Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Newsvendor Demand Chasing Revisited

Nelson Lau, J. Neil Bearden,

MANAGEMENT SCIENCE
Vol. 59, No. 5, May 2013, pp. 1245–1249
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1120.1617

Newsvendor Demand Chasing Revisited

Nelson Lau, J. Neil Bearden
Decision Sciences, INSEAD, Singapore 138676 {nelson.lau@insead.edu, neil.bearden@insead.edu}

Existing research on newsvendor behavior asserts that individuals engage in demand chasing—adjusting their order quantities towards prior demand. Several metrics have been used to identify this heuristic. By simulation, current metrics are shown to yield excessive false positives, indicating demand chasing when the true order generating process is independent of prior demand. A simple correlation measure does not suffer from this problem and is proposed here as an alternative to some of the more commonly used measures.

Key words: behavioral operations; econometric analysis; experiments; newsvendor; demand chasing
History: Received February 3, 2012; accepted June 22, 2012, by Martin Lariviere, operations management. Published online in *Articles in Advance* December 10, 2012.

1. Introduction

Extant research into newsvendor behavior has suggested that individuals engage in *demand chasing* (see, e.g., Ben-Zion et al. 2008, Bostian et al. 2008, Kremer et al. 2010, Lurie and Swaminathan 2009, Moritz et al. 2009, Schweitzer and Cachon 2000).[1] That is, they *systematically* adjust their order quantities in the direction of the previous period's demand. A number of different statistical summary measures have been used to assess demand chasing. However, some of the widely used measures are prone to false positives: they identify demand chasing even when order updating is independent of the previous period's demand. We show that a simple correlation measure does not fall victim to the same statistical risk, and we propose that it be used to measure demand chasing in newsvendor studies.

2. Results

To illustrate the performance of several measures of demand chasing, we simulate ordering decisions under two order generating processes, one that mimics demand chasing behavior and one that does not. We show that the measures differ in their Type I (false-positive) and Type II (false-negative) error rates and that they are affected by the level of noise (or random error) in the data.

2.1. Order Generating Processes

Preliminaries. Let q_t denote an agent's order quantity for period t, and let d_t represent the demand realization in that same period. In what follows, we will simulate the values of q_t according to two processes. Under the *true demand chasing* process, the q_t are generated as if an agent's ordering decisions are sensitive to the previous periods demand—specifically, that her order quantity in period t is probabilistically biased toward demand in period $t-1$. Under the *quantal choice* process, the agent's order quantity in period t is generated independent of the demand in period $t-1$. Both data generating processes are stochastic, and we vary the magnitude of noise (or variability) in the order quantities to simulate data that themselves may vary in noisiness.

The Monte Carlo results are based on simulating demand as independent and identically distributed (i.i.d.) draws from the set of integers $\{0, 1, \ldots, 100\}$, in which each quantity is sampled with equal probability. Our qualitative conclusions extend to other demand distributions. We use the uniform distribution simply to illustrate an existence result: that some currently employed measures of demand chasing can reveal "demand chasing" even when order quantities are generated independent of demand. We also present a measure that does not suffer from this shortcoming.

2.1.1. True Demand Chasing. Under our true demand chasing process, an agent's order quantity in t is mechanically adjusted toward demand in period $t-1$ but with some noise (ε). Specifically, orders are generated according to the following:

$$q_t = q_{t-1} + \beta(d_{t-1} - q_{t-1}) + \varepsilon_t, \quad \text{where } \varepsilon_t \sim \mathrm{N}(0, \sigma^2) \text{ and } q_1 = 50 + \varepsilon_1. \tag{1}$$

The magnitude of an agent's demand chasing is captured by β, and the noise in an order quantity is an i.i.d. random variable $\epsilon_t \sim \mathrm{N}(0, \sigma^2)$. When the order generating process suggests q_t values above 100 or

[1] See the textbook by Cachon and Terwiesch (2009), among others, for the newsvendor problem and its solution.

below 0, q_t is set to 100 or 0, respectively. In the simulations reported below, the mean of the order quantity in $t = 1$ was set near the mean of the demand distribution, though the qualitative results hold under alternative seeding procedures as well.[2] In all simulations, $\beta = 0.5$, while $\sigma = 10$ in the *low noise* condition and $\sigma = 20$ in the *high noise* condition.

