Stability in the inefficient 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

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

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 most 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 why the managers continued to use non-normative forecasting practices for many years despite the potential economic benefits that could be achieved through change. The reasons for the longevity of these practices are examined both from the perspective of the individual forecaster and the organization as a whole.

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

1 Introduction

Accurate forecasts are important to the success of supply-chain companies. Decisions relating to transportation, purchasing, inventory control, work-force scheduling, production planning and cash-flow planning are all dependent on them (Moon et al., 2003; Van den Broeke et al., 2018). Given this importance, the development and sales of computer-based statistical forecasting systems has become big business. Major suppliers, such as SAP and JDA have embedded 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 providers of dedicated forecasting software (see Fildes, et al. (2020) for a survey). These statistically-based systems are usually sold and purchased on the basis that their algorithms will provide accurate forecasts of future demand. Implicit in this promise of heightened accuracy is the provider's expectation that the role of human judgment in forecasting will be limited to exceptional circumstances –for example, where special events mean that managers have important information that is not available to the statistical forecast. However, there is a great deal of evidence that, in practice, the demand forecasts generated by such systems are routinely replaced by forecasts based on managers' judgments that are only seldom based on external knowledge (Fildes & Goodwin, 2007b; Fildes & Petropoulos, 2015; Franses, 2014; Arvan et al. (2018) provides a recent review). This is despite evidence that many of these judgmental interventions are usually detrimental to forecast accuracy and absorb significant amounts of management time (Fildes, et al., 2009). This paper uses an in-depth study of the forecasting process in a pharmaceutical company to explain how this sub-optimal practice can arise and persist for a considerable time.

The paper is organised as follows. First, we examine the existing literature to identify possible explanations for the prolonged operation of inefficient forecasting processes within organizations. We then explain why we have adopted a case-study based approach to address our research questions before explaining our methodology. 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 demonstrating the inefficiency of the resulting forecasts. Then we interpret our results to provide explanations of why the system was used in this way and why this state of affairs persisted for several years before changes were made that led to the adoption of new software which resulted in a revised demand forecasting process. 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 integrates complementary theoretical perspectives to contrast the technical with technology-in-practice; such a focus emphasizes the importance of software design fitting with organizational actors and the processes surrounding demand forecasting. 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. Literature review

2.1 Evidence for deviations from normative practice

The forecasting literature suggests a normative approach to the use of forecasting software, or forecasting support systems (FSSs) in organizations, which involves the following elements. First, managers should use *appropriate* statistical forecasting methods to identify regular patterns in data (assuming that sufficient data is available for this purpose (Goodwin, 2002). These methods should be designed to handle the specific characteristics of the data available to the organization. For example, simple exponential smoothing would not be suitable where the data exhibits a trend or seasonal pattern or the market is promotions driven. The use of statistical forecasting methods to identify systematic patterns is recommended because humans have a tendency to suffer from a number of cognitive biases when searching for patterns in data. For example, they tend to perceive false systematic patterns in the noise that appears in time series (O'Connor, Remus, & Griggs, 1993), to underestimate trends (Harvey & Reimers, 2013) or attach too much weight to the most recent observation (Lawrence & O'Connor, 1995). Second, if they are free to select which statistical model is appropriate in a given situation, they should regard their judgmental selection as complementary to an automatic approach (De Baets & Harvey, 2020; Petropoulos, et al., 2018) [1]. Further, judgmental adjustments should only be applied to statistical forecasts when the forecaster has important information about forthcoming events, such as sales promotions, that is not available to the statistical method (Fildes, Goodwin, & Lawrence, 2006; Sanders & Ritzman, 2001; Sroginis, Fildes, & Kourentzes, 2018). Moreover, the size of these adjustments should be accounted for and

[1] Although, Petropoulos et al. **(**2018**)** 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.

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.

