The effects of scenarios on judgmental demand forecasts and the subsequent production decisions

M. Sinan Gönül
Newcastle Business School, Northumbria University

Paul Goodwin
School of Management, University of Bath

Dilek Önkal
Newcastle Business School, Northumbria University

©M.Sinan Gönül, 19 June 2019
Overview

- Forecasts and Production Decisions
- Judgmental Forecasts and Cognitive Biases
- Research Focus
- Research Design
- Findings
- Discussion
Forecasts and Production Decisions

- Forecasts are designed to support decisions.

- Production decisions
  - **How much to produce** of each of a **number of products** when **constrained** by **manufacturing capacity**

- Forecasts for production decisions
  - **Point** forecasts to predict the central tendency
  - **Interval** forecasts to predict variability

- These forecasts allow
  - Setting **safety stocks** levels to **prevent stock-outs**
  - Achieving desired **customer service levels**
Forecasts and Production Decisions

- In production planning, managers likely access
  - *time series information* on *past demand*
  - *contextual information* relating to *demand*
    (Fildes et al. 2009; Fildes et al. 2018)

- The contextual information may take the form of *scenarios*
  - Scenarios are *powerful tools* that suggest *possible future events*
    (Godet, 1982; Goodwin and Wright, 2001)
  - Scenarios challenge managerial thinking and support strategic planning
    (Schnaars & Topol, 1987; Schoemaker, 1993; Önkal et al., 2013)

- The forecaster will have the task of *integrating* these two types of information
  to *generate forecasts* ↔ *production decisions*
Judgmental Forecasts and Cognitive Biases

- Such forecasts and the subsequent production decisions are often based on management judgment (Sanders and Manrod, 2003; Fildes and Goodwin, 2007)

- These judgmental forecasts are often inaccurate due to many cognitive biases (Lawrence et al., 2006)
  - Most recent observations may be overweighted (Bolger and Harvey, 1993; Lawrence and O’Connor, 1992)
  - Judgmental intervals tend to be too narrow, underestimating the variability

Overconfidence or hyperprecision (Arkes, 2001; Soll and Klayman, 2004; Önkal et al., 2009; Moore et al., 2015)
Judgmental Forecasts and Cognitive Biases

- **Scenarios** may help with problems due to these **cognitive biases**, especially **overconfidence** (Lawrence and Makridakis, 1989: Wright & Goodwin, 2009)

- Alternatively they **may not affect overconfidence** at all
  - **Middle ground scenarios** may **divert attention** from **extreme possibilities**
  - **Optimistic** scenarios may be preferred over **pessimistic**

(Schnaars and Topol, 1987; Newby-Clark et al., 2000)
Research Focus

- Scarcity of research on the interaction between scenarios and time series information and its effects on:
  - forecasts
  - in particular, the production decisions that follow

- This study aims to fill that research gap whether the availability of best and worst-case scenarios alongside time series information enhances or reduces the accuracy of demand forecasts → the subsequent production decisions
Participants (68 in total) were given time-series plots showing past demand over the previous 20 weeks for six products. For each product, they were asked to:

- make a point forecast
- give their confidence (probabilistic estimate) that the realized value would be within $\pm$ 5% of their point forecast
- make a production decision (i.e., decide on how many units they would order for production for a particular product)

This represented an important decision that required them to translate their forecasts (and confidence in these forecasts) into actual action given that the total production capacity was set to a fixed value (number of products x baseline demand).
Participants were randomly assigned to:

- **Group 1 – No scenarios**
  (23 participants)
  - the time-series information only

- **Group 2 – Both weak optimistic and weak pessimistic scenarios**
  (23 participants)
  - the time-series information,
  - weakly optimistic and weakly pessimistic scenarios
    (entitled as “Scenario A” and “Scenario B”)

- **Group 3 – Both strong optimistic and strong pessimistic scenarios**
  (22 participants)
  - the time-series information
  - strongly optimistic and strongly pessimistic scenarios
    (entitled as “Scenario A” and “Scenario B”)

Research Design
Scenario A:
Product K, a mobile phone with multifaceted functionality, has extremely stable demand. It has got all that is necessary to compete very successfully in its target market. It is an attractively designed phone with full-fledged features, and comes with a nicely positioned price and exceptionally encouraging promotion package. It regularly receives exceedingly positive comments in the industry magazines/websites and first-class feedback from customers. Given the recent economic conditions, we strongly expect even higher demand for this product in the periods to come.

