Journal of Operations Management 36 (2015) 33–45

Contents lists available at ScienceDirect

Journal of Operations Management

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

Effective judgmental forecasting in the context of fashion products☆

Matthias Seifert[a,*], Enno Siemsen[b], Allègre L. Hadida[c], Andreas B. Eisingerich[d]

[a] IE Business School, IE University, Spain
[b] Carlson School of Management, University of Minnesota, United States
[c] University of Cambridge Judge Business School & Magdalene College, United Kingdom
[d] Imperial College Business School, Imperial College London, United Kingdom

ARTICLE INFO

Article history:
Received 4 June 2014
Received in revised form 7 February 2015
Accepted 12 February 2015
Available online 28 February 2015
Accepted by Thomas Younghoon Choi

Keywords:
Judgmental forecasting
Fashion products
Lens model design
Demand uncertainty
Music industry
New product forecasting

ABSTRACT

We study the conditions that influence judgmental forecasting effectiveness when predicting demand in the context of fashion products. Human judgment is of practical importance in this setting. Our goal is to investigate what type of decision support, in particular historical and/or contextual predictors, should be provided to human forecasters to improve their ability to detect and exploit linear and nonlinear cue–criterion relationships in the task environment. Using a field experiment on new product forecasts in the music industry, our analysis reveals that when forecasters are concerned with predictive accuracy and only managerial judgments are employed, providing both types of decision support data is beneficial. However, if judgmental forecasts are combined with a statistical forecast, restricting the decision support provided to human judges to contextual anchors is beneficial. We identify two novel interactions demonstrating that the exploitation of nonlinearities is easiest for human judgment if contextual data are present but historical data are absent. Thus, if the role of human judgment is to detect these nonlinearities (and the linearities are taken care of by some statistical model with which judgments are combined), then a restriction of the decision support provided makes sense. Implications for the theory and practice of building decision support models are discussed.

1. Introduction

The accurate prediction of the commercial success of newly launched products or services represents a crucial managerial problem (Steenkamp et al., 1999; Stremersch and Tellis, 2004; Van den Bulte and Stremersch, 2004). Generating such forecasts can be extremely difficult, particularly in environments involving fashion-oriented consumer products (hereafter referred to as "fashion products"), where the nature of the products may contain a substantial creative, artistic component, and consumer taste constantly changes (Christopher et al., 2004; Hines and Bruce, 2007; Hirsch, 1972). In more conventional forecasting domains, for example when predicting the demand of machine spare parts in manufacturing (e.g., Sani and Kingsman, 1997) or when estimating electricity demand (e.g., Taylor and Buizza, 2003), large amounts of historical data are often available to calibrate decision support models and achieve high levels of model accuracy. Forecasts about the demand of fashion products, on the other hand, often lack such integral information as the demand pattern tends to be highly uncertain (Choi et al., 2014; Green and Harrison, 1973; Sichel, 2008; Sun et al., 2008).

Companies producing fashion products for which consumer tastes and preferences cannot be tracked continuously often perceive themselves more as trend-setters than trend-followers (Eliashberg et al., 2008). For instance, in the apparel industry, firms face the challenge of quickly commercializing new designs that are introduced during the New York or Paris Fashion Week in order to create and satisfy new consumer demand. Yet, forecasts about future sales of new clothing designs are highly volatile. They depend both on managers' ability to accurately anticipate uncertain consumer preferences and on their firm's time-to-market capability relative to its competitors. In such supply-driven environments (Moreau and Peltier, 2004), conventional time series methods typically cannot be employed to predict demand with reasonable accuracy (Eliashberg and Sawhney, 1994; Moe and Fader, 2001; Sawhney and Eliashberg, 1996). Instead, researchers have proposed several approaches to overcome the problem of model calibration when only limited and/or unreliable data are available (for an extensive review, please see Nenni et al., 2013). In particular,

☆ Please note that an article based on portions of the same empirical data sample has appeared in Organizational Behavior and Human Decision Processes (Volume 120, January 2013). Cross-references are used where appropriate.

* Corresponding author. Tel.: +34 915689600.
E-mail address: matthias.seifert@ie.edu (M. Seifert).

http://dx.doi.org/10.1016/j.jom.2015.02.001
0272-6963/

these approaches have relied on Bayesian estimation when new sales data becomes available (e.g., Green and Harrison, 1973), Fourier analysis (Fumi et al., 2013), binomial distribution models (Cachon and Fisher, 2000), the Croston method (Snyder, 2002), two-stage dynamic sales forecasting models (Ni and Fan, 2011), artificial neural networks (Au et al., 2008; Gutierrez et al., 2008; Yu et al., 2011), fuzzy logic (Thomassey et al., 2005), extreme learning machines (Sun et al., 2008; Xia et al., 2012), and hybrid intelligent models (Choi et al., 2014; Wong and Guo, 2010). Furthermore, quick response manufacturing strategies have frequently been proposed as efficient means to shorten production lead times and gather "early sales" signaling data to reduce demand uncertainty (e.g., Cachon and Swinney, 2011; Fisher and Raman, 1996; Iyer and Bergen, 1997). However, despite the growing variety of managerial approaches and quantitative models at our disposal, the practice of forecasting product and service success continues to crucially depend on human judgment (Sanders and Manrodt, 2003; Boulaksil and Franses, 2009).