There is substantial evidence that forecasters in supply chain companies do not adhere to these normative principles. A survey of company forecasters (Fildes & Goodwin, 2007a) found that on average 67% of statistical forecasts were either judgmentally adjusted or averaged with a judgmental forecast. Similarly high levels of judgemental intervention have been reported in (Fildes & Petropoulos, 2015; Franses & Legerstee, 2013; Van den Broeke, et al., 2018), while Fildes et al. (2009) found that some forecasters adjusted over 90% of their statistical forecasts. When they analysed data from four supply-chain companies, they also found that many adjustments were small – around a third less than 10% -and hence unlikely to be responses to important new information that the statistical method did not have access to. While small adjustments can reduce forecast accuracy, this damage is bound to be limited. However, making such adjustments can lead to a considerable waste of valuable management time (Fildes, et al., 2009). Although nearly 64% of forecasters in the 2007 survey reported on in Fildes & Goodwin (2007a) claimed that they documented the reasons for any judgmental adjustments, the in-company research conducted by (Fildes, et al., 2009) found that many reasons were recorded in unintelligible shorthand and none were coded into retrievable categories. Finally, Fildes and Goodwin (2007a) found that only 44% of forecasters claimed to review the extent to which their judgmental interventions improved the accuracy of their forecasts. Moreover, only 30% reported using multiple error measures -a single error measure is unlikely to capture all aspects of performance (Armstrong, 2001). Decisions on whether to adjust forecasts were therefore often uninformed by evidence on their likely effectiveness. The Fildes et al. (2009) study found that, had such reviews been conducted in the companies they studied, they would have revealed that positive (i.e. upward) adjustments were far less effective in improving accuracy than negative adjustments. Most of the adjusted forecasts (53.5%) analysed by Franses and Legerstee (2009) proved worse than the corresponding statistical model forecasts.

A consequence of these departures from prescribed practice is that forecasting systems are being used inefficiently –we define an efficient forecasting process as one where managers' interventions bring the maximum benefits to an organization given the constraints on their time. For example, in their study of four companies Fildes et al. (2009) found that managers were spending considerable amounts of time making adjustments to forecasts that reduced accuracy.

Franses and Legerstee (2013) found similar inefficiencies and these various studies are summarized in Perera et al. (2018).

2.2 Reasons why individual forecasters may deviate from normative practice

The psychology and information system literatures suggest reasons why forecasting in practice often deviates from the recommended normative procedures. Adjusting a forecast implies an absence of psychological inertia which would favour acceptance of the computer forecast as the default option unless there is a clear motive to make a change (Gal, 2006) such as improved accuracy. And accuracy has been identified in a survey of demand planners as their primary objective (Fildes & Goodwin, 2007b). The work of Payne, Bettman and Johnson (1993) indicates 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) or to have a sense of ownership or due to a lack of trust in the model (Önkal & Gönül, 2005). For example, forecasts may be deliberately overestimated by operations departments to avoid stockouts or underestimated so that sales staff will be rewarded for exceeding the forecasts (Mello, 2009). However, in the common situation where there is a genuine desire to achieve forecast accuracy, accounting for adjustments 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 a belief 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". For example, in demand forecasting managers may be subject to optimism bias through wishful thinking (Krizan & Windschitl, 2007). As a result, they adjust statistical demand forecasts upwards or make overly large positive adjustments when there is no evidence to justify such interventions (Fildes, et al., 2009). People also appear to be more tolerant of errors in human judgment compared to errors produced by algorithms. Dietvorst, Simmons, and Massey (2015) found that people soon lost confidence in an algorithm when it erred, leading to a phenomenon that they termed 'algorithm aversion'. Prahl and Van Swol (2017) attributed this to a 'perfection scheme' whereby we expect algorithms, unlike humans, to exhibit perfect accuracy. As a result, we are shocked when they err and react negatively to them. Considerations like these may account for the findings in 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 GPS navigation system, 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 that 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 perceived 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 & Konold, 1997)). When confronted with randomness, they tend to perceive patterns and causes (Heuer, Merkle, & Weber, 2017; Siegrist, Cvetkovich, & Gutscher, 2001). This leads to the belief that greater mental effort will improve the accuracy of forecasts (Kottemann, Davis, & Remus, 1994).

Finally, there is a possibility that managers have a fear of becoming deskilled if the forecasting task becomes wholly automated (Bainbridge, 1983). De Baets and Harvey (2018) have suggested that adjusting statistical forecasts gives these forecasters the opportunity to receive feedback on their performance and knowledge of their market - a learning opportunity that would be absent if they simply accepted the forecast produced by an algorithm.

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 forecasting task when they think this is required, for example by changing parameters or overriding the forecasting method's automatically selected by the system (Dietvorst, Simmons, & Massey, 2018). It might also be increased when a system provides a veneer of scientific support for one's forecasts, even though this support has not been fully employed. However, while these features may increase acceptance of the system, manipulations like these will also be associated with a tendency to reject and override the system's automatic output.