2.1.2. Quantal Choice. Under the quantal choice process, an agent's order quantity in period t is generated independent of demand in the previous period. This process is based on the model in Su (2008). He proposed that agents facing a choice among alternatives $i \in \vartheta$, each yielding utility u_i, choose alternative i with probability

$$\varphi_i = \frac{e^{u_i/\gamma}}{\sum_{i\in\vartheta} e^{u_i/\gamma}}, \qquad (2)$$

where γ is a parameter indicating the extent of cognitive and computational limitations. This is termed a bounded rationality preference function (BRPF). Applied to uniform demand, the BRPF produces an ordering distribution that is truncated normal (Su 2008, pp. 572–573).

Here, the order generating process for quantal choice is as follows:

$$q_t = X_t,$$

where

$$X_t \sim f(x) = \frac{(1/s)\phi((x-50)/s)}{\Phi((100-50)/s) - \Phi((0-50)/s)}, \qquad (3)$$

ϕ is the probability density function of the standard normal distribution, and Φ is the cumulative distribution function of the standard normal distribution.

In other words, the X_t are i.i.d. draws from a truncated normal distribution over $[0, 100]$ with a mean of 50 and a variance of s^2. In the simulations reported below, we set $s = 10$ in the low noise condition and $s = 20$ in the high noise condition. Under the quantal choice process, orders are generated independent of previous demand—by design, there is no demand chasing.

For each order generating process and noise level, we simulated the ordering decisions of 10^5 agents, each facing 31 seasons of demand. Again, for each, realized demand was generated by sampling i.i.d. draws from the integers $\{0, 1, \ldots, 100\}$. Although we only report results from two levels of noise, we ran a much more comprehensive set of numerical experiments under broader conditions of parameters. The general pattern of results we report with the two noise levels provides a concise, accurate portrait of the general pattern of results.

[2] The mean is a focal point for experimental subjects (see, e.g., Gavirneni and Xia 2009).

2.2. Metrics and Simulation Results

Next, we describe four alternative measures of demand chasing and evaluate their performance. We are particularly interested in each measure's misclassification rates.

2.2.1. Changes Towards vs. Away. This metric is defined as follows: All cases in which $d_{t-1} = q_{t-1}$ or $q_t = q_{t-1}$ are ignored. Among what remains, $q_t - q_{t-1}$ and $d_{t-1} - q_{t-1}$ being of the same sign is a change *towards*, whereas $q_t - q_{t-1}$ and $d_{t-1} - q_{t-1}$ being of opposite signs is a change *away*. A proportion of changes towards greater than 0.5 ostensibly suggests demand chasing. The statistical significance of the observed proportion is assessed by a binomial test. This metric was used in Schweitzer and Cachon (2000), Ben-Zion et al. (2008), and Kremer et al. (2010). The reported proportion of changes towards ranged from 0.64 to 0.80. All three papers claimed the number of changes towards to be significantly different from the number of changes away at the aggregate level.[3]

Table 1 summarizes the simulation results. It shows the values of the change scores for the 25th, 50th, and 75th percentiles of the aggregate distributions. Under true demand chasing, the change measures, even at the 25th percentile, tend to be well above 0.5. To classify each subject as demand chasing (or not), we used one-tailed binomial tests with no demand chasing as the null (i.e., H_0: metric value ≤ 0.5, H_A: metric value > 0.5; $\alpha = 0.025$, critical p-value $= 0.025$). At the low noise level, 98.88% of subjects from the true demand chasing process were classified as demand chasing. Thus, there was a false-negative rate of only 1.12%. At the high noise level, the false-negative rate rose to 17.33%. Hence, the change measure does quite well, especially at low noise levels, at picking up demand chasing—when demand chasing is, in fact, present.

Table 1 also shows the results on the change measure for the quantal choice order generating process, which generates orders independent of previous demand. Under this process, the change scores are substantially closer to 0.5, though the distributions still tend to be right-shifted, especially under high noise. It is important to note that at the individual level, the measure gives rise to false-positive rates that are well above the intended 0.025 Type I error rate used for the classifications. In fact, under high noise, the misclassification rate is nearly seven times the expected rate. In short, the changes towards versus away measure that has been used in some previous studies is prone to overestimation of demand chasing.

[3] Individual-level analyses are used in the current paper to establish the proportion of subjects classified as employing a demand chasing heuristic.

