Management Science

Working Paper 2020:1

Stability and innovation in the use of forecasting systems: a case study in a supply-chain company

Robert Fildes

The Department of Management Science
Lancaster University Management School
Lancaster LA1 4YX, UK

Paul Goodwin

The Management School,
University of Bath, Bath, BA2 7AY,UK.
Email: p.goodwin@bath.ac.uk

Stability and innovation in the use of forecasting systems: a case study in a supply-chain company

Robert Fildes

Department of Management Science, Lancaster University Management School,

LA1 4YX, United Kingdom.

Email: r.fildes@lancaster.ac.uk

Paul Goodwin

The Management School, University of Bath, Bath, BA2 7AY, United Kingdom.

Email: p.goodwin@bath.ac.uk

January 2020

Stability and innovation in the use of forecasting systems: a case study in a supply-chain company

ABSTRACT

Computer-based demand forecasting systems have been widely adopted in supply chain companies, but little research has studied how these systems are actually used in the forecasting process. We report the findings of a case study of demand forecasting in a pharmaceutical company over a fifteen-year period. At the start of the study managers believed that they were making extensive use of their forecasting system that was marketed on the basis of the accuracy of its advanced statistical methods. Yet the majority of forecasts were obtained by using the system's facility for judgmentally overriding the automatic statistical forecasts. Carrying out the judgmental interventions involved considerable management effort as part of an S & OP process, yet these often only served to reduce forecast accuracy. This study uses observations of the forecasting process, interviews with participants and data on the accuracy of forecasts to investigate the reasons underlying the managers' use of the system at two levels, the individual and the organizational. This evidence is then interpreted using various theories to understand the longevity of the company's forecasting process, despite potential economic benefits that could be achieved through change. However, 10 years after the original case observations radical transformations of the forecasting system were introduced. The paper concludes by considering the impetus for adopting the new system and processes, and the changes in organizational practices this has led to.

Keywords: behavioural operations forecast adjustments; forecasting support systems; judgmental forecasting; actor-network; cognitive biases; task-technology fit; organizational factors.

1 Introduction

Accurate forecasts are crucial to the success of supply-chain companies and decisions relating to transportation, purchasing, inventory control, work-force scheduling, production planning and cash-flow planning are all dependent on them. Given the importance of accurate forecasting, the development and sales of computer-based statistical forecasting systems has become big business, with major suppliers such as SAP and JDA embedding forecasting modules in their advanced planning and retail offerings, while statistically focussed software providers such as SAS also supply companies with forecasting modules that are used in demand planning. In addition there are many focused providers (see Fildes, Schaer & Svetunkov, 2018 for a survey). However there is a great deal of evidence that the demand forecasts generated by such systems are routinely replaced by forecasts based on managers' judgments (Fildes & Goodwin, 2007; Fildes et al., 2009; Fildes & Petropoulos, 2015). Despite this the designers of these systems have paid little attention to the role of judgmental interventions (Goodwin, 2015) or the organisational environment within which their products are deployed (Asimakopoulos, Dix & Fildes, 2011; Asimakopoulos & Dix, 2013; see also Arvan et al., 2019).

In the information systems literature relatively few papers have considered the on-going use of such tools as forecasting support systems (FSSs). (Examples are provided by Auer, 1998; Bagchi, Kanungo, and Dasgupta, 2003; Ruivo et al. 2014, while Venkatesh, Thong and Xu, 2016 have reviewed 'acceptance and use' including adoption, showing the primary research emphasis has been on adoption.) In the forecasting literature the role of organizational process factors in affecting how such systems are used and their on-going performance has also largely been neglected with such case-based research the exception: examples are provided by Smaros, 2007; Oliva & Watson, 2009; Phillips & Nikolopoulos, 2019. Instead, most findings about forecasting in organizations have being obtained through questionnaire-based surveys which have focussed on respondents' reports of the methods used in the responding organization and the balance between statistical methods and judgment (e.g., Fildes and Goodwin, 2007; McCarthy et al., 2006). More holistic researches have included a substantial study of the forecasting function in organizations based on case research in 16 companies in the US by Moon, Mentzer and Smith (2003). They categorized the organizational issues in forecasting as occurring in four areas: integration across management functional areas, the methods adopted, the FSS and performance measurement. Together these can lead to poor and limited processing of data, the use of inadequate algorithms, and finally, the ineffective interaction between the users and the system through both the choice of method and the adjustment of the statistical forecast (e.g., Goodwin et al., 2007; Fildes, Goodwin & Onkal, 2019). The Moon et al. study placed forecasting support systems at the heart of a successful process. This was confirmed more recently when Doering, and Suresh (2016) showed the use of 'advanced [forecasting] systems' to be a determinant of overall competence as it affects, costs, service and accuracy.

A more granular approach has examined the detailed forecasts produced by organizations. A field study of thirteen manufacturing companies by Lawrence et al. (2000) revealed that their forecasts tended to be highly inefficient but it did not investigate how or whether the companies used computer-based forecasting systems. In a study of four companies Fildes et al. (2009) found that computer-based forecasts were being used inefficiently. Franses (2013) found similar inefficiencies and these various studies are summarized in Perera et al. (2018). The inefficiencies are thought to result for a wide variety of factors, both individual and organizational, within the constraints imposed by the FSS technology.

Exactly how these factors, individual, organizational and technological, interact is the subject of this paper. The approach adopted here uses an in-depth study of forecasting in a pharmaceutical company to investigate how managers in the company used their computer-based forecasting system and why this often led to unduly inaccurate forecasts and an inefficient forecasting process.

Moreover, it seeks to explain why this ‘suboptimal’ situation persisted and how and why fundamental changes were eventually only implemented after a considerable amount of time.

The rest of the paper is organised as follows. First, we investigate the existing literature in relation to this issue of stability and subsequent change in the configured company forecasting process . We then explain why we have adopted a case-study based approach to address our research questions before explaining our data gathering process. The next section describes the company, its decision to purchase the forecasting system and the way in which the software was being used. Statistical evidence is used to present an additional perspective. Then we interpret our results to provide explanations of why the system was used in this way, why this state of affairs persisted for several years before changes were made and how the changes came about. Finally, we present our conclusions, together with the practical implications of our results.

The paper’s contribution is that it provides a novel in-depth understanding of how an inefficient demand forecasting process with a forecasting support system at its heart can exist for many years without being challenged. For researchers in forecasting and operations it offers a rare case study that contrasts the technical with technology-in-practice’; such a focus emphasizes the importance of software design fitting with organizational processes. It also demonstrates what may be required in order for changes to eventually be implemented. By providing an account of the individual and organizational processes that are present in company forecasting and the motivations and interests of the key actors it is hoped that the paper will assist those whose mission is to improve company forecasting by capitalizing on the value-added potential in the organizational processes surrounding forecasting.

2 Using forecasting support systems – a literature review

2.1 Using a forecasting support system

The research literature has paid little attention to the use of forecasting support systems

(FSSs) in organizations, a curious omission since they are part of a multimillion dollar business with many product offerings (see for example the survey by Fildes et al., 2018). Asimakopoulos, Dix, & Fildes (2011) carried out an in-depth set of interviews[1] with forecasters that decomposed how they produced their forecasts to propose a nuanced set of tasks undertaken (within a task hierarchy) that aimed to capture the observed complexity of organizational forecast activity This was compared with the standard normative view as typically presented in the forecasting literature. Much of what was captured demonstrated that tasks were undertaken that lay outside the standard forecasting processes explicit in the software design. The core tasks identified through interviews with the software designers were data exploration, the choice and fitting of statistical models and the production and evaluation of the resulting model forecasts. The forecasters required additional flexibility including the incorporation of 'special factors' such as a sales promotion or the impact of weather (perhaps a spreadsheet row permitting adjustments) and whether the forecast 'makes sense' (a graphical check).

Data exploration involved plotting historical data, changing the data length and checking fit, which is not necessarily feasible in all FSSs. The designers interviewed emphasized the individual forecaster's needs giving less weight to collaborative practices. For some users in some organizational contexts, this suggests there would be a fit with the software and for others a mismatch. Smith and Mentzer (2010) showed that the users' perceptions of the task-technology fit influenced positively reported forecast performance. A key task for users was "knowledge sharing […where] the FSS played a key role in shaping a shared forecasting meaning and in fostering interactions among relevant stakeholders" (Asimakopoulos & Dix, 2013), and insofar as the FSS facilitated these interactions, they argued the fit was improved. This, though was not explicit in the software's design.

