# POLITECNICO DI TORINO

Collegio di Ingegneria Gestionale

**Master of Science in**

**Engineering and Management**

**Master Thesis**

# Behavioral Aspects of the Newsvendors:

# A Study of their Variability


Supervisors:
Prof. Zotteri Giulio
Prof. Cantamessa Marco
Prof. Colombo Samuele

Student: Elisa Mastronardi


Year 2022/202

# Index:

Abstract .................... **Errore. Il segnalibro non è definito.**

List of Abbreviations .................... 4

1. INTRODUCTION AND NEWSVENDOR PROBLEM AND MODEL'S PRESENTATION .................... 5

1.1 Problem definition .................... 5

1.2 The objective of the research .................... 7

1.3 The Methodology .................... 7

1.2 The Structure of the Dissertation .................... 8

2. LITTERATURE .................... 9

2.1 The Newsvendor Problem and its mathematical formulation .................... 9

2.2 The main theoretical contributions to Newsvendors' behavior .................... 11

2.2 The Pull to Center effect .................... 11

2.3 The phenomenon of asymmetry .................... 14

2.4 Overview of Behavioral Theories applied to the Newsvendor .................... 15

2.4.1 Cachon and Schweitzer contribution .................... 15

2.4.2 The Prospect theory: .................... 16

2.4.3 The Learning effect .................... 18

2.4.4 The Overconfidence point of view .................... 19

2.5 Main Trends in the literature .................... 20

3. RESEARCH QUESTION: THE DRIVERS OF OVER TIME VARIANCE OF ORDERED QUANTITIES AND DEMAND FORECAST .................... 22

3.1. The main questions investigated and the underlying hypothesis .................... 22

4. METHODOLOGY AND DATA COLLECTION .................... 25

4.1 The experiment .................... 25

4.2. Measures and Variables .................... 26

4.3. Software and Statistics employed .................... 28

4.4. Data Collection Methodology .................... 29

5. RESULTS' DISCUSSION .................... 31

5.2 Results of Forecast and Purchase Explorative ANOVAs .................... 31

5.3 Analysis of the main drivers of variance ..... 42
1.3 The relationship between previously realized demand and respondent behavior ..... 46
5.5 Economic impacts of Variability ..... 49
6. CONCLUSIONS ..... 55
4.3 The results obtained ..... 55
4.4 Limitations and Future steps ..... 57
Acknowledgments ..... 59
APPENDIX A ..... 60
Figure Index: ..... 74
Bibliography ..... 76

# Abstract

In this dissertation, we propose an analysis of the variability in the answering behavior over time of respondents participating in the Newsvendor Model tasks, the so-called “newsvendors”. When posed in front of the Newsvendor task participants seems to be prone almost to every kind of mistake made (O’Keefe T. Carlson, 1969), with deep oscillation of the quantity purchased during each repetition of the model. In other words, the orders placed by the newsvendors not only differ from the optimum at the aggregate level, as already indagated by many prominent authors, but they also differ from each other in terms of how far they are from it. These points directly translate into a different level of economic inefficiency or efficiency presented by the same single participant. As put in the apt word of Daniel Kahneman, our ability to look at the past and to infer the right thing from it is sorely lacking. Humans seem to be, as a matter of fact, not prone to “rational decision-making” and, even more, unlikely to decide in a way that could be defined as “economically” efficient or rational (Kahneman D., 1981). The idea to focus on this aspect of a vast and rich topic such as the Newsvendor Model emerged from the presence of a gap in the literature on this research field, in which it seems that close to no contributions studying this characteristic are present. The investigation was focused on not only studying variability present in the orders placed by participants but also on their forecast of the expected demand level for each period. Indeed, the two variables present enough differences to be studied separately. Starting from this assumption we tried to investigate the main drivers of the variability in the answering patterns by focusing on the levels of demand forecasted and exploring, through the usage of Analysis of Variance (ANOVA), the effects of different treatments on the respondents’ behavior when forecasting. Some particular combinations of various factors held some interesting insight: specific combinations of framing and product margins can lead respondents to have more optimistic forecasts than what would have been forecasted by an ideal optimal newsvendor participant.

# List of Abbreviations

| Abbreviation | Definition |
| --- | --- |
| **PtC** | Pull to Center Effect |
| **PT** | Prospect Theory |
| **NV** | Newsvendor |
| **NVM** | Newsvendor Model |
| **HCosts** | Costs faced y participants, the so-called “human newsvendors” |
| **ICosts** | Costs that would have been incurred by the optimal Newsvendor |
| **HVC** | Human Variability in Newsvendor quantity at period t |
| **VCosts** | Variability cost |
| **ANOVA** | Analysis of Variance |

# 1. INTRODUCTION AND NEWSVENDOR PROBLEM AND MODEL'S PRESENTATION

## 1.1 Problem definition

The Newsvendor problem is a classical analytical problem featured in Inventory and Operation Management literature (Marschak, 1951). Speaking of the Newsvendor Model, we refer to a mathematical model employed to determine optimal inventory levels for the newsvendor problem.

In this problem, a vendor or seller, usually called *"newsvendor"*, is asked to determine the number of perishable goods to order for a subsequent period, given a forecast of the demand. The aim of the newsvendors during the Newsvendor's tasks is to maximize expected profits (Cachon, 2000).

This type of problem is characterized by three defining aspects: the presence of a random amount of a resource to be determined, the fact that this quantity must be selected before observing how much is needed, all the economic consequences are observable and representable by known opportunity costs (Cachon, 2000; Benzion, 2008; Porteus, 2008). Findings from its application have also been extensively employed in the design of supply chain contracts and optimal inventory systems (Bhavani Shanker Uppari, 2019).

The origins of this problem can be traced back to "A Mathematical Theory of Banking" (Edgeworth, 1888), one of the foundations of modern Inventory theory (Petruzzi, 1999). However, it was just in the Fifties this problem became a topic of extensive study by academicians (Petruzzi, 1999): primarily with Arrow and Marschak's work on Optimal Inventory policies, whose 1951 article introduced the model's first formalization, (Marschak, 1951) and then with Whitin's (1955), who was the first to illustrate margin effects in the newsvendor. The problem is still relevant today as the many contributions in terms of Scientific articles production demonstrate and, even more so, due to the growing interest in supply chain management: many commodities are seasonal or have a short life cycle and for example, replenishment decisions can be studied by applying Newsvendor model insights (Wei, 2021)

From those experimental outcomes has emerged a substantial deviation from the optimal order quantity of the Newsvendor Problem. Besides that also heterogeneity and asymmetry have been found in the individuals' answering patterns: to investigate the various factors that are being taken into account, from individuals characteristics (e.g. level of instruction, CRT score, gender, and so on) to experimental specifics specifically set such as Framing and Margin and period-wise factors. The asymmetry is intended as a significant distance between the order quantity and the optimal inventory level for different product types. One well-established source of this distortion is related to the profit margin of the products (Cachon, 2000). The analysis of these results opened a new stream of research directed toward a better understanding of the decision-making mechanisms of subjects and their managerial implication.

The main theories, which explain at the aggregate level the answers given and the sub-optimal performance of newsvendors, are mostly inspired by the field of Behavioral Economics. Evidence in this stream demonstrated that human behavior is not as rational as traditional normative theory suggests due to crucial aspects influencing the decision-making process (Shefrin, 2018), most notably what derives from Prospect Theory insights (Kahneman & Tversky, 1979). Individuals are subjected to heuristics, as in wrong simplifications and judgments made on stereotypes and biases, so to a predisposition to making errors while deciding.

The introduction of theories like, for example, bounded rationality (Su et al., 2008) and others to try to motivate the peculiar ordering (reference) gave an outstanding contribution to the investigation of what moves and hinders individuals' performance and managerial behavior. All these reasoning can be applied to inventory allocation problems and everyday business reality.

Plenty of articles have shown the presence at the aggregate level of correlation, or interestingly the absence of it, of the newsvendor ordering patterns to psychological aspects such as risk aversion and risk-seeking behavior, framing effects, waste aversion, and so on. Some of these theoretical explanations, rooted in the human characteristics of the decision-making process, have been found to elucidate better the ordering behavior than others, which have been completely ruled out from possible causes like risk seeking and risk aversion (Benzion, 2008)).

However, something that still seems to miss an explanation is the high variance, visible in the great distance present in the quantities ordered by newsvendors over time. Indeed, the orders placed by a newsvendor not only deviate from the optimum at the aggregate level so on average, but they also differ between task repetitions in terms of how far they are from the optimal values. Thus, some newsvendors will present a higher variance in the answers proposed task after task: in some cases, they will place orders further away from the optimum and incur high losses, while in other cases, they will stay closer to the optimal quantity, generating margins closer to the optimality levels.

It is evident the high potential of investigating the causes at the origin of the variability manifested: they could help determine the characteristic of a high-performing manager and even how to minimize these effects in the everyday business reality since, as the distance from the optimum increases, losses increase almost exponentially and having a decision-maker prone the less variance could be extremely advantageous.

Before addressing the implications of this topic, we should first define what drives variability in orders placed and demand forecast, both in terms of external inputs and internal experimental settings, and the implication in terms of the cost efficiency (or inefficiency) of newsvendors.

Can the variance in the newsvendors' forecast of demand be influenced by previous realizations of demand or by shock events? Does the same apply to purchased quantities?

What are the effects of an instructional lecture held on the Newsvendor? Does it reduce variability or make newsvendors more prone to errors?

## 1.2 The objective of the research

Up to this point, to the best of our knowledge, almost no papers tried to investigate this issue. It seems that most of the interest fell on the mean behavior of the newsvendor. The only other two types of investigation performed fell instead on the difference among individuals' behaviors and heterogeneity, instead of the variance manifested over time, even if it yields important insight into the level of economic efficiency or inefficiency of decision-makers (e.g., ref1, ref2, etc.).

So as stated before, the main objective of this dissertation will be to further investigate the causes and drivers of variability in newsvendors' ordering patterns, which present high oscillation over time. To identify them we will investigate which characteristics are linked to a worse performing newsvendor, one that periodically buys quantities further away from the optimal one incurring a higher loss, and could bring a great contribution to the Inventory Management fieldOur other purpose will be to determine their impacts on the cost incurred to determine how the economic inefficiency of newsvendors is attributable to irrationality.

## 1.3 The Methodology

To uncover further insights into the irrationality which seems to affect participants, we collected scientific articles to identify the best practices, and consolidated findings were collected. From this investigation, as previously stated, it also emerged the gap in the issue of variability.

By employing search engines such as Scopus, EEExplorer, and Research Gate (the first two engines were used to determine the most authoritative and notable papers, and the last one was employed to see the main trends present) we had an overview of the classic Newsvendor Problem and the Newsvendor model formulation. We selected the most reputable sources based on the type of journal/periodic in which they were published and the number of citations. From this first screening, we individuated many capital papers, such as Schweitzer and Cachon (2000), Benzion and Cohen (2008), etc.

Then, to have a clear picture of the studies dedicated to behavioral aspects of the Newsvendor, we searched for reviews collecting the main contribution taken from Behavioral Economics and then proceeded to analyze the most notable papers for each stream of research: for example, the work of Bostian, Holt, and Smith for the PtC effect description (Bostian, Holt , & Smith , 2008), Ren and Croson for Overconfidence (Croson , Ren, & Croson, 2009; Croson D., 2017)and so on.

Our last step was to determine the principal trends in this research field by looking at the most recent articles available. For this search, two trends were visible: the interest in the exploration of individual newsvendor characteristics so of the heterogeneity and the application of neuroscience and biofeedback during the performance of the experiment.

Once the scientific landscape was determined based on the main insight derived from the literature, we employed the experimental data obtained from performing the Newsvendor problem in a university setting as the basis of our analysis: once pre-processed the data were used in a series of explorative Analysis of Variance to try to answer to main open point linked to variability

## 1.4 The Structure of the Dissertation

The structure of this dissertation will thus follow a similar pattern to the methodology which we have exemplified above. First, we present the Newsvendor Problem in its mathematical characteristics and provide a summary literature review focusing on the main contributions linked to behavioral theories. Subsequently, we will propose our research questions with their underlying hypothesis and provide a deep dive into the methodology applied and the data collection process, describing how the experiment was organized and carried on and how data were processed. Then, we will discuss the results obtained for the precedent step and describe what can be inferred by the analysis executed and their economic impacts. To conclude, we will summarize the main insights from the results obtained in the final chapter of this dissertation and highlight the limitations of this research and a few ideas for future contribut

# 2. LITTERATURE

The Newsvendor Problem and its mathematical formulation

This section proposes a deep dive into the Newsvendor problem structure and its mathematical formulation. It will also focus on the expected profit-maximizing solution, the so-called Newsvendor Model, which determines the optimal inventory levels of the problem.

One of the first mathematical formulations of the problem comes from Marschak, Arrow and Harris's 1951 capital work, "Optimal Inventory policy", in which they first delineated a model under the assumption of certainty and then moved to study models under uncertainty (Marschak, 1951) .

As already stated, in the Newsvendor problem a decision-maker decides the quantity to order, q, before a single selling period, without any possibility of replenishment later on once the true demand in manifested. In the classical analytical single period problem, the only controllable variable is q, and all the other ones, such as selling price p, are approximated as fixed or assumed as non-controllable and constant, such as inventory holding costs, salvage prices, etc. In the original model the demand, D, is considered random and its distribution is uniform.

Let us call F the distribution function and f the density function.

As a rule, any inventory left over at the end of the period for which is ordered is scrapped and cannot be used at a later time. While, in case of stock-out, a loss is expected because of extra demand that is not satisfied during the period (D.F. Pyke, 1998).

The decision maker purchases each unit for cost c and sells each unit at price $p > c$. When $q > D$, each unit remaining at the end of the period can be salvaged for $s < c$.

The realized profit $\pi(D, q)$ will be a function of the realized demand and the actual quantity ordered as defined by the following equation:

$$\pi(D,q) = (p-c) \cdot \min(q,D) - (c-s) \cdot q \quad (1.1)$$

Our expected profit then will be:

$$\mathbb{E}(\pi(q,D) = \left(1 - F(q)\right) \cdot \pi(q,q) + \int_0^q f(x) \cdot \pi(q,x)dx \quad (1.2)$$

By applying Leibniz's rule to obtain the first and second derivatives, we can show that $E(\pi)$ is concave. (Khouja, 1999).

The optimal order quantity q*= arg max $\mathbb{E}$ [$\pi$ (q, D)] can be computed as the unique solution of the following equation: $F(q*) = \frac{p-c}{p-s}$ (1.3)

The abovementioned fraction is also called the critical ratio. Based on this ratio we can establish a rule on how to distinguish between a low-margin setting and a high one, for example in Schweitzer and Cachon (2000) a product was defined as high margin when:

$$\frac{1}{2} \leq \frac{p-c}{p-s} \tag{1.4}$$

Another quite common formulation of the equation for the profits per period is:

$$\pi = \begin{cases} (p-c) \cdot q - S \cdot (x-q) & if\ x \geq q \\ p \cdot x + s \cdot (q-x) - c \cdot q & if\ x < q \end{cases} \tag{1.5}$$

Where x is our realized demand, c indicates the cost of purchasing the good to sell, and S is the shortage penalty cost per unit (which in some model is considered separately from the salvage value in case of overstock).

Starting from the expected profit we can derive the expected newsvendor's utility function1:

$$E(u(w0) + \pi(q,D)) = (1-F(q))u(w0 + \pi(q,q)) + \int_0^q f(x)u(w_0 + \pi(q,q))\, dx \tag{1.6}$$

Where u($w0$) is the utility gained by the subject over its initial wealth, w is the final wealth and q is the ordered quantity.

The underage and overage costs are defined as:

cu= p -c+g

Where g is defined as the customer goodwill lost for each unit of unsatisfied demand.

co= c-s

Over time, beside the various notational difference and formulations, there has been an evolution of the newsvendor problem and many extensions of it have been proposed in the last decades: the researchers started expanding the model from a single-period decision to a multiperiod one, and they

[1] For a comprehensive list of the various utilities function utilized in these cases see Schweitzer and Cachon,2000 who provided a comprehensive list of several alternative functions

also shifted from static to dynamic models, where a change in selling prices is contemplated, and included other aspects like competition, etc.

## 2.1 The main theoretical contributions to Newsvendors' behavior studies

A question that might arise at this point of the discussion could be: why there was a push toward the study of the behavioral aspects in the newsvendor allocation decisions and why they are among the main theoretical contribution to the research in this field?

We can simply answer by looking at the everyday person's behavior when put in front of a choice under uncertainty. Decision-makers are seldom rational and usually employ mental shortcuts and intuition while deciding (Yamini, 2020), in order to reduce the difficulty of the tasks of predicting a value and evaluating probabilities into simpler operations (Tversky A., 1974).

Furthermore, individuals' mental processes are also affected by biases, so errors, which sometimes can be severe, and arise from cognitive limitations, due to their lack of information, wrongful perception of time, and so on.

This is the main reason behind the abundance of theories taken from Behavioral Sciences (in particular Behavioral Economics) as an explanation of the peculiar newsvendor ordering behavior: these cognitive limitations, proper of every human being, cannot be excluded during the analysis of the newsvendor problem, since assuming a fully rational decision maker would be deeply unrealistic.

In the following section, we will review the main theories, involving behavioral aspects and explanations of newsvendor purchasing habits, that have emerged over the years, with particular attention to some articles which represented capital contributions and influenced the main direction toward which research evolved.

## 2.2 The Pull to Center effect

One of the first Behavioral Economics' application to the Newsvendor Problem, was done to explain an experimental phenomenon, which is visible after a simple collection of data from respondents, referred to as the "Pull-to-Center effect". This effect was first identified by Schweitzer and Cachon (Cachon, 2000) and scientific papers have tried to explain it by applying many different behavioral theories, from Prospect theory to risk aversion, anchoring and demand-chasing behavior with different rates of success.

With the term Pull-to-Center ("PtC") we indicate the empirical phenomenon for which newsvendors systematically place order quantities between the optimal quantity q* and the mean of the demand distribution d (Bostian, Charles, Holt, & Smith, 2008).
The first description of this effect appeared in a 1996's study by Fisher and Raman, in which it was found that managers systematically underordered (Fisher, 1996). In this study, in which one of the main points was to analyze the effects of a corrective algorithm, it did not explain the causes of this phenomenon.
The 1996 article from Fisher was one of the first instances in which a distinctive ordering pattern was noticed, while in precedent studies it was simply highlighted that subjects made "almost every kind of mistake" (O'Keefe T. Carlson, 1969). For a better analysis of the underlying causes behind this effect, we refer to Schweitzer and Cachon's 2000 paper, which states that "subjects consistently order amounts lower than the expected profit-maximizing quantity for high-profit products and higher than the expected profit-maximizing quantity for low-profit products" (p.418), nevertheless, they order above the mean demand level in the first case and below the mean in the second (Cachon, 2000). This interval comprised between mean demand and the optimal profit-maximizing quantity is known as the "*PtC zone*" (Bhavani Shanker Uppari, 2019).
The aforementioned behavior and the PtC interval are clearly shown in the graph illustrating the experimental results obtained by Schweitzer and Cachon:

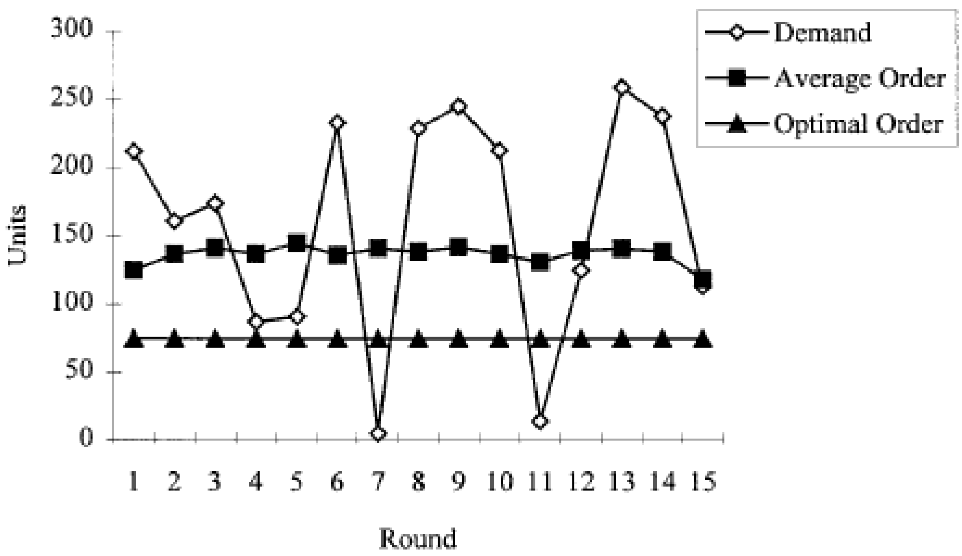



**FIGURE 1: LOW MARGIN CASE FROM SCHWEITZER AND CACHON (2000), "DECISION BIAS IN THE NEWSVENDOR PROBLEM WITH A KNOWN DEMAND DISTRIBUTION: EXPERIMENTAL EVIDENCE". *MANAGEMENT SCIENCE*, 404-420.**

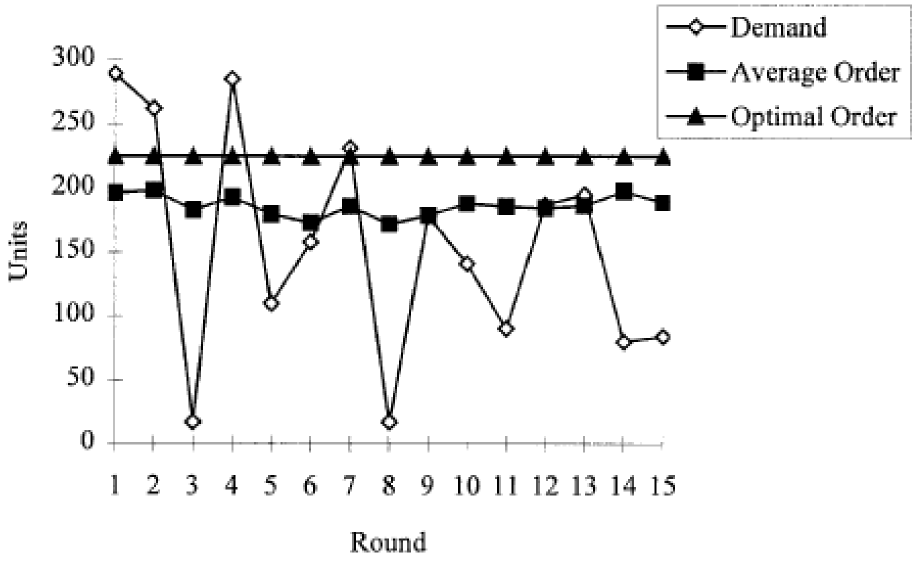
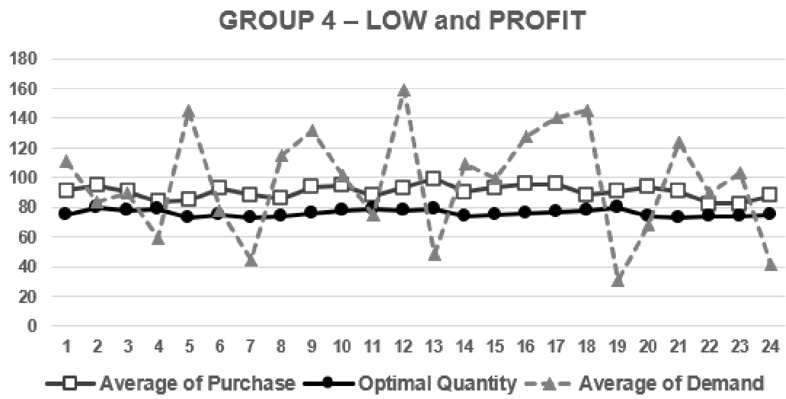



**FIGURE 2: HIGH MARGIN CASE FROM SCHWEITZER AND CACHON (2000), "DECISION BIAS IN THE NEWSVENDOR PROBLEM WITH A KNOWN DEMAND DISTRIBUTION: EXPERIMENTAL EVIDENCE". *MANAGEMENT SCIENCE*, 404-420.**

Other successive studies confirmed similar findings (Bostian, Holt , & Smith , 2008; Katok, 2008).

Again, the same pattern, perfectly attributable to the pull-to-center effect, was found in the empirical data coming from the Newsvendor Problem experiment carried out in 2019 at the Polytechnic of Turin. For comparison, the graphs plotting the results obtained are shown below.



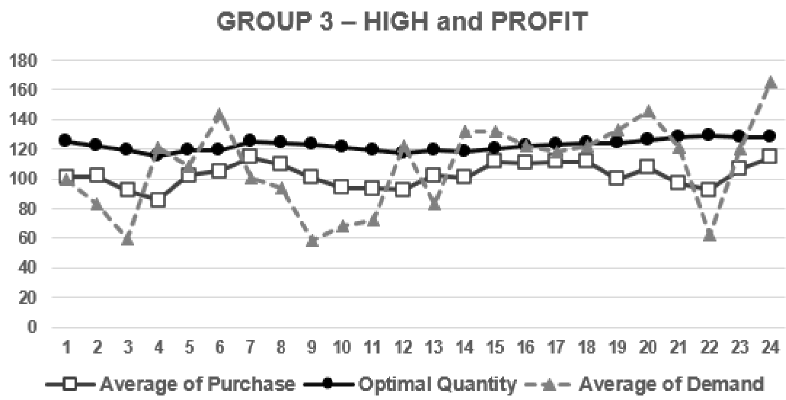



FIGURE 3: GRAPHS PLOTTED USING THE 2019 POLYTECHNIC OF TURIN'S EXPERIMENTAL DATA

The first studies proposed a more normative approach and the research focused on trying to explain the phenomenon affecting ordering decisions using mathematical modelization methods, but over time they resulted inadequate to fully explain decision-maker inventory allocation patterns (Kremer, 2010). This prompted researchers to shift their focus on the implementation of laboratory experiments designed to tackle every aspect of the newsvendor: empirical experiments allowed researchers to maintain a high degree of control on variables of interest and to better isolate and identify relationships between those and the behavior of the individual and eventual causal connections. Experiments concerning the newsvendor weren't an unseen thing: Clarkson and O'Keefe (1969) were the first to report it in an article related to buffer stocks and reaction coefficients (O'Keefe T. Carlson, 1969), even if it was in the setting of a much larger laboratory decision-making experiment (Benzion, 2008).

