019), 495–501. https://doi.org/10.1016/j.jbusres.2019.10.053.

Gregg, D. G., & Walczak, S. (2010). The relationship between website quality, trust and price premiums at online auctions. *Electronic Commerce Research, 10*(1), 1–25. https://doi.org/10.1007/s10660-010-9044-2

Gupta, P., Yadav, M. S., & Varadarajan, R. (2009). How Task-Facilitative Interactive Tools Foster Buyers' Trust in Online Retailers: A Process View of Trust Development in the Electronic Marketplace. *Journal of Retailing, 85*(2), 159–176. https://doi.org/10.1016/j.jretai.2009.02.001

Hansen, D. J., Deacon, J. H., Pascal, V., & Sethna, Z. (2020). The future is in the past: A framework for the Marketing-Entrepreneurship Interface (MEI). *Journal of Business Research, 113*, 198–208. https://doi.org/10.1016/j.jbusres.2019.05.030

Henard, D. H., & Dacin, P. A. (2010). Reputation for product innovation: Its impact on consumers. *Journal of Product Innovation Management, 27*(3), 321–335. https://doi.org/10.1111/j.1540-5885.2010.00719.x

Hinterhuber, A. (2008). Customer value-based pricing strategies: Why companies resist. *Journal of Business Strategy, 29*(4), 41–50. https://doi.org/10.1108/02756660810887079

Hinterhuber, A., & Liozu, S. M. (2014). Is innovation in pricing your next source of competitive advantage? *Business Horizons, 57*(3), 413–423. https://doi.org/10.1016/j.bushor.2014.01.002

Hung, W. T., Shang, J. K., & Wang, F. C. (2010). Pricing determinants in the hotel industry: Quantile regression analysis. *International Journal of Hospitality Management, 29*(3), 378–384. https://doi.org/10.1016/j.ijhm.2009.09.001

Hwang, M. I. (1994). Decision making under time pressure: A model for information systems research. *Information & Management, 27*, 197–203.

Hwang, S., Kim, J., Park, E., & Kwon, S. J. (2020). Who will be your next customer: A machine learning approach to customer return visits in airline services. *Journal of Business Research, 121*(August), 121–126. https://doi.org/10.1016/j.jbusres.2020.08.025

Ingenbleek, P. T. M., Debruyne, M., Frambach, R. T., & Verhallen, T. M. M. (2003). Successful new product pricing practices: A contingency approach. *Marketing Letters, 14*(4), 289–305. https://doi.org/10.1023/B:MARK.0000012473.92160.3d

Ingenbleek, P. T. M., Frambach, R. T., & Verhallen, T. M. M. (2010). The role of value-informed pricing in market-oriented product innovation management. *Journal of Product Innovation Management, 27*(7), 1032–1046. https://doi.org/10.1111/j.1540-5885.2010.00769.x

Ingenbleek, P. T. M., Frambach, R. T., & Verhallen, T. M. M. (2013). Best practices for new product pricing: Impact on market performance and price level under different conditions. *Journal of Product Innovation Management, 30*(3), 560–573. https://doi.org/10.1111/jpim.12008

Jean, R. J. "Bryan," & Kim, D. (2021). Signalling Strategies of Exporters on Internet Business-to-Business Platforms. *Journal of Management Studies*. https://doi.org/10.1111/joms.12761.

Johnston, W. J., & Mora Cortez, R. (2018). Unit pricing and its implications for B2B marketing research. *Industrial Marketing Management, 69*(February), 32–39. https://doi.org/10.1016/j.indmarman.2018.01.026

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems, 44*(2), 544–564. https://doi.org/10.1016/j.dss.2007.07.001

Kim, J. S. (2018). Measuring willingness-to-pay for mobile phone features: A multi-region study. *Journal of Research in Marketing and Entrepreneurship, 20*(2), 189–213. https://doi.org/10.1108/JRME-02-2016-0003

Kirmani, A., & Rao, A. R. (2000). No pain, no gain: A critical review of the literature on signaling unobservable product quality. *Journal of Marketing, 64*(2), 66–79. https://doi.org/10.1509/jmkg.64.2.66.18000

Kocas, C. (2002). Evolution of prices in electronic markets under diffusion of price-comparison shopping. *Journal of Management Information Systems, 19*(3), 99–119. https://doi.org/10.1080/07421222.2002.11045740

Kortge, G. D., & Okonkwo, P. A. (1993). Perceived value approach to pricing. *Industrial Marketing Management, 22*(2), 133–140. https://doi.org/10.1016/0019-8501(93)90039-A

Kromidha, E., & Robson, P. J. (2021). The role of digital presence and investment network signals on the internationalisation of small firms. *International Small Business Journal*, *39*(2), 109-129. https://doi.org/10.1177%2F0266242620958898.

Lantz, B. (2019). *Machine Learning with R: Expert techniques for predictive modeling* ((3rd ed.).). Packt Publishing.