A complementary perspective is provided by the task-technology fit model (Dishaw & Strong, 1999; Goodhue & Thompson, 1995) which asserts that information technology will only be used only if the functions available to the user support the activities of the user. Smith and Mentzer (2010) showed that the users' perceptions of the task-technology fit influenced positively-reported forecast performance. However, Bendoly and Cotteleer (2008) found that where there was a strong task-technology misfit, managers would tend to circumvent the prescribed use of the technology, particularly where such circumvention was perceived to be easy. In a forecasting system, users may perceive that automated statistical forecasts have a poor fit to their role, which is to apply their expertise and marketing knowledge to the anticipation of future demand. Where the software provides facilities for easy circumvention of automatic forecasts these will be readily adopted. Arguably, it is preferable to design an acceptable support system that improves forecasts, but is not used optimally to one which attempts to impose optimality and is rejected as a consequence This conjecture is in tune with Keen and Scott Morton's (1978) insistence nearly half a century ago that support system designers should balance descriptive realism with normative idealism.

There are a number of other circumstances where forecasting software may be perceived to have a poor task-technology fit. Designers of these systems have paid little attention to providing support for judgmental interventions, despite their widespread use (Goodwin, 2015). Also designers often do not take account of the organisational environment within which their products will be deployed (Asimakopoulos & Dix, 2013; Asimakopoulos, Dix, & Fildes, 2011); see also Arvan et al.'s review (Arvan, et al., 2018). Asimakopoulos Dix, & Fildes (2011) carried out an in-depth set of interviews[2] 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 through a spreadsheet-type –row permitting

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

adjustments) and whether the forecast 'makes sense' such as through a graphical check (Asimakopoulos, Dix, & Fildes, 2011).

As Asimakopoulos, Dix and Fildes (2011) describe, 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 whilst giving less weight to collaborative practices. For some users, this suggests there would be a fit with the software and for others a mismatch: The second task highlighted was choosing an 'appropriate' statistical model from a pre-defined set to match the circumstances expected. (Alternatively, the model could be chosen automatically permitting user overrides.) For users the evaluation phase was seen as both graphical and also based on forecast error measures. Again, this would require a range of error measures to meet business needs including different levels of aggregation (e.g. product to product group) and different forecast horizons to correspond to the decision context. The error evaluation could form part of an iterative cycle where exploration and analysis of fit, model selection and forecast error analysis lead to further interventions by stakeholders. Finally, the FSS could potentially facilitate knowledge sharing, seen as a key task identified by users […where] the FSS played a key role in shaping a shared forecasting meaning and in fostering interactions among relevant stakeholders" (Asimakopoulos & Dix, 2013). Insofar as the FSS facilitated these interactions, they argued, the fit was improved. This, though, was not typically explicit in the software's design but rather formed a part of the organization's usage in practice (Assimakopoulos et al., 2011).

2.3 Organizational factors that may lead to deviations from normative practice

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. The forecasters thought that improvements, ~~–~~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, Mentzer, and Smith 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. Their study 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', [3] have examined how forecasters, carrying out their tasks, interact with other organizational actors through the FSS (Tuomikangas & 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" (Fishbein & 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 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 (McGovern & Hicks, 2004; Orlikowski, 1992). While this perspective acknowledges that there is duality in that people and technology interact reciprocally it has been criticised for

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

understating the role of technology and its characteristics in this process (Hanseth, Aanestad, & Berg, 2004).

Actor-network theory (ANT) (Czarniawska, 2017; Latour, 2005) avoids the dichotomy between human and non-human entities by using the concept of an actor (or actant). An actor is any element which has the power to stimulate action and can be either a human, a collection of humans or an item such as a mobile app that encourages users to report their health and encourages them to quarantine if necessary, or a forecasting system. Actors all have interests -in the case of non-human actors these reflect those of their designers - and they try to enrol other actors in order to create an alignment of the other actors' interests with their own. The result of successful enrolment is an actor network. This network relates, not to the static relationships between the actors, but to the processes in which they are involved (Zackariasson & Wilson, 2004).

Three key concepts associated with ANT are inscription, translation and irreversibility. When a technical artefact, such as a forecasting system, is created by an actor certain functions are inscribed into it that are designed to protect the actor’s interests. Designers define artefacts with ‘specific tastes, competencies, motives, aspirations, political prejudices and the rest’ (Akrich, 1992). Translation refers to the process of interacting with other actors ‘to build heterogeneous networks of human and non-human actors forming alliances and mobilising resources’ (McGrath, 2002).

