Analyzing learning effects in the newsvendor model by probabilistic methods

BY Jonas Andersson, Kurt Jörnsten, Jostein Lillestøl and Jan Ubøe

DISCUSSION PAPER

NHH

Institutt for foretaksøkonomi
Department of Business and Management Science

FOR 13/2019
ISSN: 1500-4066
October 2019

Analyzing learning effects in the newsvendor model by probabilistic methods

Jonas Andersson, Kurt Jörnsten, Jostein Lillestøl, and Jan Ubøe*
Norwegian School of Economics,
Helleveien 30, N-5045 Bergen, Norway.

October 11, 2019

Abstract

In this paper, we use probabilistic methods to analyze learning effects in a behavioral experiment on the newsvendor model. We argue why we should believe that suggested orders follow a multinomial logit distribution, and use the single parameter in that model to extract information on learning effects. We revisit the data, analyzed previously by Bolton et al. (2012), and show that our model predicts the pull-to-center effect in these experimental data very well.

Keywords: Behavioral OR, experimental economics, bounded rationality, probabilistic methods, pull-to-center effect.

*Corresponding author: jan.uboe@nhh.no, phone: 004755959978, fax: 004755959650

1 Introduction

Discrete choice models emerged in the 1970's, with the pioneering work of Daniel McFadden on random utility maximization; see McFadden (1974) and Train (2003). The theory has been applied with success within several different fields in economics, and has obvious relevance to newsvendor behavior. Retailers often base their decisions on partial and incomplete information leading to a certain type of randomness in ordering. Managers should seek to understand the nature of this randomness and use their knowledge to improve performance. It is hence of some surprise that this approach is largely ignored in the literature on the newsvendor problem. A notable exception is Su (2008).

In the single-period newsvendor model, a retailer wishes to order a quantity q from a manufacturer. Demand D is a random variable, and the retailer selects an order quantity q maximizing his expected profit. When the distribution of D is known, the problem of determining an optimal quantity is easily solved. The basic problem is very simple, but it appears to have endless variations. There is now a very large body of literature on such problems; for further reading, refer to the reviews by Cachón (2003) and Qin et al. (2011) and the numerous references therein.

In their seminal empirical paper Schweitzer and Cachón (2000) observe what they call the "pull-to-center" effect, i.e., that the agents order too little in high profit cases and too much when the profit is low. They suggest that this effect can be explained by anchoring and adjustments to previously observed demand. The paper inspired several other researchers to do similar experiments, we mention Bostian et al. (2008), Kremer et a. (2010), Wachtel and Dexter (2010), Bolton et al. (2012) and also Rudi and Drake (2014). All these papers discuss similar types of experiments and pursue different kinds of explanations for the "pull-to-center" effect. While there is probably some truth in all these arguments, we suggest that there may be a much simpler explanation. We suggest that this effect is exactly what one would expect when agents choose orders via a discrete choice distribution. This kind of explanation was first suggested by Su (2008).

Important work of researchers like Daniel Kahneman and Amos Tversky have put psychology in the forefront of decision theory. While it is certainly true that normal probability theory breaks down in cases with nested structures, transparent choice sets like the ones we discuss in this paper are not expected to suffer from such weaknesses. When the choice set is a normal probability space, a multinomial logit distribution is expected in most cases. The reason behind this is basically the Fisher-Tippet-Gnedenko theorem which describes the limiting distribution for the maximum of a sequence of iid random variables. These models are robust in the sense that deviations from the ideal assumptions may not significantly change the statistical choice distribution. The models are non-linear, but the effect of the non-linearity often disappears under aggregation. This means that we expect these models to work even in cases with relatively small choice sets where the choice distributions vary considerably across agents.

Su (2008) suggests that agents can choose any order, and that orders are selected from a probability distribution based on, e.g., random utility maximization. If they do, a good model fit should be obtained using the multinomial logit model. He proves that the “pull-to-center” effect can be derived analytically from the multinomial logit model. Ubøe et al. (2017) takes this argumentation one step further in that they define bounded rationality as follows: Agents are boundedly rational if and only if states with less total cost are more probable. Drawing on theory from Smith (1978) and Erlander (2010) they arrive at the same conclusion as Su (2008), i.e., that orders are chosen from a multinomial logit model.

In this paper, we offer a new analysis of the data from the paper Bolton et. al (2012). Bolton et. al (2012) examine a laboratory experiment where the agents suggest order quantities in a newsvendor context. The parameters in the newsvendor problem are fixed, and the agents level of information progress over time. By analysis of variance (ANOVA), we analyze changes in the cost sensitivity parameter when subjects obtain new information. Comparison is made with respect to such changes between subjects with different experience and level of training. In doing so we use tests that control for mass significance. The method that we use enables us to show that:

- Learning effects in the newsvendor model can be measured by the cost sensitivity parameter in a single-parameter multinomial logit model.
- The costs sensitivity parameter, studied over the course of the experiment, can be used to obtain precise predictions of average orders of the subjects.
- The pull-to-center effect observed in the experiment is captured by the proposed method.

The paper is organized as follows: In Section 2 we discuss our definition of bounded rationality and the rationale for using models of this kind. In Section 3, we give a brief summary of the experimental design from Bolton et. al (2012). In Section 4, we describe how the parameter in our model can be used to extract information on learning effects. In Section 5, we use our model to predict the average order quantity in each phase of the experiment, and find that our theoretical model predicts the observed average order well. In Section 6, we offer some concluding remarks.