Scenario B:
This product has been serving its purpose and target market for a long time. Its customers seem to be satisfied with it and its sales performance is stable within a band. It could have continued like this for some time. However, our company has been experiencing vital problems with a major supplier, which happens to be the producer of a key part for this model. If this dispute cannot be solved shortly, we certainly will not be able to produce Product K until we find another supplier with equally good credentials. While it is very difficult to replace the existing one, it will certainly take some time until (a) we find such a supplier, and (b) it starts delivering the required parts. If customers learn about this problem, there is a very high possibility that we will be faced with significantly lower demand in the next period.
YOUR FORECAST:

What is your *point forecast* for period 21: ..................

What is your confidence (probabilistic estimate) that

the realized value would be within ± 5% of your point forecast: ................. (between 0% and 100%)

YOUR PRODUCTION DECISION

How many units will you order for production? .................. (between 0 and 750)

(Please note that *total production capacity* for period 21 is 750 units. Therefore your production orders for all six products should add up to a maximum of 750. Please keep in mind that there are different costs associated with over-production vs. under-production and make your decisions accordingly. Please use the checklist in the end for production plans)
Findings – The time-series for product demands

- Artificially created to control the levels of uncertainty and trend - similar to previous studies on judgmental forecasting (e.g. Gönül, Önkal & Lawrence, 2006; Önkal, Gönül & Lawrence, 2008; Önkal, Sayım & Gönül, 2013)

- Six untrended series, half with high noise and half with low noise

\[ y(t) = 125 + \text{error}(t) \quad t = 0, 1, \ldots, 20 \]

- error(t) was normally distributed with zero mean and a standard deviation of:
  - 10% (i.e., 0.1 \times 125 = 12.5) for low noise
  - 20% (i.e., 0.2 \times 125 = 25) for high noise
### Findings – Accuracy of Point Forecasts

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of participants</th>
<th>MAE</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 - No scenarios</td>
<td>23</td>
<td>16.88</td>
<td>7.60</td>
</tr>
<tr>
<td>Group 2 - Weak scenarios</td>
<td>23</td>
<td>24.27</td>
<td>6.98</td>
</tr>
<tr>
<td>Group 3 - Strong scenarios</td>
<td>22</td>
<td>23.37</td>
<td>8.81</td>
</tr>
</tbody>
</table>

- Statistically produced forecasts (by *Forecast Pro*) on the series had MAE = 10.67
- So the software’s forecasts were substantially *more accurate* than those produced by human judgment.
Findings – Calibration of Confidence Assessments

Bias = True SD – Implied estimate of SD (from Confidence assess.)

Almost all variability estimates were **too low** compared to true levels of variation – suggesting **overconfidence**.

Statistically similar across all groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean Bias (low noise)</th>
<th>SD</th>
<th>Mean Bias (high noise)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 - No scenarios</td>
<td>7.66</td>
<td>1.40</td>
<td>19.60</td>
<td>1.43</td>
</tr>
<tr>
<td>Group 2 - Weak scenarios</td>
<td>6.96</td>
<td>3.43</td>
<td>18.66</td>
<td>4.24</td>
</tr>
<tr>
<td>Group 3 - Strong scenarios</td>
<td>7.11</td>
<td>2.00</td>
<td>19.66</td>
<td>1.77</td>
</tr>
</tbody>
</table>
Findings – Production Decision Quality

- Mean **customer service levels** (\% of demand that could be fulfilled) across the products vs. the **expected level of total sales** in week 21.
Findings – Production Decision Quality

- Mean **customer service levels** (% of demand that could be fulfilled) across the products vs. the **expected level of total sales** that in week 21.

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean service level</th>
<th>Expected total sales (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1 - No scenarios</td>
<td>44.90%</td>
<td>665.2</td>
</tr>
<tr>
<td>Group 2 - Weak scenarios</td>
<td>42.20%</td>
<td>628.4</td>
</tr>
<tr>
<td>Group 3 - Strong scenarios</td>
<td>42.60%</td>
<td>642.9</td>
</tr>
</tbody>
</table>
Discussion – Main Findings

- Providing scenarios to judgmental forecasters worsened forecast accuracy.
- Judgmental point forecasts of future demand (with or without scenarios) were less accurate than software produced ones.
- Judgmental forecasters perceived the demand variability to be much lower than its true value – demonstrating overconfidence.
- Scenarios did not reduce the tendency of forecasters to be overconfident.
- The production level decisions had a greater deviation from optimality when they also received best-case and worst-case scenarios.
Discussion – What Next?

- These findings raise two questions.

  - Why were *many of the decisions* so far from *the efficient frontier*?
    - Was this due to *inaccurate point forecasts*?
    - Was this due to *underestimation* of the variance of the probability distribution of demand?
    - Was this due to an *inability to handle* the need to *allocate* the total units of production capacity?

  - Why did those who *did not receive scenarios* make *‘better’* decisions?
THANK YOU