Although the inscription of an artefact is an embodiment of the designer’s view of how the system will be used, there is no guarantee that users will follow the prescribed pattern of use. The possibility of deviation from these anticipated patterns depends upon the strength of the inscription. For example, an inscription for a forecasting system may take the form of restrictions on the use of judgmental interventions within the program (~~Monteiro, 2000;~~ Goodwin et al., 2011). However, if users are able to circumvent this restriction by, for example, changing the statistical forecast itself then the inscription is weak. The strength of inscriptions is related to the irreversibility of the actor-network. Irreversibility is the degree to which it is subsequently impossible to return to a point where alternative possibilities exist (Walsham, 1997). Actor-networks with high irreversibility have accumulated a resistance to change so it is very difficult, if not impossible, for alternative translations to be made. ANT’s symmetrical treatment of humans and non-humans is not without its critics (Sayes, 2014). For example, Rose and Jones (2005) argue that ‘Human agents have purposes and forms of awareness that machines do not’.

Despite these criticisms, ANT possesses features that have the potential to illuminate the forecasting process at the pharmaceutical company. Its symmetry means that it will not downplay the role of technologies, such as the forecasting system, in the formation and continued existence of the complex process of interactions that lead to a set of forecasts. Moreover, by considering the mutual interaction of all of the actors, it guides the researcher towards an assessment of their interests and how these might be aligned. It may also help to explain how these alliances ensure various organisational-technological processes endure.

2.4 Research questions

In the information systems literature, relatively few papers have considered the on-going use of systems. Examples are provided in (Auer, 1998; Bagchi, Kanungo, & Dasgupta, 2003; Ruivo, Oliveira, & Neto, 2014), while Venkatesh, Thong and Xu (2016) have recently 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 case-based research being the exception –rare examples have focussed on action research approaches to improve the forecasting process (Oliva & Watson, 2009; Phillips & Nikolopoulos, 2019) while others (Caniato, Kalchschmidt, & Ronchi, 2011; Smaros, 2007) focussed on how forecasting software could be used to integrate statistical methods with the company’s judgmental forecasts. More generally, Kaipia, et al., (2017) and Smaros, (2007) have provided evidence on how systems may facilitate collaboration, while Moon, Mentzer and Smith (2003), building on extensive case-based researchplaced forecasting support systems at the heart of a successful process As did Phillips and Nikolopoulos (2019) in their action research. More recent evidence has been provided in Doering and Suresh (2016) showing the use of -‘advanced [forecasting] systems’ to be a determinant of overall competence as it affects, costs, service and accuracy. However, none of these studies focused on on-going use. 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 their organizations, with particular emphasis on the balance between statistical methods and judgment e.g. Fildes & Goodwin, 2007a; McCarthy, et al., 2006.

The objectives of the research we report on here are quite different from those of earlier studies in that we seek to understand in depth: (i) how managers in an organization produce their demand forecasts, integrating advice and information from diverse sources, (ii) the extent to which the processes they use deviate from normative practices and (iii) why any deviations from normative practices exist and persist, despite the benefits that could potentially be achieved through change. Essentially, our aim is to understand the components that affect the sustained use of a FSS despite deleterious economic consequences and inefficiencies and how these interact. We also examine evidence of how change disrupted the stable system.

The literature review suggests the following propositions which we explore empirically.

Proposition 1: *A forecasting system that permits the manipulation of statistical forecasts and judgmental interventions will be acceptable to, and in the perceived interests of, all actors even when the resulting forecasting process is inefficient.*

Proposition 2: *In the absence of external forces the forecasting process will prove to be stable and resilient.*

In the next section, we explain how we set about achieving our objectives and evaluating the above propositions.

3. Methodological issues and data collection

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; De Baets, forthcoming; Fildes & Goodwin, 2007a; Fildes & Petropoulos, 2015; Klassen & Flores, 2001; Mady, 2000; McCarthy et al., 2006; Sanders & Manrodt, 1994). 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 actors 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 (Easterby-Smith, et al., 2020; Nandhakumar & Jones, 1997; Walsham, 1995) 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 collection

The company we studied was, and remains, a cost-conscious UK subsidiary of an American pharmaceutical company that embraces a number of business units and 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 medical products for treating both animals and humans. It was chosen as part of a larger study of manufacturing and retail companies and software suppliers with forecasting processes similar to other manufacturers (Fildes & Goodwin, 2007b; Fildes, et al., 2009). 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 in the role of 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 that 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 (Benbasat, Goldstein, & Mead, 1987; Nardulli, 1978). This consisted of a ample of over 3000 forecasts that were supplied by the company and for which the actual outcomes were known.