2 Bounded rationality and probabilistic models

The main message from discrete choice theory can be summarized as follows: When members of a group of fairly similar agents pick the best option from a choice set S which is a normal probability space, we expect that the aggregate distribution satisfies

$$\text{Probability of choosing item } i = \frac{e^{v_i}}{\sum_{j\in S} e^{v_j}}, \tag{1}$$

where v_i is the expected utility of choosing item i. This can be explained by the Fisher-Tippet-Gnedenko theorem which is an analogue to the central limit theorem for the maximum distribution, see, e.g., Andersson and Ubøe (2012). The distribution given by (1) is unusually robust in that it can be derived from several very different lines of reasoning. It can be derived from random utility theory, Manski (1977). It can be derived from maximum entropy, e.g., Wilson (1967), Anas (1983), Erlander and Stewart (1990). It is the solution of the maximum utility problem, Erlander and Stewart (1990). The rational inattention approach, Matejka and McKay (2015), leads to the same model under a uniform prior.

The assumption that S is a normal probability space is in fact important. Some of the most interesting applications of discrete choice theory occur in contexts where this condition is violated. As the problem we consider in this paper is safely within the context of normal probability theory, there is no need to draw on these more advanced constructions.

Erlander (2010) offers a particularly nice angle of attack. Assume that there is a cost (negative utility) c_i associated with choosing item i. Seek any probability distribution on S with the property that if an allocation of choices leads to a larger aggregate cost than another, it will be less probable. If this monotonicity principle holds for any allocation of arbitrary length, there exist a non-negative constant β such that

$$\text{Probability of choosing item } i = \frac{e^{-\beta c_i}}{\sum_{j \in S} e^{-\beta c_j}}. \tag{2}$$

These distributions are hence the only probability distributions compatible with the monotonicity property stated above, and they are called probabilistically cost efficient. The constant β can be interpreted as the sensitivity to cost. If β is very small, costs do not matter and the resulting distribution will be very close to a uniform distribution. At the other extreme a large β will imply that only the objects with the smallest cost, possibly more than one, will be chosen.

2.1 Applications to the newsvendor model

In this paper, we will examine experimental data for a single-period newsvendor model. This model is specified as follows.

$$\begin{aligned}
W &= \text{ wholesale price per unit (fixed)} \\
q &= \text{ order quantity (rate chosen by the retailer)} \\
R &= \text{ retail price per unit (fixed)} \\
D &= \text{ demand (random rate)} \\
S &= \text{ salvage price per unit (fixed)}
\end{aligned}$$

A retailer is trading a commodity and orders q units from a manufacturer. He hopes to sell enough of these units to make a profit. We assume that the manufacturer offers a wholesale price W, and that the retail price R is exogenously given. Unsold items can be salvaged at the exogenously given salvage value S. The retailer's profit is denoted by $\Pi(q)$, and it is easily shown that

$$\Pi(q) = (R - S)\min[D, q] - (W - S)q. \tag{3}$$

A straightforward computation shows that the retailer maximizes expected profit when

$$P(D \leq q) = \frac{R - W}{R - S} \;\; \Rightarrow \;\; q = F_D^{-1}\left[\frac{R - W}{R - S}\right], \tag{4}$$

where F_D denotes the cumulative distribution of D. As pointed out by Su (2008) and Ubøe et al (2017) a newsvendor will not always order the optimal q given by (4). The newsvendor may instead select several different order quantities, and the inclination to select any particular q is defined in terms of a probability distribution. If the newsvendor is boundedly rational, he should have an inclination towards optimal choices. A suboptimal choice is associated with a cost, which is the loss in expected profit in comparison with the optimal choice. We can hence specify a cost function

$$c(q) = \Pi(q_{\text{opt}}) - \Pi(q). \tag{5}$$

The only probabilistically cost efficient distributions compatible with this specification of costs are given by

$$\text{Probability of ordering } q \text{ units} = \frac{e^{-\beta c(q)}}{\sum_{j=0}^{d_{\max}} e^{-\beta c(j)}}. \tag{6}$$

To examine how the experimental data develop over time, we fit (in the sense of maximum likelihood) a parameter β_t to the observations recorded at time t. If learning progresses over time, we expect to see that the β-parameter increases, reflecting that the agents more frequently pick orders with smaller costs.

2.2 Robustness with respect to agent heterogeneity

The mathematical reasoning above requires that agents within each group are similar. In real life experiments, we must expect some variation of sensitivity across a group. If we blend agents with different sensitivity to costs (i.e. with different β parameters), it is easy to see that the exact choice distribution is not given by (6). Even in cases where the β-parameter varies considerably over the group, this blending effect may be surprisingly small. To simplify notation we define $f(q, \beta)$ to be the right hand side of (6). Assuming that agent preferences across the group is specified by a density ψ in β, the true mixture distribution $\overline{f}(q)$ is given by

$$\overline{f}(q) = \int f(q, \beta)\psi(\beta)d\beta. \tag{7}$$

To examine the effect of a non-linear aggregation, we may compare $\overline{f}(q)$ with a multinomial logit distribution $f(q, \overline{\beta})$ where $\overline{\beta} = \int \beta\psi(\beta)d\beta$ is the expectation of the β parameter across the group. In general these distributions are difficult to compare, but in the experiments carried out in this paper, we are in the position where everything except β is known in the model. To be precise, the experiment in Bolton et. al (2012) uses uniform demand with the specific values

$$R = 12, \qquad S = 0, \qquad W = 3, \qquad d_{\text{min}} = 0, \qquad d_{\text{max}} = 100.$$

which implies $q_{\text{opt}} = 75$. The typical value reported in the experiment is $\beta = 0.01$. Figure 1 compares the exact mixture distribution with the logit distribution corresponding to the average β in two cases:

i) β uniform over the interval $[0.005, 0.015]$, i.e. sensitivity varies with a factor 3.
ii) β uniform over the interval $[0.0002, 0.0198]$, i.e. sensitivity varies with a factor 100.