Qualitative interactions and forecast adjustments have thus been established as a key set of tasks within supply chain forecasting, whether or not they enhance performance. The existing

[1] System designers were also interviewed. Two FSS users interviewed were part of the forecasting team in the case organization we study here.

psychological literature points to a number of reasons for this heavy reliance on the forecaster's judgmental interventions in the company's forecasting process. Clearly, some interventions are easy to justify. These will be where the forecaster has information, not available to the statistical system, about future events that are highly likely to have a large impact on demand. The work of Payne, Bettman, and Johnson (1993) suggested that people seek to balance cognitive effort with accuracy considerations when making judgments and decisions. Making an adjustment involves more effort than the simple acceptance of a statistical forecast so a forecaster making judgmental interventions must perceive that there are benefits to be gained through this extra effort. In many cases these benefits will be political in that forecasters may deliberately bias their forecasts to try to gain advantage in the organisation (Fildes & Hastings, 1994; Galbraith & Merrill, 1996; Oliva & Watson, 2009). However, accounting for adjustments where there is a genuine desire to achieve forecast accuracy, but where such adjustments prima facia do not enhance performance, requires a more elaborate explanation.

Kleinmuntz (1990) has suggested that one reason why people prefer to use their heads (i.e. judgment) rather than formulae, is 'deluded self confidence', which he defines as confidence that you will beat the odds because you have real expertise in a domain. Kleinmuntz concludes that "people are indeed not as good as they think they are at using their heads". People also appear to be more tolerant of errors in human judgment compared to errors produced by algorithms. Dietvorst. Simmons and Massey (2014) found that people soon lost confidence in an algorithm when it erred, leading to a phenomenon that they termed 'algorithm aversion'. Considerations like these may account for the findings in a study by Önkal, et al. (2009) where people made larger adjustments to forecasts they thought emanated from an algorithm than those they thought came from a human expert, even though the forecasts were identical.

Unlike the indications of systems such as a satnav it is inevitable that the forecasts produced by a demand forecasting system will be perceived to be inaccurate, not the least because of the noise associated with demand. The psychological literature on advice-taking suggests another reason why the imperfection of a support system might be exaggerated in the eyes of users. This research suggests that the weight attached to advice is dependent on the reputation of the adviser, but negative information about an adviser is perceived to be more diagnostic than positive (Yaniv & Kleinberger, 2000). If we regard a statistical forecast as a form of advice, albeit from a machine, then errors arising from noise and special events may likely diminish the system's reputation. Indeed, Kaplan, Reneau, and Whitecotton (2001) found that people were more likely to rely on a support system when its accuracy was not disclosed.

This apparent imperfection in forecasting systems is also likely to be overemphasised because the environment may be thought of as largely predictable (Dawes ,1979). In particular, there is much evidence that humans have a poor conception of randomness (e.g. Falk and Konold, 1997). When confronted with randomness, they have a tendency to perceive patterns and causes (Siegrist, Cvetkovich, & Gutscher, 2001; Heuer, Merkle & Weber, 2016). This leads to the belief that greater mental effort will improve the accuracy of forecasts (Davis & Kottemann, 1994).

These considerations imply that it is important to distinguish between the acceptance of a forecasting support system per se and acceptance of its automated output. In Davis's widely cited technology acceptance model (TAM) (Davis, 1989) the perceived usefulness of a system (defined as "the degree to which a person believes that using a particular system would enhance his or her job performance") is a key driver of its acceptability. Perceived usefulness is likely to be to be increased when the system allows users to manipulate aspects of the task, for example by changing parameters or overriding the forecasting method's automatically selected by the system. However, actions like these will also be associated with a tendency to reject and override the system's automatic output. While involvement with the task is likely to be associated with greater acceptance of the system, it can also lead to an illusion of control, which leads to further overconfidence (Davis & Kottemann, 1994). Illusion of control occurs when factors that are usually associated with good performance in skilled situations are found in tasks where the outcome is largely or wholly governed by chance. This increased confidence and acceptability is not, however, necessarily associated with greater accuracy. The manipulations that the user is permitted to perform might substantially reduce the accuracy of the system's recommendations, when compared to the advice it would have produced automatically (Lawrence, Goodwin & Fildes, 2002).

In summary, the literature suggests that individual forecasters will embrace an FSS when the system design is flexible enough to accommodate the various tasks the forecaster and stakeholders engage in around the production of supply chain forecasts. However, they are likely to use the forecasting system inefficiently because of their over-confidence in their supposed expertise, their propensity to see patterns and causes when none exist, their exaggerated distrust of a system known to be imperfect and their need to control the outcome.

2.2 Organizational interactions potentially influencing FSS use

Individual forecasters in organizations do not usually work in isolation. Their job requires that they interact with other people from both inside and outside the organization in order to acquire information or to explain their forecasts. As a result, it is also important to consider the potential effects of political, social and other influences on the way that they use forecasting support systems and the fit between the FSS and its stakeholders, as the technology is interpreted in practice. The forecasting literature has seen very limited organizational analysis. An early review of forecasting practice touched on process issues but failed to find any in-depth studies (Winklhofer, Diamantopoulos, & Witt, 1996). In a wide-ranging case study, Fildes and Hastings (1994) identified the tasks forecasters undertook, their intra-organizational interactions and the limited information shared between functions, and the lack of such information in the FSS as important limitations on forecasting effectiveness. The case organization was analysed in terms of the credibility and importance of forecasting and the organizational motivation to improve. Improvements, the forecasters thought, were to be found in enhanced data (in the FSS), better software including an enhanced role for judgment as well as better methods. The major cross-organization study by Moon et al. (2003) identified 'degree of communication, coordination and collaboration' across functional areas and the development of a consensus forecasting as important characteristics in effective organizational forecasting. They highlighted the role of the forecasting system in making information available: the FSS, the information it contains, the forecasters and other organizational actors through their interactions all contributed to the effectiveness of the forecasting process. However, none of these studies, nor the practitioner literature focussing on 'Sales and Operations Planning' [2] have examined <u>how</u> forecasters, carrying out their tasks, <u>interact</u> with other organizational actors through the FSS (Tuomikangas and Kaipia, 2014).

Beyond the field of forecasting, a range of models have been developed to try to explain how the interplay between the technology itself and these social interactions determines the way in which technology is used (Orlikowski,1992). For example, the Unified Theory of Acceptance and Use of Technology (UTAUT) model (Venkatesh et al., 2003) posits that the way technology is used will be influenced by "the person's perception that most people who are important to him [or her] think he [or she] should perform or not perform the behaviour in question" (Fishbien and Ajzen, 1975). These models vary in the relative importance that they place on the role of humans and technology in influencing each other and in shaping working practices. For example, at one extreme, the technological imperative model (e.g., Siegel et al., 1986) implies that technology is an exogenous influence on human behaviour and organisational attributes. This view discounts the actions of humans in developing, appropriating and changing technology and assumes that people operate and behave like machines. Under this assumption, forecasters would "compliantly carry out the orders and commands they received, making efficient and effective use of all information and all technologies available to them" (Davis et al., 1992).

This viewpoint can be contrasted with models which adopt a 'social construction of technology' perspective. These models recognise that forecasting is carried out within a complex

[2] Tuomankangas and Kaipia (2014) give a synthesis of the existing literature, both practitioner and academic.

social context, that the forecasting system will be understood differently by different individuals and that the meaning attached to the forecasting system will be determined by a shared understanding arising from social interaction. In this perspective the design, shaping and use of a forecasting system would be seen as resulting from political actions and negotiations between a multiplicity of stakeholders (Orlikowski, 1992; McGovern & Hicks, 2004). While this perspective acknowledges that there is duality in that people and technology interact reciprocally it has been criticised for understating the role of technology and its characteristics in this process (Hanseth, 2004). The dichotomy between human and non-human entities is avoided by using the concept of an actor, being any element which has the power to initiate action and can be either a human, a collection of humans or an item, as here, such as a forecasting system (e.g. Latour, 2005 in proposing his Actor-Network theory). We have adopted a social construction of technology perspective in understanding the forecasting processes we have observed while recognizing Hanseth's criticisms.