Several years after, the initial fieldwork, 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.[4] This has allowed us to discover why the changes had occurred and how they had been made possible.

[4] 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.

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 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 forecast accuracy has always been seen as the primary objective (Fildes & Goodwin, 2007b).

At the start of the study, there were three logistics managers who were responsible for the initial forecasts for around 350 stock-keeping 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[5] 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.

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 on sales (these were known from data on orders) so that the series represented the level of demand. They then used the system to produce the 'base-line'

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

forecasts. These were forecasts which took 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 for a typical pattern). From this perspective, in the early years of a product's life it took time for demand to build up as doctors or veterinarians 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. Figure 1 also shows six 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 only designed to extrapolate linear trends as shown in Fig. 1. 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 [where] 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 were usually based on, at most, two years' 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 (see Önkal et al. (2008) for a laboratory study of how people adjust previously adjusted forecasts)

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 forecasts, supplied by the company, We first estimated the effect on accuracy of adjusting the FSS's automatic baseline forecasts. Because managers had kept no record of the original automatic statistical baseline forecasts that were 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 where there was sufficient data (136 from the total of 214 SKUs, a total of 3264 forecasts for years 2003 and 2004). 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 was available to us to fit the past data, we may have underestimated the system's ability to produce accurate baseline forecasts. Our analysis suggested there was little difference in accuracy between the company's judgmental baseline forecasts and those automatically calculated, despite the extra effort entailed in producing them - their median absolute percentage error (MdAPE) was 14.5% while the MdAPE for the automatic simulated baseline forecasts was 14.8%. 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 rather than serious damage to forecasting accuracy.

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. ~~However, e~~Earlier analysis has shown that the effectiveness of the adjustment depends on whether the system forecast is adjusted downward (where the balance of MI opinion of the demand planning team is negative) or adjusted upward in response to positive MI. Table 1 shows the effectiveness of the MI adjustments as they depend on the direction of adjustment and the size of the adjustment. Nine adjusted forecasts were set equal to zero where the actual demand proved positive and these have been excluded.~~.~~

	No. of Observations	Value Added: Bias (%)	Overall FVA (%)	Size of adjustment (no. of observations)			
	3264 (2034 adjusted)	15.5%	5.2%	<10%	10% to 50%	50% to 100%	≥100%
Positive information	1409	-11.0%	-10.7%	-10.6% (570)	-2.8% (620)	-9.9% (120)	-8.7% (99)
Negative information	1384	44.3%	20.1%	3.4% (683)	25.1% (613)	69.5% (79)	316%. (9)

Table 1: Forecast Value Added: Overall and by adjustment size. (Full sample of observations with 2 years or more of history, small observations excluded).

Forecast accuracy was measured using the Average Relative Mean Absolute Error (AvgRelMAE) defined by Davydenko and Fildes (2013). This measure is calculated as the geometric mean over SKUs of the ratio of MAE (final) to MAE (system) for a particular adjustment size. Table 1 reports the forecast value-added (FVA) measure calculated as (1-AvgRelMAE)×100%. This is a measure shown to be robust and readily interpretable (see Davydenko and Fildes, 2013). So for example, for positive adjustments between 50% and 100% the FVA is as –26.3%, showing that the adjustment leads to a decrease of 26.3% when comparing mean absolute error of the adjusted forecasts with the MAE of the corresponding system forecasts. Forecast bias was measured using the Average Relative Absolute Mean Error (AvgRelAME) (Davydenko, 2012, p. 64). In the AvgRelAME the ratio of the absolute mean errors is averaged across series using the geometric mean. Overall, the FVA and bias show an improvement from the adjustment, a conclusion confirmed by considering alternative measures of FVA.

We see from the results shown in the table that only -negative information adds consistent value overall and in particular the larger the negative adjustment the greater the value added by the MI process. However, the 100% adjustment where the final forecast is set equal to zero has led to substantial errors. The size of the adjustment is probably a measure of the strength of the market intelligence possessed by the members of the review meeting (despite this specific adjustment error).. It seems therefore that only when the proposed adjustment was substantial was the effort of making the judgmental adjustment potentially worthwhile and over-optimism when MI is positive typically led to a deterioration with (46% improvements compared to 59% for negative adjustments). Note that earlier analysis reported in Table 7 of Fildes et al. (2009) found that this company and two others were making inefficient use of the information contained in the previous period's demand figure and that this applied when adjustments were either positive or negative In essence, the activity of integrating information efficiently into the final forecast proved inadequate.