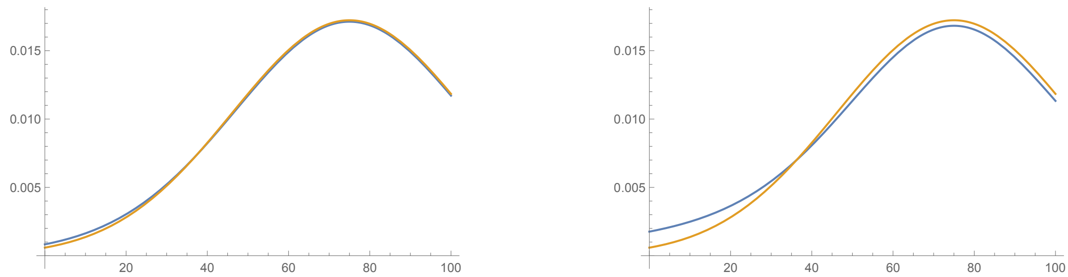

Figure 1: Comparisons of $\overline{f}(q)$ and $f(q, \overline{\beta})$ in two cases.

We see that the aggregation bias is hardly noticeable in both cases. In the extreme case where sensitivity varies by a factor 100 across agents, the effect is noticeable but still relatively small. Bolton et. al (2012) only make use of about 25 agents in each group, and we do not expect that variation of sensitivity across agents within the groups is an issue. As we mentioned in the introduction, we can expect that the model performs well even in cases where the sensitivities to costs vary considerably across members within each group.

3 The experimental design

The data set we are analyzing in this paper is the same as the one used in Bolton et. al (2012), and a complete description of the experiment can be found in that paper. We will here briefly summarize the main features of the experiments. The agents are divided into 3 groups, freshmen students, graduate students, managers. Each of these groups were split into two equally large subgroups. Within each group, one subgroup watched a one-hour video on the newsvendor problem. The other subgroup received no such training. This leads us to consider six different groups:

- Freshmen, basic
- Freshmen, trained
- Graduates, basic
- Graduates, trained
- Managers, basic
- Managers, trained

The experiment involves a 100 period newsvendor game, and involves an initial phase and 3 consecutive phases. In the initial phase, the subjects read a two-page briefing, including a graph showing the demand of the previous 50 periods.

- In the first phase, the subjects place 40 orders, receiving feedback (on their earnings) after each order.
- The subjects then receive a handout stating that demand in uniformly distributed between 1 and 100 and is uncorrelated.
- In the second phase, the subjects place 40 orders, receiving feedback after each order.
- The subjects then receive a handout with a graph showing how expected profit depends on the order quantity.
- In the third phase, the subjects place 20 orders, receiving feedback after each order.

The experiment is described and analyzed in detail by Bolton et. al (2012). For our purposes, it is sufficient to say that the experiment is carried out using state of the art methodology, and we see no reason to question the quality of the data.

4 Using probabilistic models to extract information on learning effects

In this section we will use the development of the sensitivity to cost parameter β during the course of the experiment to study how the level of information develops over time. We hypothesize two potential sources of changes in β:

- Increases due to information given to the subjects at time 40 and 80, when first the distribution of demand is given and then a graph of expected profit as a function of order quantity. These changes should be in the form of jumps in β at time 40 and 80.
- Increases due to continuous learning of the demand distribution. This would be observed as a systematic trend in β over time. We assume that this trend is linear.

Each β is determined as the maximum likelihood value, based on the data and the probabilistic model described above. They turned out as in Figure 2 for each of the six groups:

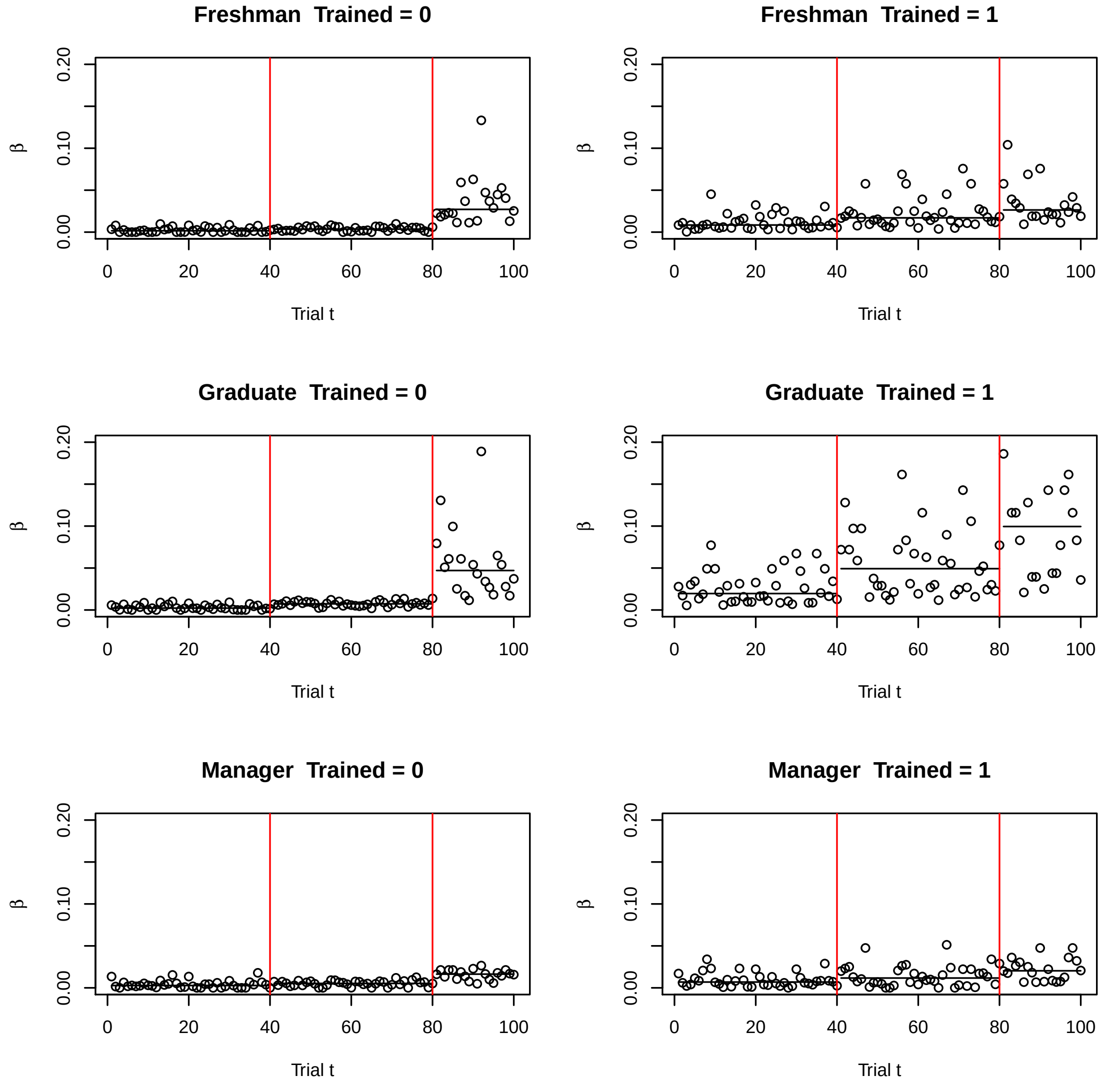

Figure 2: Observed β-values plotted against time for each of six groups. Vertical lines divide the three time periods $[1, 40]$,$[41, 80]$ and $[81, 100]$ and horizontal lines are the median of the period.

Descriptive statistics (mean and median) are given in Table 1.
For several groups, we observe upward changes in level from one period to the next, in particular from period 2 to period 3. The managers, trained or not, seemingly learned the least, and the graduate students learned the most. Initially the trained graduates started out in period 1 at a slightly higher β-level than the other groups, but still learns. However, the question is: Are the changes statistically significant?

It is also notable that the β-variation is increasing from period to period, and differs between the groups. The three trained groups have larger variation than the corresponding untrained

Group	β	Period 1	Period 2	Period 3
Freshman	mean	0.0026	0.0038	0.0362
Not trained	median	0.0016	0.0036	0.0271
Freshman	mean	0.0117	0.0223	0.0347
Trained	median	0.0085	0.0170	0.0264
Graduate	mean	0.0034	0.0076	0.0551
Not trained	median	0.0022	0.0073	0.04704
Graduate	mean	0.0265	0.0595	0.2409
Trained	median	0.0196	0.0492	0.0995
Manager	mean	0.0041	0.0053	0.0158
Not trained	median	0.0028	0.0053	0.0163
Manager	mean	0.0094	0.0142	0.0218
Trained	median	0.0069	0.0117	0.0203

Table 1: Descriptive statistics for β: Mean and median

groups. For the untrained groups, the major increase in variation comes in the third period. Largest variation all along comes for the Graduate-Trained group. We have no good explanation for this time pattern across the individuals of each group. [1]

A number of different modes of analysis are available for exploratory analysis and judging signifcance. They range from pairwise comparisons to joint analysis based on factorial models. Common standard models are usually based on strict assumptions, like independence, homogeneity of variances and normality. In the current data we have to recognize the possibility of trend and serial correlation within periods, heterogeneity of variances between periods and non-normality with extreme outliers. A preliminary exploratory study is based on a linear regression model, formulated as

$$\beta_t = a_0 + a_1 \cdot t + a_2 \cdot D_{2,t} + a_3 \cdot D_{3,t} + \epsilon_t \tag{8}$$

where $t = 1, ..., 100$, $D_{2,t} = 1$ if $t \in [41, 80]$ and zero otherwise and $D_{3,t} = 1$ if $t \in [81, 100]$ and zero otherwise. Within this model $a_1 > 0$ corresponds to positive trend (continuous learning) and $a_i > 0$, for $i = 2, 3$ to positive shifts, i.e. learning from new information given prior to new phase. The possibility of serial correlation may be accounted for by assuming ϵ_t auto-correlated. It turned out that the standard t-test for no trend ($a_1 = 0$, i.e., no continuous learning) against the alternative ($a_1 > 0$) gave non-significant a_1 for all six groups. Moreover, a likelihood ratio test for no auto-correlation turned out not significant as well. One may also test the hypothesis of no shift within the model as is. Alternatively, one could re-estimate the regression with no trend and auto-correlation. Then the regression model with only 0-1 variables is just a representation of a one-factor ANOVA model. It is preferable to switch to this framework, as it offers a wider range of opportunities (i.e. multiple comparisons), and testing with weaker assumptions (i.e. non-parametric tests).