Table 1 Simulation Results: Changes Towards vs. Away

Model	Noise level	25th percentile	50th percentile	75th percentile	Critical p-value	% classified as demand chasing	False-positive rate (%)	False-negative rate (%)
True demand chasing	Low	0.82	0.86	0.90	0.025	98.88	—	1.12
True demand chasing	High	0.72	0.77	0.83		82.67	—	17.33
Quantal choice	Low	0.50	0.56	0.62		6.25	6.25	—
Quantal choice	High	0.55	0.61	0.67		17.11	17.11	—

2.2.2. Adjustment Score. Following Schweitzer and Cachon (2000), the *adjustment score* is defined as $(q_t - q_{t-1})/(d_{t-1} - q_{t-1})$. These scores are computed separately for moves toward and away from the previous period's demand. Schweitzer and Cachon (2000, pp. 412–413) found in one experimental condition that "average adjustment scores toward and away from prior demand were 0.31 and -0.11, respectively, $t(38) = 2.75$, $p < 0.01$," but that in another condition, the difference was not significant. A significant difference between adjustment scores towards and adjustment scores away was taken as an indicator of demand chasing.

Table 2 displays the average difference in the simulated agents' adjustment scores (towards minus away) at the 25th, 50th, and 75th percentiles of the aggregate distribution for low and high noise under the two order generating processes. Following Schweitzer and Cachon's (2000) basic approach, agents were classified as demand chasing using (one-tailed) t-tests (H_0: adjustment score towards less than or equal to adjustment score away, H_A: adjustment score towards greater than adjustment score away; $\alpha = 0.025$, $t_{\text{critical}} = 2.045$). The false-negative rate with true demand chasing was 16.40% under low noise and 13.57% under high noise. There was no substantial pattern of variation in the false-negative rate with noise level. Under quantal choice, the observed false-positive rates were more than 35 times greater than the expected 0.025 Type I error rate in both noise conditions. There was no meaningful pattern of variation in the false-positive rate with noise level.

2.2.3. Regression β. Bostian et al. (2008) operationalized demand chasing by what they called a *partial adjustment model*:

$$q_t = q_{t-1} + \beta(d_{t-1} - q_{t-1}) + \varepsilon_t, \tag{4}$$

where β reflects the extent of demand chasing and ε is an error term assumed to be i.i.d. normal (Bostian et al. 2008, p. 594). Analyzing aggregate data, the authors estimated $\beta = 0.15$ (see p. 598).

Moritz et al. (2009) used the same model, running the equivalent regression

$$q_{i,t} - q_{i,t-1} = \beta_i(d_{i,t-1} - q_{i,t-1}) + \varepsilon_{i,t}, \tag{5}$$

where the subscript i emphasizes that the analyses were done at the individual level. In one experimental condition, the mean estimated β was 0.34, which was statistically significantly different from 0 (Moritz et al. 2009, p. 18). Lurie and Swaminathan (2009) used the same basic regression approach with additional control variables and concluded that they found evidence of demand chasing.

Table 3 summarizes the results from estimating the regression β as per Equation (5) on the simulated agents' ordering decisions. Based on the previous use of this approach, a positive β value purportedly indicates demand chasing. Under true demand chasing, regression recovered the β parameter quite efficiently—the true β was 0.5, and the 25th and 75th percentile estimates are close to 0.5. Higher noise resulted in estimates that were further from the true β. One-tailed t-tests (H_0: $\beta \leq 0$, H_A: $\beta > 0$; $\alpha = 0.025$, $t_{\text{critical}} = 2.045$) were used to classify subjects as demand chasing. The false-negative rate was 0.00% at low noise and increased to 0.05% at high noise. Under quantal choice, the majority of estimated β values were positive. Even at low noise, the 25th percentile of the estimated β was 0.05. As noise increased, the estimated β increased. Whereas the expected false-positive rate was 2.5%, the observed false-positive rates were notably higher—23.00% at low noise and 59.22% at high noise. The false-positive rate increased as noise levels rose.