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 that prescribed in the research literature. This raises the question: why did managers in a company operating in a highly competitive environment adopt an inefficient sub-optimal 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, Fildes, & Kourentzes (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

As introduced in the literature review, 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. In what follows we adopt two lenses to understand what we have observed: the individual forecaster perspective and the organisation perspective.

5.1 An individual forecaster 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” (quoted in ~~(~~Asimakopoulos, et al., (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. 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, can 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 those 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-Szalanski, & Barnes, 1987). This emphasis on case-specific information meant that 'base-rate' information, like long term trends, was underweighted (Hoch & Schkade, 1996; Tversky & Kahneman, 1974). 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 sales 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 on offer 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.

To understand 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. 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. Ironically, the strength of the system in the network was the weakness of the inscription, allowing users the flexibility they valued.

However, the provision of an easy-to-use facility for judgmental intervention was also in the interests of the company's middle managers (see also Dietvorst et al. (2018)). 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 encouraging the participation of the product managers whose involvement 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 the 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 accurate forecasts to support the annual planning cycle as their objectives, 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. No evidence was available to contradict this view since error measures were rudimentary. 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, nor, by 2004, would it have had the necessary credibility. 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. The product managers would have faced the threat that their knowledge of their markets would be less valued and their ability to game the forecasts to their advantage might be restricted. 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 as the system designers had noted, generally FSSs are not designed around the concept of collaborative work.

6. How the forecasting system eventually 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.

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 diverse forecasting processes across the company and what 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. 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". However, a forecaster, who was one of those interviewed and most involved in using the old system in the subsidiary, commented that the new system lacked the graphical flexibility of the old, and was initially regarded as 'not as easy to use'.

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. Nevertheless, the developing role of the Centre of Excellence has the potential to innovate further through central 'advice' and 'guidance'.

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, in particular in the mis-match between the available models and the company's demand profile. 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, overriding the intentions of the system’s designers. As a result, consistent with Proposition 1, all those involved in the forecasting process accepted the limitations of the system, failing to explore its possible(and actual) inefficiencies o: its flexibility had been moulded to the tasks they faced and its maintenance was perceived to be in the various actors’ interests.

Secondly, the research has focussed on the on-going operation of an IS, where we observed the individual and organizational drivers for sub-optimal (economic) use with consistent biases, for example with the reduction in accuracy associated with positive information adjustments. From our observations of the original system and latterly, the changes that have taken place we answered the key question behind this study as to why these patterns of behaviour persisted. 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? Consistent with Proposition 2, no elements of the established network facilitated process improvement or provided an incentive for individuals to change their mental models of the forecasting task. 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. Yet the vested interests that many actors had in continuing to make heavy use of judgmental involvement remained in the new organizational processes, despite their known limitations. Further central involvement may well present the opportunity to limit damaging and time-wasting interventions.

Inevitably, our research has limitations. It is based on a single organization -though other research suggests that the forecasting process we observed is typical of other companies (e.g., Fildes et al, 2009; Franses and Legerstee, 2013). In addition, our conclusions are derived partly from our interpretation of the responses to interview questions and our observations of the forecasting process. This interpretation is bound to be influenced by existing theories and earlier findings in the literature. Finally, whilst we observed the company at two points in time, these did not permit a detailed analysis of the forces that led to such substantial changes as occurred in 2011-2014.

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' through software (Gilliland, 2008) where this is potentially valuable, into the forecasting process. Key elements in this are the design of encompassing data bases and user interfaces to support the S&OP activity (see the discussion in (Kaipia, et al., 2017)). As the empirical evidence we have presented here all too clearly demonstrates, there was neither the expertise nor the in-company appraisal to effectively incorporate the knowledge generated by the company's complex S&OP process. For such information to be used effectively through the FSS, customers will need to create a demand for more sophisticated systems and processes, 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, neither component of which was initially present in the case we have described. New developments becoming available in commercial software such as in machine learning (Fildes, Schaer, Svetunkov, & Yusupova, 2020) offer potential improvements but present particular problems of interpretation and implementation, such as the requirement for expert in-house knowledge. Nor does the value added in the established process, limited though it may be, necessarily carry over to these more advanced and complex models. However, even where they offer the potential for adding value, there may still be behavioural, organizational and political barriers that preclude efficient forecasting. Further engaged research is needed to investigate how these barriers can be overcome.