Based on empirical results various attempts by the scientific community have been made to explain the pull-to-center effect through the analysis of data obtained via experiments.

Amid these empirical experiments academics were faced with a fascinating aspect of the Newsvendor problem: the pull-to-center effect cannot be explained by the classical risk and loss aversion theories (Bostian, Charles, Holt, & Smith, 2008).

Hence, numerous models based on decision-making heuristics such as mean anchoring and insufficient adjustment, minimization of the expected *ex-post* inventory costs, and demand chasing behavior have been proposed (Cachon, 2000).

In addition to that, Prospect theory has been advanced as an explanation of the ordering behavior at first with mixed results: only a few authors have shown some success at replicating a pull-to-center ordering pattern (Chirag Surti A. C., 2020). Meanwhile, better results have been achieved by applying bounded rationality theories (Su, 2008).

## 2.3 The phenomenon of the asymmetry

Another fundamental finding by Schweitzer and Cachon, besides the PtC, was the asymmetry in the ordering pattern between high-profit and low-profit settings of the experiments. It appears that orders are closer to the expected demand for the low-profit scenario than for the high-profit one (Cachon, 2000), in other words, the extent of the deviation of the purchase from the mean demand are different among the two types of products. Also, on this aspect plenty of theories have been developed adapting and employing aspects coming from Behavioral Economics. Plenty of research was done on the topic, mostly after 2010, to explain the asymmetry in a normative way and to demonstrate that its presence can still be compatible (in the aggregate) and be explained by the aforementioned behavioral models, both for the reference-dependent newsvendor and in the mean preserving one (Jammernegg W., 2021; Croson D., 2017; Benzion, 2008). It was demonstrated that both models are well suited for heterogeneous ordering and at the same time allow overordering, as well as underordering, for the high-margin/low-margin product.

The asymmetric behavior is explained by the incidence of other two factors: stockout and waste aversion. In the case of the reference-dependent newsvendor, if she/he overorders both products stockout aversion is to be blamed, while in the case of underordering waste aversion is the dominating factor. Moreover, a mean preserving newsvendor overorders both products if the demand variance of the low-profit product is underestimated, in other words, if we are in the presence of overconfident behavior, and the demand variance of the high-profit product is overestimated (underconfident behavior) (Jammernegg W., 2021).

## 2.4 Overview of Behavioral Theories applied to the Newsvendor

To exemplify the analysis of the literature performed and the results obtained by investigating the application of behavioral theories on the classical Newsvendor framework, we have presented below a brief overview of key authors' contributions, starting from the cornerstone 2000's paper from Cachon and Schweitzer, *"Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence"*.

### 2.4.1 Cachon and Schweitzer contribution

Cachon and Schweitzer were the first academics who tried to disentangle all the possible biases imputable as the cause of the pull-to-center effect in the newsvendor context and, more generally, of the overall ordering decisions: the experimental evidence and the results obtained were the setting stones for all the following research on the newsvendor.

They started from the assumption that individuals' choices could be moved from other facts besides profit maximization. Participants can have:

1. Different preferences other than profit maximization.
2. They can apply heuristics in their decision-making process.

3. They can have biased forecasts of demand distribution (even if this factor was excluded from the study and the demand distribution was deemed as known).

A summary of the main aspects investigated in their 2000's paper is proposed in Table 1.

One of their main hypotheses linked the inventory decision pattern of the newsvendor to a mean anchoring and insufficient adjustment heuristics: the decision maker selects an anchor, be it the mean demand or the prior period demand realization, etc., and then adjusts away from that amount (Tversky A., 1974). In the 2000 article, consistency was found between the ordering behavior and the anchoring on mean demand, while it was found only weak support for the *chasing demand* heuristics (Cachon, 2000). Ultimately, they concluded also that the phenomenon could be explained by the desire to minimize the difference between the quantity purchased and the actual demand realized, the so-called ex-post inventory error.

| ***Behavioral Theories and Biases*** | **Does it explain the Pull-to-center effect?** |
|---|---|
| *Risk-seeking and risk aversion* | No |
| *Prospect theory* | Mixed results |
| *Loss aversion* | No |
| *Waste aversion* | No |
| *Stock-out aversion* | No |
| *Minimizing Ex-Post Inventory Error* | Yes |
| *Anchoring and insufficient adjustment* | Yes |

**TABLE 1: SCHWEITZER AND CACHON 'S (2000) ANALYSIS RESULTS**

The experimental evidence was collected via two experiments: one utilizing the classical newsvendor setting with uniform distribution and one with a high demand distribution, to have only positive profits. From this second experiment, they found out that if the problem is framed entirely in the domain of gains (negative profits are not possible) PT won't explain newsvendor behavior: it will predict in fact that subjects will be risk averse and always order below q*, which does not happen (Cachon, 2000).

### 2.4.2 The Prospect theory:

Some of these results were successfully disproved by later efforts: most notably the ones on Prospect Theory (for short it will be indicated as PT from now on).

Given that PT is one of the most used and popular frameworks for modeling decision-making under uncertainty (Nagarajan, 2014), the earlier results from Schweitzer and Cachon were found somewhat surprising. In fact, by further research, the prospecting behavior was found to perfectly explain the newsvendor purchasing choice under specific assumptions (Chirag Surti A. C., 2020).

The Prospect theory was developed by Nobel prize Daniel Kahneman and Amos Tversky in 1979 using controlled studies: it describes how "prospects" matter, in other words, how individual risk preference changes asymmetrically if a problem under uncertainty is framed in terms of gains or in terms of losses (Kahneman D., 1981), which is totally against the rational decision-maker behavior under the Expected Utility theory assumptions. More precisely, PT suggests that individuals don't choose the highest expected utility solution but:

- Individuals do not consider the "utility" of absolute outcomes, but the positive or negative deviations (gains or losses), from a reference point.
- When faced with a risky choice framed in terms of gains, they become risk averse favoring more secure and certain outcomes, even if they have lower expected utility.
- While they become risk-seeking when the choice is framed in terms of losses preferring to potentially avoid losses over higher utilities.

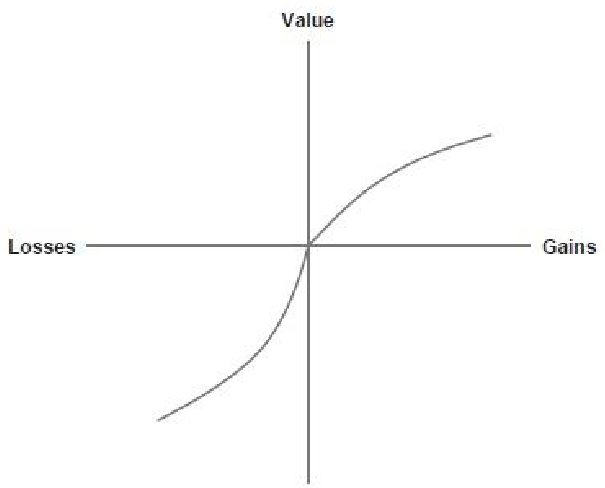



FIGURE 4:PROSPECT THEORY'S OUTCOME VALUES ASYMMETRY. IMAGE TAKEN FROM PSYCHOLOGY.COM HTTP://PSYCHOLOGY.IRESEARCHNET.COM/PAPERS/PROSPECT-THEORY/

It is fundamental to note that Shweitzer and Cachon's results were obtained by considering as a reference point the zero profits situation: with this assumption they obtained that the PT could not predict the PtC effect in the case of exclusively positive payoffs as a result of inventory decisions. This argument was later on refuted by the 2015 study by Long and Nasiry that, by using as a reference point the weighted average of the maximum and minimum expected profits obtainable with a particular order quantity, demonstrated that the PT can correctly predict the *PtC.* The only point under discussion remains the arbitrariness of the chosen reference point since the average they employed is somewhat unjustified both on an empirical and theoretical basis. The same cannot be said for models proposed later on, some of which also managed to incorporate individual-level heterogeneity in the purchased quantities, for which a more comprehensive approach was preferred

(Bhavani Shanker Uppari, 2019). A reference point should be *a salient point* within one's cognitive norm[2] (Kahneman D. , 1992): under this assumption, the authors of the 2019 paper "*Modelling Newsvendor Behavior: A Prospect Theory Approach"* adopted first the maximum, minimum and sure-shot payoffs as a reference and then, as a stochastic reference, the demand-related information, due to the fact that in the Newsvendor experiment subjects are usually presented with the demand distribution and not the payoff distribution. In particular, the predictive performance of the model obtained by using the mean demand as a reference point outperforms all the others with a substantial reduction of the prediction errors when compared with other models results.

### 2.4.3 The Learning effect

The presence of learning behavior across multiple sessions of the game was another main point of focus in the research on the Newsvendor. Feedback and information frequency were examined as drivers of newsvendors' adjustment towards or away from the Newsvendor problem optimum, as done by Bolton and Katok (Bostian, Charles, Holt, & Smith, 2008), or with respect to generic decision-making processes (Lurie & Swaminathan, 2009). Interestingly, from this contribution emerged, more generally, that by looking at inventory management experiments in environments characterized by random noise, frequent feedback on previous tasks/decisions leads to decline in performance (Lurie & Swaminathan, 2009).

While, as already stated, Bostian and Holt focused more on specific effects visible in the Newsvendor by creating various learning models based on the three main heuristics proposed by Schweitzer and Cachon (Bostian, Holt , & Smith , 2008): (1) anchoring on mean demand and partially adjusting towards the optimal order, (2) anchoring on the most recent order quantity and partially adjusting towards the most recent demand observation (usually referred to as *demand-chasing behavior*), and (3) maximizing a utility function subjected to regret for *ex- post* inventory errors, modelized by including a penalty for errors in either direction.

These learning models were investigated with different results (Bostian, Holt , & Smith , 2008): the one that best unifies the effect observed in the experimental data collected was the experience-

[2] With the term *Normative Cognition,* we define a theoretical construct for which human cognition is governed by a set of social conventions, values, and norms shared by groups of individuals.

weighted attraction model (for short *EWA*) (Bostian, Holt , & Smith , 2008; Cramer & Ho, 1999). This model, which incorporates memory, and reinforcement effects to explain individuals' choices, permits noisy adjustments towards the optimum with **recency effects**, for which decisions are considered to be affected more by recent events, and **reinforcement effects**, for which strategies are thought to be 'reinforced' by their previous payoffs. In Bostian, Holt, and Smith's work the propensity to choose a strategy depends in some way on its reinforcement, so it is influenced mostly by payoff strategies yielded in the past (Cramer & Ho, 1999). Not only do the predictions obtained by applying the model exhibit similar behavior to the collected data, but they also replicate the pull-to-center effect.

### 2.4.4 The Overconfidence point of view

Another point of interest that emerged from a screening of the available literature was the introduction of the concept of overconfidence.

So far, the few papers which have investigated this aspect, overconfidence, concerning the Newsvendor obtained convincing results at the aggregate level. In one of the most notable ones, it appears that overprecision explains about one-third of the observed ordering errors and that the effect of overprecision is robust to learning (Ren & Croson, 2013).

While in a later contribution of the same authors the theoretical model appeared to be also quite in line with the responses observed by Schweitzer and Cachon (Croson D., 2017).

Those first experimental studies tried to elicit particularly the over-precision aspect of the multifaceted overconfidence phenomenon in the respondents (Moore D., 2008) through a pre-task test (Ren C. , 2013): over-precisions manifests itself when individuals believe and thus act as their information or their estimate to be more precise (accurate) than it actually is (Croson , Ren, & Croson, 2009) Thus, overconfidence is defined as a biased belief that the distribution of demand has a lower variance than its true variance (Croson , Ren, & Croson, 2009). By keeping this notion in mind, the authors tried to derive incentives that unbiased managers can offer to overconfident newsvendors to induce optimal ordering behavior.

Nevertheless, we should point out that they mostly focused on one of the three declinations of overconfidence: the over-precision feature, applied to the demand forecast of an overconfident individual. This was done at the expense of other two fundamental aspects: overestimation of performances and over-placement (Schatz, 2017), the exaggerated belief of superiority of individuals concerning others, which both appear to be left out of all the studies proposed up to now.

For example, both Overestimation and the Dunning-Kruger effect, for which a person lacking knowledge and skills in a certain area tends to overestimate its competence (Dunning & Kruger, 2000), could be easily monitored during the administration of the problem, as it has already been done in various studies (Feld, 2017). Considering these facts, we can state that considering overconfidence as a unitary construct is not only imprecise but could prevent the analysis to reach further depts (Schatz, 2017) , leading to a somewhat partial and myopic conclusion.

An important point to make is that, as far as we know, Overconfidence was studied among the possible ways to explain heterogeneity, even if quite successfully, only in the aforementioned stream of papers and still in a theoretical way, so it could be interesting to investigate if its link to asymmetry and heterogeneity still holds at an experimental level.

## 2.5 Main Trends in the literature

Through a thorough exploration of the available articles on the subject two main trends emerged. On one side it appears a clear interest in the application of neuroscience and the employment of biofeedback, mainly brain electrical signals (Zhang, 2014) or eye movement, to support decision-making analysis: a few recent applications to Newsvendor problem's experimental findings analysis were found. In particular, starting from 2013, contributions started to appear to focus on the study of brain activity in correlation to Newsvendor's tasks: one of the main findings was the activation of the Dorsolateral Prefrontal cortex and OFC while deciding during the Newsvendor problem (Akash, 2019).

In addition to that, studies of neural responses via EEG experiments during decision-making tasks were used to investigate attention states during the tasks of the problem (Truong, 2020).

On the other hand, another trend in the literature can be distinguished from the classical Newsvendor stream of research: the study of heterogeneity, the presence of individuals ordering outside the PtC range, and inside with high variance between quantities purchased. This phenomenon was found to be quite diffused among all the experimental data gathered (Jammernegg W., 2021), but still, articles undertaking its analysis seem to be few and far in between.

However, a substantial part of this misalignment and heterogeneity in the result obtained cannot be fully explained by normative aggregate models. This fact further cemented the need to go beyond them and explore all the possible causes of the individual and cognitive characteristics which are proper of the single respondent. This need factually opened a new stream of research focused on individuals' inner aspects such as cultural background, education levels, professional experience, and, so on.

One of the most notable ones is the 2009 paper by Moritz et all. "Cognition and individual difference in the Newsvendor Problem: Behavior under dual process theory", which posits the Dual Process theory cognitive explanation for why individuals deviate from optimality and specifically explores the relationship between individual performance and cognitive reflection, measured by the Cognitive Reflection Test (CRT) (Moritz, Hill, & Donohue, 2009). Furthermore, we saw one study linking the performance difference between human newsvendors and the normative agent in the Newsvendor problem to gender risk appetite (de Véricourt, 2013).

# 3. RESEARCH QUESTION: THE DRIVERS OF OVER TIME VARIANCE OF ORDERED QUANTITIES AND DEMAND FORECAST

*"The illusion that we understand the past fosters overconfidence in our ability to predict the future."* (Kahneman, 2011)

As put in the apt word of Daniel Kahneman, our ability to look at the past, and to infer the right thing from it is sorely lacking. Humans seem to be, as a matter of fact, not prone to "rational decision-making" and, even more, unlikely to decide in a way that could be defined as "economically" rational, by reasoning in an appropriate way (Zenker, 2012). Interestingly enough, this insight can be applied perfectly to the behavior of respondents over time when put in front of a Newsvendor allocating problem: purchase decisions manifest a deep oscillation in each task not only between individuals participating but also in the quantities allocated by the same respondent in time. Moreover, to the best of our knowledge, it seems that studies on the presence of this variability are somewhat missing from the current scientific literature landscape, even if the concept of variability, which represents a key aspect of the economic underperformance of players. Indeed, this dissertation will try to fill the gaps in the literature on the variability that affects newsvendors over time by applying to the Newsvendor Problem experiment assumptions inference inability when focusing on participants' forecast of demand. Variability and thus irrationality does not reside only in planned inventory quantities, but also in the demand forecasted for that period by our human newsvendor.

Indeed, respondents can be subject to biases and errors when put in front of a forecast of future demand levels, like over-precision in their estimation or anchoring, and it is also easily assumable that a newsvendor may evaluate its option by looking at past demand realization.

Furthermore, they can also be affected by irrationality created by external inputs, like the unexpected realization of demand for a certain period and related bad performances in terms of costs.

## 3.1. The main questions investigated and the underlying hypothesis

The main open points highlighted in the precedent chapter when researching the literature currently available lead us to explore the data to try to identify the main external drivers of the variability affecting newsvendors' behavior.

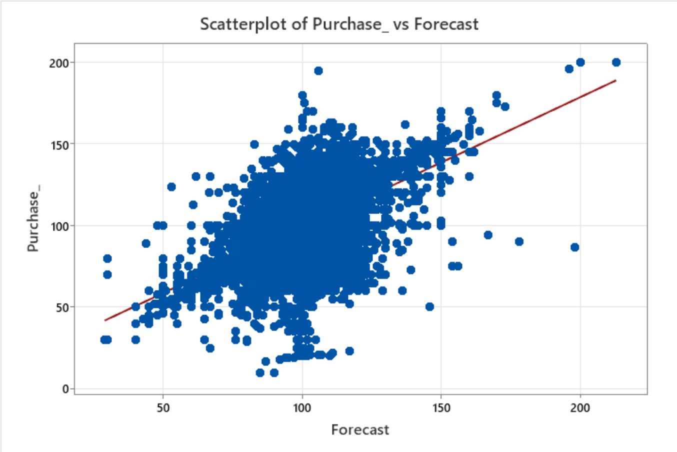
This dissertation will try to propose a thorough exploration of the data collected and, to a certain degree, to provide an answer to questions arising from the analysis of those experimental results. Our attention will be focused mostly on four outstanding questions:

1. First and foremost, what drives variability over time in both the quantity purchased and the demand forecasted by subjects in the problem?
2. Can the variance in the demand level predicted to be influenced by the previous realization of demand? So, do newsvendors anchor on previous demand levels when forecasting, and in particular do they employ all data visible on the demand? Does the same apply to orders placed (the variable purchase correlates to the mean of the manifested demand shown for the previous 5 periods)?
3. Does the shock in the realized demand, which we supposed as one of the main drivers of variability, influence the respondents' behavior in terms of forecasted demand and amount purchased? What are its impacts on cost and thus on newsvendors' performance?
4. Does providing a lecture on the Newsvendor problem improve respondents' performance and reduce respondents' variance?

It is very important to note the distinction between inventory quantities and demand forecasts done in our study. This choice was taken since, even if orders and demand levels forecasted by respondents seem to be very close concepts and common sense would lead us to think that subjects could feasibly interpret them as a unitary concept, the two variables also show substantial differences and thus it is correct to suppose that they can give a very different contribution to variability manifested by subjects.

Thus, investigating them separately could lead us to a deeper understanding of the Newsvendor decision-making process.

To demonstrate the abovementioned point, a Pearson correlation was applied to Purchase and Forecast, and the results show a slight positive linear correlation between the two sets of data.



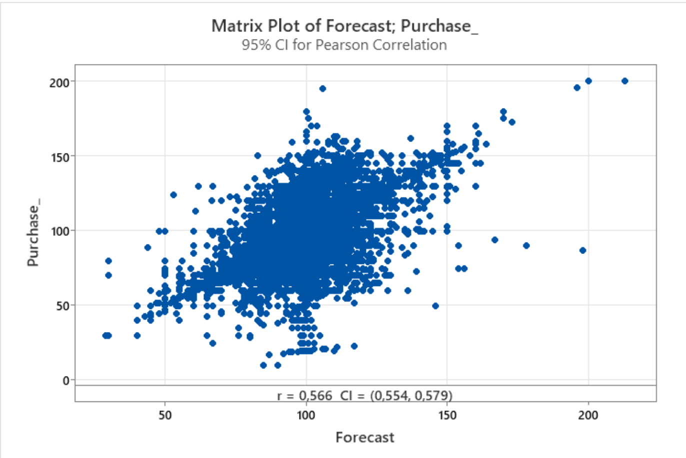




**FIGURE 5: PEARSON CORRELATION BETWEEN THE VARIABLE PURCHASE AND FORECAST AND FIGURE 5 SHOWING THE TREND LINE**

# 4. METHODOLOGY AND DATA COLLECTION

The experiment was designed to investigate combinations of the two conditions Framing, and Margin in a classic Newsvendor Problem setting, as proposed by Schweitzer and Cachon in their 2000 cornerstone paper "Decision bias in the Newsvendor problem with a known demand distribution: experimental evidence".

During each task, participants played for 20 rounds. After each round of the simulation, they received feedback on their performances, either in terms of the cost incurred, if the framing was negative, or in terms of profits achieved in case of positive feedback.

In the course of the experiment, they were asked to answer two distinctive questions:

1)How many items will they place in their order? Those answers were used to populate the variable "Purchase".

2) How much demand for the item will we have for that period? The answers collected fall under the variable "Forecast".

## 4.1 The experiment

A Newsvendor problem experiment was performed in the Politecnico di Torino premises in 2019, involving 218 subjects, of which 137 were male (62,9%) and 81 were female (37,1%). The experiment was divided into two sessions, 149 individuals participated in both. All the participants were third-year students of the Ingegneria Gestionale course.

Th experiment was structured in two sessions held on different days, respectively called "Session one", consisting of two tasks, in which the Newsvendor problem was presented with different configurations of Framing and Margin variables, and three questionnaires, one of which is the CRT Test (Cognitive Reflection Test) and "Session two", again with consisting of two tasks, but only two questionnaires.

The CRT (Frederick, 2005) was the first questionnaire to be administered. It consists of a series of numeric-answer questions aimed at emphasizing the distinction between two types of cognitive processes: System 1 thinking, which doesn't require too much attention and occurs quickly, and System 2, slower and more reflective.

The abovementioned configuration depended on the different combinations of Framing and Margin independent experimental variables: the first one, Framing, determines whether the problem is presented in terms of gains maximization (identified under the label 1) or cost minimization (identified with a 0), while the Margin variable represents if the product is a high margin, with each unit being priced 180, or low margin one, priced 120, as similarly done by many outstanding authors in the Newsvendor literature when structuring the experiment (Cachon, 2000;AJ A. Bostian, 2008; Benzion, 2008; Bhavani Shanker Uppari, 2019 etc.).