Past research has therefore emphasized the importance of understanding how and when managerial judgment contributes to improving forecasting accuracy. For example, in the context of newsvendor decision making, accurately estimating product demand before the selling period is extremely important in order to minimize inventory costs and avoid lost sales (Bolton and Katok, 2008; Bostian et al., 2008; Bendoly et al., 2006; Schweitzer and Cachon, 2000). Lee and Siemsen (2015) showed that decomposing the forecasting task from the ordering task in these contexts can improve judgmental accuracy under certain conditions. Task decomposition can therefore be seen as a possible means to *debias* the so-called pull-to-center effect that has often been found in behavioral studies of the newsvendor problem (e.g., Ho et al., 2010; Ren and Croson, 2013; Su, 2008). Moreover, Blattberg and Hoch (1990) demonstrated in the context of catalog fashion sales how buyers' demand predictions could be improved by relying on an equally weighted combination of model and expert forecasts. Gaur et al. (2007) emphasized the effectiveness of using dispersion among expert judgments to predict the actual uncertainty of demand for fashion products. Eliashberg et al. (2008) also concluded that managerial judgment is essential for the scheduling of motion pictures and the prediction of box office results. Furthermore, Lawrence et al. (2006) argued that managerial judgment is likely to prove valuable whenever the "ecological validity" of formal models is low, which is indicated by a low fit between model and forecasting environment.

The purpose of this paper is to study the conditions that determine the effectiveness of judgmental forecasting in environments involving fashion products. The sense-making mechanism underlying expert judgment is often viewed as a pattern-matching process during which forecasters perceive informational stimuli of a forecasting event and make their inference by comparing them to similar situations experienced in the past. As forecasting performance is limited by judges' ability to retrieve comparable situations from their memory, providing more *historical cases* to increase the possibility of finding a good match may therefore support their judgmental processes (Hoch and Schkade, 1996). A second approach to aid decision making is to support the strengths of expert judgment by providing forecasters with better *contextual data*. Numerous studies have demonstrated that the strengths of expert judges lie in particular in their ability to diagnose new variables, recognize abnormal "broken leg cues", and evaluate information that is otherwise difficult to quantify (Blattberg and Hoch, 1990; Einhorn, 1974). Providing judges with more contextual data may therefore enable them to make better sense of a forecasting event. Such contextual data might comprise product-specific information relative to promotional activities, manufacturing data, and more general domain knowledge including competitor data or macroeconomic forecasts (Lawrence et al., 2006).

We extend Hoch and Schkade's (1996) work by offering empirical insights into the usefulness of decision support approaches relying on historical cases and on contextual data in a setting characterized by high uncertainty. Whereas research on the general accuracy and appropriateness of judgmental forecasts has a long-standing tradition, only few studies have attempted to decompose the forecasting context into different types of knowledge components in order to examine their effect on judgmental performance. Among these, Blattberg and Hoch (1990), Stewart et al. (1997), and Seifert and Hadida (2013) have studied whether judgmental forecasts can add value beyond the predictions of linear models. However, these studies rest on the assumption that human judgment is capable of approximating the linear regression model of the environment fairly well. This assumption appears questionable, at least in some forecasting contexts, when considering that the information processing capacity of forecasters is limited. In addition, while Lee and Siemsen (2015) as well as Seifert and Hadida (2013) acknowledge that differences in forecasting effectiveness may depend on task structure, a more systematic decomposition of the forecasting environment is required to fully understand how task characteristics fundamentally influence judgmental performance and how decision support systems should be designed to improve predictive accuracy.

To study how historical demand anchors and contextual anchors interactively influence the performance of *human judgment*, we employ a judgment analysis approach (Cooksey, 1996; Hammond, 1996). Judgment analysis allows us to analyze managerial predictions beyond forecasting accuracy by decomposing judgments into two different components. First, we focus on the degree to which a manager's interpretation of a forecasting event matches the efficiency of a linear model. Second, we measure the extent to which a managerial judgment can reduce the residual variance of the linear model by interpreting contextual knowledge surrounding a forecasting event and, hence, add predictive value over and beyond the linear model. Our empirical context is the music industry, which, due to its creative and artistic nature, can be understood as a typical market for fashion products (Santagata, 2004). Specifically, we study forecasts about the Top 100 chart entry positions of upcoming pop music singles. While past research has primarily focused on sales predictions in the domain of fashion retailing, the music sector appears to be a particularly interesting forecasting domain where product success is highly contingent on the subjective evaluation of a multitude of independent industry gatekeepers (Vogel, 2007).[1] Our study reveals that when the primary concern is to maximize predictive accuracy and forecasts are based on human judgment *only*, providing both types of decision support data is likely to minimize forecast errors. However, more interestingly, our results also showcase the ambivalent nature of decision support anchors as the presence of both types of data appears to improve forecasters' ability to exploit nonlinearities, while impairing their effectiveness of interpreting linearities in the task environment. In fact, our field data indicate that the exploitation of nonlinearities is easiest for human judgment if contextual data are present but historical data are absent. Thus, if the objective is to use human judgment to maximize the explanation of nonlinearities surrounding the forecasting task, we suggest that decision support should be restricted

[1] In the music industry, rank predictions are highly correlated with actual sales forecasts. Specifically, our secondary data analysis of Top 40 *album* chart ranks in the UK between 1996 and 2003 and their associated average weekly sales levels reveal a strong negative correlation of $r = -0.74$, indicating that smaller ranks are associated with higher sales levels. A corresponding analysis of the Top 40 *singles* charts resulted in a similarly strong negative correlation of $r = -0.68$.