[1]The estimation of β provides opportunity to calculate (asymptotic) standard errors to each β, reflecting the variation between the individuals. They show a similar increasing pattern.

With the one-factor analysis of variance (ANOVA) model, no learning corresponds to constant mean through the three periods. The ordinary F-test gave P-values less than 5% for all six groups. Most often they were tiny, except for the Graduate-Trained group (P=0.02). Based on these findings we may reject the hypothesis of no learning. Wondering whether the departure from strict assumptions may twist the P-calculations, one may instead use the non-parametric Kruskal-Wallis test. This gave tiny P-values for all groups, also for the Graduate-Trained. Our understanding of this, is that the most serious violation of standard assumption is the outliers in the the last period, which will give some undeserved favor to the no-significance conclusion.

The performed tests rejected no learning effects, but do not tell which of the differences are significant. Here one may make three comparisons: period 2 versus period 1, period 3 versus period 2, and period 3 versus period 1. At the outset one cannot rule out that the first two are statistically not significant, but the latter is. A common post hoc test of pairwise differences, extending the two-sample t-test, is the TukeyHSD-test (”Honest Significant Differences”), controlling the joint risk of false significance. Here we report in Table 2 the observed differences in period means and the 95% joint confidence intervals of (expected) differences, which may be checked whether they include zero (no learning effect) or not. Moreover, we report P-values, similar to the two-sample t-test, but adjusted for multiple comparison. A similar non-parametric alternative is based on the two-sample Wilcoxon test. The corresponding adjusted P-values for this test are reported as well. The criteria for significance is taken as 5 %.

The differences from period 1 to period 2 (tabled as 2 - 1) are smaller throughout, than the differences from period 2 to period 3, which add up the the differences from period 1 to period 3 (tabled as 3 - 1). The latter represents the joint effect of learning from both packages of information given, i.e. first the demand distribution and then how the expected profit depends on the order quantity. The learning effects from the first package of information (2 - 1) are not statistically significant judged by the parametric test (T), except for the Freshman-Trained group. However, judged by the non-parametric test (W), the learning effects were significant for all groups, except barely non-significant for the Manager-Trained group (P=0.06).The learning effect from the second package of information (3 - 2) are statistically significant for all groups, regardless of test method, parametric (T) or non-parametric (W). Then the joint learning effect (3 - 1) will be statistically significant in all groups as well, regardless of test method.

The discrepancy between the parametric test (T) and the non-parametric test (W) may be due to the larger variation in the third period, causing an inflated joint variance estimate. However, the algorithm claims to account for heterogeneity. One should be aware of that the two tests are designed for slightly different conceptions of level. While the T-method is linked to the mean, the W-method may be linked to the median, with null hypothesis equal medians, and alternative upward shift. As we have seen, the mean and the median of β turned out different in our data.

Group	Periods	Difference	Lower	Upper	P-adjust-T	P-adjust-W
Freshman	2 - 1	0.0013	-0.0054	0.0080	0.4472	0.0065
Not trained	3 - 2	0.0324	0.0242	0.0406	0.0000	0.0000
	3 - 1	0.0337	0.0255	0.0419	0.0000	0.0000
Freshman	2 - 1	0.0106	0.0015	0.0196	0.0087	0.0004
Trained	3 - 2	0.0124	0.0014	0.0235	0.0116	0.0034
	3 - 1	0.0230	0.0120	0.0340	0.0000	0.0000
Graduate	2 - 1	0.0042	-0.0061	0.0146	0.2978	0.0000
Not trained	3 - 2	0.0475	0.0349	0.0602	0.0000	0.0000
	3 - 1	0.0517	0.0391	0.0644	0.0000	0.0000
Graduate	2 - 1	0.0330	-0.1183	0.1842	0.4312	0.0001
Trained	3 - 2	0.1814	-0.0038	0.3667	0.0281	0.0028
	3 - 1	0.2144	0.0292	0.3996	0.0095	0.0000
Manager	2 - 1	0.0012	-0.0012	0.0035	0.2255	0.0115
Not trained	3 - 2	0.0106	0.0077	0.0135	0.0000	0.0000
	3 - 1	0.0118	0.0089	0.0146	0.0000	0.0000
Manager	2 - 1	0.0048	-0.0012	0.0107	0.0714	0.0607
Trained	3 - 2	0.0076	0.0003	0.0149	0.0190	0.0160
	3 - 1	0.0124	0.0051	0.0197	0.0002	0.0003

Table 2: Confidence intervals (joint 95% confidence within group) and one-sided P-values for testing β-level differences, parametric (T) and non-parametric (W)

In general, the difference in medians from period to period may be just as relevant as the mean. Among existing methods for constructing confidence intervals for the difference of theoretical medians based on the empirical medians, the resampling method seems preferable, given the nature of our data. We show this by giving in Table 3 the confidence limits for the difference of medians for Graduate-Trained group. [2]

Group	Periods	Difference	Lower	Upper
Graduate	2 - 1	0.0296	0.0017	0.0575
Trained	3 - 2	0.0503	-0.0055	0.1061
	3 - 1	0.0799	0.0280	0.1318

Table 3: Differences of β-medians with joint 95% comfidence limits by resampling

In comparison with the confidence interval for differences in means above, we see that the difference from period 1 to period 2 came out significant. This was already uncovered by the paiwise Wilcoxon-test, but masked by the standard TukeyHSD-test.

This non-standard approach could of course be taken for all groups, but will not change anything with respect to the conclusions about the learning arrived at by more standard methods.