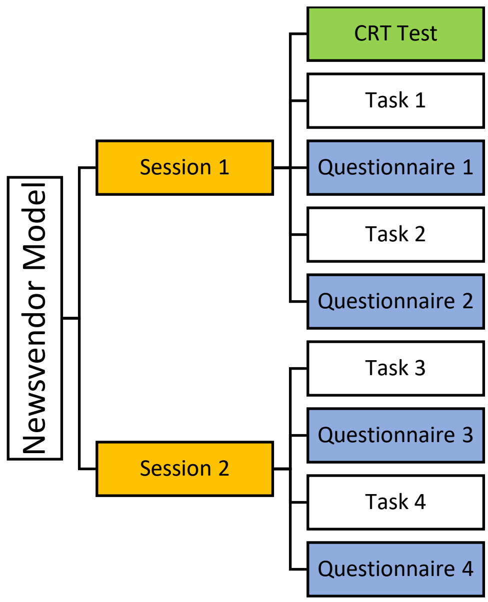



**FIGURE 6: GRAPHIC REPRESENTATION OF THE EXPERIMENT STRUCTURE UTILIZED DURING THE 2019 EXPERIMENT**

The table reported below summarizes the various combinations of Framing and Margin. It is Important to note how the Newsvendor will purchase each unit for a fixed price of 100.

| ***Margin*** ***Frame*** | | **High Margin** | **Low Margin** |
|---|---|---|---|
| *Negative* | Stock Out Cost | 80 | 20 |
| | Over Stock Cost | 20 | 80 |
| *Positive* | Selling Price | 180 | 120 |
| | Selling Price Overstock | 120 | 180 |
| *Purchasing Cost* | | 100 | 100 |

**TABLE 2: SUMMARY OF FRAMING AND MARGIN COMBINATIONS AND THEIR RELATIVE COST OF STOCK OUT/OVERSTOCK AND SELLING PRICES**

## 4.2. Measures and Variables

The independent variables utilized are listed below.

| *Variable Name* | **Description** | **Formula** |
|---|---|---|
| *Session* | This variable identifies the two sessions of the model held. Session 1 was performed before a lecture on the Newsvendor problem, Session 2 after | |
| *Framing* | Identifies one of the two settings (0,1) in which the problem was defined. In the first setting the model is presented in term of costs incurred and the respondents aimed to minimize those cost. While, with a Frame= 0 the problem was presented in terms of profits maximization. | |
| *Margin* | This variable identifies the product characteristics. With Margin= 0 the product is a High Margin product priced at €80. When Margin=0 the product has a low margin and its priced €20. Stock-out and Over-stock costs are respectively €80 and €20. | |
| *Group* | The group is identified by the specific sequence of the experimental conditions | |
| *Task* | | |
| *Repetition* | Time buckets in which the respondents play (from 6 to 25) | |
| *Demand (t)* | Effective demand manifested in each repetition the respondent plays. This is an experimental condition randomly defined beforehand and made know to participants only after they provide an answer for the period t. | |
| *Shock t-1* | | |
| *Shock t-1 Corrected* | This variable defines a shock level of demand in the period t-1 on three levels (-1, 0, 1) instead of 2. With -1 we define a negative shock in the demand so a demand level which is lower than the sum of the mean minus the standard deviation (SD=30). While with 1we define a positive shock in the demand, so a demand level higher than the mean of the demand plus the standard deviation (SD=30): | $Shock\ t-1\ corretto = \begin{cases} 1 & se\ D_{t-1>=130} \\ 0 & \\ -1 & \text{se } D_{t-1<=70} \end{cases}$ |
| *NVM (t)* | Optimal purchase quantity in the NVM | |

**TABLE 3: SUMMARY OF THE INDEPENDENT VARIABLES EMPLOYED IN THE MODEL**

The quantitative variables obtained are:

| *Quantitative Variables Names* | **Description** |
|---|---|
| *Forecast(t)* | Demand prevision elaborated by the participant at time t |
| *Purchase(t)* | Participants ordered quantities at time t |

| | |
|---|---|
| *SS (t)* | Security Stock at time t |
| *Costs(t)* | Costs related to unsatisfied demand due to stockouts and/or overstocking |
| *Delta_Demand_Purchase(t)* | Difference between realized demand at time t and quantity purchases and thus sold |
| *Delta_NVM_Purchase(t)* | Difference between the optimal quantity ordered by the Ideal Newsvendor and the participants' orders |

Some additional variables, derived from the experimental results, and a naming convention were also introduced.

| ***Derived Variables Names*** | | ***Formula*** |
|---|---|---|
| ***Human Variability in Newsvendor quantity at period t, per participant i, per condition j ($HVN_{i,j}$ (t))*** | *(2.1)* | $HVN_{i,j}(t) = q*_j + \left(\frac{1}{20}\sum_{t=6;i;j}^{25;i;j} NVMP_{i,j}(t)\right) - \frac{1}{20}\sum_{t=6;j}^{25;j} q*_j(t)$ |
| ***Human Cost at period t, per participant i, per condition j ($HCost_{i,j}$ (t))*** | *(2.2)* | $HCost_{i,j}(t) = MAX[m*(D_j(t) - NVM\_P_{ij}(t)); c*(NVM\_P_{i,j}(t) - D_j(t))]$ |
| ***Ideal Cost at period t, per condition j ($ICost_j$ (t))*** | *(2.3)* | $ICost_j(t) = MAX[m*(D_j(t) - q*_j(t)); c*(q*_j(t) - D_j(t))]$ |
| ***Variability Cost at period t, per participant i, per condition j ($VCost_{i,j}$ (t))*** | *(2.4)* | $VCost_{i,j}(t) = MAX[m*(D_j(t) - HVN_{i,j}(t)); c*(HVN_{i,j}(t) - D_j(t))]$ |
| ***Delta Cost at period t, per participant i, per condition j ($DC\_IN_{i,j}$ (t))*** | *(2.5)* | $DC\_IN_{i,j}(t) = HCost_{i,j}(t) - ICost_j(t)$ |
| ***Delta Variability Cost at period t, per participant i, per condition j ($DVC\_IN_{i,j}$ (t))*** | *(2.6)* | $DVC\_IN_{i,j}(t) = HCost_{i,j}(t) - VCost_{i,j}(t)$ |
| ***Demand level at period t, per condition j ($DL_j$ (t))*** | *(2.7)* | $DL_j(t) = \begin{cases} Low & if\ D_j < \mu - \sigma \\ High & if\ D_j > \mu - \sigma \\ Average & else \end{cases}$ |
| ***Profit at period t, per participant i, per condition j ($Profit_{i,j}$ (t))*** | *(2.8)* | $P_{i,j}(t) = p*MAX[D_j(t); NVM\_P_j(t)] + MAX[m*(D_j(t) - NVM\_P_{ij}(t)); c*(NVM\_P_{i,j}(t) - D_j(t))]$ |

**TABLE 4: SUMMARY OF THE DERIVED VARIABLES COMPUTED FROM THE DATA COLLECTED**

## 4.3. Software and Statistics employed

To support the data analysis we have employed various software: starting with Excel for the dataset structuring and maintenance, while Matlab was used for the exploratory data analysis (e.g. n- way ANOVA, Interaction plots etc.) and so on. We also utilized Minitab as an imaging tool to introduce better-quality graphs to support our investigation and facilitate the reader. The choice of computational software has fallen on the well-known Matlab, instead of R, for a reason a familiarity with the tool. The scripts utilized will be reported in the Annex section and numbered accordingly.

We will now briefly sze the statics performed and the underlying statistical theory.

Two-way and N-way ANOVA ("Analysis of Variances") have been amply employed in the data analysis section. The ANOVA, which is simply a collection of statistical models to study the difference among means of different groups, is based on the law of total variance for the variance of variables is deconstructed in contribution attributable to different sources of variation (Scheffé & Henry, 1959).

The differences between a one-way and n-way ANOVA reside in the number of independent variables employed by the model: one for the first, while up to n- for the other.

After identifying with the ANOVAs variables or interaction of variables that hold statistically significant effects on our variable of interest, Forecast, and Purchase, we proceeded to analyze how they influence them. To do so we have employed Interaction Plots to have a graphical representation of the variables' mean behavior concerning the factors proposed and their confidence intervals.

## 4.4. Data Collection Methodology

As already mentioned, during the experiment participants were firstly subjected to Critical Thinking Test, a simple questionnaire aimed at determining the level of reflective thinking of its respondents. Once the rounds of the Newsvendor task itself started players were asked to provide two types of answers and populate the allocated spots on the screen: one related to the forecast of the demand for the current period, t, having visible the realized demand for the 5 precedent periods ( while for each first round, the demand amount defaulted to 87 units), then were required to place an order for the units they wanted to purchase for period t. This distinction was made to clearly distinguish the effects linked to the previsions in time and the number of units they instead expected to sell, to identify the safety stock amounts, so the additional level of stocks expressly purchased to mitigate the risk of any eventual stockout (Monk, 2009). Demographic data, like age, gender, number of exams passed, and average score, were also requested.

Once collected from the experiment, the data were cleaned from incomplete or systematically random answers and typos, by removing values outside a range equal to $\mu \pm 3\sigma$ and respondents with less than

50% of valid responses. After the data pre-processing step, 1460 records from the 218 participants were deemed sound to be analyzed.

After the data collection phase and pre-processing activities, we proceeded with a thorough analysis, first, of the overall results collected, without any distinction between the two sessions of the experiment performed and then we reproduced the same investigation step by segregating the data session by session. This was done to study if irrational behavior transcended the presence of the different sessions and a specific lecture dedicated to the Newsvendor problem held between the two.

The first explorative analysis was performed by executing ANOVAs on Forecast and Purchase data under various treatments generated by different combinations of Framing, Margin, and Session variables. The effectsofn these treatments on inventory quantity are already well known in the literature, among the most striking regularities present in ordering behavior we have the already mentioned Pull to Center effect, for which average order quantities are too low when they should be high and vice versa (Bostian, Holt , & Smith , 2008). While demand level predictions remained least studied.

In addition to that, both variables were compared with their ideal, and thus rational equivalents and their differences were subjected to analysis of variance to pinpoint exactly what external factors contribute to enlarging or reducing the effects of the irrationality, mainly by analyzing their effects on costs.

When we say costs we mean, more precisely, the difference between the costs incurred by an ideal and thus optimal newsvendor and the real costs faced by participants. This was done to better understand how this delta, taken as a proxy of the economic inefficiencies of the so-called *Human Newsvendors,* grows or gets smaller with respect to drivers of irrationality.

# 5. RESULTS’ DISCUSSION

As previously mentioned, this investigation's prime topic and objective will be to try to understand the phenomenon of variability in the Newsvendor’s answer pattern over time and its drivers.

However, differently from what proposed so far in the literature studying biases and heuristics in the Newsvendor, which manifest in the average behavior of subjects during the experiment, all the phenomena linked to individual differences in the response will be studied by looking at the variability of the data collected participant by participant.

To achieve a more in-depth exploration of the causes that lie at the basis of the variance shown by participants, the analysis of the quantity purchased, and the demand forecast will be performed separately since the variability affects not only the inventory levels but also the demand estimation for a period. To prove the presence of this variability on the variable Forecast, respondents’ forecasts of demand for each period t were compared with the ideal rational behavior: an optimal forecaster at time t would have based its prevision on the mean of the precedent periods realized demands.

## 5.2 Results of Forecast and Purchase Explorative ANOVAs

An exploratory n-way ANOVA was performed with Matlab on the data collected across the two Sessions, to verify the influence of the various combinations of treatments such as Session, Framing, and Margin and to investigate if the results obtained differed greatly between the two chosen response variables: the Purchased Quantity and the Forecast of Demand.

The proposed analysis of variance, ANOVA, derives from the law of total variance, where the observed variance in a particular variable is partitioned into components attributable to different sources of variation (Lars St»hle, 1990).

Indeed, the results obtained from the analysis clearly show different effects on the two variables generated by the combinations of Framing and Margin.

## Analysis of Variance

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 11282.6 | 1 | 11282.6 | 48.42 | 0 |
| Margin | 17591.4 | 1 | 17591.4 | 75.49 | 0 |
| Framing | 1209.6 | 1 | 1209.6 | 5.19 | 0.0227 |
| Session:Margin | 21552.3 | 1 | 21552.3 | 92.49 | 0 |
| Session:Framing | 23305 | 1 | 23305 | 100.01 | 0 |
| Margin:Framing | 139.4 | 1 | 139.4 | 0.6 | 0.4393 |
| Session:Margin:Framing | 226.6 | 1 | 226.6 | 0.97 | 0.3241 |
| Error | 2677978.8 | 11492 | 233 | | |
| Total | 2757056.8 | 11499 | | | |

Constrained (Type III) sums of squares.

FIGURE 6: ANALYSIS OF VARIANCE (ANOVA) RESULTS FOR THE VARIABLE FORECAST

The demand forecast is strangely enough not influenced by the combination of the Framing and Margin (p> 0,05) and sensitive only to a minor extent to the framing effects in general. One reason behind that could be linked to the fact that their effect is covered entirely by the effects of the variable Session, differently from what happens for Purchase.

## Analysis of Variance

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Framing | 82978.2 | 1 | 82978.2 | 244.73 | 0 |
| Margin | 1212625.2 | 1 | 1212625.2 | 3576.37 | 0 |
| Session | 22632.3 | 1 | 22632.3 | 66.75 | 0 |
| Framing:Margin | 7257.9 | 1 | 7257.9 | 21.41 | 0 |
| Framing:Session | 95922.5 | 1 | 95922.5 | 282.9 | 0 |
| Margin:Session | 217625 | 1 | 217625 | 641.84 | 0 |
| Error | 3896880.7 | 11493 | 339.1 | | |
| Total | 5504240.9 | 11499 | | | |

Constrained (Type III) sums of squares.

FIGURE 7: ANALYSIS OF VARIANCE (ANOVA) RESULTS FOR THE VARIABLE PURCHASE

The assumption made at the start of this chapter, regarding the collection of results across two Session of the experiment, were made to achieve more precise conclusions and to avoid bringing unnecessary noise to the table.

Moreover, the clear effect of the variable Session is also pointing out the potential interest of exploring the data of the two Sessions separately, since the lecture held in between and the learning effects, incurred due to participants' increased adaptation to the game dynamics, could potentially hide some deeper insights on the demand variability.

To confirm this point we decided to look at the difference between respondents' *i* and optimal Newsvendors' forecasts at time t by having both session data and for each session separately: in the first case, we can clearly see that even if the mean of the two groups doesn't show any significant difference ($p > 0,001$), the gap between the variances of the two groups is statistically significant ($p < 0,001$). In fact, the results of the two-variance test, Bonett's and Lavene's tests, are in favor of the alternative hypothesis, for which the variance of the respondents' forecast is dissimilar to the variance of the optimal newsvendor prediction of demand.

### Test

| | |
|---|---|
| Null hypothesis | $H_0: \sigma_1 / \sigma_2 = 1$ |
| Alternative hypothesis | $H_1: \sigma_1 / \sigma_2 \neq 1$ |
| Significance level | $\alpha = 0,05$ |

| Method | Test Statistic | DF1 | DF2 | P-Value |
|---|---|---|---|---|
| Bonett | 1654,09 | 1 | | 0,000 |
| Levene | 2191,08 | 1 | 22998 | 0,000 |

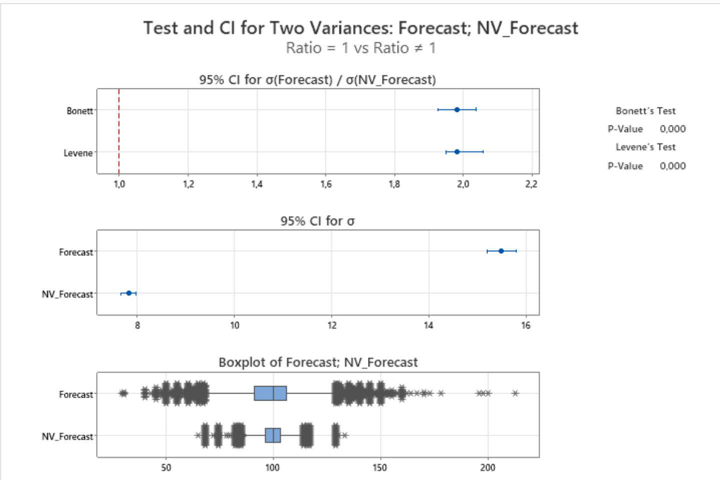



**FIGURE 8: TWO-VARIANCE TEST RESULTS**

### Descriptive Statistics

| Variable | N | StDev | Variance | 95% CI for σ |
|---|---|---|---|---|
| Forecast | 11500 | 15,484 | 239,765 | (15,188; 15,789) |
| NV_Forecast | 11500 | 7,818 | 61,129 | (7,662; 7,979) |

**FIGURE 9: DESCRIPTIVE STATISTIC FOR THE 2 CATEGORIES OF DATA COLLECTED**

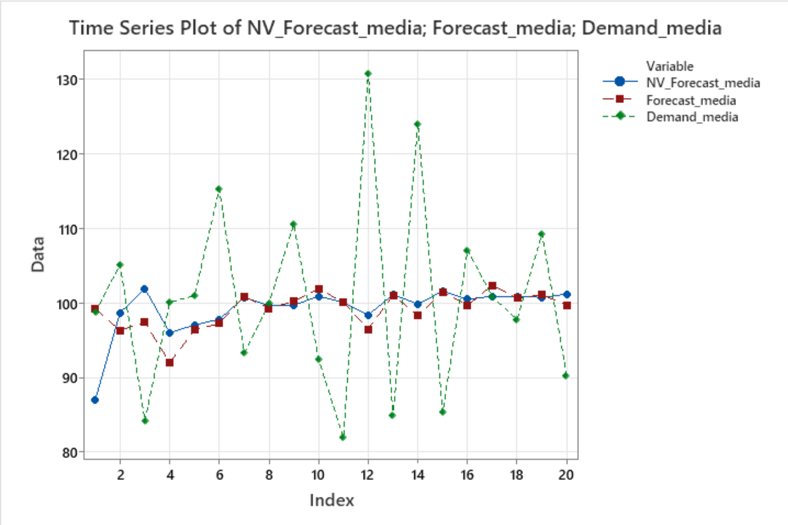



**FIGURE 10: SESSION 1 AND 2 MEANS OF THE FORECAST, REALIZED DEMAND, AND NEWSVENDOR'S OPTIMAL FORECAST**

Thus, a certain degree of irrationality is present, making the variance of the respondents' previsions substantially higher than the one they would have if they were perfectly rational, represented in the picture by the blue dotted line (labeled in the legend as "NV_Forecast_mean").

Very similar to what happens when looking at the mean of the inventory level for each repetition.

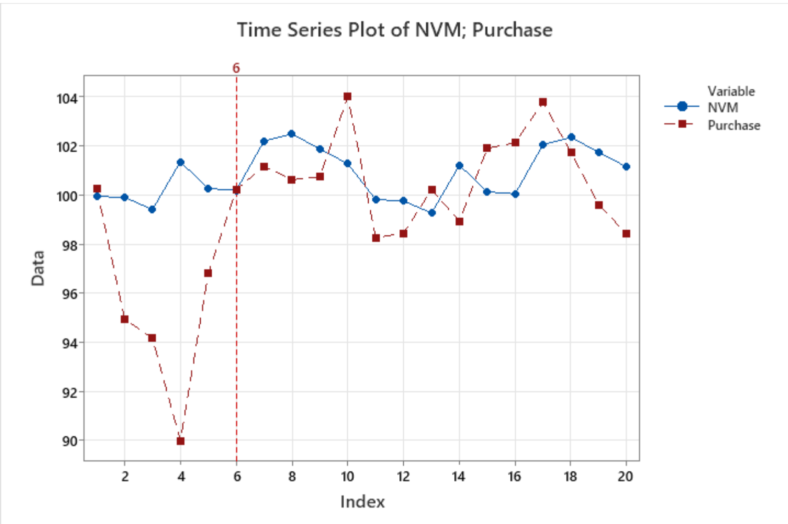



**FIGURE 11: PURCHASE'S MEAN OVERTIME[3]**

The forecast's means for each Session and repetition, from 1 to 20, are pictured in the below plots besides the optimal newsvendor forecast and the realized demand mean.

[3] Figure 11's graph shows the sixth repetition highlighted in red since it represents the "visibility" point for the participants

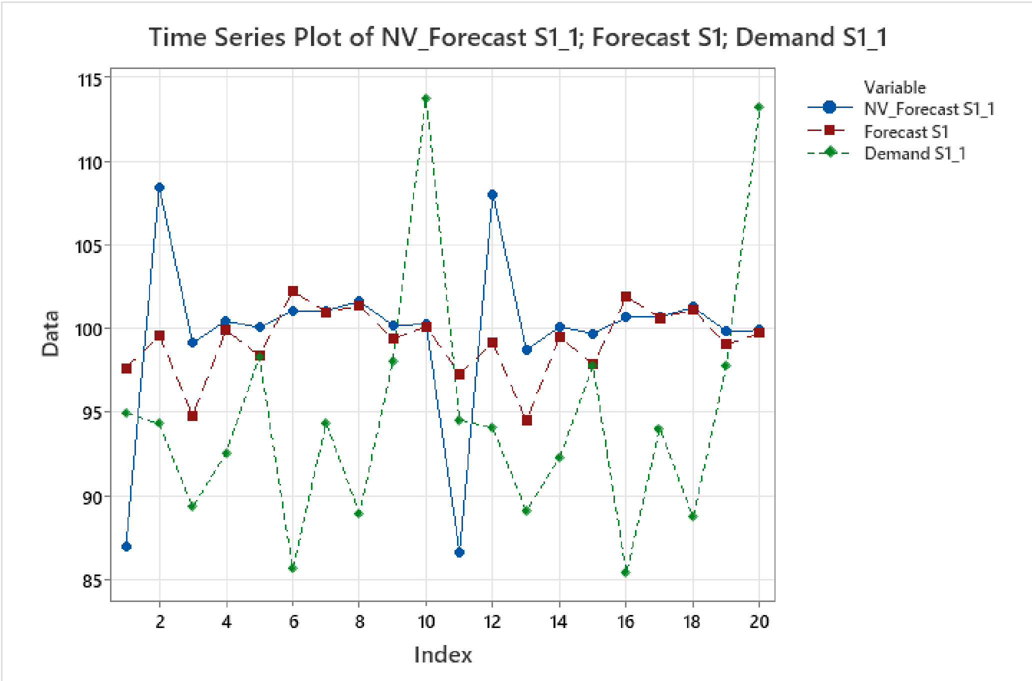



FIGURE 12: SESSION 1 PLOT OF THE MEANS OF FORECAST AND NV_FORECAST FOR EACH ORDER REPETITION

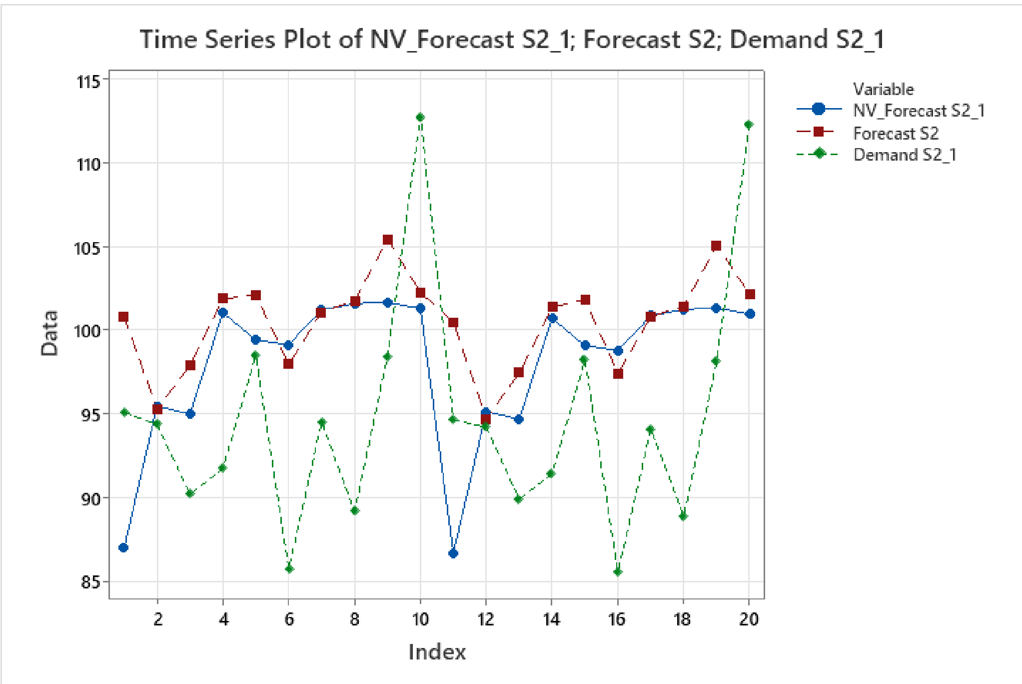



FIGURE 13: SESSION 1 PLOT OF THE MEANS OF FORECAST AND NV_FORECAST FOR EACH ORDER REPETITION

Again, the human newsvendor forecasts' population mean seems to be graphically very close to the mean of the previsions made by the ideal newsvendor. However, the results coming from the two-variance test and a paired t-test[4] between the derived variables NV_Forecast and Forecast, performed on Session 1 and Session 2 data separately, hint at other conclusions: Session 1 data's Variance ($\sigma^2$) of Forecast is equal to 207.571, while the NV_Forecast variance is 53.266 with a 95% Confidence Interval. Indeed, the variables' variance differs in a statistically significant way between each other ($p<0,001$). The same can be also obtained for the population's mean: the Null hypothesis is rejected ($p<0,001$) in favor of the alternative hypothesis for which the NV_Forecast mean is different from the Forecast's one.

### Test

| | |
|---|---|
| Null hypothesis | $H_0$: $\sigma_1 / \sigma_2 = 1$ |
| Alternative hypothesis | $H_1$: $\sigma_1 / \sigma_2 \neq 1$ |
| Significance level | $\alpha = 0,05$ |

| Method | Test Statistic | DF1 | DF2 | P-Value |
|---|---|---|---|---|
| Bonett | 797,78 | 1 | | 0,000 |
| Levene | 1105,12 | 1 | 11558 | 0,000 |

### Descriptive Statistics

| Variable | N | StDev | Variance | 95% CI for σ |
|---|---|---|---|---|
| NV_Forecast | 5780 | 7,298 | 53,266 | (7,099; 7,506) |
| Forecast | 5780 | 14,407 | 207,571 | (14,012; 14,819) |

**FIGURE 14: TWO-VARIANCE TEST ON SESSION 1 NV_FORECAST AND FORECAST VARIABLES**

### Test

| | |
|---|---|
| Null hypothesis | $H_0$: μ_difference = 0 |
| Alternative hypothesis | $H_1$: μ_difference ≠ 0 |

| T-Value | P-Value |
|---|---|
| 9,93 | 0,000 |

### Estimation for Paired Difference

| Mean | StDev | SE Mean | 95% CI for μ_difference |
|---|---|---|---|
| 1,893 | 14,497 | 0,191 | (1,519; 2,267) |

*μ_difference: population mean of (NV_Forecast - Forecast)*

### Descriptive Statistics

| Sample | N | Mean | StDev | SE Mean |
|---|---|---|---|---|
| NV_Forecast | 5780 | 99,988 | 7,298 | 0,096 |
| Forecast | 5780 | 98,095 | 14,407 | 0,190 |

**FIGURE 15: PAIRED T-TEST ON SESSION 1 NV_FORECAST AND FORECAST VARIABLES**

[4] The paired sample t-test or dependent sample t-test is a statistical procedure used to decide whether the mean difference between two sets of observations is zero. It was applied instead of the two-sample t-test since the data sample employed are dependent on one another (Ross & Willson, 2017).

Then this process was also repeated on Session 2 data with similar results (Appendix A). This aptly demonstrated why the two-session data behavior should be taken into consideration. Indeed by not taking it into account we could risk losing information or misinterpreting it.

Taking this as an ulterior proof of the effects of the variable Session across repetitions, we studied the behavior of the demand predictions across sessions and under various treatment combinations.

The variable Session was taken into consideration when performing the first exploratory Analysis of Variance, along with Margin and Framing whose effects have been already extensively studied.

From the ANOVA on the variable Forecast, it emerged that the treatments with Session and Margin have statistically significant effects, while the interaction of Framing and Margin (Framing* Margin) doesn't yield the same results.