3 Research Method

As our literature review revealed issues of organizational practice have enjoyed little research attention (for a summary see e.g. Fildes, 2017): the focus of forecasting research has primarily been on modelling methods and the evaluation of different methods. However, the objectives of the research we report on here are quite different in that we seek to understand how forecasters go about their organizational tasks when using these methods through a forecasting support system. We therefore need a more in-depth approach.

3.1 The Case Study Approach

For a deep account of the reasons underlying the way managers use and interact with their forecasting systems we needed to understand their perspectives, beliefs and motives because it is reasonable to assume that, according to these characteristics, they are behaving rationally as individuals (Kanter, 1977). Such an understanding is unlikely to be achieved through a 'traditional' survey device like a postal questionnaire, though to date this approach has been the predominant method for investigating forecasting practices in companies (e.g. Dalrymple, 1987; Klassen & Flores, 2001; Mady, 2000; Sanders & Manrodt, 1994; Fildes & Goodwin, 2007; Fildes & Petropoulos, 2015). As the literature review makes clear, what is required is a study of a real forecasting process undertaken at two levels –first at the level of the individual forecaster and then at the level of the network of all the participants in the process. Our research also required a methodology that could reveal a deep understanding of reasons for the managers' behavior and of the social and organizational context in which they operate. We therefore adopted an interpretive approach (Walsham, 1995; Nandhakumar & Jones, 1997; Easterby-Smith et al., 2011) based on direct observations of the forecasting process and semi-structured interviews with participants. We also triangulated the results obtained with statistical data relating to the forecasts. From this we have developed an explanation of the way that individual managers used and perceived the system and how these individual psychological factors were combined with other forces that existed both within and outside the organisation so that any pressures to change the existing use of the system would be likely to be supressed.

Although we have observed similar behaviour to that described below in several companies, in common with many other case-based studies we decided to focus on one company. There are a number of advantages to this single organization case study approach (Walsham, 1995). In particular, it allows a situation to be studied in depth and from a range of perspectives.

3.2 Data gathering

The company we studied was, and remains in 2019, a cost-conscious UK subsidiary of an American pharmaceutical company embracing a number of business units, which prides itself on its application of modern management methods: Aspects of this organization have been disguised for reasons of confidentiality, both organizational and personal. The company supplies products for treating both animals and humans. Initially, the research study involved visits to the company's local headquarters by either two or three researchers over a period of around six months. The first visit included a presentation by the company's managers, followed by a general question and answer session. The researchers also observed two meetings where the forecasts were finalised and they conducted semi-structured interviews with the main participants in the forecasting process: two Logistics Managers, a Product Manager, a Marketing Manager, a Finance Manager, a placement student who was acting as a Commercial Manager, and a Stock Replenishment Planner. Independent sets of notes were taken by each of the researchers at all meetings and interviews, which were also tape-recorded. The research team sent a summary of their understanding of the organisation and its forecasting process to the company for validation and any necessary corrections. In addition, members of the team attended two user conferences run by the company which supplied the forecasting system and also interviewed two of their software developers. Note that, while our approach is predominantly interpretive, we have also had access to statistical evidence and have used this to triangulate and strengthen our findings (Nardulli, 1978; Benbasat, Goldstein, & Mead, 1987). This consisted of a sample of 3264 forecasts that were supplied by the company and for which the actual outcomes were known. Following the initial field work, after major changes had been made to the company's forecasting process, we interviewed managers who had implemented the new system or who had been involved with the process during the change.[3] This has allowed us to discover why the changes had occurred and how they had been made possible.

4 The forecasting system and forecasters

At the start of our study in 2004 the UK subsidiary, responsible for a number of European countries, was using a forecasting support system that it had bought seven years earlier. Other regions were using different software including relying on Microsoft Excel. These FSSs were used to forecast the monthly demand for its products worldwide. The UK system went through various

[3] We do not suggest the study was initially seen as longitudinal, but after reviewing the continuing lack of research in this area, we sought out new contacts.

upgrades and changes in the supplier's ownership, remaining in use until 2015. The original system had been bought 'off-the-shelf' (as opposed to being an in-house development), with some 'personalised' settings tailored to the company by the supplier based on an analysis of the company's sales data. At the time of the purchase it was thought that a forecasting system was needed "to do the job properly", as one logistics manager explained. Before this, individuals had made their own forecasts, often using a ruler to fit a line to paper copies of sales graphs. The system was chosen by a group of middle managers over two alternatives (including an Enterprise Resource Planning (ERP) system) with a 9 to 1 vote in its favour. The choice was primarily driven by the perceived 'user-friendliness' of the system, and the marketing and sales people commented that they particularly liked the ease with which the forecasts could be changed to reflect managerial judgment.

The system was perceived as being "fairly extensively used" [in the words of one logistics manager] in producing forecasts. Its use was regarded as a big improvement on the previous approach and managers felt that forecasting accuracy had also improved –though no empirical data existed to support this. The system was regarded as "the best available" [this quote is from the same logistics manager] and, while some users had complaints about particular facilities, most were generally satisfied with it. Its perceived central role in the forecasting process was never questioned and no one, in the meetings we had with participants, suggested switching to an alternative system or making other fundamental changes to the existing forecasting process.

The system consisted of a database and query language, various statistical forecasting algorithms, graphical facilities and an interactive component which permitted the user to adjust the statistical forecast that the software has generated. The adjustment could be performed directly by changing the forecast that the method had produced, by changing the parameters of the algorithm or the number of historic observations used to fit the forecasting model. Changing the parameters of the algorithm allowed the users to select different models, such as simple exponential smoothing or smoothing with trend and seasonal components. Systems like this are marketed by vendors largely on the basis of the accuracy and sophistication of their in-built automatic *statistical* methods. For example, at the time SAP claimed that its Forecasting and Replenishment for Retail (SAP F&R) software provided a “sophisticated demand forecast” with automated calculations. These marketing postures fairly reflect (and reflected) the motivation of demand planners where accuracy has always been seen as the primary objective (Fildes & Goodwin, 2007).

At the start of the study, there were three logistics managers who were responsible for the initial forecasts for around 350 stock keep units (SKUs) and for managing the inventory. The forecasting process also involved fifteen product managers, who looked after the sales of groups of products, as well as financial and marketing managers. In addition, a placement student was spending a year using the system to produce forecasts for products with ‘well behaved’ demand patterns. Interestingly, the marketing and sales staff took a keen interest in the forecasts for their products, which contrasted with the attitude of staff in some other companies visited by the research team. In these other companies sales personnel apparently saw their objectives as maintaining customer relationships and making deals and, in consequence, had little interest in numbers, forecasts and computer systems.

Three years before our study, the company had adopted the Six Sigma approach[4] to managing. Two logistics managers had achieved ‘green belt’ status, while another manager was in the process of achieving the higher level of ‘black belt’. This indicates that they had achieved a proficiency in the application of statistical tools to management problems, although they had no training in statistical forecasting methods and minimal training in the use of the FSS (the manual had been lost a long time earlier). One of the Six Sigma projects that coincided with our study concerned the company’s forecasting process.

[4] Six Sigma is a data driven method for eliminating defects in any process – including those used in manufacturing and service industries.

Forecasting had been selected for the Six Sigma treatment because, i) in the words of one manager: "it took an enormous amount of time, effort and resources and pain to produce the various forecasts" and ii) there were concerns about forecast accuracy. At a rough estimate, forecasting was taking around 80 person-hours of managers' time each month in meetings alone (see later) and the managers clearly wanted to see this effort rewarded with higher levels of accuracy. Because of this the forecasting improvement project was regarded as "a big strategic project".

4.1 How the forecasting support system was used

The primary objective of the forecasters was to produce forecasts of demand looking forward two months, reflecting the production planning requirements of the company's manufacturers. In *theory*, the derivation of the forecasts involved two main tasks. First, a logistics manager cleaned the sales history to remove the effects of stockouts (these are known from data on orders) so that the series represented the level of demand. They then used the system to produce the 'base-line' forecasts. These were forecasts which take no account of market intelligence (MI) and were simply based on an extrapolation of past demand patterns. Secondly, these base forecasts were presented at a forecast review meeting where they were judgmentally adjusted for MI to produce the final forecast. As we will show, the actual practice of producing the forecasts involved some blurring of these stages. MI was sometimes used in setting the baseline forecasts at stage 1, while recent past patterns in the demand were sometimes used as a reason for adjustment at review meetings rather than MI.