One may also want to make comparisons of the level differences in the three periods across

[2]Bonferroni correction is used to obtain the overall 95% guarantee.

groups. This can be done within an ANOVA model with three factors: Group (F=Freshman, G=Graduate, M=Manager), Training (0=not trained, 1=Trained) and Period (1=Trial 1-40, 2=Trial 41-80, 3=Trial 81-100). The model then expresses the β-level by the three main effects, one for each factor, and three pairwise interactions. Analysis provides estimation and significance testing for each factor and for each interaction. As above, one may do estimation and significance testing of differences (post hoc testing). The multiple comparison aspect is again taken care of using TukeyHSD. In this case all three main effects and all three interactions were significant, and so were all differences going from one category on a factor to another category on the factor. Time trend may be added to the model as a covariate, but turns out not significant, as realized before. The main findings in this section may be summarized in graphs of the main effects and interactions from the ANOVA.

In the main effect graph (Figure 3) one see, for each factor, the (on average) changes of going from one level of a factor to another level, without taking into account possible interactions.

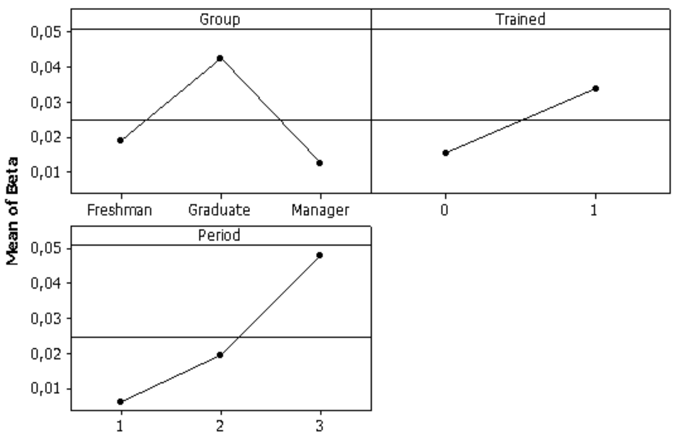

Figure 3: Main effects obtained from three-factor ANOVA

The interaction graph (Figure 4) shows how the relationship between β and the levels of one factor may differ for different levels of another factor. The plot displays means for the levels of one factor (on the x-axis) with separate lines for each level of the other factor. The more nonparallel the lines are, the greater the strength of the interaction. Parallel lines correspond to no interaction.

Recall that β measured the sensitivity to costs. From the graphs one clearly see the following: Graduates are the most sensitive, and managers the least sensitive. Trained particioants are more sensitive than the untrained. The sensitivity increases from period to period, the largest increase comes after the second package of information, prior to period 3, when the subjects

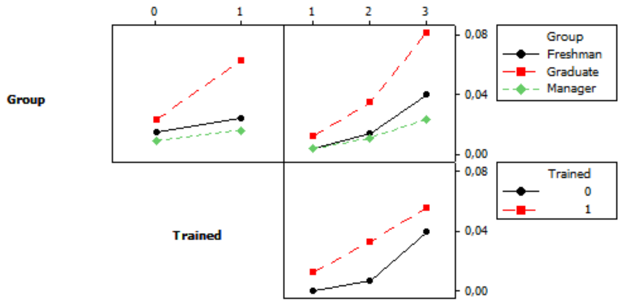

Figure 4: Interaction effects obtained from three-factor ANOVA

receive a handout with a graph showing how the profit depends on the order quantity. The interactions show that particularly sensitive to costs are the trained graduates in period 3.

5 Pull-to-center

Many authors have discussed the pull-to-center effect, and many explanations have been offered. As noted by Su (2008) and Ubøe et al (2017) a pull-to-center effect will be present because of random choice when orders are drawn from a multinomial logit distribution. All experiments in our data set were carried out with a critical fractile of 0.75. This means underorders are expected, which is also what is observed in the data set. In the theoretical model, the amount of underorder will depend on the information level β.

If we let $c(q)$ denote loss (in comparison with the optimal choice) if the newsvendor orders q units, and random orders Q are drawn from the probability distribution given by (6), the expected order can be computed by the formula

$$\mathrm{E}[Q] = q_{\mathrm{opt}} + \frac{\int_{d_{\min}-q_{\mathrm{opt}}}^{d_{\max}-q_{\mathrm{opt}}} q \exp[-\beta\, c(q+q_{\mathrm{opt}})]dq}{\int_{d_{\min}-q_{\mathrm{opt}}}^{d_{\max}-q_{\mathrm{opt}}} \exp[-\beta\, c(q+q_{\mathrm{opt}})]dq}, \tag{9}$$

where $[d_{\min}, d_{\max}]$ is the support of the demand D, see Ubøe et al (2017). In our experiment, D is uniformly distributed, and in that case it is easy to see that

$$c(q) = \frac{R-S}{2(d_{\max}-d_{\min})}(q-q_{\mathrm{opt}})^2 \tag{10}$$

The experiment uses the values $R = 12, S = 0, d_{\min} = 0, d_{\max} = 100, q_{\mathrm{opt}} = 75$. Note that in this case $c(q+q_{\text{opt}}) = \frac{6}{100}q^2$, and (9) takes the form

$$\mathrm{E}[Q] = 75 + \frac{\int_{-75}^{25} q \exp\left[-\frac{6\beta}{100}q^2\right] dq}{\int_{-75}^{25} \exp\left[-\frac{6\beta}{100}q^2\right] dq}. \tag{11}$$

By anti-symmetry

$$\int_{-75}^{25} q \exp\left[-\frac{6\beta}{100}q^2\right] dq = \int_{-75}^{-25} q \exp\left[-\frac{6\beta}{100}q^2\right] dq = -\int_{25}^{75} q \exp\left[-\frac{6\beta}{100}q^2\right] dq < 0, \tag{12}$$

and hence

$$\mathrm{E}[Q] = 75 - \frac{\int_{25}^{75} q \exp[-\frac{6\beta}{100}q^2]dq}{\int_{-75}^{25} \exp[-\frac{6\beta}{100}q^2]dq} < 75. \tag{13}$$

The expected order is then a function of the information level β only, see Figure 5. Strictly speaking D has a discrete distribution, but with a resolution at unit level a continuous approximation makes no difference.