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 11282.6 | 1 | 11282.6 | 48.42 | 0 |
| Margin | 17591.4 | 1 | 17591.4 | 75.49 | 0 |
| Framing | 1209.6 | 1 | 1209.6 | 5.19 | 0.0227 |
| Session:Margin | 21552.3 | 1 | 21552.3 | 92.49 | 0 |
| Session:Framing | 23305 | 1 | 23305 | 100.01 | 0 |
| Margin:Framing | 139.4 | 1 | 139.4 | 0.6 | 0.4393 |
| Session:Margin:Framing | 226.6 | 1 | 226.6 | 0.97 | 0.3241 |
| Error | 2677978.8 | 11492 | 233 | | |
| Total | 2757056.8 | 11499 | | | |

Constrained (Type III) sums of squares.

### FIGURE 16: ANOVA FOR THE VARIABLE FORECAST

The same steps were also proposed for Purchase obtaining quite different outcomes: as expected from the many examples found in the literature, all the 3 grouping variables and their interaction appear to have substantial effects.

**Analysis of Variance**

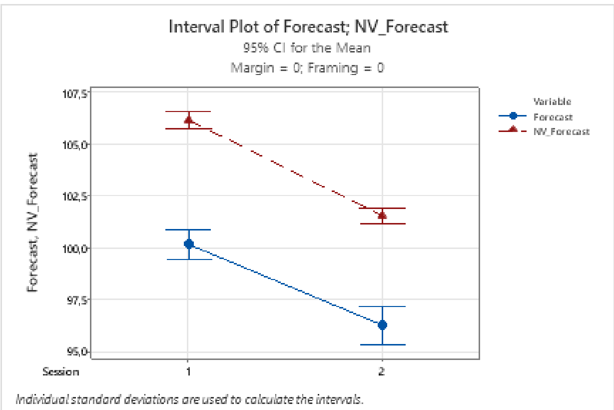
| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Framing | 82978.2 | 1 | 82978.2 | 244.73 | 0 |
| Margin | 1212625.2 | 1 | 1212625.2 | 3576.37 | 0 |
| Session | 22632.3 | 1 | 22632.3 | 66.75 | 0 |
| Framing:Margin | 7257.9 | 1 | 7257.9 | 21.41 | 0 |
| Framing:Session | 95922.5 | 1 | 95922.5 | 282.9 | 0 |
| Margin:Session | 217625 | 1 | 217625 | 641.84 | 0 |
| Error | 3896880.7 | 11493 | 339.1 | | |
| Total | 5504240.9 | 11499 | | | |

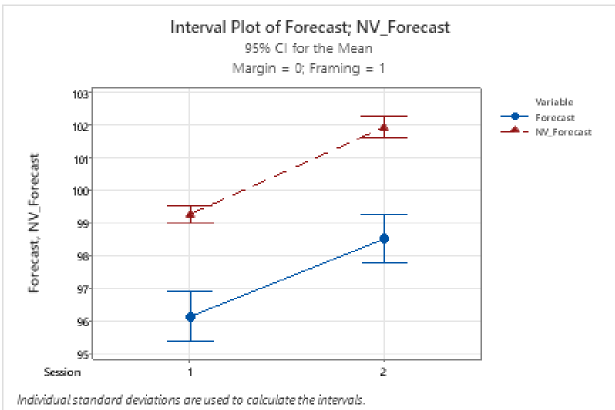
Constrained (Type III) sums of squares.

### FIGURE 17: ANOVA FOR THE VARIABLE PURCHASE

The second step was then to have a look at how those variables influenced the variable Forecast and thus subjects' answers. To do so Interval Plots were utilized as a graphical way to show any interesting pattern found.

As shown in the interval plots[5] with a 95% Confidence Interval reported below, in the low-margin setting independently from the Framing, coherently with what emerged from the ANOVA, the normative *newsvendor* forecast is higher than the forecasted by the participants. Conversely, in the high-margin setting, they will predict lower values than the human newsvendor. These findings seem to be pretty in line with the Pull-to-Center effect, even if the concept is related to the purchased quantities instead of the forecast of demand. It is also important to note how the grouping for the variable Session doesn't show any significant effect with respect to the combination of Framing and Margin in accordance with what resulted from the ANOVA. While if we look at the graph only taking in to consideration the influence of Session and Margin is evident how with high margins, Margin set to 1, the subjects strongly diverge from the optimal behavior when switching from Session 1 to Session 2 data. We could interpret this as  negative influence of the lecture held in between the phases: participants when facing higher selling prices are more prone to misinterpret the contents of the focus session on the Newsvendor problem performing worse than in the first instance.




[5] Interval plots are used to assess and compare confidence intervals of the means of groups by identifying whether the two groups' populations have similar mean. It provides also a comparison of the variation present in each group.

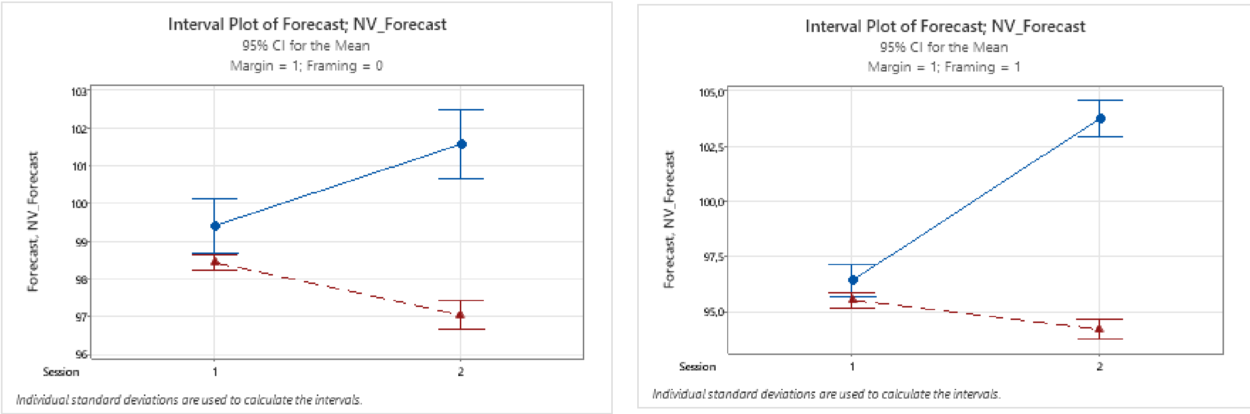
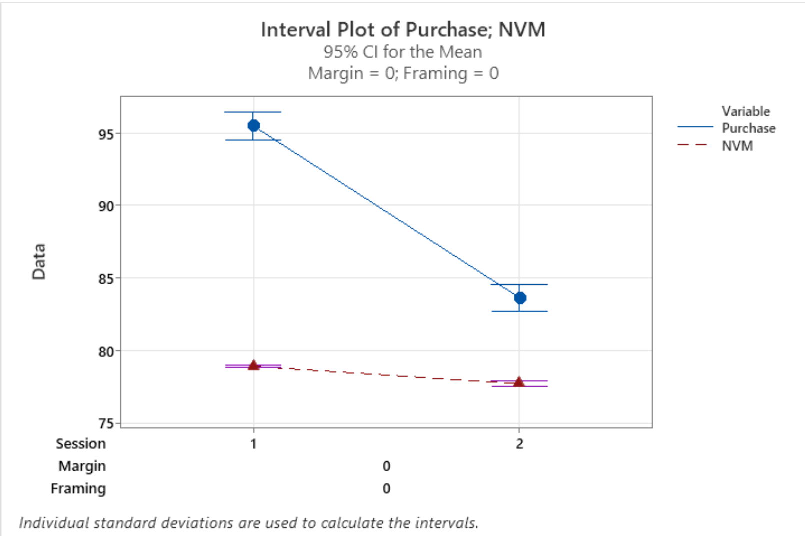



**FIGURE 18 A), B), C), D) THE INTERVAL PLOTS FOR FORECAST AND NV_FORECAST WITH A 95% CONFIDENCE INTERVAL**

The same disruptive effect can be seen on Purchase, so inventory levels, but with an important difference: with high margins newsvendors purchase less than what their optimal counterpart would din the second session, while respectively forecasting more. In addition to that the worsening of performance is visible with low margins instead of higher ones, with participants misinterpreting the insights coming from the lecture and diminishing their optimism.



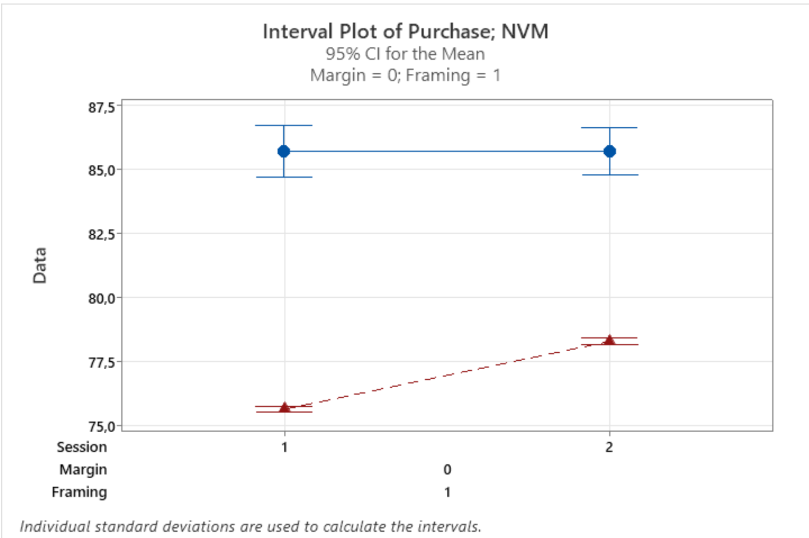
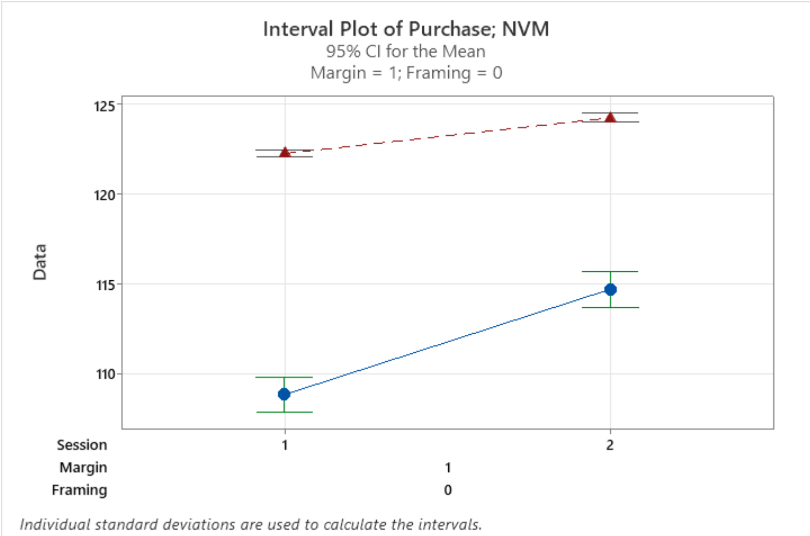



**FIGURE 19 A) AND B) : INTERVAL PLOT FOR PURCHASE AND NVM_PURCHASE FOR LOW MARGIN SETTING**



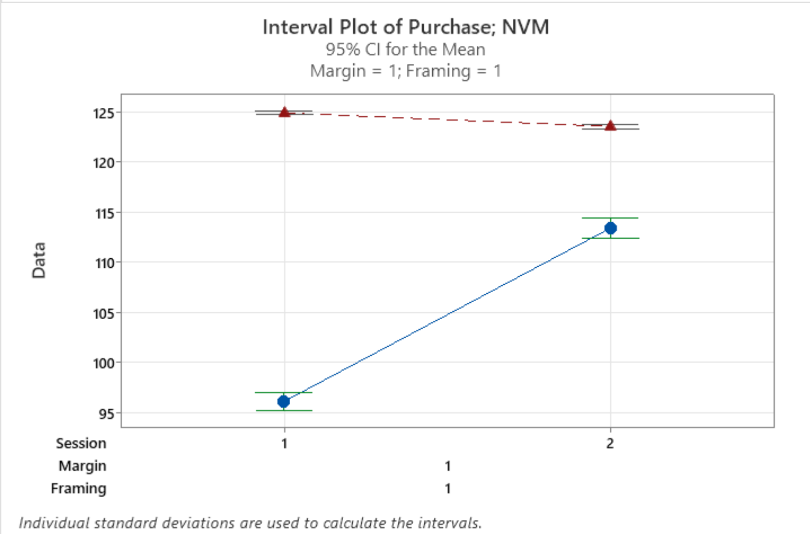




**FIGURE 20 A) AND B): INTERVAL PLOT FOR PURCHASE AND NVM_PURCHASE FOR HIGH MARGIN SETTING**

Our investigation then proceeded by individuating and inserting other factors in the ANOVA's treatments in order to identify the main drivers of variability. In the process of doing so, we incurred in some interesting results: it appears that the external factors that influence the forecast of demand in the Newsvendor problem, besides the previously realized demand, $D_{t-1}$, is the presence of a shock in the demand realizations, $Shock_{t-1}$, as defined in Paragraph 4.2.

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 9754 | 1 | 9754 | 41.94 | 0 |
| Margin | 17231.6 | 1 | 17231.6 | 74.09 | 0 |
| Framing | 1156.9 | 1 | 1156.9 | 4.97 | 0.0257 |
| Shock t-1 | 1003.5 | 1 | 1003.5 | 4.31 | 0.0378 |
| Session:Margin | 21476 | 1 | 21476 | 92.33 | 0 |
| Session:Framing | 20175.5 | 1 | 20175.5 | 86.74 | 0 |
| Session:Shock t-1 | 472.2 | 1 | 472.2 | 2.03 | 0.1542 |
| Margin:Framing | 4.5 | 1 | 4.5 | 0.02 | 0.8891 |
| Margin:Shock t-1 | 0 | 1 | 0 | 0 | 0.9924 |
| Framing:Shock t-1 | 111.2 | 1 | 111.2 | 0.48 | 0.4894 |
| Session:Margin:Framing | 259.7 | 1 | 259.7 | 1.12 | 0.2907 |
| Session:Margin:Shock t-1 | 369.8 | 1 | 369.8 | 1.59 | 0.2074 |
| Session:Framing:Shock t-1 | 1599.6 | 1 | 1599.6 | 6.88 | 0.0087 |
| Margin:Framing:Shock t-1 | 3090.5 | 1 | 3090.5 | 13.29 | 0.0003 |
| Session:Margin:Framing:Shock t-1 | 471.8 | 1 | 471.8 | 2.03 | 0.1544 |
| Error | 2671079.4 | 11484 | 232.6 | | |
| Total | 2757056.8 | 11499 | | | |

Constrained (Type III) sums of squares.

**FIGURE 21: EXPLORATORY ANOVA RESULTS FOR FORECAST WITH MARGIN, FRAMING, SESSION AND SHOCK T-1 AS FACTORS**

From this first exploration of new variables, one could assume, albeit as we will demonstrate later on, incorrectly that shock realizations of demand have a small influence on subjects' predictions.

However, this point was disproved by simply correcting the way Shock t-1 was modelized: instead of having a binary function to describe both excessively high/low demands manifested for a period (Shock t-1 = 1 ) and normal realizations (Shock t-1=0), 3 levels were defined. Once adjusted the strong effect of the variable on Forecast appear more clearly.

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 10840.4 | 1 | 10840.4 | 48.92 | 0 |
| Margin | 11482 | 1 | 11482 | 51.81 | 0 |
| Framing | 3303.7 | 1 | 3303.7 | 14.91 | 0.0001 |
| Shock t-1 Corretto | 119504.3 | 2 | 59752.2 | 269.63 | 0 |
| Session:Margin | 15071.5 | 1 | 15071.5 | 68.01 | 0 |
| Session:Framing | 13825.2 | 1 | 13825.2 | 62.38 | 0 |
| Session:Shock t-1 Corretto | 98.9 | 2 | 49.4 | 0.22 | 0.8 |
| Margin:Framing | 6.4 | 1 | 6.4 | 0.03 | 0.8647 |
| Margin:Shock t-1 Corretto | 352.8 | 2 | 176.4 | 0.8 | 0.4512 |
| Framing:Shock t-1 Corretto | 1069.6 | 2 | 534.8 | 2.41 | 0.0896 |
| Session:Margin:Framing | 146.5 | 1 | 146.5 | 0.66 | 0.4162 |
| Session:Margin:Shock t-1 Corretto | 921.2 | 2 | 460.6 | 2.08 | 0.1252 |
| Session:Framing:Shock t-1 Corretto | 1529 | 2 | 764.5 | 3.45 | 0.0318 |
| Margin:Framing:Shock t-1 Corretto | 2073.4 | 2 | 1036.7 | 4.68 | 0.0093 |
| Session:Margin:Framing:Shock t-1 Corretto | 212 | 2 | 106 | 0.48 | 0.6199 |
| Error | 2543213 | 11476 | 221.6 | | |
| Total | 2757056.8 | 11499 | | | |

Constrained (Type III) sums of squares.

**FIGURE 22: EXPLORATORY ANOVA RESULTS FOR FORECAST WITH MARGIN, FRAMING, SESSION AND SHOCK T-1 "CORRETTO" AS FACTORS**

Even more so, it was visible for Purchase. The distortion created by not segregating positive and negative shocks hided also the effects of the interaction of Shock t-1 and Margin, which seemed in the first instance to be insignificant. For the ANOVA results see the Annex.

## 5.3 Analysis of the main drivers of variance

To better understand what influences the way participants define their prediction during the experiment, respondents were also specifically asked which strategy they choose to employ: it emerged that the vast majority of participants, more precisely the 54.3%, based their forecast of demands on a moving average of the previous periods' demand.

This interesting insight was at the basis of our choice to investigate the effects of both the direct precedent period demand level and the mean of the past realized demand for the previous five precedent periods.

In addition to that, from the preliminary analysis of variance presented in Paragraph 5.2, it appears that the external factors that influence the forecast of demand in the Newsvendor problem, besides the previously realized demand, $D_{t-1}$, is the presence of a shock in the demand realizations, $Shock_{t-1}$.

The significance of previous demand levels fits well with the presence of the so-called "Pull-to-Center effect": for which in both the regimes of high-profit margin and low-profit margins, the average choices converged towards the mean demand (Sharma & Nandi, 2018) and thus suggests that in fact, the respondents tend to be influenced by the effective demand.

While the new variable introduced to represent sharp fluctuations of the levels of desired quantities in our fictitious market, from what we know, represents a novelty in the field.

By plotting the Interaction Graphs for both Forecast, Purchase and their respective ideal counterparts we proceed to interpret participants' behavior when put in front of shocks in the realized demand levels and the interaction with other factors such as Margin and Framing, also based on the results coming from the ANOVA.

The results obtained show a reversal of patterns between high margins and low margins treatments: forecasted quantities grow when *newsvendors* are faced with positive shocks in the demand level, but overall participants remain less optimistic in their prevision with respect to the ideal newsvendors which settles on generally higher quantities. This could be a symptom of the perception of higher losses created by the high-margin product which will push respondents to be more conservative in their demand predictions.

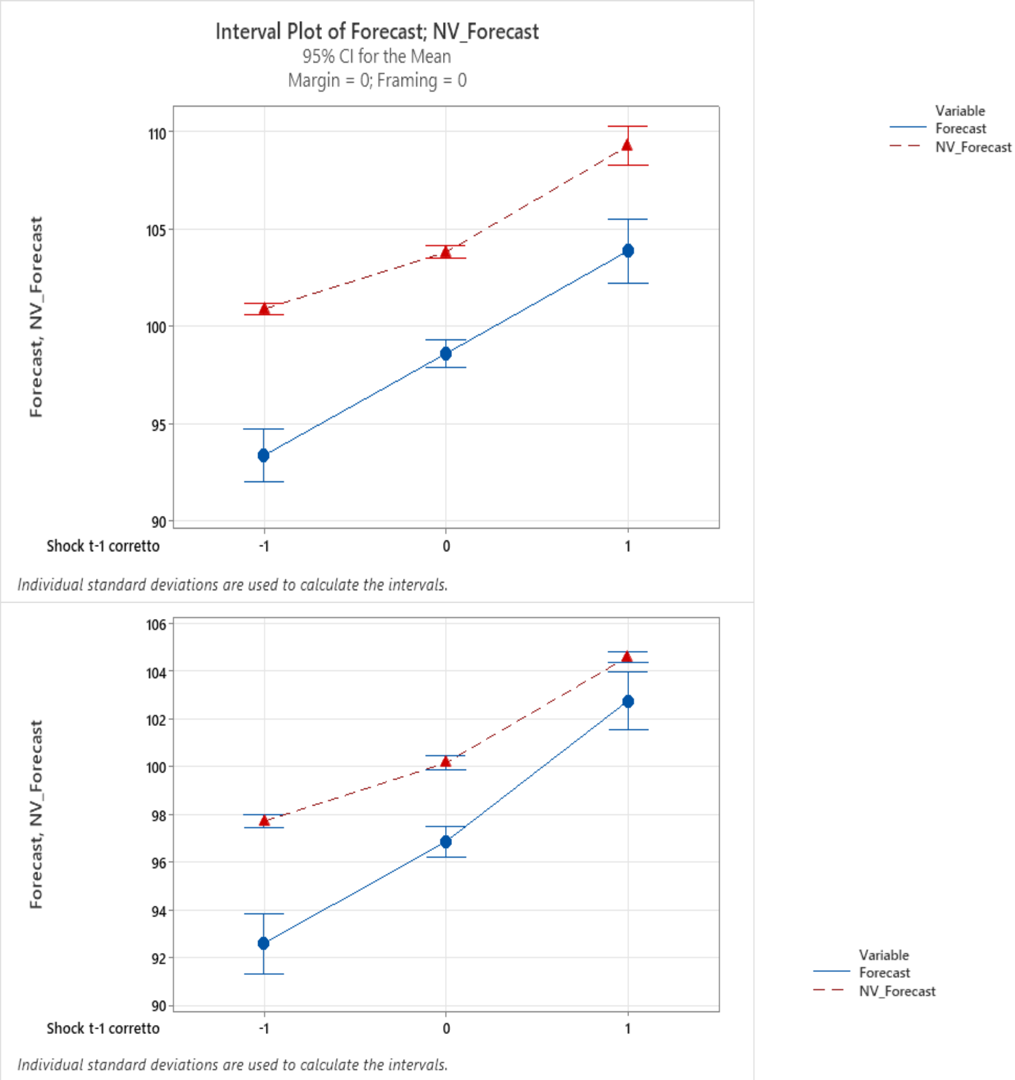
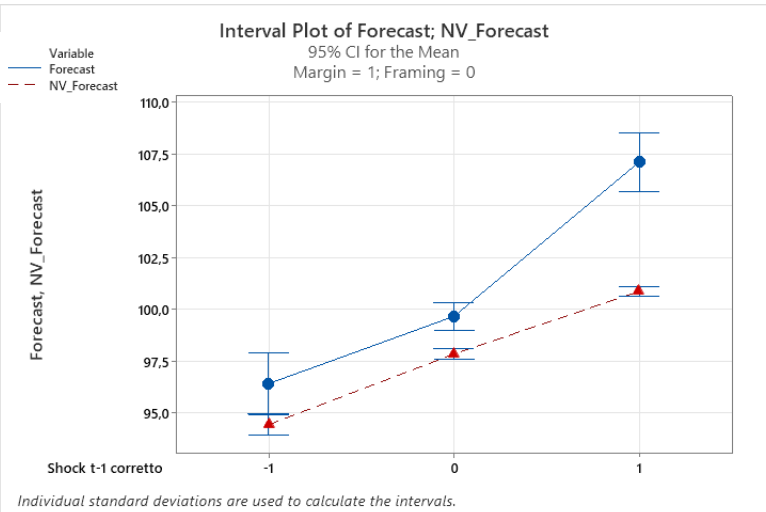



**FIGURE 23 A) AND B): INTERVAL PLOTS FOR FORECAST AND NEWSVENDOR FORECAST WITH A LOW MARGIN TREATMENT IN RELATION TO SHOCK$_{T-1}$**

While the opposite is clearly visible for low margins: respondents appears to be more optimistic and forecast more than their ideal counterpart, but still responding to the shocks as expected.