A particular difficulty in producing accurate forecasts arose because of the effects of cross border trade (CBT) where customers buy the company's products from overseas subsidiaries, usually at a lower price. This resulted in many unforeseen fluctuations in the demand data that were used when producing the forecasts. The degree to which CBT had impacted on the most recent observations was also difficult to ascertain as it took time to obtain information on the level of this activity. Apart from CBT, there were many other uncertainties in the market, such as the outcomes of tenders, competitors' actions and consumers' behaviour. For example, with animal medicines, farmers might switch brands when the drug ceased to be effective because bacteria had become resistant to its effects.

Task 1 Obtaining the baseline forecasts

The patterns of the demand history varied according to the product types. However, the forecasters explained that most products had a life-cycle which caused their underlying pattern to have a non-linear trend (see fig 1). From this perspective, in the early years of a product's life it took time for demand to build up as doctors needed to be persuaded to prescribe the drug. Following this, the product experienced a mature phase of demand, before finally losing its patent protection. This caused sales to decline as generic products were marketed at a lower price. The figure also shows two months ahead forecasts at various stages of the product life cycle where a forecasting method similar to that embedded in the company's FSS, ETS (Hyndman & Khandakar, 2008), has been used to produce them.

Insert figure 1 about here

Notwithstanding the demand forecasters' belief in the product life cycle, the statistical methods embedded in the system were designed to extrapolate linear trends as the figure shows. Despite the adequacy of these linear forecasts for parts of the cycle, one forecaster explained that, to try to adapt the system's automatic extrapolations so that they matched the perceived life-cycle pattern, they proceeded as follows**:**

a) They selected an 'appropriate' length of demand history, for a given product so that the system generated a trend line that gave the best fit to the selected data, using the least squares criterion. Usually, two years of past data were used, but it could be much less (e.g. six months) and by manipulating the length of the demand history, a more acceptable trend line could often be obtained. The two-month ahead forecasts were then calculated from an extrapolation of this trend line.

b) To further improve the apparent fit of the trend line to the past data and also to obtain forecasts that "looked right", the forecasters often used their judgment to override the forecasts obtained in (a). This could be simply achieved by using a mouse to reposition the trend line on the graph. For example one forecaster stated "I'd actually re-model it [the statistical forecast] using different lengths, different levels, different trends and try and make … or use [the] system to generate, a more, what's the word I'm looking for, a more reasonable forecast". One forecaster admitted: "I don't know how it [the forecasting system] calculates" and said that sometimes s/he was surprised at the system's extrapolations judging them to be unrealistic. In some cases the forecasters were ostensibly trying to model the non-linear trends resulting from the perceived product life-cycle by fitting and adjusting linear trends to relatively short sequences of past data. (We will discuss later whether the perception that linear extrapolations were inadequate for short-term forecasts, was correct or, indeed, whether this was merely a pretext for intervention.)

Task 2 Incorporating the effects of market intelligence (MI)

The system's displays of the forecasts resulting from Stage 1 were presented on a large screen at one of 17 monthly product group review meetings. As mentioned above, the main purpose of these meetings was officially to allow the forecast to take into account market intelligence (MI). One forecaster stated: "I guess the most important task in terms of the forecasting system[s] is actually the meetings that we have where we use it and we produce the new forecast". The attendees at the forecast review meetings were the relevant product manager, whose role was to adjust the forecast for MI, the relevant forecaster, who might challenge these adjustments, and representatives of the market research, finance and commercial functions. One of the logistics managers said "years ago we [Logistics] owned the [forecasting] process; [Marketing] owned the forecasts". Since then, senior management had insisted that all parties at the review meetings had to jointly own and agree the forecasts.

The review meetings that were observed differed in character. For example, the first meeting concerned forecasts for animal products. Here, the forecasts that were agreed were based almost exclusively on the product manager's intimate knowledge of his market. These were never challenged. A meeting to forecast the demand for a human medicine had a number of contrasting characteristics. In particular, there was great emphasis on very recent demand history. As stated earlier, the forecast initially presented usually were based on, at most, two year's past data because "further back the trends tend to be different" [Quote from a Logistics Manager]. The appropriateness of this forecast was then assessed in a forensic discussion of very recent demand patterns, with particular emphasis on the last three months. An explanation was sought for every movement in the graph over these months, though reasons for these movements were usually unknown or highly speculative (e.g. "Why was October low and November high?" Answer: "…November is normally part of the wholesaler's build …. We always do better in November. Having said that we didn't last year, did we?") [Quotes are taken directly from tape recordings of the meeting].

The actual forecasting process can thus be summarised as:

Automatic statistical baseline forecast

→ Replacement with judgmentally derived baseline forecast

→ Further judgmental adjustment at Review Meeting to obtain final forecast.

4.2 The accuracy of the judgmental interventions

To investigate the effect of the judgmental interventions on forecast accuracy we carried out an analysis of the sample of 3264 forecasts that were supplied by the company. Because managers had kept no record of the original automatic statistical baseline forecasts that would have been generated by their system and because we had restricted access to this software we simulated these forecasts by applying the *Forecast Pro* forecasting system (www.forecastpro.com), in automatic mode to 24 consecutive months of past demand data. These simulated forecasts provided plausible estimates of the automatic statistical baseline forecasts produced by the company FSS as they were based on a similar algorithm. However, because only 24 months of past data was available to us, we may have underestimated the system's ability to produce accurate baseline forecasts.

First, we estimated the effect of replacing the system's automatic statistical baseline forecasts with judgmental baseline forecasts. We were unable to do this for baseline forecasts that had been subsequently adjusted for market intelligence as a record of the original automatic system-generated baseline forecast did not exist. Our analysis was therefore confined to the 37.7% of baseline forecasts that were recorded as unchanged –even though they may have differed from the original software's automatic forecasts. We assumed that a baseline forecast had been changed when the recorded baseline forecast differed from the baseline forecast produced by Forecast Pro. For *these* forecasts, the company's judgmental baseline forecasts were, on average, slightly less accurate than those automatically supplied, despite the extra effort entailed in producing them - their median absolute percentage error (MdAPE) was 12.0%, while the MdAPE for the automatic simulated baseline forecasts was 11.7%. Thus, to the extent that this difference in accuracy is not a result of our estimation process, it appears that the cost of ignoring the system's recommendations was primarily one of wasted management effort and time in their various meetings rather than serious damage to forecasting accuracy. These results also cast doubt on the life cycle as the demand forecasters conjectured.

62.3% of all the recorded baseline forecasts were subsequently judgmentally adjusted, ostensibly for MI. Did *these* adjustments lead to improved accuracy? Analysis of the sample indicated that moderate improvements were sometimes achieved: the MdAPE of the baseline forecasts was 17.3%, while that of the adjusted forecasts was 14.3%. However, only 51.3% of forecasts were improved through MI adjustment and the most successful adjustments tended to be larger. Less than 45% of the smallest adjustments (below the first quartile) improved accuracy while over 58% of the largest adjustments (above the third quartile) resulted in improvements. The size of the adjustment is probably a measure of the strength of the market intelligence possessed by the members of the review meeting. It seems therefore that only when the proposed adjustment was substantial was the effort of making the judgmental adjustment worthwhile.

How did managers perceive the quality of their forecasting process? The perception of one of the logistics managers was that they were "good on reporting error levels, but not good on using the data that they have to improve forecast accuracy" (e.g. stock level data that were available for some customers were not used). In particular, this manager thought that there was potential for improving their ability to learn from past forecast errors. However, despite the forecasting improvement project managers saw little need for *fundamental* changes in their use of the statistical system that they had purchased. Although the system's statistical methods played a limited role in the forecasting process, the final forecasts were largely perceived and presented by the managers as being the output of an advanced modern system –indeed they were referred to as the 'System forecasts'. Carrying out the judgmental interventions involved considerable management effort and time, which could only be justified economically if they had led to improved accuracy. But the company forecasters instead relied on their beliefs that their interventions were valuable without seeking any evidence that this was the case.