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

Acknowledgments
The authors would like to thank the anonymous contributors to this case study for giving their time to explaining the processes in the case organization. They 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. The research was supported by grants GR/60181/01 and GR/60198/01 from the Engineering and Physical Sciences Research Council (EPSRC).

References

Akrich, M. (1992).The description of technical objects. In W.E. Bijker & J. Law (Eds), *Shaping Technology/Building Society*, Boston: MIT Press, 205-234.

Armstrong, J. S. (2001). Evaluating forecasting methods. In J. S. Armstrong (Ed.), *Principles of forecasting: A handbook for researchers and practitioners*. Norwell, MA: Kluwer. 443-472.

Arvan, M., Fahimnia, B., Reisi, M., & Siemsen, E. (2018). Integrating human judgement into quantitative forecasting methods: A review. *Omega*. *86*, 237-252.

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.

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.

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.

Bainbridge, L. (1983). Ironies of automation. In *Analysis, design and evaluation of man–machine systems* (pp. 129-135): Elsevier.

Beach, L. R., Christensen-Szalanski, 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* (pp. 49-62): Chichester: Wiley.

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

Bendoly, E., & Cotteleer, M. J. (2008). Understanding behavioral sources of process variation following enterprise system deployment. *Journal of Operations Management, 26*, 23-44.

Caniato, F., Kalchschmidt, M., & Ronchi, S. (2011). Integrating quantitative and qualitative forecasting approaches: Organizational learning in an action research case. *Journal of the Operational Research Society, 62*, 413-424.

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

Czarniawska, B. (2017). Actor-network theory. In Langley, A. and Tsoukas, (eds.) *The SAGE*

Handbook of Process Organization Studies, 160-173Dalrymple, 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*, 319-340.

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

Davydenko, A. (2012). Integration of judgmental and statistical approaches for demand forecasting: Models and methods (doctoral dissertation). Lancaster University, UK, https://doi.org/10.13140/RG.2.2.31788.62083

Davydenko, A., & Fildes, R. (2013). Measuring forecasting accuracy: The case of judgmental adjustments to sku-level demand forecasts. *International Journal of Forecasting, 29*, 510-522.

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

De Baets, S. (forthcoming). Surveying forecasting: a review and directions for future research. *International Journal of Information and Decision Sciences*.

De Baets, S., & Harvey, N. (2020). Using judgment to select and adjust forecasts from statistical models. *European Journal of Operational Research, 284*, 882-895.

De Baets, S., & Harvey, N. (2018). Forecasting from time series subject to sporadic perturbations: Effectiveness of different types of forecasting support. *International Journal of Forecasting, 34*, 163-180.

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*, 114-126.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science, 64*, 1155-1170.

Dishaw, M. T., & Strong, D. M. (1999). Extending the technology acceptance model with task–technology fit constructs. *Information & Management, 36*, 9-21.

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. R., & Jaspersen, L. J. (2020). *Management and business research* Los Angeles: Sage.

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

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

Fildes, R., & Goodwin, P. (2007b). Good and bad judgment in forecasting: Lessons from four companies. *Foresight: The International Journal of Applied Forecasting*, 5-10.

Fildes, R., Goodwin, P., & Lawrence, M. (2006). The design features 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., & Hastings, R. (1994). The organization and improvement of market forecasting. *Journal of the Operational Research Society, 45*, 1-16.

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

Fildes, R., Schaer, O., Svetunkov, I., & Yusupova, A. (2020). Software survey: Forecasting 2020. *OR/MS Today, 47*.

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human perception and performance, 1*, 288.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behaviour: An introduction to theory and research. Reading. Mass: Addison-Wesley.

Franses, P. H. (2014). *Expert adjustments of model forecasts: Theory, practice and strategies for improvement*: Cambridge: Cambridge University Press.

Franses, P. H., & Legerstee, R. (2009). Properties of expert adjustments on model-based sku-level forecasts. *International Journal of Forecasting, 25*, 35-47.

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.

Gal, D. (2006). A psychological law of inertia and the illusion of loss aversion. *Judgment and Decision Making, 1*, 23-32.

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

Gilliland, M. (2008). Forecast value added analysis: Step-by-step. SAS White Paper.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 213-236.

Goodwin, P. (2002). Integrating management judgment and 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? *Foreside: 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., & Stephens, G. (2011). Restrictiveness and guidance in support systems. *Omega*, *39*, 242-253.

Hanseth, O., Aanestad, M., & Berg, M. (2004). Guest editors' introduction: Actor-network theory and information systems. What's so special? *Information Technology & People*, 116-123.