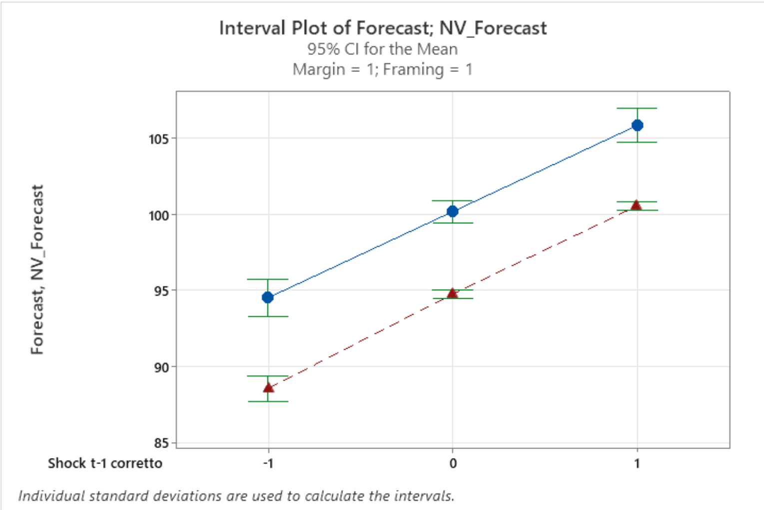
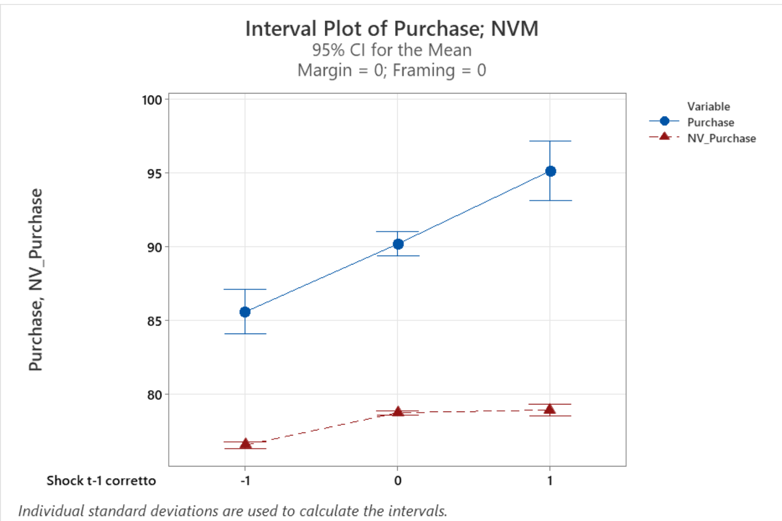



**FIGURE 24 A) AND B): INTERVAL PLOTS FOR FORECAST AND NEWSVENDOR FORECAST WITH A LOW MARGIN THREATMENT IN RELATION TO SHOCKT-1**

The same principle could be applied also to the Inventory levels, described by the variable Purchase as done below:



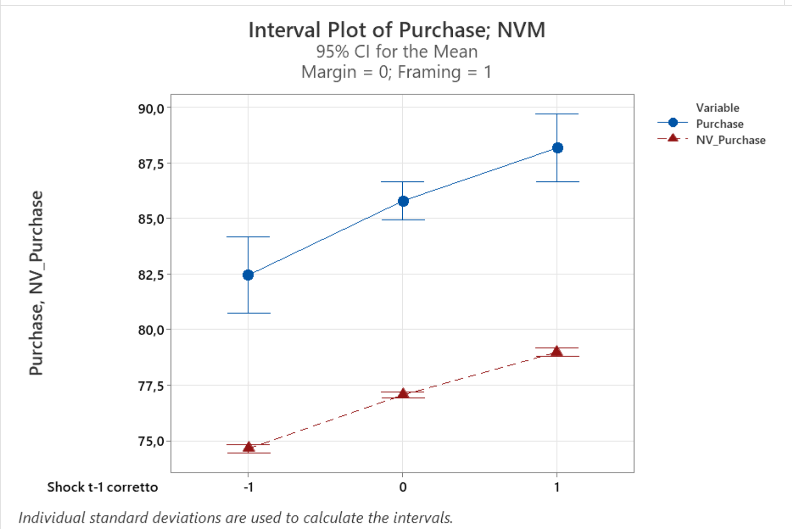
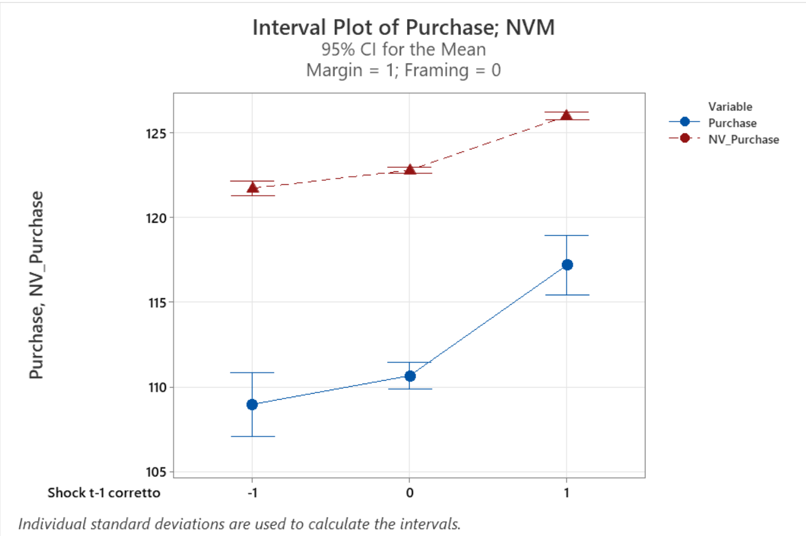



**FIGURE 25 A) AND B): INTERVAL PLOTS FOR PURCHASE AND NV_PURCHASE WITH A LOW MARGIN TREATMENT IN RELATION TO SHOCK$_{T-1}$**

It clearly appears how the variable purchase is strongly influenced by the effects of Shock t-1: respondents seem to purchase more or less in response to the shock realization of demand. Indeed, they tend to blindly take these single out-of-the-ordinary realizations as telling of higher or lower future demand levels. Interestingly enough, this fits also really well with Schweitzer and Cachon's theories on *ex*-post inventory error for which subjects behave as if their utility function incorporates a preference to reduce the absolute difference between the chosen quantity and realized demand after it is already incurred (Cachon, 2000), so the aforementioned inventory error.

Moreover, it is also visible the effect of the interaction between Margin and Shock t-1 factors. With low margins parties tend to order on average more than what would order an ideal participant, while the contrary seems to be valid for higher margins. We could interpret this as participants being biased by low margins, which are assumed, incorrectly, as less risky than higher ones: we can thus say that they are more optimistic in one case than in the other.



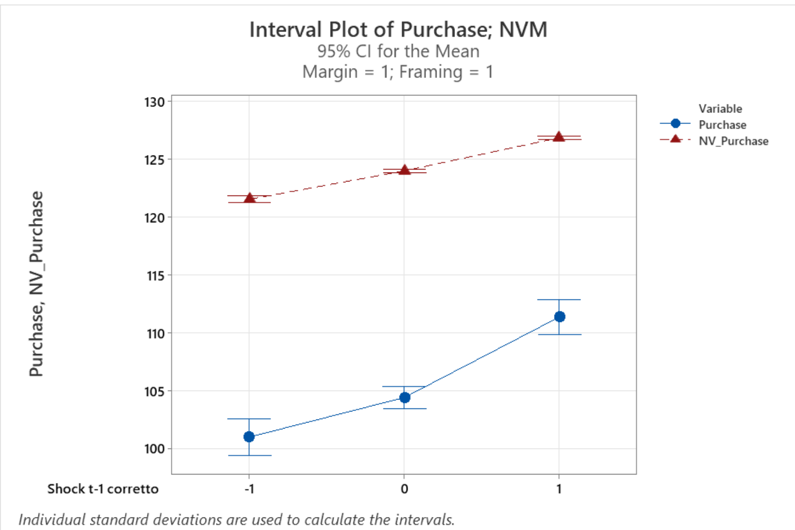



FIGURE 26 A) AND B): ): INTERVAL PLOTS FOR PURCHASE AND NV_PURCHASE WITH A HIGH MARGIN TREATMENT IN RELATION TO SHOCK$_{T-1}$

To further sustain this theory we had a look at the delta between those variables to see if, indeed, their distance grows when faced with shocks and thus this deep oscillation in demands enhances irrational behavior in the respondents. Starting, again, from the results obtained from an analysis of variance we first asserted which factors have significant effects on Delta_Forecast_NV, then proceeded to draw an Interval plot.

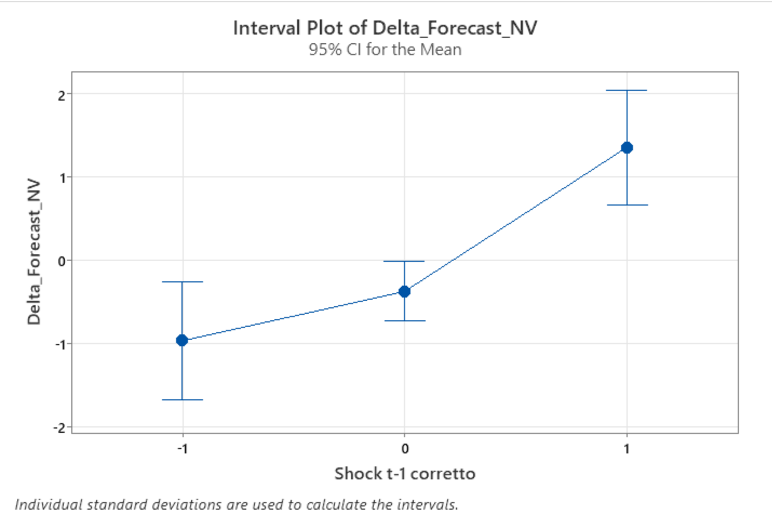



FIGURE 27: DELTA_ FORECAST_NV WITH RESPECT TO SHOCK$_{T-1}$

The same analysis were then replicated separating Session 1 and Session 2 data to avoid including in the analysis any distortion generated by the informational lecture held on the Newsvendor between the two.

## 1.3 The relationship between previously realized demand and respondent behavior

Then we proceed to explore the relationship between subjects' forecasts/ordered quantities and previously realized demand in two ways:

1) By studying the effects of the directly previous demand at time t, treated as a continuous variable.
2) By studying the effects of the mean of the 5 precedent periods demand, which is shown to respondents during each task of the Newsvendor problem.

This was done to determine if and how visive information inputs and feedback influence respondents. Indeed, by giving them information about previous levels of demand we can assume with good probability that participants will use it to adjust their estimations. This hypothesis fits well with the anchoring and insufficient adjustment heuristics which was employed with success in various scientific articles to model newsvendor behavior. From the notable papers on the topic that have been taken as a reference to support this type of investigation we reference, in particular, an article from D’Urso et all that demonstrated that decision-makers during the Newsvendor Problem tasks follow two different modalities of reasoning to find an anchor according to the demand information received (D'Urso, 2017): when knowing the demand distribution they will find a fixed anchor based on the mean of the distribution and then proceed to make small adjustments, conversely in the absence of this information respondents will be guided by the historic data and anchor on a progressive mean of the demand levels making larger adjustments and thus will show higher variance over time.

From the two explorative ANOVAs performed it emerged that participants’ Forecast is influenced by realizations of demand of precedent periods be it the precedent period t, modelized by $D_{t-1}$, or the mean of the five periods preceding t, depicted by the variable $D_{t-5}$.

Interestingly, it appears that the effects of $D_{t-1}$ on the variable Forecast, combined with the interaction of Margin and Framing are not only statistically significant but tend to cover the influence of Session. We could translate this point as a tendency of participants to be more influenced, when forecasting and/or placing orders, by events related to the demand realization than the lecture held between the sessions.

This point could have huge practical implications on how to improve allocation problems performance of managers, for example, in real life: since subjects, even after a specific focus session on the nature of the problem at hand, are still easily influenced by real outcomes of the market, it is easy to assume that maybe they should be exposed with less frequency to the high amount of information on markets trends that now a day is available (Lurie & Swaminathan, 2009). Just by thinking about how recent advances in information technology in our globalized and overtly connected world enable decision-makers to have information almost in real time and to receive frequent feedback on the outcomes of their actions, it appears evident that the performance decline individuated by Lurie and Swaminathan could be easily exacerbated.

Furthermore, this opens the discussion to the more general topic of how decision-making supports systems influence and sometimes can even improve newsvendors’ performances (D’Urso D, 2021) , which won’t be further investigated to not diverge from the main focus of the dissertation.

## Analysis of Variance

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 763.9 | 1 | 763.9 | 3.47 | 0.0626 |
| Margin | 739.1 | 1 | 739.1 | 3.36 | 0.067 |
| Framing | 4.5 | 1 | 4.5 | 0.02 | 0.8861 |
| D t-1 | 140467.1 | 1 | 140467.1 | 637.89 | 0 |
| Session:Margin | 2085.5 | 1 | 2085.5 | 9.47 | 0.0021 |
| Session:Framing | 2774.3 | 1 | 2774.3 | 12.6 | 0.0004 |
| Session:D t-1 | 19.7 | 1 | 19.7 | 0.09 | 0.7647 |
| Margin:Framing | 2481.2 | 1 | 2481.2 | 11.27 | 0.0008 |
| Margin:D t-1 | 147.7 | 1 | 147.7 | 0.67 | 0.4128 |
| Framing:D t-1 | 70.7 | 1 | 70.7 | 0.32 | 0.571 |
| Session:Margin:Framing | 12.1 | 1 | 12.1 | 0.06 | 0.8145 |
| Session:Margin:D t-1 | 14.2 | 1 | 14.2 | 0.06 | 0.7996 |
| Session:Framing:D t-1 | 103.5 | 1 | 103.5 | 0.47 | 0.4929 |
| Margin:Framing:D t-1 | 2939.4 | 1 | 2939.4 | 13.35 | 0.0003 |
| Session:Margin:Framing:D t-1 | 1.2 | 1 | 1.2 | 0.01 | 0.9411 |
| Error | 2528858.1 | 11484 | 220.2 | | |
| Total | 2757056.8 | 11499 | | | |

Constrained (Type III) sums of squares.

**FIGURE 28: ANOVA RESULTS FOR FORECAST WITH RESPECT TO TREATMENTS CONTAINING SESSION, MARGIN, FRAMING, AND $D_{T-1}$**

The same findings are also visible in the variable Purchase, which includes all the quantities ordered in each task by players. Again, the effect of D t-1 seems to perfectly cover the influence of the variable Session as shown in the ANOVA results reported below.

## Analysis of Variance

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 281.9 | 1 | 281.9 | 0.85 | 0.3568 |
| Margin | 74522.5 | 1 | 74522.5 | 224.48 | 0 |
| Framing | 5371.3 | 1 | 5371.3 | 16.18 | 0.0001 |
| D t-1 | 78791.7 | 1 | 78791.7 | 237.34 | 0 |
| Session:Margin | 217585.9 | 1 | 217585.9 | 655.41 | 0 |
| Session:Framing | 94013.4 | 1 | 94013.4 | 283.19 | 0 |
| Session:D t-1 | 836.4 | 1 | 836.4 | 2.52 | 0.1125 |
| Margin:Framing | 7860 | 1 | 7860 | 23.68 | 0 |
| Margin:D t-1 | 2287.4 | 1 | 2287.4 | 6.89 | 0.0087 |
| Framing:D t-1 | 114 | 1 | 114 | 0.34 | 0.5579 |
| Error | 3814168 | 11489 | 332 | | |
| Total | 5504240.9 | 11499 | | | |

Constrained (Type III) sums of squares.

**FIGURE 29: ANOVA RESULTS FOR PURCHASE WITH RESPECT TO TREATMENTS CONTAINING SESSION, MARGIN, FRAMING, AND $D_{T-1}$**

To further prove the influence also of the visive feedback provided on newsvendors behavior we performed the same analysis but instead of using D t-1 the ANOVA was repeated for both Forecast and Purchase experimental values with Dt-5, so the mean of the 5 previous realizations of demand visible to the players during the game. The main results are reported below.

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 731.4 | 1 | 731.4 | 3.47 | 0.0624 |
| Margin | 44.1 | 1 | 44.1 | 0.21 | 0.6472 |
| Framing | 0.5 | 1 | 0.5 | 0 | 0.96 |
| D t-5 | 208659.4 | 1 | 208659.4 | 991.19 | 0 |
| Session:Margin | 320.5 | 1 | 320.5 | 1.52 | 0.2173 |
| Session:Framing | 2250 | 1 | 2250 | 10.69 | 0.0011 |
| Session:D t-5 | 1416.3 | 1 | 1416.3 | 6.73 | 0.0095 |
| Margin:Framing | 69.3 | 1 | 69.3 | 0.33 | 0.5662 |
| Margin:D t-5 | 654.4 | 1 | 654.4 | 3.11 | 0.0779 |
| Framing:D t-5 | 12.4 | 1 | 12.4 | 0.06 | 0.8081 |
| Session:Margin:Framing | 80.8 | 1 | 80.8 | 0.38 | 0.5355 |
| Session:Margin:D t-5 | 11.1 | 1 | 11.1 | 0.05 | 0.8183 |
| Session:Framing:D t-5 | 1268.7 | 1 | 1268.7 | 6.03 | 0.0141 |
| Margin:Framing:D t-5 | 60.9 | 1 | 60.9 | 0.29 | 0.5908 |
| Session:Margin:Framing:D t-5 | 75 | 1 | 75 | 0.36 | 0.5505 |
| Error | 2417547.5 | 11484 | 210.5 | | |
| Total | 2757056.8 | 11499 | | | |

Constrained (Type III) sums of squares.

**FIGURE 30: ANOVA RESULTS FOR FORECAST WITH RESPECT TO TREATMENTS CONTAINING SESSION, MARGIN, FRAMING AND $D_{T-5}$**

We would like to highlight how for Purchase we have excluded to study the interaction between Session* Framing* $D_{t-5}$ since both the variables don't yield any statistically significant effect.

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 705.1 | 1 | 705.1 | 2.17 | 0.1411 |
| Margin | 3077.6 | 1 | 3077.6 | 9.45 | 0.0021 |
| Framing | 713.3 | 1 | 713.3 | 2.19 | 0.1388 |
| D t-5 | 139674.1 | 1 | 139674.1 | 429.07 | 0 |
| Session:Margin | 216623.7 | 1 | 216623.7 | 665.45 | 0 |
| Session:Framing | 82962.4 | 1 | 82962.4 | 254.85 | 0 |
| Session:D t-5 | 1843.9 | 1 | 1843.9 | 5.66 | 0.0173 |
| Margin:Framing | 7753.7 | 1 | 7753.7 | 23.82 | 0 |
| Margin:D t-5 | 3714.4 | 1 | 3714.4 | 11.41 | 0.0007 |
| Framing:D t-5 | 3.7 | 1 | 3.7 | 0.01 | 0.9149 |
| Error | 3740001 | 11489 | 325.5 | | |
| Total | 5504240.9 | 11499 | | | |

Constrained (Type III) sums of squares.

**FIGURE 31: ANOVA RESULTS FOR PURCHASE WITH RESPECT TO TREATMENTS CONTAINING SESSION, MARGIN, FRAMING AND $D_{T-5}$**

## 5.5 Economic impacts of Variability

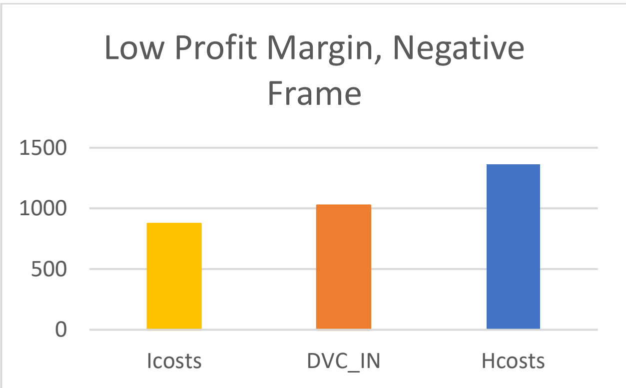
Starting from one of the main drivers of variability individuated, Shock $_{t-1}$, we will have a look into its effects on the costs incurred by respondents' and we will try to identify how much of those costs are directly imputable to variability. But first let us determine how much of the actual costs are are attributable to variability. To do so we have proceeded to compute the following three amounts:

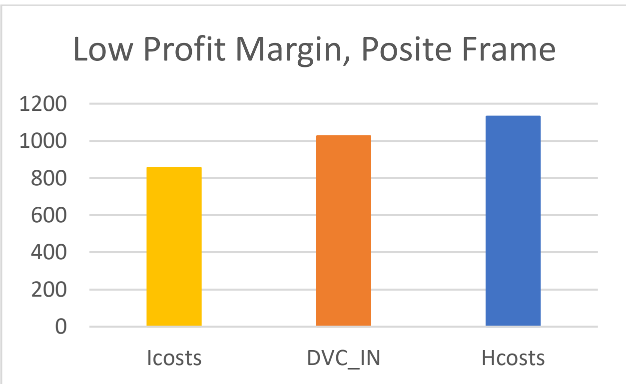
1. The ideal costs which would have been faced by an optimal newsvendor, Icosts;
2. The costs that were actually faced by human newsvendors, Hcosts;
3. The amount of costs generated by the variability, Vcosts, computed as follows:

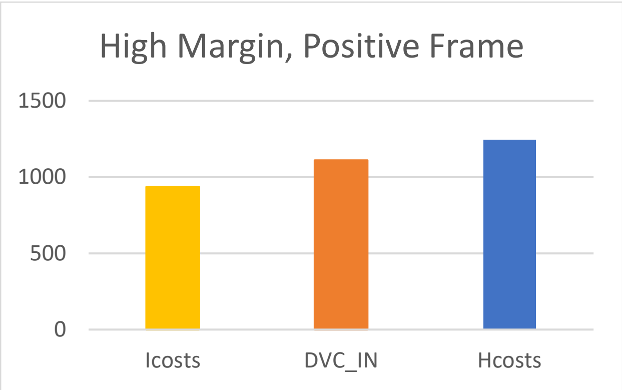
$$VCost_{i,j}\ (t) = MAX\ [m*(D_j\ (t) - HVN_{i,j}\ (t));c*(\ HVN_{i,j}\ (t) - D_j\ (t))] \quad (1.7)$$

Where all the quantities employed are descripted in Chapter 4, Paragraph 4.1. The results obtained show how a part of the cost incurred by human newsvendors is linked to variability and

how Margins influence those amounts.





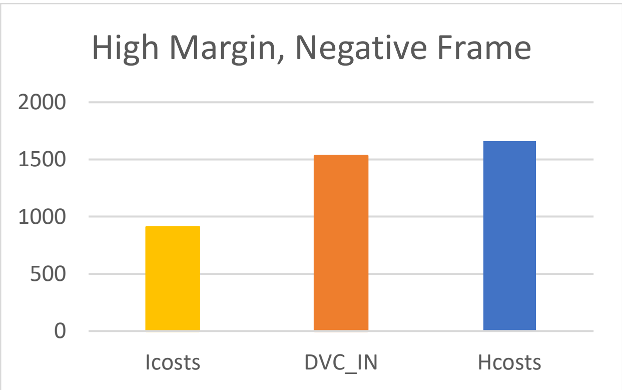




**FIGURE 32 A), B), C), D) : HISTOGRAMS OF THE VARIOUS TYPOLOGIES OF COSTS COMPUTED**

It appears that high-margin products not only generate more costs overall, but also possess higher Vcosts with respect to low-margin ones. This is further corroborated by the interval plot for VCosts with respect to Margins and Framing interaction: with a negative frame when switching from high to low margin products we face a significant reduction of costs linked to variability, while the contrary is true for positive framed problems.

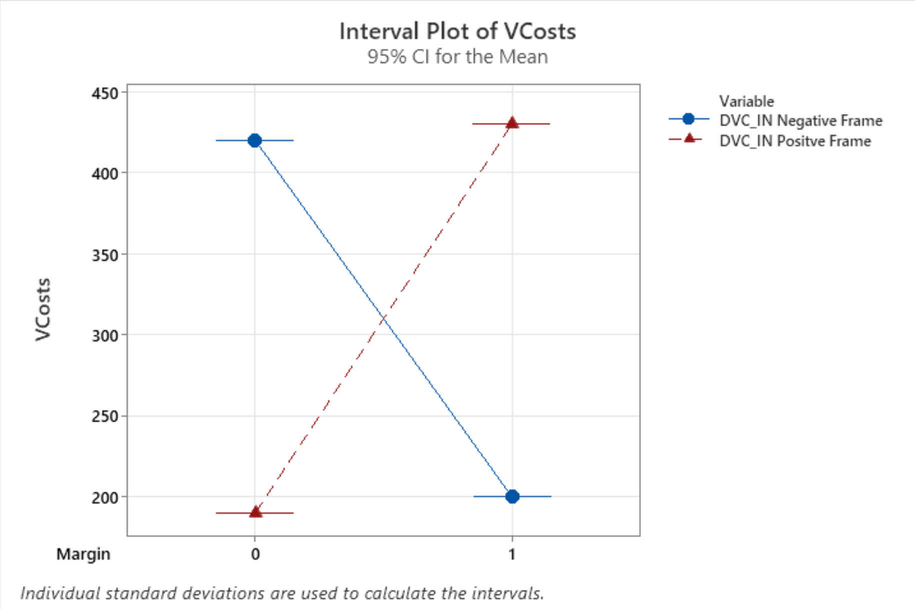



**FIGURE 33: VCOSTS INTERVAL PLOT FOR NEGATIVE AND POSITIVE FRAMINGS**

By plotting the Interval graph with Shock $_{t-1}$ as grouping factor we see a steady decline of the difference between real and ideal *newsvendor* costs, but the confidence intervalsslight overlapping suggest that this decline in quantities is not significative.

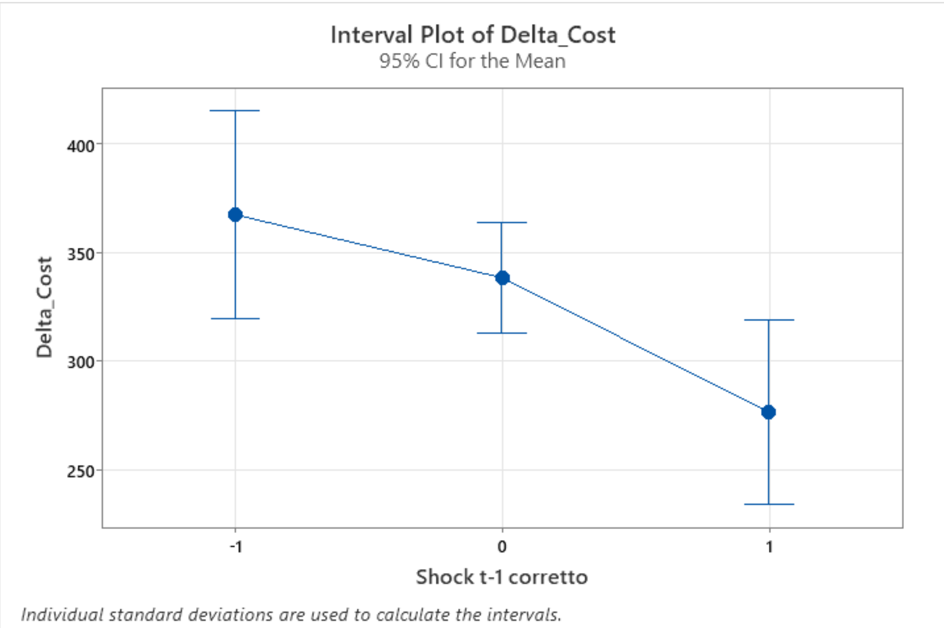



**FIGURE 34: DELTA_COSTS' INTERVAL PLOT WITH RESPECT TO SHOCK$_{T-1}$**

However, by employing the findings from our exploratory three-way ANOVA, not only the Shock $_{t-1}$ but also Framing and Margin interaction (Framing* Margin) seems to yield significant effects on the Delta Cost, so the difference between costs incurred by participants and the one faced by ideal *newsvendor*. With positive Margins, the delta quantity reduces when in the presence of positive shocks in the demand levels, defined as realizations higher than the mean of the effective demand, μ, plus the standard deviation of the demand distribution.