In many respects the observed forecasting process in the pharmaceutical company was contrary to the 'normative' approach suggested by forecasting research. In general, this indicates that managers should allow *appropriate* statistical forecasting methods to identify regular patterns in data (assuming that sufficient data is available for this purpose) (Goodwin, 2002) or alternatively, regard their judgmental approach to model selection as complementary to an automatic approach (Petropoulos et al., 2018). (Although, Petropoulos et al. found that judgmental model selection could lead to more accurate forecasts, the selection in their study was restricted to a choice between 'optimised' forecasting models. Forecasters did not have the ability to change parameters or the length of data history to which a model was fitted.) As we have discussed, manipulating a statistical model, that is designed to identify regular patterns may lead to forecasters reading false patterns into the noise that appears in time series (O'Connor, Remus, & Griggs, 1993). Further, judgmental interventions should only be applied to statistical forecasts when the forecaster has important information about forthcoming events that is not available to the statistical method. (Sanders & Ritzman, 2001; Fildes, Goodwin, & Lawrence, 2006; Sroginis, Fildes and Kourentzes, 2019)., Moreover, the size of these adjustments should be accounted for and their rationale recorded (Goodwin, 2002) and they should, in general, be true adjustments to statistical forecasts –accounting only for the extra information- rather than replacements of these forecasts (Goodwin & Fildes, 1999). Finally, any choice of method, or any decision on what length of series history to employ should be informed, as far as possible, by analysis of the historical data and past forecast accuracy.

This raises the question: why did managers in a company operating in a highly competitive environment adopt such an inefficient approach to an activity as crucial as demand forecasting? The question is important because this situation continues to persist in many forecasting processes in other companies (for a recent summary of its prevalence see Sroginis, et al., 2019). While the literature may provide part of the answer, it is important to examine whether other factors are involved.

5 Explaining the use of the system

Various theories purport to explain how individuals adopt, use and modify technology (here the FSS) so that their individual and organizational collaborative requirements are met. Orlikowski (2000) emphasized that such technologies need to be understood through the lens of 'technologies-in-practice' rather than immutable artefacts. The tasks that forecasters undertake (Asimakopoulos et al., 2011), as we have described in the previous section, lead to the system being 'bent' so that the normative models implicit in the system's design are modified to meet these individual and organizational requirements. We adopt these two lenses to understand what we have observed.

5.1 An individual perspective

The statistical forecasting system used in this company was designed to filter out the random noise that is associated with demand time series in order to identify the underlying systematic patterns so that this could be extrapolated into the future. However, the managers exhibited an intolerance of randomness and, consistent with Dawes (1979), they appeared to believe that almost every movement in their graphs had a predictable cause: "if you go into the numbers, look at the grid……open the plot , go back…and then understand why in 2001 there was a different pattern" (Asimakopoulos, Dix,, & Fildes, 2011, from an interview with a user). This tendency to see causes and explanations for random changes was apparently exacerbated when individual managers were regarded as experts in the factors that underlay the behaviour of a time series. For example, it was clearly difficult for a marketing manager to admit they could not account for all of the month-to-month increases or decreases in the demand for a product, even though many of these movements were probably inherently unpredictable. In addition, in seeking to explain these movements, hindsight bias (Fischhoff, 1975) is likely to increase the belief that the random movements could have been predicted.

As we have indicated, the facility in the forecasting system that allowed the judgmental manipulation of the base-line forecasts using a mouse was highly regarded by the forecasters and, consistent with studies on participatory design, was a major factor in the acceptability of the system. However, as we discussed earlier, such participation can be associated with an illusion of control which would further enhance belief in the predictability of demand.

In this company the devaluing of the automatic forecasts was exacerbated by the fact that some of the movements in the time series, which a statistical method will discount as noise, could be foreseen, at least in part. These movements were caused by special events for which there may have been little or no past data, thereby precluding statistical estimation. In these circumstances, the human forecaster, who is aware of the impending event, will usually improve on the statistical forecast by intervening (Goodwin & Fildes, 1999). However, the observable deficiency of the statistical forecast on these occasions apparently contaminated belief in the automatic forecasts on other occasions, when its errors were genuinely unpredictable (Goodwin & Fildes, 1999), an example of algorithm aversion (Dietvorst et al.,2015).

This belief that all or much of the variation in time series is explainable appeared to have another important effect. While a statistical method will usually characterise a time series as having a relatively simple systematic pattern overlaid with noise, the managers seemed to perceive the series as a set of individually explainable outcomes. This is associated with a propensity to use epistemic logic (where the focus is on the underlying causes of an *individual* event) rather than aleatoric logic (where the focus is on the *set* of observations and element specific information is ignored (Beach, Christensen-Szalanaski, & Barnes, 1987). This emphasis on case-specific information meant that 'base-rate' information, like long term trends, was underweighted (Tversky & Kahneman, 1974; Hoch & Schkade, 1996). It also meant that the forecasters' interest was usually confined to recent observations which were perceived as being the result of current ongoing or recently concluded events. Their attempts to get the statistical forecasts to provide as close a fit as possible to a few recent observations was symptomatic of this. In any case, recalling the many events and circumstances that were perceived to have shaped the past history would have put too great a load on memory so there was a natural bias towards recency. Against this background, the automatic forecasts of a statistical time series method were bound to be regarded with scepticism. The focus on recent patterns and individual outcomes meant that the system's ability to detect longer term systematic underlying movements was generally undervalued.

The psychological literature on accepting advice also provides insights into why the automatic statistical forecasts were often changed. Research by Yaniv and Kleinberger (2000) suggested that people are more likely to trust their own beliefs, rather than the advice because they have greater access to the rationale for these beliefs. The statistical forecasting system did not provide an explanation for its forecasts and the advice it provided was therefore mute and unsupported.

5.2 The organizational forecasting perspective

The individual perspective that we have just adopted does not provide a complete explanation for the way that the forecasting system was used. For example, why were managers apparently happy with a system that was unable to explain movements in time series that they judged to be largely predictable and which produced only linear extrapolations when they perceived the underlying trends in demand to be non-linear? Also, there were pressures in the organisation to improve forecast accuracy, through for example the Six Sigma initiative, so why were the fundamental aspects of the forecasting process and the way the system was used never questioned?

. We first set about classifying the interests of the actors associated with the forecasting process. The actors we identified as having important roles were the senior managers and accountants, the marketing and product managers, the logistics managers who produced the forecasts and the software vendor. We also considered the role of the FSS itself.

For the purpose of understanding the networked forces that create stability, it is useful to start with the perspective of a single actor. This actor will be referred to as the 'focal actor' and we examine how other actors' alignment with the focal actor's interests led to the formation of a stable network of aligned interests (Sidorova & Sarka, 2002). In our case, we designate the software vendor as the focal actor, though we could have taken the perspective of another actor as our starting point and we would still have derived the same rationale for the formation of the network.

The vendor was interested in obtaining sales of the forecasting system. This interest was served by advertising the accuracy and sophistication of the system's inbuilt statistical methods and its facilities for incorporating judgmental intervention, together with the system's ease of use (evidence for the highlighting of these attributes was found on the software company's web site).

The vendor also wanted to maximise the profit on the sale. This would be achieved by selling a system containing a *standard* (rather than a customised) set of statistical forecasting methods in order to spread the system's development costs. In the words of one software developer: "We live in a commercial reality, you see, and the customer will come along and say I would like something [a new facility] and you say I can't do this unless you co-fund the development" [this quote has been slightly re-worded to improve clarity]. The provision within the software of easy-to-use facilities for judgmental intervention would thus serve the vendor's interests in a second way because it would effectively place the costs of any local adaptation (or customisation) of the *forecasts* upon the user. This could also reduce the chances of the system being blamed for forecast errors, so ensuring continued use. Continued use was in the vendor's interests because users would pay for the maintenance of the system and would attend user conferences and purchase upgrades. Also, the existence of an active body of existing users was likely to attract new customers.

However, the provision of an easy-to-use facility for judgmental intervention was also in the interests of the company's middle managers. They could be seen to be using an advanced system containing reportedly sophisticated and accurate statistical methods, while at the same time being easily able to control the forecasts. The existence of these facilities for intervention was particularly useful in the enrolment of the product managers whose participation in forecasting was seen as crucial because of their market intelligence. It allowed them to derive prestige by demonstrating their expertise in their markets at forecast review meetings and gave them the opportunity of attempting to push the forecasts in directions that suited the balance of their interests. For example, one product manager, commenting on the system, said: "It's there, it's useful, but it needs to be managed since no way can it have the market intelligence".