Harvey, N., & Reimers, S. (2013). Trend damping: Under-adjustment, experimental artifact, or adaptation to features of the natural environment? *Journal of Experimental Psychology-Learning Memory and Cognition, 39,* 589-607.

Heuer, J., Merkle, C., & Weber, M. (2017). Fooled by randomness: Investor perception of fund manager skill. *Review of Finance, 21*, 605-635.

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

Hyndman, R., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software, 27*, 1 – 22.

Kaipia, R., Holmström, J., Småros, J., & Rajala, R. (2017). Information sharing for sales and operations planning: Contextualized solutions and mechanisms. *Journal of Operations Management, 52*, 15-29.

Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.

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

Keen, P. G. W., & Morton, M.S.S. (1978). MS (1978). *Decision Support Systems: An Organizational Perspective*. Reading MA: Addison-Wesley.

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

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

Kottemann, J. E., Davis, F. D., & Remus, W. E. (1994). Computer-assisted decision making: Performance, beliefs, and the illusion of control. *Organizational Behavior and Human Decision Processes, 57*, 26-37.

Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological Bulletin, 133*, 95.

Latour, B. (2005). Reassembling the social. Oxford: Oxford University Press.

Lawrence, M., & O'Connor, M. (1995). The anchor and adjustment heuristic in time-series forecasting. *Journal of Forecasting, 14*, 443-451.

Mady, M. T. (2000). Sales forecasting practices of Egyptian public enterprises: Survey evidence. *International Journal of Forecasting, 16*, 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*, 303-324.

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 & Management, 42*, 243-257.

McGrath, K. (2002). The Golden Circle: a way of arguing and acting about technology in the London Ambulance Service. *European Journal of Information Systems*, 11, 251-266.

Mello, J. (2009). The impact of sales forecast game playing on supply chains. *Foresight: The International Journal of Applied Forecasting*, 13, 13-22.

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*: Ballinger Publishing Company Cambridge, MA.

O'Connor, M., Remus, W., & Griggs, K. (1993). Judgmental 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., & Gonul, M. S. (2005). Judgmental adjustment: A challenge for providers and users of forecasts. *Foresight: The International Journal of Applied Forecasting*, 13-17.

Önkal, D., Gönül, M. S., & Lawrence, M. (2008). Judgmental adjustments of previously adjusted forecasts. *Decision Sciences, 39*, 213-238.

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.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science, 3*, 398-427.

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

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.

Petropoulos, F., Kourentzes, N., Nikolopoulos, K., & Siemsen, E. (2018). Judgmental selection of forecasting models. *Journal of Operations Management, 60*, 34-46.

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

Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, *36*, 691-702.

Rose, J., & Jones, M. (2005). The double dance of agency: A socio-theoretic account of how machines and humans interact. *Systems, Signs and Actions*, 1, 19-37.

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*, 166-184.

Sanders, N. R., & Manrodt, K. B. (1994). Forecasting practices in United-States corporations - survey results. *interfaces, 24*, 92-100.

Sanders, N. R., & Ritzman, L. P. (2001). Judgmental adjustments of statistical forecasts. In J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Norwell, MA: Kluwer.

Sayes, E. (2014). Actor–Network Theory and methodology: Just what does it mean to say that nonhumans have agency?. *Social Studies of Science, 44*(1), 134-149.

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

Siegrist, M., Cvetkovich, G. T., & Gutscher, H. (2001). Shared values, social trust, and the perception of geographic cancer clusters. *Risk Analysis, 21*, 1047-1054.

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

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. (2018). Interpreting algorithmic and qualitative information when making judgmental forecast adjustments. In *International Symposium on Forecasting*. Boulder, Colorado: International Institute of Forecasters.

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

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.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 85,* 1124-1131.

Van den Broeke, M., De Baets, S., Vereecke, A., Baecke, P., & Vanderheyden, K. (2018). Judgmental forecast adjustments over different time horizons. *Omega*, 87, 34-45.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425-478.

Venkatesh, V., Thong, J. Y. L., & 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 is research: Nature and method. *European Journal of Information Systems, 4*, 74-81.Walsham, G. (1997). Actor-network theory and IS research: Current status and future prospects. In

A.S. Lee, J. Liebenau &.I. DeGross (Eds.). *Information Systems and Qualitative Research* London: Chapman and Hall, 466-480.

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.

Zackariasson, P., & Wilson, T. L. (2004). Internetworked after-sales service. *Industrial Marketing Management*, *33*, 75-86.

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