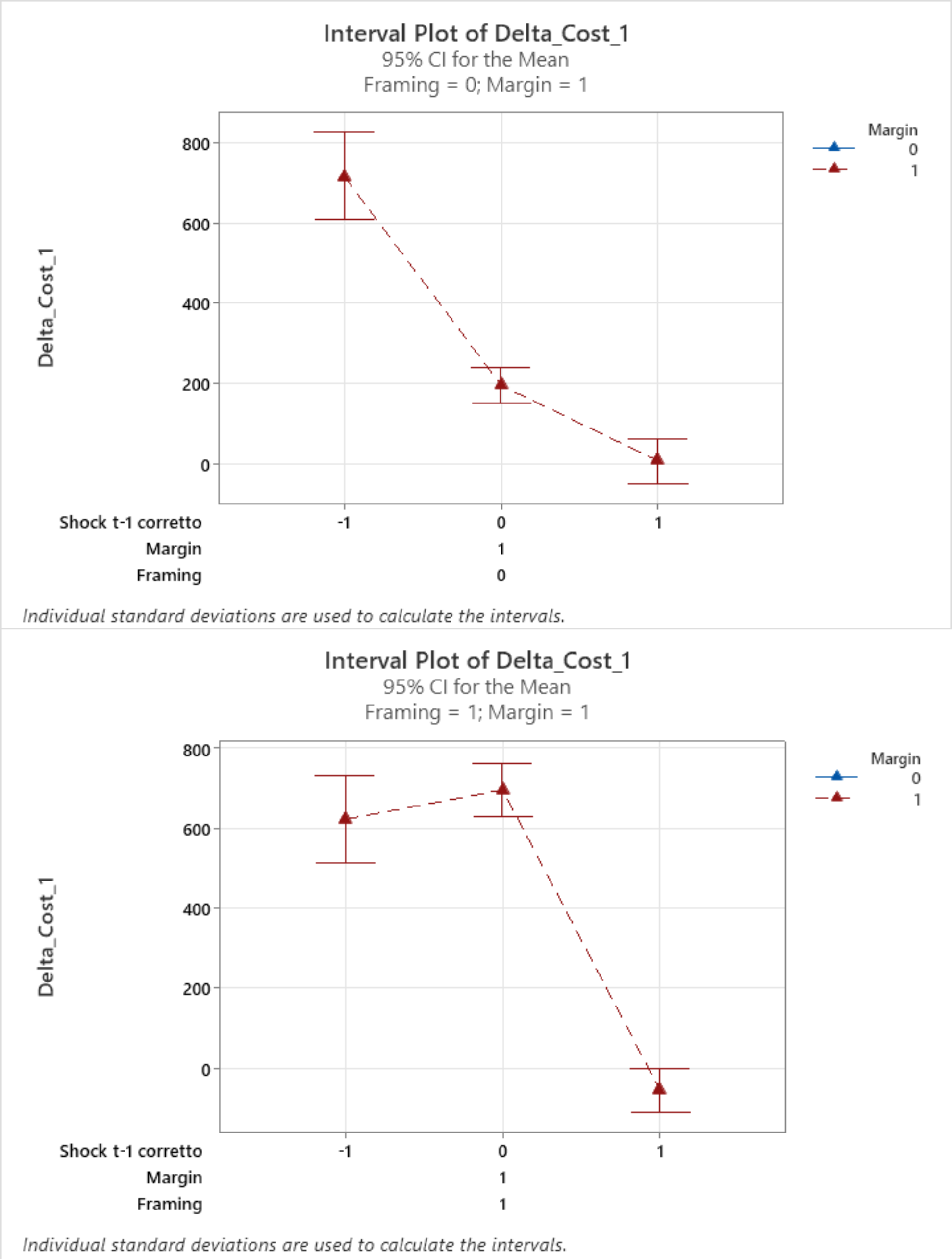



**FIGURE 35 A) AND B) : DELTA COSTS' INTERVAL PLOTS WITH HIGH MARGIN TREATMENTS**

We could translate this point as human newsvendor being less optimistic when faced by higher margin and thus purchasing more items, but in an insufficient way, even when facing positive shocks, while being more prone to under-purchase even less in case of negative shock levels thus increasing their losses. Conversely, with a low margin, it appears that participants are more prone to over-purchase, probably pushed by the margin which seems to yield lower losses for the untrained eye.

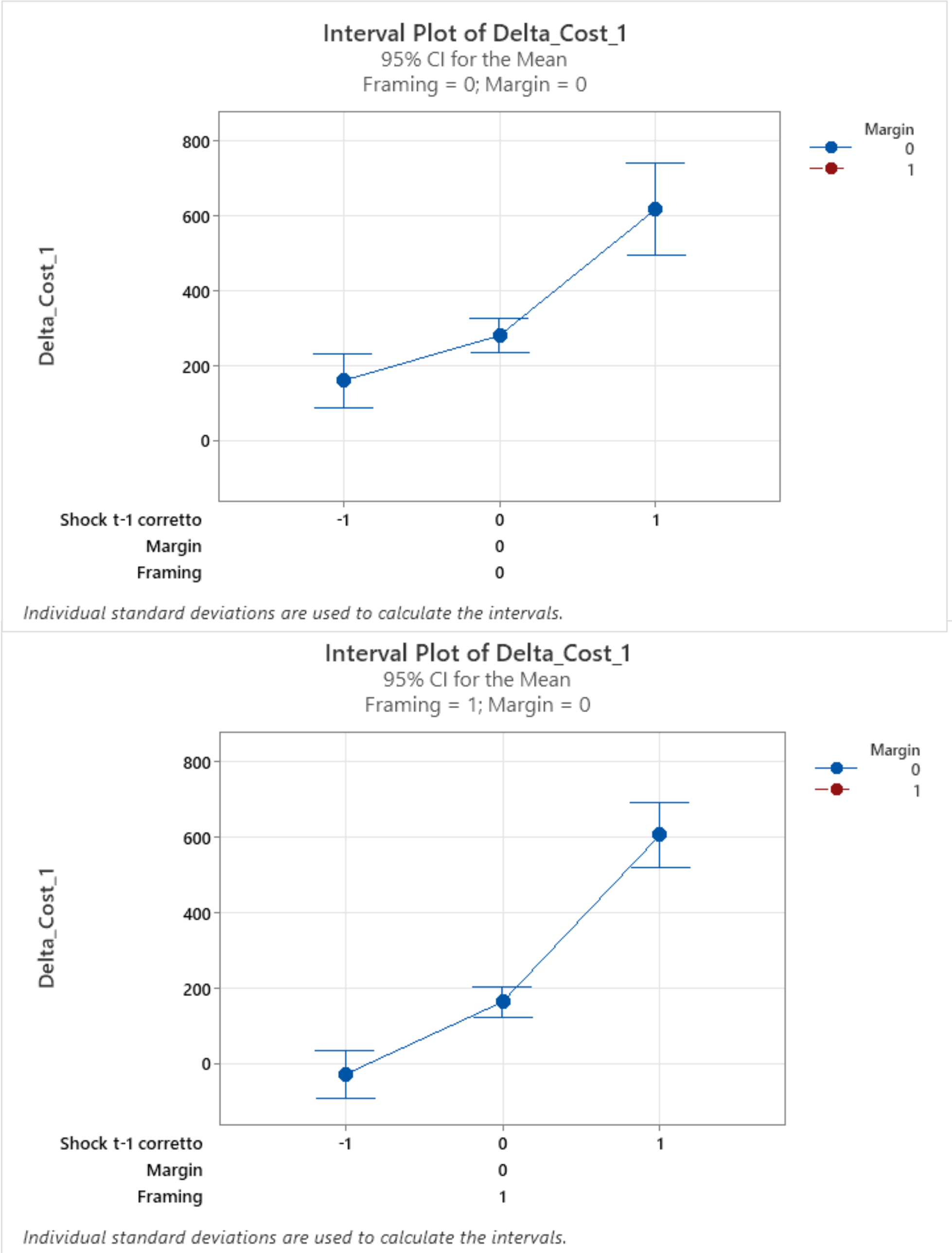



**FIGURE 36 A) AND B) : DELTA COSTS' INTERVAL PLOTS WITH LOW MARGIN TREATMENTS**

This answers to one of our initial questions: shock realization of demand foster variability in the behavior of participants and thus irrationality with a cost, which differs depending on the margin of the product in exam.

# 6. CONCLUSIONS

Our aim in this dissertation was to uncover what drives the variability that afflicts singular respondents over time, which seems that, at the best of our knowledge, a few papers have tried to study in depth in the Newsvendor problem available literature.

With our analysis of the experimental data collected from the execution of the Newsvendor problem in a university setting we tried to explain what drives the variability which seems to afflict each respondent over time. Indeed, we obtained a few interesting results. Starting from a separate analysis of the inventory levels and the demand forecast, we determined the drivers of variability for both experimental values: it appears that newsvendors are influenced by both experimental factors, appositely set, and by external inputs such as shocks in the demand levels. In particular, newsvendors seem to blindly follow those strong changes in the realized demand, irrationally interpreting those random occurrences as signals of growing or shrinking demand.

Before going in the detail of the results obtained, it is important to note how this investigation stemmed from a gap in the literature on the behavioral aspects of Newsvendor Problem solvers. From a through screening of the available scientific articles it emerged that almost none was dedicated to the investigation of the aspects linked to the variability manifested by newsvendor during the problem’s timeframe. This would seem quite disconcerting, if we think of how this variability is indeed a contributor to cost incurred by participant and one of the main determinants of the economic efficiency or inefficiency of newsvendors. Placing an order closer to the optimum yields lower costs than purchasing very far away from it: it is to the high variance demonstrated by newsvendor answers during each task to which a substantial part of the costs is attributable.

The approach employed for the analysis of the data collected was structured, first, by segregating the two variables of interest, Forecast, and Purchase, starting from the fact that they do behave differently. Our assumption was thus that they also possess different relations with newsvendors’ variability over time.

Then we proceeded to explore the data collected by performing some ANOVAs on both variables to uncover their behavior under different treatments. Once determined what influences them, so what generates statistically significant effects, be it an experimental factor or an external input, we plotted the Interaction Graphs to have a visual proof of those effects.

## 4.3 The results obtained

Hence, in this chapter we will discuss in detail the results obtained and try to understand the deeper meanings held by them, particularly their contribution to the understanding of the behavioral aspects of the Newsvendor Problem.

In detail from the explorative ANOVAs it was clear the influence of Margin, Framing and Session exercises on both Forecast and Purchase: for Forecast, it is evident the disruptive effects generated by the interaction between a positive margin and the variable session. Indeed, newsvendors seem to be lulled into a more optimistic state after the lecture and not only anticipate higher demand levels overall but diverges from what would have been the optimal behavior, so to foresee lower levels.

The same effects even more evident have been found from the analysis of Purchase, but with a significant reversion of trends: with low margins the newsvendors appear to purchase on average more than their ideal counterpart, but manifesting a downward trend between sessions, contrarily from what an optimal participant would do. It is interesting to note that this happens while at the same time forecasting more! This perfectly confirms what was assumed in the third chapter of this thesis: that Forecast and Purchase are substantially different and thus should be studied separately since they are influenced in different ways and thus contribute differently to newsvendor's variability over time.

Then we proceeded with this parallel analysis between the two variables also to individuate possible external inputs, not linked to the experiment setup, which influence and drive newsvendors' variability. We identified shock levels of demand, which can be modeled by the variable Shock $_{t-1}$, as one of the main contributors to variability for both Forecast and Purchase. Here, again we found the effect on the Forecast of the variable Margin, but this time it was also visible a strong interaction between the Margin and the variable Shock $_{t-1}$. Again, it was observable that with high margins participants behave in an optimistic manner in relation to demand levels previsions with respect to the ideal newsvendor, while at the same time, they tend to follow the strong shocks in the effective demand, by adapting to higher or lower demands levels in case of upwards positive shocks or negative ones.

Again, a similar pattern can be found for Purchase, but with a twist: there is again an inversion of tendencies, low margins bring newsvendor to purchase more than what the optimum would dictate. This can be interpreted as a tendency to distinguish between what they forecast and what they would actually purchase when put in front of different margins: participants will be led toward being less optimistic for low margins' products when forecasting due to their less attractive demand, while when placing orders they will be more prone to over purchasing due to the perception of low margins as more secure than high margins' products.

The same reasoning appears to be valid also for the inventory levels: the variable Purchase in the various analysis of Variance performed is influenced by the D t-1 and D t-5. This demonstrates that information not only has effects on a prediction but also alters newsvendors' orders.

But how to link all these findings with variability? In a very simplistic way, we have taken the distance between human newsvendor forecasts and the ideal newsvendor ones and had a look at how it behaves when facing the three levels of Shock $_{t-1}$. By plotting the interaction graphs it emerged that indeed the distance between the two values and thus the delta grows more when put in front of a positive, upward shock than with a downward one. A more sophisticated approach could be to perform a regression to see how much of the delta is explained by the shocks in the demand levels.

This adaptation to external inputs enhances newsvendors variability and thus moves them from the optimal levels of both Forecast and Purchase, generating a lot of additional costs for participants. Interestingly, the effects in terms of economic efficiency are quite variable and dependent on the particular treatment newsvendors are subjected. More generally, we found out that newsvendors not only look and infer information from the demand realized in the direct precedent period but follow the mean of the 5 decision-making periods visible on the screen during each task of the Newsvendor Problem.

## 4.4 Limitations and Future steps

The analysis, however, presents various limitations. First and foremost, it is not clear how much of the variability is imputable to Purchase or Forecasts. It would be interesting to investigate this aspect in order to isolate the higher contributor to irrationality in the Newsvendor decision-making process and try to create the right setting to manage and dampen its effects. This will represent an insightful contribution to the Inventory management world.

Another point to bring to the attention would be a further investigation of the effects of the variable Session and of the learning by doing on the Variability: indeed we didn't had the change to study the variable Task in this setting thus neglecting to look the effects linked to experience and learning processes.

Following the main trends now present in the scientific literature on the Newsvendor Problem application, it could be also insightful to employ the new stream of studies on brain activation during NVM tasks to better characterize what determines respondents' irrational behaviour. Evidence which areas of the brain activate when faced with external inputs, such as shocks in the demand level, or to track-eye movements and skin-conductance while subjects actively participate in the Newsvendor tasks. It could be interesting to link this stream of research with Individuals heterogeneity in the answering pattern.

One more alternative could be the joint study of heterogeneity and variability: Do different categories of individuals show different variability in the orders placed overtime?

This could be easily linked to studies on the individuals heterogeneity present in newsvendors and ,in particular, to the studies related to gender and the results obtained from the CRT ("Cognitive Reflection Test"). On the last point it is indeed easily observable how higher ranking participants in the CRT tend to perform on average economically better in the newsvendor than ones with lower scores( for reference the Interval plot for Delta Cost in relation to CRT has been reported in the Annex) . So why stopping here and not having a look at their variability?

# Acknowledgements

I would like to express my deepest gratitude to Professor Zotteri, Colombo and Cantamessa for their guidance and support during the whole inception, analysis and drafting of this dissertation. Without their insights, profound knowledge and expertise on the topic this endeavor would not have been possible and I would probably still be lost in the amount of information available.

A special mention to Professor Samuele Colombo who, despite the different time zone, advised me and inspired me through the whole process.

Words cannot express my gratitude to my family, for their utmost patience and never-ending moral support during my academic journey. Big thanks also to my little sister, Aurora, who never failed to put me on the spot when I sorely needed it.

Lastly, I would like to mention my friends, who kept me motivated during the highs and lows of my University years. From Trieste to Turin, they have been my rocks, always making me laugh even in the face of adversity.

Some say that friends come and go, but I am lucky to have found so willing to stay.

# APPENDIX A

Figure 1: Paired T-test for Session 2 NV_Forecast and Forecast means.

### Descriptive Statistics

| Sample | N | Mean | StDev | SE Mean |
|---|---|---|---|---|
| NV_Forecast | 5720 | 98,556 | 8,250 | 0,109 |
| Forecast | 5720 | 100,166 | 16,437 | 0,217 |

### Estimation for Paired Difference

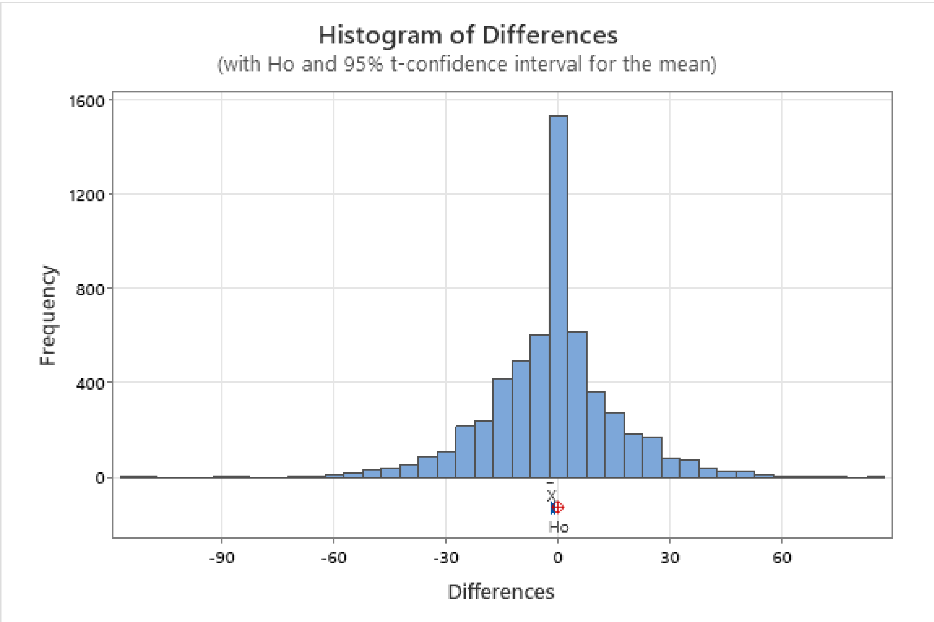
| Mean | StDev | SE Mean | 95% CI for μ_difference |
|---|---|---|---|
| -1,610 | 17,117 | 0,226 | (-2,053; -1,166) |

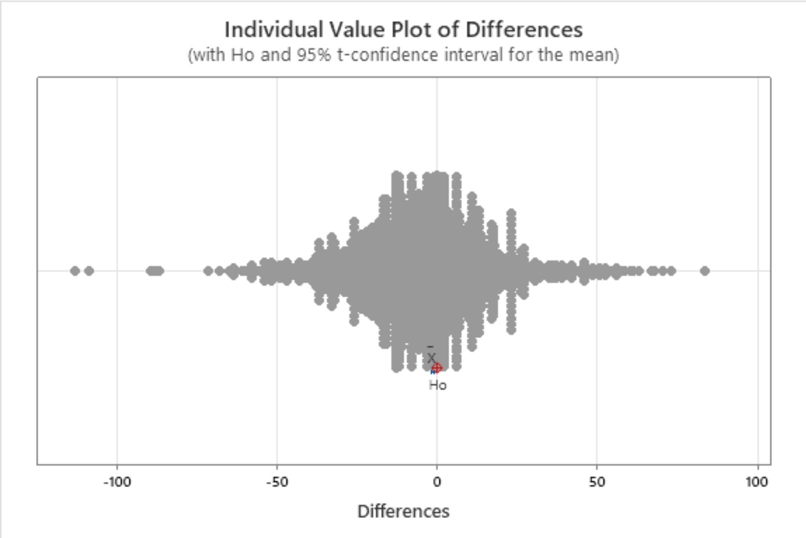
*μ_difference: population mean of (NV_Forecast - Forecast)*

### Test

| | |
|---|---|
| Null hypothesis | $H_0$: μ_difference = 0 |
| Alternative hypothesis | $H_1$: μ_difference ≠ 0 |

| T-Value | P-Value |
|---|---|
| -7,11 | 0,000 |





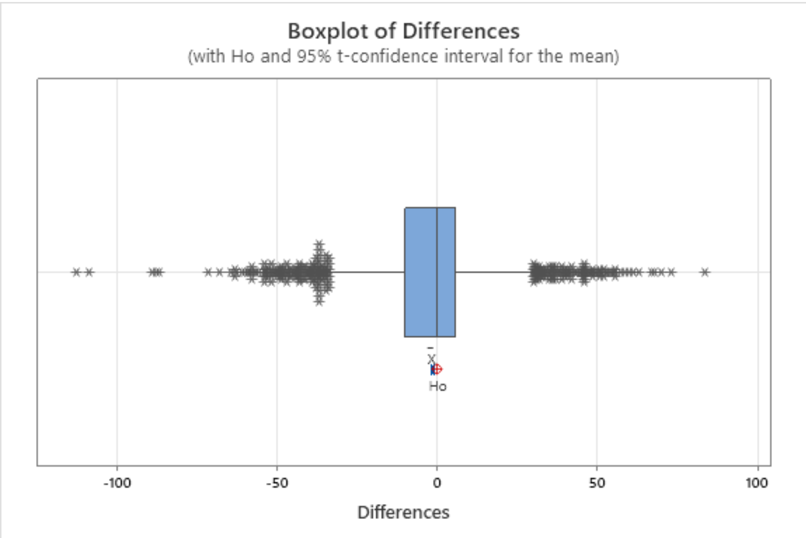




Figure 2: Two-Variance Test on Session 2 NV_Forecast and Forecast

### Method

$\sigma_1$: standard deviation of NV_Forecast
$\sigma_2$: standard deviation of Forecast
Ratio: $\sigma_1/\sigma_2$
The Bonett and Levene's methods are valid for any continuous distribution.

### Descriptive Statistics

| Variable | N | StDev | Variance | 95% CI for σ |
|---|---|---|---|---|
| NV_Forecast | 5720 | 8,250 | 68,055 | (8,027; 8,481) |
| Forecast | 5720 | 16,437 | 270,182 | (16,006; 16,886) |

### Ratio of Standard Deviations

| Estimated Ratio | 95% CI for Ratio using Bonett | 95% CI for Ratio using Levene |
|---|---|---|
| 0,501881 | (0,483; 0,522) | (0,475; 0,513) |

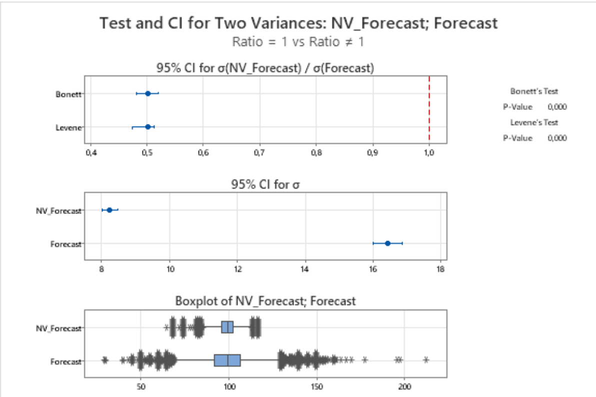
## Test

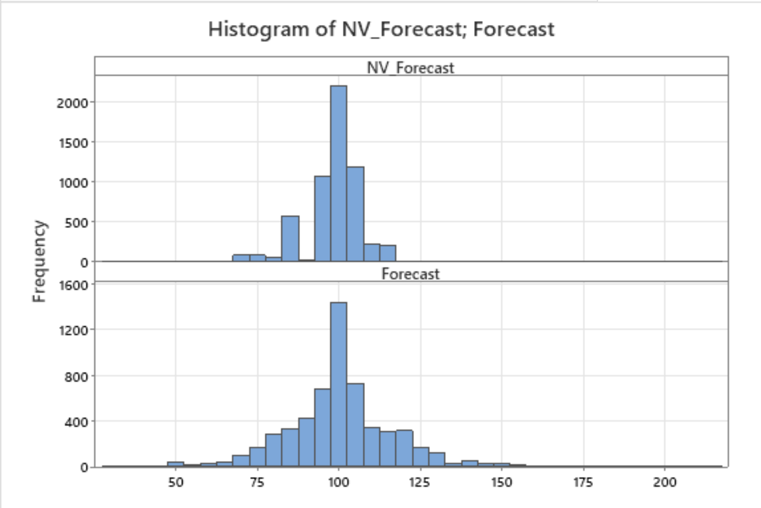
Null hypothesis: $H_0$: $\sigma_1 / \sigma_2 = 1$

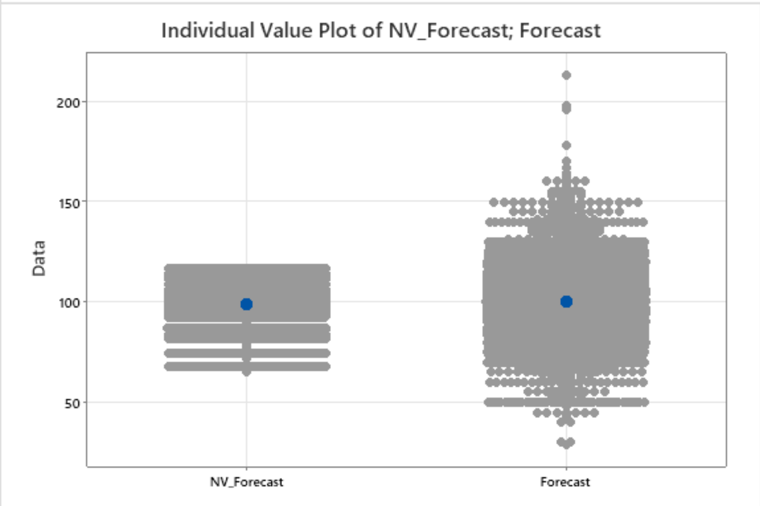
Alternative hypothesis: $H_1$: $\sigma_1 / \sigma_2 \neq 1$

Significance level: $\alpha = 0,05$

| Method | Test Statistic | DF1 | DF2 | P-Value |
|---|---|---|---|---|
| Bonett | 881,42 | 1 | | 0,000 |
| Levene | 1106,46 | 1 | 11438 | 0,000 |



## Figure 3: Forecast and Purchase 2-way ANOVA with Framing and Margin

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Margin | 101.8 | 1 | 101.8 | 0.5 | 0.4805 |
| Framing | 17796.1 | 1 | 17796.1 | 87.04 | 0 |
| Margin:Framing | 365.4 | 1 | 365.4 | 1.79 | 0.1813 |
| Error | 1181012.5 | 5776 | 204.5 | | |
| Total | 1199550 | 5779 | | | |

Constrained (Type III) sums of squares.