The fact that FSS produced linear extrapolations, when the managers perceived the underlying trends to be non-linear, was paradoxically a factor that assisted in securing its acceptance. It provided a pretext for interventions, allowing users to make adjustments for other reasons. To maintain their own standing, the logistics managers needed to produce baseline forecasts which looked credible at review meetings where colleagues had an intolerance of noise in the time series. To achieve this they could use the intervention facilities to fit and refit past trends to different lengths of past history until a close fitting trend was achieved. One logistics manager described the system as being “quite good” because it allowed the graphical fit of the trend line to be easily assessed when judgmental changes were made to it or the length of the demand history altered.

The senior managers, including accountants, had as their objectives accurate forecasts to support the annual planning cycle, and the avoidance of costs arising from forecast errors (such as excess inventory). It was in senior managers’ interests to receive timely forecasts that they perceived to be from an advanced, modern forecasting system yielding baseline forecasts that were as accurate as possible, given the then current technology. It was also in their interests to ensure the inclusion of all relevant middle managers in the process. The FSS served these interests because, it produced graphical and tabular displays that could be used in review meetings involving groups of managers and allowed forecasts to be easily and publicly changed during these meetings. The old ‘ruler and paper’ system would not have been compatible with such meetings. Senior Managers also wanted to be able to exercise some control and monitoring of this process. From their perspective the system also had a facility which allowed for adjustment for market intelligence and, by requiring documentation of these adjustments, they perceived that control over the process could be exercised though this was not implemented in practice. Also, in relation to total turnover, the cost of the system was small (though it was large enough to be regarded as a serious tool).

As we have discussed, all the actors had a stake in resisting any change. For the middle managers it would involve the risk of losing the benefits of control over the forecasts, disruption and (in the case of the managers with direct forecasting responsibility) the need to learn a new system. For the senior managers changing to another system would have involved purchasing and other costs, disruption and probably resistance from middle managers. Nor was any evidence collected on the value-added arising from the different tasks that contributed to the final forecast. Had this been collected, it might have signalled a need for change. All of this served to consolidate the alignment of interests of the vendor and middle and senior managers and helped to ensure the stability of the network configuration. Although managers indicated that they felt their forecasts' accuracy could be improved (this was part of their main motive for inviting us into the company) they evidently wished to make these improvements within the existing structure. A suggestion by one of the researchers at the end of the interviews that the company might be using an inappropriate system, and that what was needed was a model that supported extrapolations based on product life cycles, was received sceptically. It was apparent that the company would have liked to find ways of making better use of available information generated by the many forecast review meetings in order to improve the quality of their judgmental interventions, but the role of the forecasting system would, they said, remain unchanged. As the system designers had noted, generally FSSs are not designed around the concept of collaborative work.

6 Why and how the forecasting system changed

We returned to the case organization 14 years after our previous visits to find that what had seemed a stable system had changed dramatically with new software and new processes at the core. Within the various regional subsidiaries there had been a wide range of forecasting processes and software; forecasts were produced from the purely judgmental to the more sophisticated system observed in the case subsidiary. Managers who were external to the UK subsidiary perceived this assortment of methods to be inefficient and in 2011 a centralized demand management team, a small 'Centre of Excellence', was set up 'to validate and consolidate the global pharma demand and represent the link between the local affiliates and the global supply chain planning organizations'. Centres were also established in some of the other business units, which typically employed staff with technical (statistical) expertise. They considered that the centralization of processes and software would allow quality control to be exercised, accuracy targets set and the sharing of information between regional units. In addition, the organization as a whole had become a user of SAP which in 2013 led, after consideration of some limited alternatives, to the adoption across all regions of SAP-APO. SAP as implemented offered major additional capabilities beyond forecasting including support for logistics and manufacturing operations.

In trying to understand the disruptive forces that led to changes in the system and its processes, we have interviewed two major actors, the company-wide manager in charge of demand planning and a software provider of supporting forecasting software. In addition, we were able to interview one of the original forecasters in the UK subsidiary. Factors internal to the organisation (though external to the subsidiary), rather than changes in the environment appeared to provide the stimulus for change in the company we studied.

In his critique of theories of change Todnem By (2005) classifies change in organisations in terms of (i) its rate of occurrence (e.g. continuous or discontinuous and infrequent), (ii) its scale (e.g. department based or strategic) and (iii) how it comes about (e.g. through careful planning or as a reaction to external events). In this case, the initial change was discontinuous and isolated, company-wide –though not affecting the company's strategic direction – and it was planned, rather than reactive.

The punctuated equilibrium model has been used to explain why some organizations experience long period of stability followed by occasional sudden changes (Gersick, 1991). In an application of the model to the adoption of software in organizations, Lassila and Brancheau (1999) indicate that it posits that stable or 'equilibrium' periods are "sustained by inertia resulting from various factors of cognition, motivation and obligation". . This characterization is consistent with observations in this company. Disruption can result from 'internal changes that cause structures and activities to go out of alignment or from changes outside the organization that impair the system's ability to adjust with its environment'. More specifically, internal factors that can lead to change include: (i) employee turnover (Lassila and Brancheau, 1999) -where new employees bring novel perspectives to the organization; (ii) a subjectively perceived misfit between what the current process is delivering and what it ought to be delivering and (iii) a desire for the restoration of control by senior managers (Jacobs et al. 2013). These elements all figure in the changes in the company's demand planning processes.

The outstanding driver of change here was the top-down requirement for standardisation. The company-wide introduction of SAP as a platform proved the opportunity: the UK subsidiary's satisfaction with the established FSS was of limited counter-weight. The change was demanded by an allied network of actors at the top of the multinational organization who perceived a misfit between the existing process across the company and what it could be achieving, and who, wished to establish some central control over the process. By virtue of their position, these actors had the power to push forward their agenda, aided by the centralized technological change. It also appears that their arguments for change were sufficient to establish 'the balance of opinion' amongst those who would be directly affected by the innovation. One manager at the centre of the organization reported that there was 'curiosity' and 'lots of interest' in the proposed new system when the rationale for it was explained to them. However, the new system lacked the graphical flexibility of the old, the forecaster in the subsidiary commented, and was initially regarded as 'not as easy to use'.

With new staff technically trained in post, the centralized demand have continued to innovate and in 2016 implemented a model selection routine bolted on to APO., where the initial baseline forecasts were automatically generated by the system (APO enhanced by the bolt-on forecasting software, iqast (www.iqast.de), which had been shown to lead to improved accuracy). These baselines were then used to highlight products needing review by the regional subsidiaries. The senior analyst, through analysis from the enhanced software, committed time to discuss with the subsidiaries the recommended models. However the subsidiaries continued to have final responsibility for their forecasts with their S&OP process remaining essentially unchanged: the major change was in the construction of the base-line. The forecast users remained much the same as in earlier years, including production and operations but also finance. The earlier emphasis on product life-cycle, which justified many of the adjustments, has to a certain extent been absorbed into a more extensive S&OP process where more information is available to the participants. Trust in the forecasts, the senior global demand analytics manager remarked, was gained by “sitting down with the affiliates” to show the value of the new base line and of working together on the final forecast: “with this approach you can gain trust [and] cause the local team to feel empowered…. Not just receiving a number from the top”.

From a forecasting perspective it is unclear to what extent the innovation led to improvements in the process in the UK subsidiary: dramatic improvements were claimed elsewhere. In selecting SAP, no comparative testing of alternative algorithms was carried out and in fact, SAP-APO is known to have limited forecasting capability with method selection heuristics that are poorly designed (Chockalingam, 2010). Moreover, while the new system precluded the manipulation of the system’s parameters, ex-post judgmental adjustment of its forecasts was still permitted –one manager who led the innovation estimates that forecasts for established products are “70% statistics and 30% management judgment”, while for new products the contribution of judgment was 50%. He also estimated that the innovation had reduced the mean absolute percentage error (MAPE) of forecasts by around 7 percentage points after the introduction of SAP, but no empirical evidence was available to substantiate this.