Lee, B. C., Ang, L., & Dubelaar, C. (2005). Lemons on the Web: A signalling approach to the problem of trust in Internet commerce. *Journal of Economic Psychology, 26*(5), 607–623. https://doi.org/10.1016/j.joep.2005.01.001

Li, H., Jiang, J., & Wu, M. (2014). The effects of trust assurances on consumers' initial online trust: A two-stage decision-making process perspective. *International Journal of Information Management, 34*(3), 395–405. https://doi.org/10.1016/j.ijinfomgt.2014.02.004

Liozu, S. M. (2017). State of value-based-pricing survey: Perceptions, challenges, and impact. *Journal of Revenue and Pricing Management, 16*(1), 18–29. https://doi.org/10.1057/s41272-016-0059-8

Lockyer, T. (2005). The perceived importance of price as one hotel selection dimension. *Tourism Management, 26*(4), 529–537. https://doi.org/10.1016/j.tourman.2004.03.009

Magno, F., Cassia, F., & Ugolini, M. M. (2018). Accommodation prices on Airbnb: Effects of host experience and market demand. *TQM Journal*. https://doi.org/10.1108/TQM-12-2017-0164

Mavlanova, T., Benbunan-Fich, R., & Lang, G. (2016). The role of external and internal signals in E-commerce. *Decision Support Systems*. https://doi.org/10.1016/j.dss.2016.04.009

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review, 20*(3), 709–734.

Morris, M. H., & Paul, G. W. (1987). The relationship between entrepreneurship and marketing in established firms. *Journal of Business Venturing, 2*(3), 247–259. https://doi.org/10.1016/0883-9026(87)90012-7

Pappas, N. (2016). Marketing strategies, perceived risks, and consumer trust in online buying behaviour. *Journal of Retailing and Consumer Services, 29*, 92–103. https://doi.org/10.1016/j.jretconser.2015.11.007

Pohland, L., & Kesgin, M. (2018). Pricing determinants in hotels: The case of luxury, upscale, and mid-scale price segments. *Journal of Revenue and Pricing Management, 17*(4), 218–230. https://doi.org/10.1057/s41272-017-0133-x

Raja, J. Z., Frandsen, T., Kowalkowski, C., & Jarmatz, M. (2020). Learning to discover value: Value-based pricing and selling capabilities for services and solutions. *Journal of Business Research, 114*, 142–159. https://doi.org/10.1016/j.jbusres.2020.03.026

Reuer, J. J., Tong, T. W., & Wu, C. W. (2012). A Signaling theory of acquisition premiums: Evidence from IPO targets. *Academy of Management Journal, 55*(3), 667–683. https://doi.org/10.5465/amj.2010.0259

Salminen, J., Yoganathan, V., Corporan, J., Jansen, B. J., & Jung, S. G. (2019). Machine learning approach to auto-tagging online content for content marketing efficiency: A comparative analysis between methods and content type. *Journal of Business Research*, *101*(September 2018), 203–217. https://doi.org/10.1016/j.jbusres.2019.04.018.

Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly: Management Information Systems, 35*(3), 553–572. https://doi.org/10.2307/23042796

Simon, H. A. (1996). *The sciences of the artificial* ((3rd ed.).). MIT Press.

Singh, J. P., Irani, S., Rana, N. P., Dwivedi, Y. K., Saumya, S., & Kumar Roy, P. (2017). Predicting the "helpfulness" of online consumer reviews. *Journal of Business Research, 70*, 346–355. https://doi.org/10.1016/j.jbusres.2016.08.008

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics, 87*(3), 355–374.

Stangl, B., Inversini, A., & Schegg, R. (2016). Hotels' dependency on online intermediaries and their chosen distribution channel portfolios: Three country insights. *International Journal of Hospitality Management*. https://doi.org/10.1016/j.ijhm.2015.09.015

Teubner, T., Hawlitschek, F., & Dann, D. (2017). Price Determinants on Airbnb: How Reputation Pays Off in the Sharing Economy. *Journal of Self-Governance and Management Economics, 5*, 53–80. https://doi.org/10.22381/JSME5420173.

Wang, D., & Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com. *International Journal of Hospitality Management, 62*, 120–131. https://doi.org/10.1016/j.ijhm.2016.12.007

Wardell, C. L., Wynter, L., & Helander, M. (2008). Capacity and value based pricing model for professional services. *Journal of Revenue and Pricing Management, 7*(4), 326–340. https://doi.org/10.1057/rpm.2008.18

Windisch, G. (2019). Better profit formulas for new business models: Capital goods industries in a fragmented market. *Journal of Business Strategy, 41*(2), 29–38. https://doi.org/10.1108/JBS-09-2018-0159

Yip, N. K. T. (2012). Making qualitative decisions from quantitative cues: Understanding the customers' willingness to pay. *Journal of Revenue and Pricing Management, 11*(5), 562–566. https://doi.org/10.1057/rpm.2012.18

Tatjana Christen is Ph.D. Candidate at the University of St.Gallen, Institute of Technology Management, Switzerland. She is a consultant at McKinsey Company, currently in her educational leave for her Ph.D. studies.

Manuel Hess is Assistant Professor at the Global Center for Entrepreneurship and Innovation, University of St.Gallen, Switzerland. He holds a Ph.D. from the University of St. Gallen, and leads the St.Galler Startup Navigator Research Lab.

Dietmar Grichnik is full Professor and Director of the Institute of Technology Management, University of St.Gallen, Switzerland.

Joakim Wincent is Professor at Hanken School of Economics, Finland, and at the Global Center for Entrepreneurship and Innovation, University of St.Gallen, Switzerland.