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Margin | 38546.2 | 1 | 38546.2 | 147.18 | 0 |
| Framing | 6859.5 | 1 | 6859.5 | 26.19 | 0 |
| Margin:Framing | 5.2 | 1 | 5.2 | 0.02 | 0.8878 |
| Error | 1496966.3 | 5716 | 261.9 | | |
| Total | 1545172.6 | 5719 | | | |

Constrained (Type III) sums of squares.

## Figure 4: Delta_Forecast_NV ANOVA results

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Session | 29458.1 | 1 | 29458.1 | 130.1 | 0 |
| Margin | 151934.8 | 1 | 151934.8 | 671.01 | 0 |
| Framing | 7637.7 | 1 | 7637.7 | 33.73 | 0 |
| Shock t-1 Corretto | 7073.3 | 2 | 3536.7 | 15.62 | 0 |
| Session:Margin | 17089.5 | 1 | 17089.5 | 75.47 | 0 |
| Session:Framing | 2440.9 | 1 | 2440.9 | 10.78 | 0.001 |
| Session:Shock t-1 Corretto | 2310.5 | 2 | 1155.2 | 5.1 | 0.0061 |
| Margin:Framing | 699.1 | 1 | 699.1 | 3.09 | 0.0789 |
| Margin:Shock t-1 Corretto | 525.9 | 2 | 262.9 | 1.16 | 0.3131 |
| Framing:Shock t-1 Corretto | 1440.9 | 2 | 720.5 | 3.18 | 0.0415 |
| Session:Margin:Framing | 7777.8 | 1 | 7777.8 | 34.35 | 0 |
| Session:Margin:Shock t-1 Corretto | 4599.8 | 2 | 2299.9 | 10.16 | 0 |
| Session:Framing:Shock t-1 Corretto | 973.5 | 2 | 486.7 | 2.15 | 0.1166 |
| Margin:Framing:Shock t-1 Corretto | 4464.7 | 2 | 2232.3 | 9.86 | 0.0001 |
| Session:Margin:Framing:Shock t-1 Corretto | 682.5 | 2 | 341.3 | 1.51 | 0.2216 |
| Error | 2598473.4 | 11476 | 226.4 | | |
| Total | 2925452.8 | 11499 | | | |

Constrained (Type III) sums of squares.

## Figure 5: PURCHASE ANOVA RESULTS

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Framing | 84282.5 | 1 | 84282.5 | 248.67 | 0 |
| Margin | 1145120.4 | 1 | 1145120.4 | 3378.6 | 0 |
| Session | 18571.6 | 1 | 18571.6 | 54.79 | 0 |
| Shock t1 | 591.1 | 1 | 591.1 | 1.74 | 0.1867 |
| Framing:Margin | 7089.6 | 1 | 7089.6 | 20.92 | 0 |
| Framing:Session | 98027.9 | 1 | 98027.9 | 289.22 | 0 |
| Framing:Shock t1 | 1669.5 | 1 | 1669.5 | 4.93 | 0.0265 |
| Margin:Session | 215804.1 | 1 | 215804.1 | 636.72 | 0 |
| Margin:Shock t1 | 40.8 | 1 | 40.8 | 0.12 | 0.7286 |
| Session:Shock t1 | 683.7 | 1 | 683.7 | 2.02 | 0.1555 |
| Error | 3894000.8 | 11489 | 338.9 | | |
| Total | 5504240.9 | 11499 | | | |

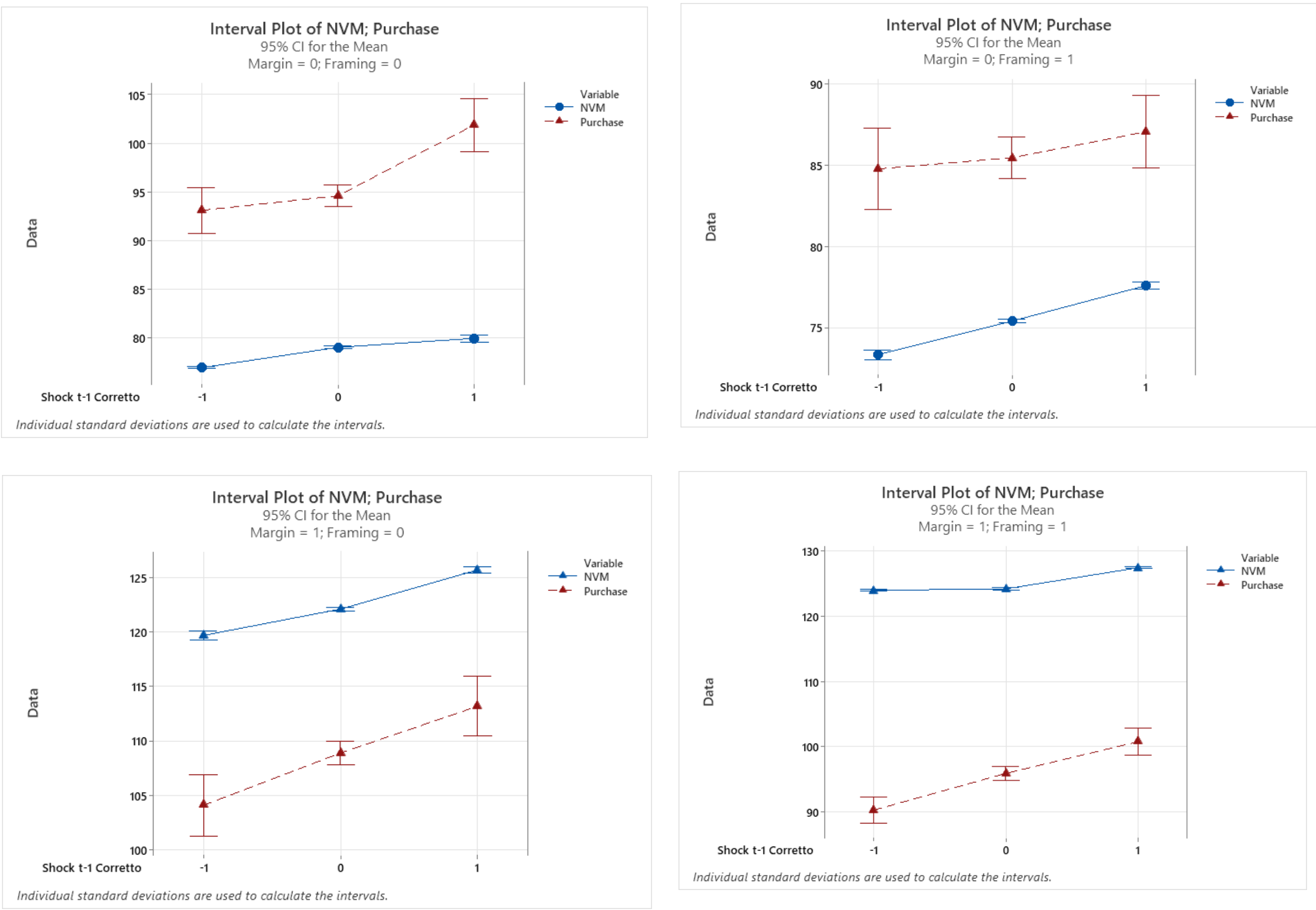
Constrained (Type III) sums of squares.

## Analysis of Variance

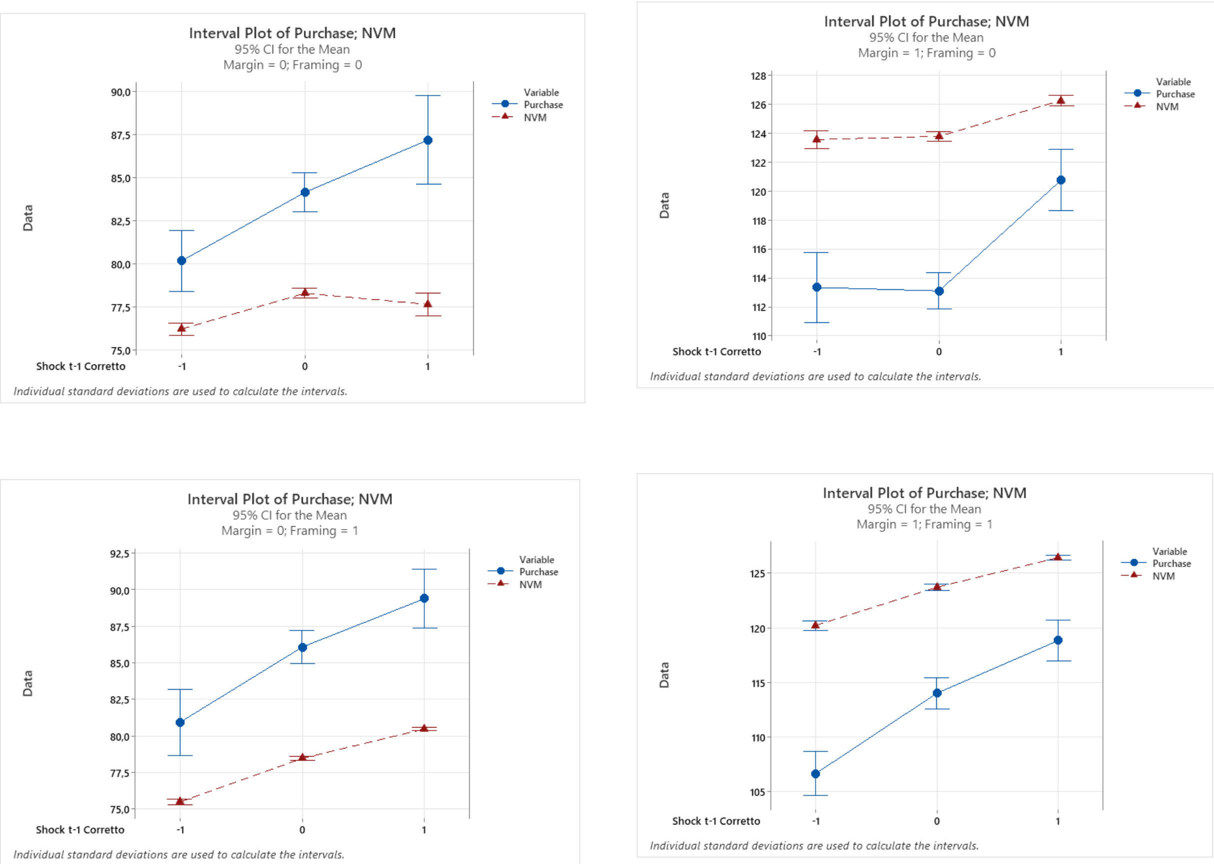
| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Framing | 81588.5 | 1 | 81588.5 | 245.94 | 0 |
| Margin | 887522.4 | 1 | 887522.4 | 2675.36 | 0 |
| Session | 17168.5 | 1 | 17168.5 | 51.75 | 0 |
| Shock t-1 Corretto | 75053.5 | 2 | 37526.8 | 113.12 | 0 |
| Framing:Margin | 6136.8 | 1 | 6136.8 | 18.5 | 0 |
| Framing:Session | 99430.9 | 1 | 99430.9 | 299.73 | 0 |
| Framing:Shock t-1 Corretto | 3161.6 | 2 | 1580.8 | 4.77 | 0.0085 |
| Margin:Session | 210835.7 | 1 | 210835.7 | 635.55 | 0 |
| Margin:Shock t-1 Corretto | 2866.6 | 2 | 1433.3 | 4.32 | 0.0133 |
| Session:Shock t-1 Corretto | 704.3 | 2 | 352.2 | 1.06 | 0.346 |
| Error | 3810031.7 | 11485 | 331.7 | | |
| Total | 5504240.9 | 11499 | | | |

Constrained (Type III) sums of squares.

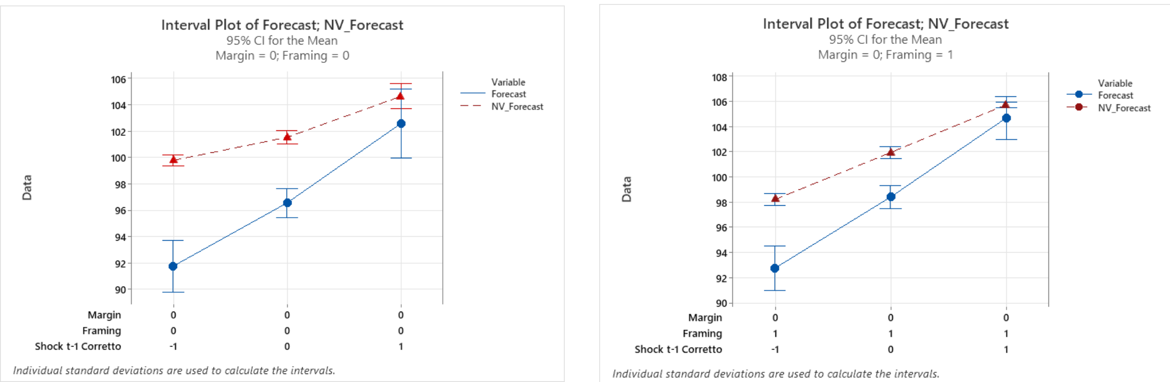
## Figure 6: Session 1 data, Purchase vs Shock t-1

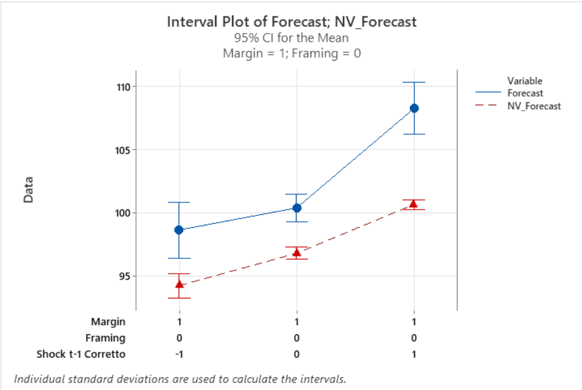


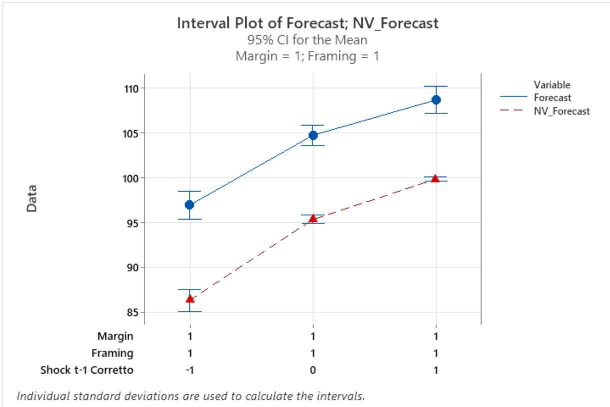
## Figure 7: Session 2 data, Purchase vs Shock t-1



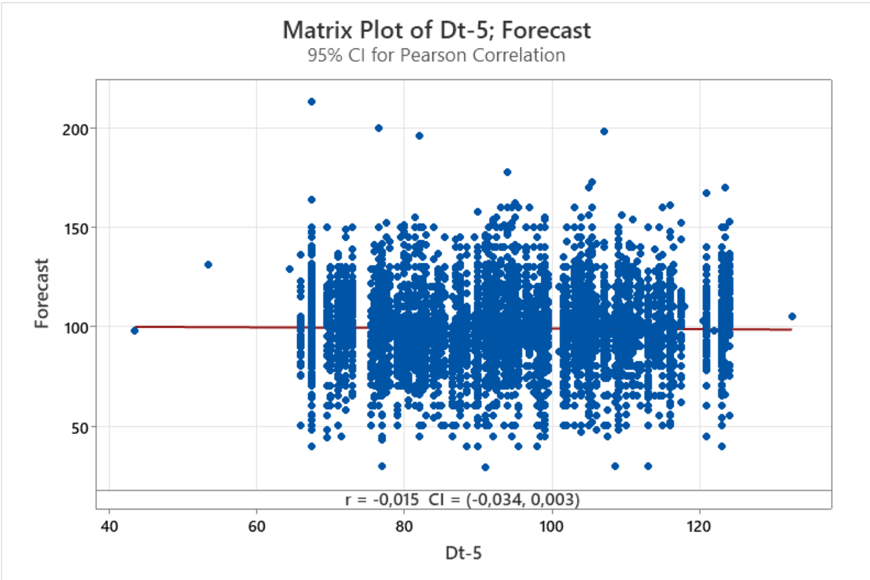

## Figure 8: Session 2 data, Forecast vs Shock t-1






## Figure 9: Session 1, Forecast vs $D_{t-5}$

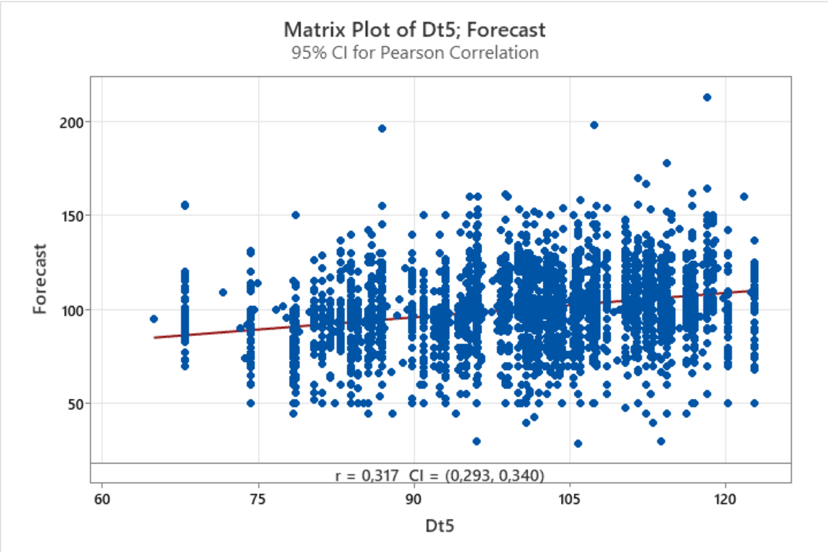



### Analysis of Variance

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Margin | 1133 | 1 | 1133 | 6 | 0.0143 |
| Framing | 333.4 | 1 | 333.4 | 1.77 | 0.184 |
| D t-5 | 70214.1 | 1 | 70214.1 | 414.22 | 0 |
| Margin:Framing | 0.3 | 1 | 0.3 | 0 | 0.9679 |
| Margin:D t-5 | 1328 | 1 | 1328 | 7.03 | 0.008 |
| Framing:D t-5 | 78.9 | 1 | 78.9 | 0.42 | 0.5181 |
| Margin:Framing:D t-5 | 0.6 | 1 | 0.6 | 0 | 0.9543 |
| Error | 1089877.9 | 5772 | 188.8 | | |
| Total | 1199550 | 5779 | | | |

Constrained (Type III) sums of squares.

Figure 10: Session 2, Forecast vs $D_{t-5}$




**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Margin | 79.2 | 1 | 79.2 | 0.34 | 0.5588 |
| Framing | 1447.5 | 1 | 1447.5 | 6.25 | 0.0124 |
| D t-5 | 153164.5 | 1 | 153164.5 | 661.34 | 0 |
| Margin:Framing | 187.1 | 1 | 187.1 | 0.81 | 0.3688 |
| Margin:D t-5 | 310.1 | 1 | 310.1 | 1.34 | 0.2473 |
| Framing:D t-5 | 959.9 | 1 | 959.9 | 4.14 | 0.0418 |
| Margin:Framing:D t-5 | 169.8 | 1 | 169.8 | 0.73 | 0.3918 |
| Error | 1322883.8 | 5712 | 231.6 | | |
| Total | 1545172.6 | 5719 | | | |

Constrained (Type III) sums of squares.

Figure 12: Forecast vs Dt-1 for both Session 1 and Session 2 data

**Analysis of Variance**

| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Margin | 288.18 | 1 | 288.18 | 1.42 | 0.2331 |
| Framing | 741.88 | 1 | 741.88 | 3.66 | 0.0557 |
| D t-1 | 7679.11 | 1 | 7679.11 | 37.9 | 0 |
| Margin:Framing | 1926.56 | 1 | 1926.56 | 9.51 | 0.0021 |
| Margin:D t-1 | 427.77 | 1 | 427.77 | 2.11 | 0.1463 |
| Framing:D t-1 | 136.98 | 1 | 136.98 | 0.68 | 0.411 |
| Margin:Framing:D t-1 | 2663.59 | 1 | 2663.59 | 13.15 | 0.0003 |
| Error | 1169375.21 | 5772 | 202.59 | | |
| Total | 1199550 | 5779 | | | |

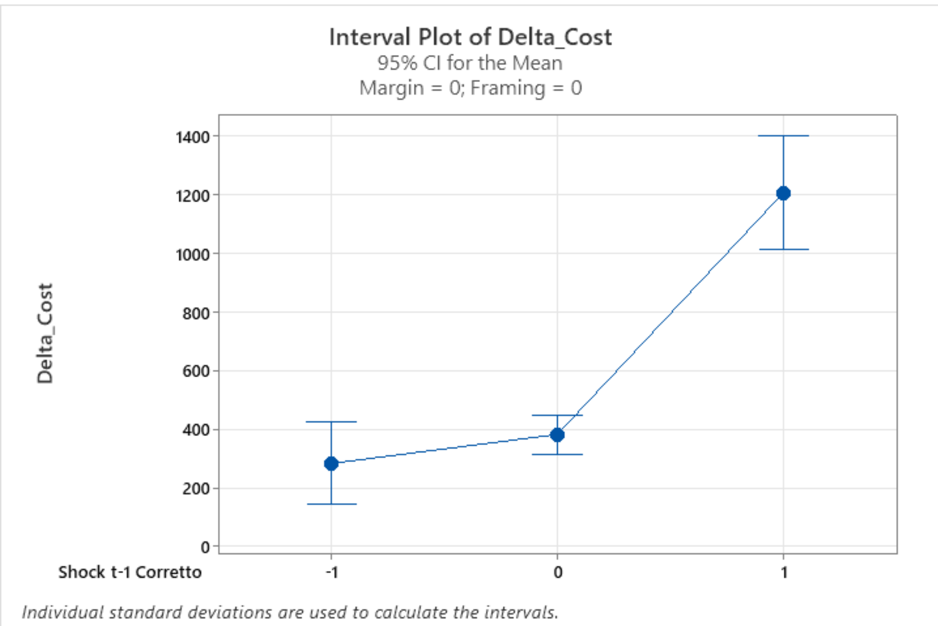
Constrained (Type III) sums of squares.

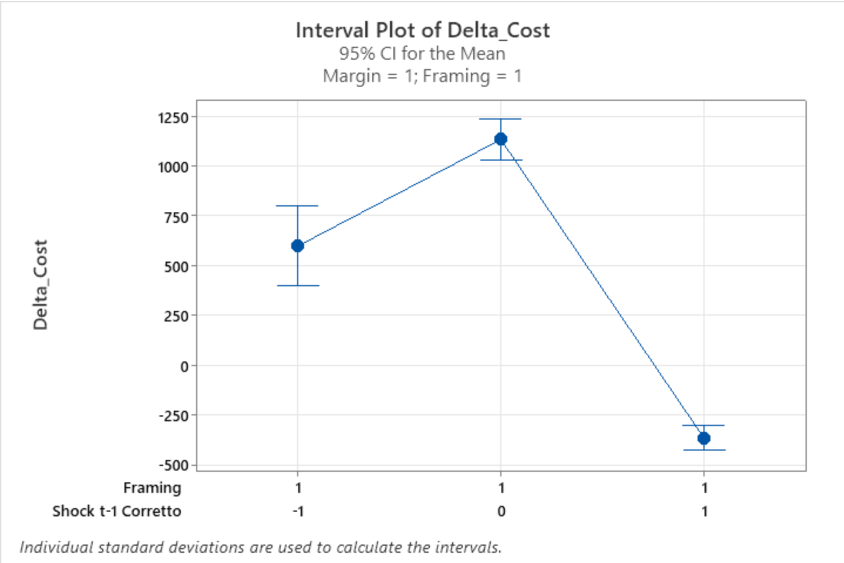
**Analysis of Variance**

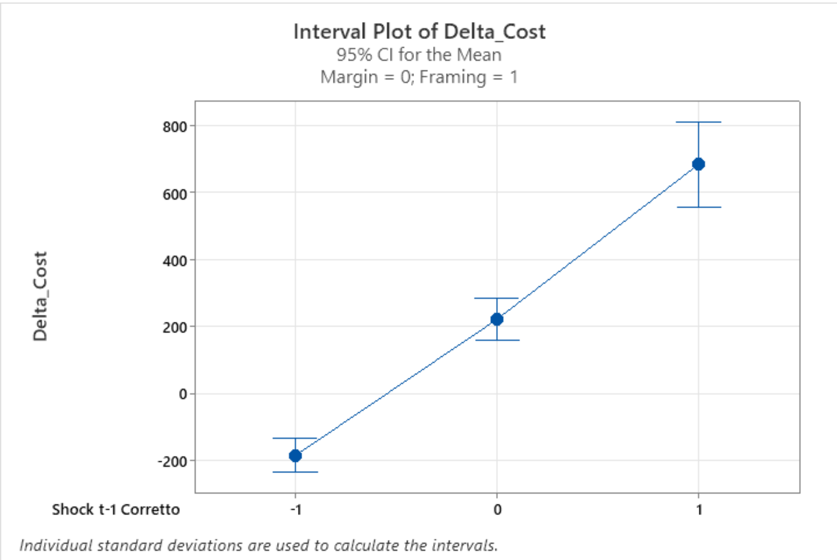
| Source | Sum Sq. | d.f. | Mean Sq. | F | Prob>F |
|---|---|---|---|---|---|
| Margin | 2678.2 | 1 | 2678.2 | 10.78 | 0.001 |
| Framing | 1289.3 | 1 | 1289.3 | 5.19 | 0.0228 |
| D t-1 | 72148.9 | 1 | 72148.9 | 290.28 | 0 |
| Margin:Framing | 1433.2 | 1 | 1433.2 | 5.77 | 0.0164 |
| Margin:D t-1 | 35.3 | 1 | 35.3 | 0.14 | 0.7063 |
| Framing:D t-1 | 173.3 | 1 | 173.3 | 0.7 | 0.4038 |
| Margin:Framing:D t-1 | 1415.6 | 1 | 1415.6 | 5.7 | 0.017 |
| Error | 1419733.1 | 5712 | 248.6 | | |
| Total | 1545172.6 | 5719 | | | |

Constrained (Type III) sums of squares.