The new forecasting process has proved stable for many of the same reasons that the original process, which had stayed in place for around 20 years. The key new actors coming into play that changed the equilibrium were an alternative software system (SAP) with a forecasting module (SAP-APO), those senior managers looking for a consistent process across regions and the newly formed central demand management team. Its centralized adoption left the regional forecasting roles and tasks unscathed and the forecast value added (FVA) (Gilliland, 2008) of the task of judgmental adjustment to support market intelligence was initially unanalysed. However, in one important respect, the new network has been designed so that change in one key component, the generation of the base line forecasts through novel software extensions, could be organic with future innovations such as the introduction of machine learning methods easy to implement. As a second example, in 2019/20 based on the earlier success, the team was expanded to continue to work towards better base line forecasts and “explore usage of new technologies”. Also, a new process whereby the subsidiaries only focus on the products where the FVA has been estimated as positive leaving the remaining products untouched, is being introduced. These innovations have been made possible through the centralized demand management team’s remit and the trust established between the small centre and the subsidiaries, an aspect not present in all the business units. However, SAP-APO has remained a significant constraining force, an ‘actor’ that requires others to work around it. Features in the old system such as changing the data used in choosing the baseline model have disappeared but the top-down model selection routine has proved sufficiently convincing that those producing the final consensus forecast can work successfully with the changed organizational network. The senior analyst commented, “The Global Demand Analytics Team however will focus in the next years in evaluating the latest state of the art software to further increase the accuracy of the baseline.”

7 Conclusions

From the technological imperative perspective, people make rational economic decisions in adopting and using technology. The forecasting processes in this company cannot easily be interpreted through such a lens. Instead with the complementary perspectives of individual forecasters’ cognitive processes, the tasks they undertake and an organizational analysis , the decision to use the original forecasting system despite its poor task-technology fit becomes understandable. The case study evidence we have presented has highlighted a number of key lessons about support systems-in-use.

Firstly, managers used the system in a way which did not accord with its design and advertised purpose. They effectively moulded the system so that the collaborative tasks they undertook and their shared understanding of the market could be 'inscribed' into the FSS. The new software system (APO) in its production of baseline forecasts was outside the local forecasters' influence and not amenable to their direct intervention.

Secondly, the research has focussed on the on-going operation of an IS, the individual and organizational drivers for sub-optimal (economic) use and why these patterns of behaviour persisted and subsequently changed. This research poses a stark question to those seeking to improve the quality of forecasting in supply-chain companies: how can individual cognitive biases *and* the organisational and personal barriers embodied in stable networks like the one described be overcome to achieve more efficient but equally acceptable forecasts? No elements of the established network facilitated process improvement or provided an incentive for individuals to change their mental models of the forecasting task.

Thirdly, the research shows that change was difficult and long in coming even when there were external interventions such as the company-wide 6-sigma project or the introduction of researchers such as us, analysing forecast errors and the processes and software from which they derived. Change in the end came about by a centralized initiative, external to the network of actors we had observed and it was discontinuous. It was driven by software standardization across the whole organization. The new system had a poorer task-technology fit, nor was it appraised on the grounds of offering improved accuracy. However, the establishment of the Centre of Excellence in Demand Management provided the means by which accuracy improvements were achieved. The vested interests that many actors had in continuing to make heavy use of judgmental interventions remains in the new organizational processes, despite their known limitations. But once again, centralized changes are expected to lead to less frequent and more effective interventions.

Our continued involvement with a wide range of supply-chain companies and software providers, together with the findings of recent surveys, lead us to expect that judgmental interventions, despite their limitations, will continue to meet the individual and collaborative organizational needs of the forecasters, ensuring a consensus around which the organization can plan. The challenge for researchers and software designers alike is to develop FSSs that can meet such organizational requirements while at the same time improving forecast accuracy. Such a novel research agenda places equal weight on innovative statistical methods and the effective incorporation of 'forecast value added', where this is potentially valuable, into the forecasting process. For this to happen customers will need to create a demand for more sophisticated FSSs, despite the perceived threat to their autonomy as forecasters, as well as a willingness to pay for these improved facilities (Goodwin, 2015). This will require a recognition of the operational benefits that increased accuracy can bring and a demonstration that improved FSSs can achieve this.

Acknowledgements: We would like to thank the anonymous contributors to this case study for giving their time to explaining the processes in the case organization. We would also like to thank researchers who were involved with aspects of the case at an early stage, Dr. Stavros Asimakopoulos, Andrea Franco and Professor Konstantinos Nikolopoulos.

8 References

Arvan, M., Fahimnia, B., Reisi, M., & Siemsen, E. (2019). Integrating human judgement into quantitative forecasting methods: A review. *Omega*, 86, 237-252.

Asimakopoulos, S., Dix, A., & Fildes, R. (2011). Using hierarchical task decomposition as a grammar to map actions in context: Application to forecasting systems in supply chain planning. *International Journal of Human-Computer Studies*, *69*, 234-250.

Asimakopoulos, S., & Dix, A. (2013). Forecasting support systems technologies-in-practice: A model of adoption and use for product forecasting. *International Journal of Forecasting, 29*, 322-336.

Auer, T. (1998). Quality of IS use. *European Journal of Information Systems*, 7, 192-201.

Bagchi, S., Kanungo, S., & Dasgupta, S. (2003). Modeling use of enterprise resource planning systems: a path analytic study. *European Journal of Information Systems,* 12, 142-158.

Beach, L.R., Christensen-Szalanaski, J., & Barnes, V. (1987). Assessing human judgment: has it been done, can it be done, should it be done? In G. Wright & P. Ayton (Eds.), *Judgmental Forecasting,* Chichester: Wiley, 49-62.

Benbasat, I., Goldstein, D.K., & Mead, M. (1987). The case research strategy in studies of information systems. *MIS Quarterly,* 11, 369-386.

Chockalingam, M (2010). Forecast modeling capabilities in SAP APO vs other statistical tools https://demandplanning.net/Newsletters/DPnewsletter_November2010.pdf.

Other Statistical Tools. https://demandplanning.net/Newsletters/DPnewsletter_November2010.pdf

Dalrymple, D.J. (1987). Sales forecasting practices, results from a United States survey. *International Journal of Forecasting*, 3, 379-391.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13, 318-339.

Davis, G.B., Lee, A.S., Nickles, K.R., Chatterjee, S., Harthung, R., & Wu, Y. (1992). Diagnosis of an information systems failure. A framework and interpretive process. *Information and Management,* 23, 293-318.

Davis, F.D., & Kottemann, J.E. (1994). User perceptions of decision-support effectiveness - 2 production planning experiments. *Decision Sciences*, 25, 57-78.

Dawes, R.M. (1979). Robust beauty of improper linear-models in decision-making. *American Psychologist,* 34, 571-582.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144, 1140126.

Doering, T., & Suresh, N. C. (2016). Forecasting and performance: Conceptualizing forecasting management competence as a higher-order construct. *Journal of Supply Chain Management*, 52, 77-91.

Easterby-Smith, M., Thorpe, R., & Jackson, P.. (2011). *Management Research*, 4th edition, London: Sage Publications.

Falk, R., & Konold, C. (1997). Making sense of randomness: implicit encoding as a basis for judgment, *Psychological Review*, 104, 301-318.

Fildes, R. (2017). Research into forecasting practice. *Foresight: The International Journal of Applied Forecasting* (44 Winter), 39-46.

Fildes, R. & Goodwin, P. (2007). Against your better judgment? How organizations can improve their use of management judgment in forecasting. *Interfaces*, 37, 570-576.

Fildes, R., Goodwin, P., & Lawrence, M. (2006). The design of forecasting support systems and their effectiveness. *Decision Support Systems*, 42. 351-361.

Fildes, R., Goodwin, P., Lawrence, M. & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, 25, 3-23.

Fildes, R., Goodwin, P., & Önkal, D. (2019). Use and misuse of information in supply chain forecasting of promotion effects. *International Journal of Forecasting*, 35, 144-156.

Fildes, R., & Hastings, R. (1994). The organisation and improvement of market forecasting. *Journal of the Operational Research Society*, 45, 1-16.

Fildes, R. and Petropoulos, F. (2015). Improving forecast quality in practice. *Foresight: The International Journal of Applied Forecasting* 36, 5-12.

Fildes, R., Schaer, O., & Svetunkov, I. (2018). Software survey: Forecasting 2018. *OR/MS Today*, *45*.

Fischhoff, B. (1975). Hindsight not equal to foresight. Effect of outcome knowledge on judgment under uncertainty, *Journal of Experimental Psychology –Human Perception and Performance*, 1, 288-299.

Fishbein, M., and Ajzen, I. (1975). *Belief, Attitude. Intention and Behavior: An Introduction to Theory and Research.* Addison-Wesley, Reading, MA.

Franses, P. H., & Legerstee, R. (2013). Do statistical forecasting models for SKU-level data benefit from including past expert knowledge? *International Journal of Forecasting*, 29 80-87.