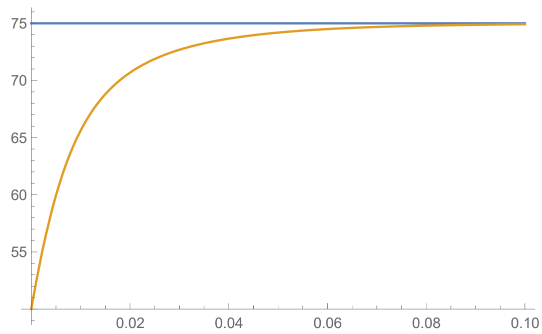

Figure 5: Expected orders as a function of the information level β.

From Figure 5 we see that underorders are expected at all information levels, but that the tendency to underorder decreases to zero with increasing β. For a rigorous proof of this monotonicity, see Ubøe et al (2017).

5.1 Predicting expected orders

The average values of β from Table 1 have been inserted in (11) to predict the expected order for each group and phase. The results are reported in Table 4 together with the observed averages.

Group	Phase 1		Phase 2		Phase 3	
	Predicted	Observed	Predicted	Observed	Predicted	Observed
Freshmen/basic	56	49	58	51	73	69
Freshmen/trained	67	62	71	66	73	70
Graduates/basic	57	52	63	55	74	71
Graduates/trained	72	68	74	72	75	73
Managers/basic	58	53	60	54	69	66
Managers/trained	65	60	68	61	71	67

Table 4: Predicted and observed values for the 6 different groups

As we can see from Table 4, the theoretical model predicts the observed average orders well. In Figure 6 we have displayed the observed average orders for each subgroup. For comparison we present the corresponding values derived from our probabilistic model in Figure 7.

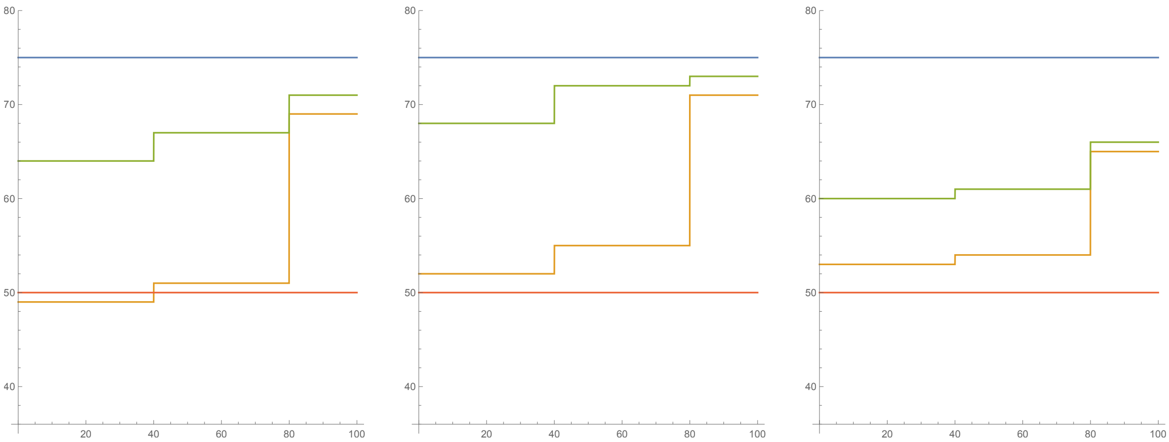

Figure 6: Average observed orders for freshmen, graduates and managers as reported in Bolton et. al (2012).

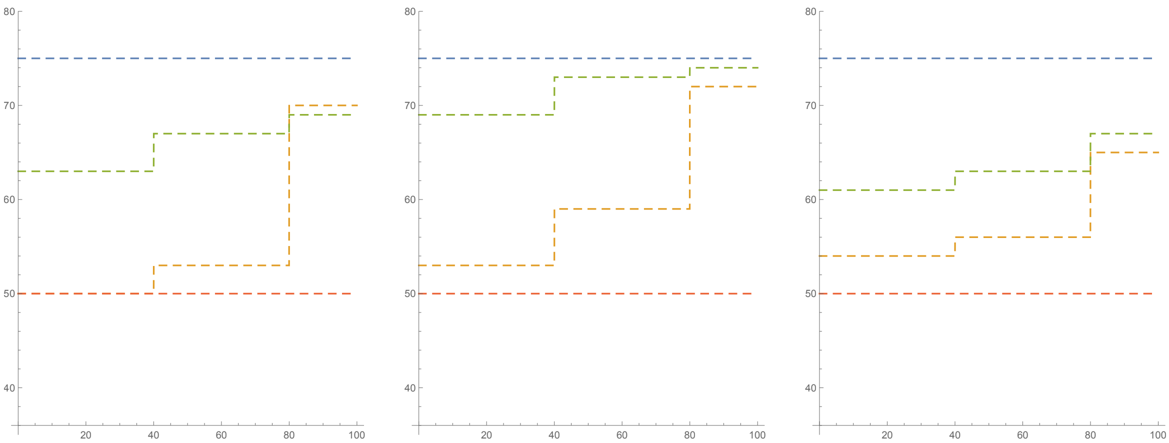

Figure 7: Predicted orders for freshmen, graduates and managers using our multinomial logit model.