Figure 13: Session 1 Delta Costs vs Shcok, Framing and Margin




Interval Plot of Delta_Cost
95% CI for the Mean
Margin = 1; Framing = 1
Delta_Cost
1250
1000
750
500
250
0
-250
-500
Framing
Shock t-1 Corretto
1
1
1
-1
0
1
Individual standard deviations are used to calculate the intervals.


Interval Plot of Delta_Cost
95% CI for the Mean
Margin = 0; Framing = 1
Delta_Cost
800
600
400
200
0
-200
Shock t-1 Corretto
-1
0
1
Individual standard deviations are used to calculate the intervals.

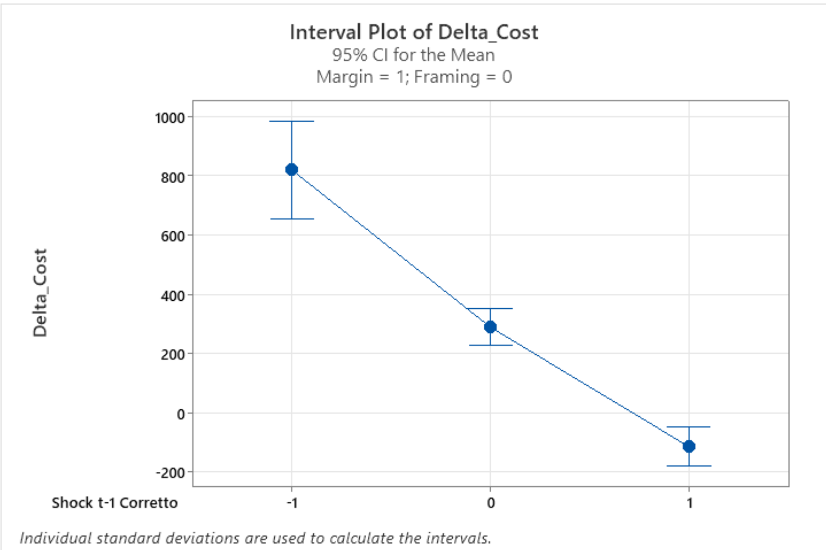
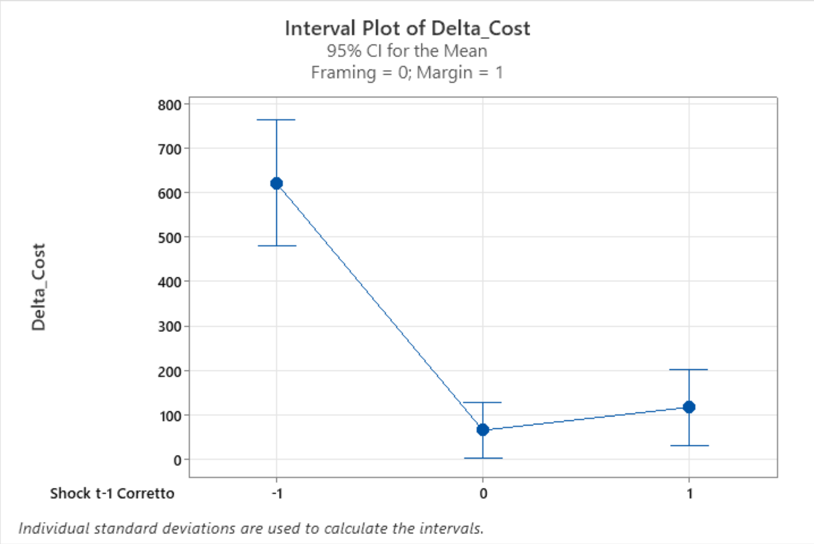
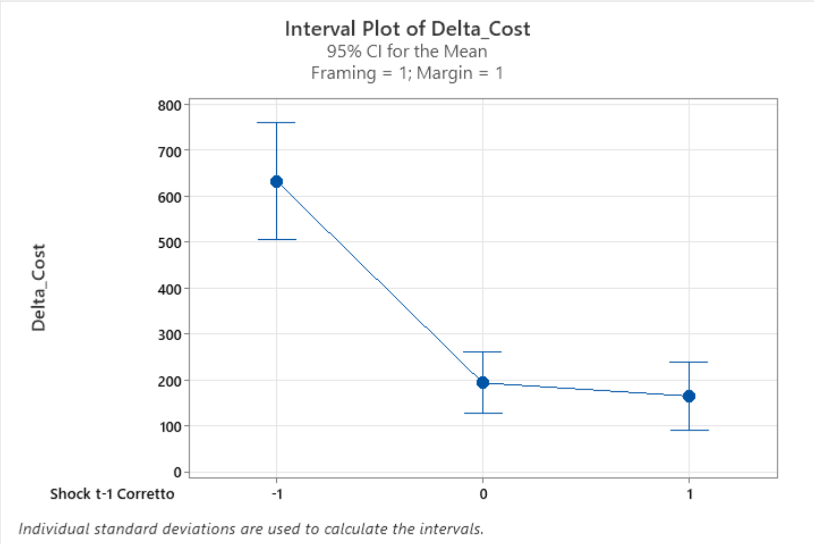
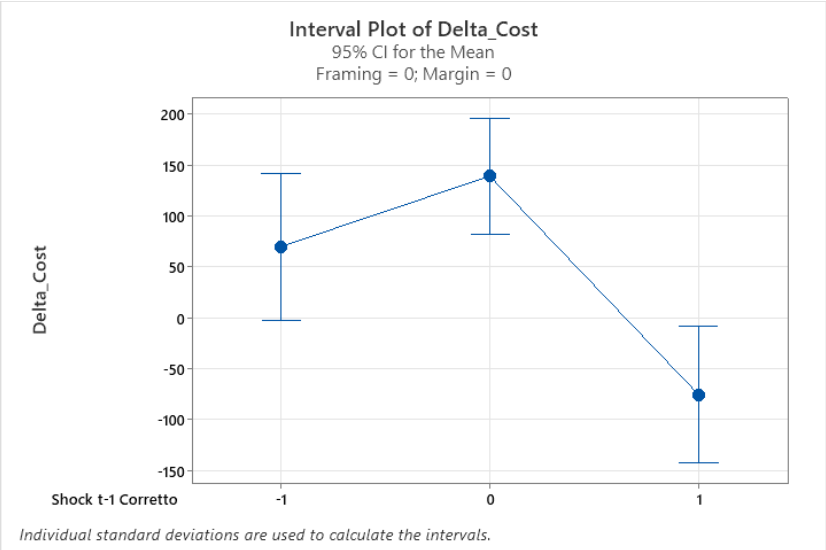

Interval Plot of Delta_Cost
95% CI for the Mean
Margin = 1; Framing = 0
Delta_Cost
1000
800
600
400
200
0
-200
Shock t-1 Corretto
-1
0
1
Individual standard deviations are used to calculate the intervals.

Figure 14: Session 2 Delta Costs vs Shcok, Framing and Margin



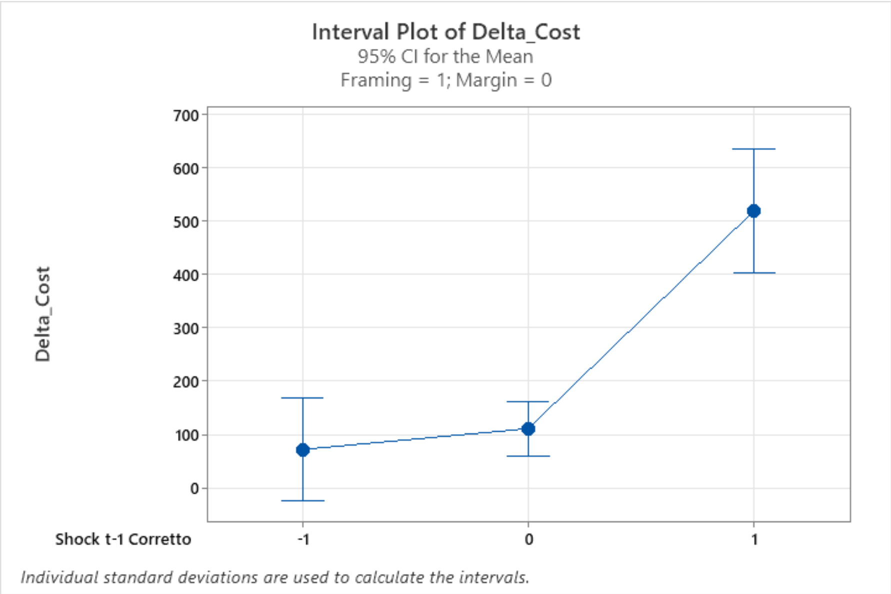
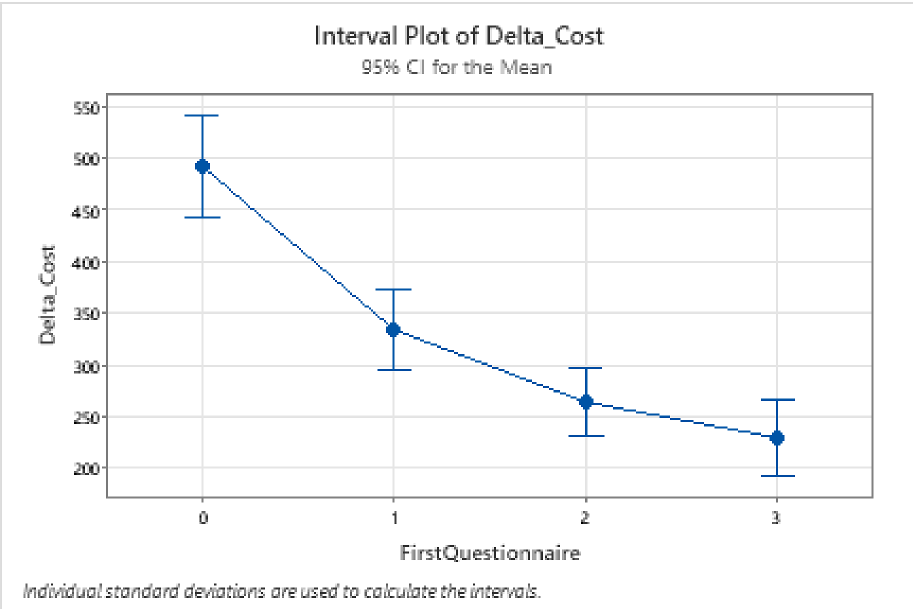



Figure 15: Delta cost vs CRT Scores



Scripts employed for the ANOVA on both Forecast and Margin Variables:

```
clear all, close all, clc

D=readmatrix("S1Dt5.xlsx");

Forecast_1=D(:,3);
Framing_1=D(:,6);
Margin_1=D(:,7);
Shock_Corretto1=D(:,8);
Purchase1=D(:,9);
Dt1=D(:,10);
Dt5=D(:,11);
Delta1=D(:,12);
Delta_C1=D(:,13);


d=anovan(Delta_C1,{Margin_1,Framing_1,Shock_Corretto1},'model','full','varnames',{'Marg
in','Framing','Shock t-1 Corretto'});

e=anovan(Forecast_1,{Margin_1,Framing_1},'model','full','varnames',{'Margin','Framing'}
);

E=readmatrix("S2Dt5.xlsx");

Forecast_2=E(:,3);
Framing_2=E(:,6);
Margin_2=E(:,7);
Shock_Corretto2=E(:,8);
Purchase2=E(:,9);
Dt_1=E(:,10);
Dt_5=E(:,11);
Delta2=E(:,12);
Delta_C2=E(:,13);

f=anovan(Delta2,{Margin_2,Framing_2,Shock_Corretto2},'model','full','varnames',{'Margin
', 'Framing', 'Shock t-1 Corretto'});
f=anovan(Forecast_2,{Margin_2,Framing_2},'model','full','varnames',{'Margin','Framing'}
);
group =[Margin_2.*Framing_2 Shock_Corretto2];
interactionplot(Forecast_2,group,"varnames",{'Margin*Framing', 'Shock t-1 corretto'});
F=readmatrix("s1+s2Dt5.xlsx");

Sessione=F(:,2);
Forecast_=F(:,3);
Framing_=F(:,6);
Margin_=F(:,7);
Shock_Corretto=F(:,8);
Purchase=F(:,9);
Dt1_=F(:,10);
Dt5_=F(:,11);
Shock_1=F(:,12);
Delta=F(:,13);
Delta_Purchase=F(:,14);
Delta_Cost=F(:,15);
h=anovan(Forecast_,{Sessione,Margin_,Framing_},'model','full','varnames',{'Session','Ma
rgin','Framing'});
```

```
g=anovan(Purchase,{Sessione,Margin_,Framing_},'model','full','varnames',{'Session','Mar
gin','Framing'});
```

# Figure Index:

Figure 1: Low margin case from Schweitzer and cachon (2000), “Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence”. *Management Science*, 404-420........12

Figure 2: High margin case from Schweitzer and cachon (2000), “Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence”. *Management Science*, 404-420........13

Figure 3: graphs plotted using the 2019 polytechnic of Turin’s experimental data........14

Figure 4:Prospect Theory’s Outcome values asymmetry. Image taken from Psychology.com http://psychology.iresearchnet.com/papers/prospect-theory/........17

Figure 5: Pearson correlation between the variable Purchase and Forecast and Figure 5 showing the trend line........24

Figure 6: Analysis of Variance (ANOVA) results for the variable Forecast........32

Figure 7: Analysis of Variance (ANOVA) results for the variable Purchase........32

Figure 8: Two-Variance test results........33

Figure 9: Descriptive statistic for the 2 categories of data collected........33

Figure 10: Session 1 and 2 MEANS of the Forecast, Realized Demand, and Newsvendor’s Optimal forecast........34

Figure 11: Purchase’s mean overtime........34

Figure 12: Session 1 plot of the means of Forecast and NV_Forecast for each order repetition........35

Figure 13: Session 1 plot of the means of Forecast and NV_Forecast for each order repetition........35

Figure 14: Two-variance test on Session 1 NV_Forecast and Forecast variables........36

Figure 15: Paired T-test on Session 1 NV_Forecast and Forecast variables........36

Figure 16: ANOVA for the variable Forecast........37

Figure 17 : ANOVA for the variable Purchase........40

Figure 18 A), B), C), D): The Interval Plots for Forecast and NV_Forecast with 95% Confidence interval........41

Figure 19 A) and B): Interval Plot for Purchase and NVM_Purchase for Low Margin settings........42

Figure 20 A) and B): Interval Plot for Purchase and NVM_Purchase for Low Margin setting........42

Figure 21: Exploratory ANOVA results for Forecast with Margin, Framing, Session and Shock $_{t-1}$ .......43

Figure 22: Exploratory ANOVA results for Forecast with Margin, Framing, Session and Shock $_{t-1}$ Corretto..........................................................................................................................................44

Figure 23 A) AND B): Interval Plots for Forecast and Newsvendor Forecast with a low margin treatment in relation to Shock$_{t-1}$ .........................................................................................................46

Figure 24 A) AND B): Interval Plots for Forecast and Newsvendor Forecast with a high margin treatment in relation to Shock$_{t-1}$..........................................................................................................47

Figure 25 A) AND B): Interval Plots for Purchase and Newsvendor Purchase with a low margin treatment in relation to Shock$_{t-1}$ .........................................................................................................48

Figure 26 A) AND B): Interval Plots for Purchase and Newsvendor Purchase with a high margin treatment in relation to Shock$_{t-1}$..........................................................................................................49

Figure 27: Delta_Forecast_NV with respect to Shock t-1 ..........................................................51

Figure 28: ANOVA results for Forecast with respect to treatments containing Session, Margin, Framing and D$_{t-1}$........................................................................................................................52

Figure 29: ANOVA results for Purchase with respect to treatments containing Session, Margin, Framing and D$_{t-1}$........................................................................................................................52

Figure 30: ANOVA results for Forecast with respect to treatments containing Session, Margin, Framing and D$_{t-5}$........................................................................................................................52

Figure 31: ANOVA results for Purchase with respect to treatments containing Session, Margin, Framing and D$_{t-5}$........................................................................................................................54

Figure 32 A), B), C), D): Histograms of Various typologies of cost computed.....................................55

Figure 33: Vcosts Interval Plot for Negative and Positive framings.......................................................55

Figure 34: Delta_Costs' Interval Plot with respect to Shock$_{t-1}$ ..........................................................56

Figure 35 A) and B) : Delta Costs' interval plots with High Margin treatments...................................57

Figure 36 A) and B): Delta Costs' Interval Plots with Low Margin treatments.....................................58

# Bibliography


AJ A. Bostian, C. A. (2008). The Newsvendor “Pull-to-Center Effect”: *Manufacturing & Service Operations Management*.

Akash, M. H. (2019). *Understanding Scalp EEG in response to Newsvendor Decision-Making.* University of Texas at Arlington.

Benzion, C. Y. (2008). Decision-Making and the Newsvendor Problem: an Experimental Study. *The Jurnal Of the Operational Research Society*, 1281-1287.

Bhavani Shanker Uppari, S. H. (2019). Modeling Newsvendor Behavior: A Prospect Theory Approach. *Manufacturing & Service Operations Management, 21*(3), 481-500.

Bostian, Holt , & Smith . (2008). Newsvendor "Pull-to-Center" Effect: Adapting Learning in a Laboratory Experiment . *Manufacturing & Service Operations Management* , 590-608.

Cachon, S. (2000). Decision Bias in the Newsvendor Problem with a Known Demand Distribution: Experimental Evidence. *Management Science*, 404-420.

Chirag Surti, A. C. (2020). The newsvendor problem: The role of prospect theory and feedback. *European Journal of Operational Research*, 251-261.

Chirag Surti, A. C. (2020). The newsvendor problem: The role of prospect theory and feedback. *European Journal of Operational Research*, 251-261.

Cramer, C., & Ho, T.-H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 827-874.

Croson , Ren, & Croson. (2009). How to manage an overconfident newsvendor. Retrieved from esearchgate.net/publication/228434339_How_to_manage_an_overconfident_newsvendor

Croson D., C. R. (2017). The overconfident newsvendor . *Journal of the Operational Research Society*, 496-506.

D.F. Pyke, R. P. (1998). *Inventory Management and Production.* NY: Wiley.

D’Urso D, C. F. (2021). Measuring How Decision Support Systems Improve Newsvendors’ Performance: The Subjects’ Version. *Sustainability*, 10251.

de Véricourt, K. J. (2013). Sex, risk and the newsvendor. *Journal of Operations Management*, 86-92.

Dunning, D., & Kruger, J. (2000). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology*, 1121-1134.

D'Urso, D. M. (2017). A behavioural analysis of the newsvendor game: Anchoring and. *Computers & Industrial Engineering, 111*, 552-562.

Feld, J. &. (2017). Estimating the Relationship between Skill and Overconfidence. *journal of Behavioral and Experimental Economics*, 18-24.

Fisher, A. R. (1996). Reducing the cost of demand uncertainty. *Operations Research* , 87-99.

Frederick. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19*, 25-42.

Ho, Lim, & Cui. (2010). Reference Dependence in Multilocation Newsvendor models: a structural analysis. *Management Sciences, 56*(11), 1891-1910.

Hutchinson, A. (2000). Knowledge calibration: What Consumers Know and What They Think They Know. *Journal of Consumer Research*, 123-156.

Jammernegg W., K. P. (2021). Heterogeneity, asymmetry and applicability of behavioral newsvendor. *European Journal of Operational Research*, 638-646.

Kahneman. (2011). *Thinking, Fast and Slow.* Penguin Non Fiction Press.

Kahneman D., T. A. (1981). The Framing of Decisions and the Psychology of Choice. *Science*.

Kahneman, D. (1992). Reference points,anchors, norms and mixed feelings. *Organization Behavior Human Decision Process* , 296-312.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263-291.

Katok, B. (2008). Learning by Doing in the Newsvendor Problem: A Laboratory Investigation of the Role of Experience and Feedback. *Manufacturing & Service Operations Management* , 519-538.

Khouja, M. (1999). The single-period (news-vendor) problem: literature review. *The international Journal of Management Sciences*, 537-553.

Kremer, M. M. (2010). Do random errors explain newsvendor behavior? *Manufacturing and Service Operations Management, 12*(4), 673-681.

Lars St»hle, S. W. (1990). Multivariate analysis of variance (MANOVA). *Chemometrics and Intelligent Laboratory Systems*, 127-141.

Li, Chen Gotao, & Chen J. (2019). Individual and Cultural Differences in the Newsvendor Decision Making. *International Journal of Operations & Production Management* , 164-186.

Lurie, N., & Swaminathan, J. (2009). Is timely information always better? The effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes, 108*(2), 315-329.

Marschak, K. J. (1951). Optimal Inventory Policy. *Econometrica, 19*(3), 250--272. Retrieved from http://www.jstor.org/stable/1906813

Meng Li , Petruzzi , & Zhang . (2016). Overconfident Competing Newsvendors. *Management Science*, 2637-2646.

Monk, E. a. (2009). *Concepts in Enterprise Resource Planning* (3rd ed.). Boston: Course Technology Cengage Learning.

Moore D., H. P. (2008). The Trouble with Overconfidence. *Psychological Review, 115*(2), 502-517.

Moritz, B., Hill, A., & Donohue, K. (2009). *Cognition and Individual Differences in the Newsvendor Problem:.* WORKING PAPER .

Moritz, Hill, & Donohue. (2013). Individual differences in the newsvendor problem: Behavior and cognitive reflection. *Journal of Operations Management*, 72-85.

Nagarajan, S. (2014). Prospect Theory and the Newsvendor Problem . *Management Science*, 1057-1062.

O'Keefe T. Carlson, J. (1969). Buffer stocks and reaction coefficients:. *Review of Economic Studies, 36*, 467–484.

Petruzzi, M. D. (1999). Pricing and the Newsvendor Problem: A Review with Extensions. *Operations Research*, 183-194.

Porteus, E. (2008). The Newsvendor Problem. In D. L. Chhajed, *Building Intuition* (pp. 115-134). Boston, MA: Springer.

Ren, C. (2013). Overconfidence in Newsvendor Orders: An Experimental Study. *Management Sciences*, 2502-2517.

Ren, Y., & Croson, R. (2013). Overconfidence in newsvendor orders: An experimental study. *Management Science, 59*(11), 2502–2517.

Ross, A., & Willson, V. (2017). *Basic and Advanced Statistical Tests.* Brill | Sense.

Schatz, D. A. (2017, June 12). The Three Faces of Overconfidence. *Social and Personality Psychology Compass*, 1-12.

Scheffé, & Henry. (1959). *The Analysis of Variance.* New York: Wiley.

Sharma, & Nandi. (2018). Review of Behavioral Decision Making in the Newsvendor Problem. *OPERATIONS AND SUPPLY CHAIN MANAGEMENT, 11*(4), 200-2013.

Shefrin, H. (2018). *Behavioral Corporate Finance* (International Edition ed.). McGraw-Hill Ed.

Su, X. (2008). Bounded Rationality in Newsvendor Models. *Manufacturing & Service Operations Management*, 566-589.

Truong, W. (2020). Analysis of correlation in Neural Responses across multiple subject or trial during decision-making for Newsvendor problem. *PUB. IEEE 20th BIBE CONFERENCE.*

Tversky A., K. D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 1124-1131.

Wei, C. (2021). Supply Chain Replenishment decision for Newsvendor Products with Multiple Periods and a Short Life Cycle. *Sustainability* .

*Wikipedia* . (n.d.). Retrieved from https://en.wikipedia.org/wiki/Prospect_theory

Yamini, S. (2020). Behavioral perspective of newsvendor ordering decisions: review, analysis and insights. *Management Decision, 59*(2), 240-257.

Zenker. (2012). Review of Thinking, Fast and Slow by Daniel Kahneman. *Inquiry Critical Thinking Across the Disciplines*.

Zhang, G. e. (2014). Linking brain electrical signals elicited by current outcomes with future risk decision-making. *FRONTIERS IN BEHAVIORAL NEUROSCIENCE, 8*(2), 84.