Galbraith, C.S. & Merrill, G.B. (1996). The politics of forecasting: Managing the truth, *California Management Review*, 38, 29-43.

Gersick, C.J. (1991). Revolutionary change theories: a multilevel exploration of the punctuated equilibrium paradigm *Academy of Management Review*, 16, 10-36.

Gilliland, M. (2008). Forecast value added analysis: Step-by-step. *SAS Institute whitepaper*.

Goodwin, P. (2002). Integrating management judgment with statistical methods to improve short-term forecasts. *Omega, International Journal of Management Science*, 30, 127-135.

Goodwin, P. (2015) “Commentary: Where is the support for judgment?” *Foresight: the International Journal of Applied Forecasting* ,39, 14-15.

Goodwin, P., & Fildes, R. (1999). Judgmental forecasts of time series affected by special events: Does providing a statistical forecast improve accuracy? *Journal of Behavioral Decision Making*, 12, 37-53.

Goodwin, P., Fildes, R., Lawrence, M., & Nikolopoulos, K. (2007). The process of using a forecasting support system. *International Journal of Forecasting*, 23, 391-404.

Hanseth, O.(2004). Actor-network theory and information systems: what's so special? *Information Technology and People*, 17, 116-123.

Heuer, J., Merkle, C., & Weber, M. (2016). Fooled by randomness: Investor perception of fund manager skill. *Review of Finance*, *21*(2), 605-635.

Hoch, S. J., & Schkade, D.A. (1996). A psychological approach to decision support systems. *Management Science*, 42**,** 51-64.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 27(3), 1–22. https://doi.org/10.18637/jss.v027.i03

Jacobs, G., Van Witteloostuijn, A., & Christe-Zeyse, J. (2013). A theoretical framework of organizational change. *Journal of Organizational Change Management*, 26, 772-792.

Kanter, R.M. (1977). *Men and Women of the Corporation*. New York,Basic Books.

Kaplan, S.E., Reneau J.H., & Whitecotton, S. (2001). The effects of predictive ability information, locus of control, and decision maker involvement on decision aid reliance. *Journal of Behavioral Decision Making*, 14, 35-50.

Klassen, RD., & Flores, B.E. (2001). Forecasting practices of Canadian firms: survey results and comparisons. *International Journal of Production Economics*,70, 163-174.

Kleinmuntz, B. (1990). Why we still use our heads instead of formulas - toward an integrative approach. *Psychological Bulletin*, 107, 296-310.

Lassila, K. S., & Brancheau, J. C. (1999). Adoption and utilization of commercial software packages: Exploring utilization equilibria, transitions, triggers, and tracks. *Journal of Management Information Systems*, 16, 63-90.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor Network Theory.* Oxford: Oxford University Press.

Lawrence, M., Goodwin, P., & Fildes, R. (2002). Influence of user participation on DSS use and decision accuracy. *Omega, International Journal of Management Science*, 30, 381-392.

Lawrence, M., O'Connor, M., & Edmundson, B. (2000). A field study of sales forecasting accuracy and processes. *European Journal of Operational Research*, *122*(1), 151-160.

Mady, M.T. (2000). Sales forecasting practices of Egyptian public enterprises: survey evidence. *International Journal of Forecasting,* 6, 359-368.

McCarthy, T. M., Davis, D. F., Golicic, S. L., & Mentzer, J. T. (2006). The evolution of sales forecasting management: A 20-year longitudinal study of forecasting practices. *Journal of Forecasting, 25*(5), 303-324. doi:10.1002/for.989

McGovern, T., & Hicks, C. (2004).How political processes shaped the IT adopted by a small make-to-order company: a case study in the insulated wire and cable industry. *Information and Management*, 42, 243-257.

Moon, M. A., Mentzer, J. T., & Smith, C. D. (2003). Conducting a sales forecasting audit. *International Journal of Forecasting*, 19, 5-25.

Nandhakumar, J., & Jones, M. (1997). Too close for comfort? Distance and engagement in interpretive information systems research. *Information Systems Journal*, 7, 109-131.

Nardulli, P.F. (1978). *The Courtroom Elite: An Organizational Perspective on Criminal Justice*. Cambridge, MA: Ballinger Press.

O'Connor, M., Remus, W., & Griggs, K. (1993). Judgemental forecasting in times of change. *International Journal of Forecasting*, 9, 163-172.

Oliva, R., & Watson, N. (2009). Managing functional biases in organizational forecasts: A case study of consensus forecasting in supply chain planning. *Production and Operations Management, 18*, 138-151.

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, *22*, 390-409.

Orlikowsi, W.J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organizational Science*, 11, 404-428.

Orlikowsi, W.J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science*, 11, 404-428.

Payne, J.W., Bettman, J.R., & Johnson. E.J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.

Perera, H. N., Hurley, J., Fahimnia, B., & Reisi, M. (2018). The human factor in supply chain forecasting: A systematic review. *European Journal of Operational Research*., 274, 574-600. doi:https://doi.org/10.1016/j.ejor.2018.10.028

Petropoulos, F., Kourentzes, N., Nikolopoulos, K., & Siemsen, E. (2018). Judgmental selection of forecasting models. *Journal of Operations Management, 60*, 34-46. doi:https://doi.org/10.1016/j.jom.2018.05.005

Phillips, C. J., & Nikolopoulos, K. (2019). Forecast quality improvement with action research: A success story at pharmaco. *International Journal of Forecasting, 35*, 129-143.

Ruivo, P., Oliveira, T., & Neto, M. (2014). Examine ERP post-implementation stages of use and value: Empirical evidence from Portuguese SMEs. *International Journal of Accounting Information Systems*, *15*(2), 166-184.

Sanders, N.R., & Manrodt, K. B. (1994). Forecasting practices in US Corporations: Survey results. *Interfaces*, 24, 92-100.

Sanders, N.R., & Ritzman, L.P. (2001). Judgmental adjustment of statistical forecasts. In. J.S. Armstrong (Ed.). *Principles of Forecasting*. Norwell:MA: Kluwer Academic Publishers, 405-416.

Siegrist, M., Cvetkovich, G.T., & Gutscher, H. (2001). Shared values, social trust and the perception of geographic cancer clusters. *Risk Analysis*, 21, 1047-1053.

Sidorova, A., & Sarker, S. (2002). *Unearthing Some Causes of BPR Failure: An Actor network Theory Perspective*. School of Accounting, Information Systems and Business Law, Washington State University.

Siegel, J., Dubrovsky, V., Kierler, S., & McGuire, T.W, (1986). Group processes in computer-mediated communication. *Organizational Behavior and Human Decision Processes*, 37, 157-187.

Smaros, J. (2007). Forecasting collaboration in the European grocery sector: Observations from a case study. *Journal of Operations Management*, 25, 702-716.

Sroginis, A., Fildes, R., & Kourentzes, N. (2019). Use of contextual and model-based information in behavioural operations. Working Paper. Lancaster.

Smith, C. D., & Mentzer, J. T. (2010). Forecasting task-technology fit: The influence of individuals, systems and procedures on forecast performance. *International Journal of Forecasting*, 26, 144-161.Sroginis, A., Fildes, R., & Kourentzes, N. (2020). *Use of contextual and model-based information in behavioural operations*. Working Paper. Lancaster.

Todnem By, R. (2005). Organisational change management: A critical review. *Journal of Change Management*, 5, 369-380.

Tuomikangas, N., & Kaipia, R. (2014). A coordination framework for sales and operations planning (S&OP): Synthesis from the literature. *International Journal of Production Economics*, 154, 243-262. doi:http://dx.doi.org/10.1016/j.ijpe.2014.04.026

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, 185, 1124-1131.

Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: towards a unified view. *MIS Quarterly*, 27, 425-478.Venkatesh, V., Thong, J. Y., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17, 328-376.

Walsham, G. (1995). Interpretive case-studies in research - nature and method. *European Journal of Information Systems*, 4, 74-81.

Winklhofer, H., Diamantopoulos, A., & Witt, S. F. (1996). Forecasting practice: A review of the empirical literature and an agenda for future research. *International Journal of Forecasting*, *12*, 193-221.

Yaniv I., & Kleinberger E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior And Human Decision Processes*, 83,: 260-281.

Figure 1 A typical product life cycle, as hypothesized, with ETS forecasts (from forecast origins 6, 12, 18 and 24)