It is straightforward to include more parameters, reflecting, e.g., risk aversion, in models of this kind. As a single parameter model seems sufficient to capture the main effects related to these data, we leave such extensions to future work.

6 Concluding remarks

The Fisher-Tippet-Gnedenko theorem suggests that the maximum distribution of iid random variables will lead to an aggregate distribution of multinomial logit type regardless of the choice distribution for the potential choices. For this to work we need to assume that the agents are fairly similar and that they choose between many alternatives. In real world applications these conditions are of course never met, but these models are robust in the sense that even strong violations of the assumptions do not necessarily create much change in the aggregate distribution. There will always be second order effects, but such effects will hardly matter in cases with only a moderate number of observations.

In the paper we have demonstrated how probabilistic models can be used to extract information about learning effects. The basic idea is that the β parameter in our model represents a quantification of learning effects. When this parameter increases, the agents choose less costly alternatives more often. We have demonstrated that the average values increase systematically over time, but that only the transition from phase 2 to 3 represents a statistically significant effect.

The discussion in Section 5 showed that our theoretical model predicts average order quantities very close to the observed values. These results would have been unsurprising if the model had been equipped with lots of parameters. In our model, there is only one parameter, and we find it quite remarkable that this single parameter is capable of carrying the information in the system to such high levels of precision. In our opinion, this demonstrates that probabilistic methods are just the right tools to analyze data sets of this kind.

Acknowledgement
The authors are very grateful to Professor Ulrich Thonemann for giving us access to and permission to use his experimental data from Bolton et al. (2012).

References

[1] Anas, A. 1983. Discrete choice theory, information theory and the multinomial logit and gravity models. *Transportation Research*, 17, 13–32.

[2] Andersson, J. and J. Ubøe. 2012. Some aspects of random utility, extreme value theory, and mutinomial logit models. *Stochastics*, 84, 2-3, 425–435.

[3] Bolton, G., A. Ockenfels, and U. Thonemann. 2012. Managers and students as newsvendors: how out-of-task experience matters. *Management Science*, 58, 12, 2225–2233.

[4] Bostian, A. J., C. A. Holt, and A. M. Smith. 2008. Newsvendor pull-to-center effect: adaptive learning in a laboratory experiment. *Manufacturing & Service Operations Management*, 10, 4, 590–608.

[5] Cachón G. P., 2003. Supply chain coordination with contracts. In: *The Handbook of Operations Research and Management Science: Supply Chain Management: Design, Coordination and Operation.* Chapter 6. A. G. de Kok and S. C. Graves (eds.), Elsevier, Amsterdam, pp. 229–340.

[6] Erlander, S. 2010. *Cost-Minimizing Choice Behaviour in Transportation Planning*, Springer, Berlin Heidelberg.

[7] Erlander, S. and N. F. Stewart. 1990. The Gravity Model in Transportation Analysis: Theory and Extensions, VSP, Utrecht.

[8] Kremer, M., S. Minner, and L. N. Van Wassenhove. 2010. Do random errors explain newsvendor behavior? *Manufacturing & Service Operations Management*, 12, 4, 673–681.

[9] Manski, C. 1977. The structure of random utility models. *Theory and Decision*, 8, 229–254.

[10] Matejka, F., and A. McKay. 2015. Rational inattention to discrete choices: a new foundation for the multinomial logit model. *American Economic Review*, 105, 1, 272-298.

[11] McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, 105–142. Academic Press, New York.

[12] Rudi, N., and D. Drake. 2014. Observation bias: the impact of demand censoring on newsvendor level and adjustment behavior. Management Science 60 (5), 1334-1345.

[13] Qin, Q., R. Wang, A. J. Vakharia, Y. Chen, and M. Seref, 2011. The newsvendor problem: Review and directions for future research. *European Journal of Operational Research* 213, 361–374.

[14] Schweitzer, M. E., and G. P. Cachón. 2000. Decision bias in the newsvendor model with a known demand distribution: experimental evidence. *Management Science*, 46, 3, 404–420.

[15] Smith, T. 1978. A cost-efficiency principle of spatial interaction behaviour. *Regional Science and Urban Economics*, 8, 313–337.

[16] Su, X. 2008. Bounded rationality in newsvendor models. *Manufacturing & Service Operations Management*, 10, 4, 566–589.

[17] Train, K. E. 2003. *Discrete Choice Methods with Simulation.* Cambridge University Press.

[18] Ubøe, J., J. Andersson, K. Jörnsten, J. Lillestøl, and L. Sandal. 2017. Statistical testing of bounded rationality with applications to the newsvendor model. *European Journal of Operational Research* 259, Issue 1, 251–261.

[19] Wachtel, R. E., and F. Dexter. 2010. Review of behavioral operations experimental studies of newsvendor problems for operating room management. *Anesthesia & Analgesia*, 110, 6, 1698–1710.

[20] Wilson, A.G. 1967 A Statistical Theory of Spatial Distribution Models, *Transportation Research*, 1, 253–269.

NHH

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no