Table 2 Simulation Results: Adjustment Score (Difference in Towards and Away)

Model	Noise level	25th percentile	50th percentile	75th percentile	t-test critical value	% classified as demand chasing	False-positive rate (%)	False-negative rate (%)
True demand chasing	Low	1.37	1.91	2.82	2.045	83.60	—	16.40
True demand chasing	High	2.04	2.77	3.89		86.43	—	13.57
Quantal choice	Low	1.43	1.90	2.57		89.61	89.61	—
Quantal choice	High	2.63	3.52	4.85		89.26	89.26	—

Table 3 Simulation Results: Regression β

Model	Noise level	25th percentile	50th percentile	75th percentile	t-test critical value	% classified as demand chasing	False-positive rate (%)	False-negative rate (%)
True demand chasing	Low	0.47	0.50	0.54	2.045	100.00	—	0.00
True demand chasing	High	0.45	0.51	0.57		99.95	—	0.05
Quantal choice	Low	0.05	0.11	0.16		23.00	23.00	—
Quantal choice	High	0.22	0.30	0.39		59.22	59.22	—

Table 4 Simulation Results: Correlation

Model	Noise level	25th percentile	50th percentile	75th percentile	t-test critical value	% classified as demand chasing	False-positive rate (%)	False-negative rate (%)
True demand chasing	Low	0.68	0.73	0.77	2.045	99.98	—	0.02
True demand chasing	High	0.44	0.53	0.61		88.64	—	11.36
Quantal choice	Low	−0.13	0.00	0.13		2.22	2.22	—
Quantal choice	High	−0.13	0.00	0.13		2.34	2.34	—

This regression-based approach adopted in several previous studies to assess demand chasing has a substantially higher than assumed Type I error rate. The actual Type I error rate in our simulations went as high as more than 20 times the presumed rate.

2.2.4. Correlation. Bolton and Katok (2008) used a straightforward measure of demand chasing: using subjects as the unit of analysis, they took the correlation between q_t and d_{t-1}. In their ad hoc hierarchical classification of newsvendors, 30% of the subjects in the low margin condition (i.e., 6 of 20 subjects) and 10% in the high margin condition (i.e., 2 of 20 subjects) displayed a significant positive correlation, suggesting demand chasing in some individual subjects (Bolton and Katok 2008, p. 528). (At the 5% significance level, one would expect around one subject per condition to be classified as demand chasing under the null of no demand chasing for all subjects.) Note that whenever q_t and d_{t-1} are independent, the expected correlation between them is 0 (see statistics textbooks such as Newbold et al. 2010). Furthermore, when orders are placed independent of previous demand, this correlation measure does not suffer from inflated Type I error rates.

To illustrate, Table 4 summarizes the results from the correlation measure. Under true demand chasing, a high positive correlation was obtained for the majority of subjects. At low noise, the 25th percentile correlation value was 0.68. As expected, the correlation decreased with increasing noise; but it remained high even under high noise; the 25th percentile correlation value was 0.44. One-tailed t-tests of the transformation of the correlation[4] (H_0: correlation ≤ 0, H_A: correlation > 0; $\alpha = 0.025$, $t_{\text{critical}} = 2.045$) were used to classify subjects as demand chasing. At low noise, the false-negative rate was just 0.02%, and at high noise, it was 11.36%. The false-negative rate increased with noise, but its absolute level suggests that the correlation measure is quite powerful.

Under quantal choice ordering, the correlation metric values were close to 0. The 25th and 75th percentile values were −0.13 and 0.13, respectively. There was no trend in the correlation values with variation in noise levels. Most important, the observed false-positive rates were in line with the expected rate of 2.5%. In the low noise condition, the false-positive rate was 2.22%; in the high noise condition, it was 2.34%. There was no association between the false-positive rate and the noise level.

3. Conclusion

Our simulations show that three statistics used in extant literature to assess demand chasing can suffer from inflated Type I error rates. However, a simple measure of the correlation between the previous demand and the current order quantity does not suffer from this problem, and it is also a reasonably powerful test (when there is true demand chasing). The false-positive problem can significantly impact research conclusions. As one demonstration, experimental data from Study 1 of Bolton and Katok (2008) were reanalyzed. Using the three metrics described in §§2.2.1–2.2.3, between 58% and 100% of subjects in the low margin condition and between 53% and 93% of subjects in the high margin condition were classified as demand chasing. Using correlation, only 22% of the subjects in the low margin condition and 20% of the subjects in the high margin condition were

[4] That is, $t_{n-2} = (r\sqrt{n-2})/\sqrt{1-r^2}$. Again, see textbooks such as Newbold et al. (2010).

classified as demand chasing. The relative paucity of demand chasing was confirmed in a new experiment using a mid-margin newsvendor problem. In it, the three existing metrics classified between 44% and 96% of subjects as demand chasing, whereas correlation identified just 9% of subjects as demand chasing.

This paper demonstrates that even when demand chasing is absent, some extant measures deviate from their *presumed value* under the null of no demand chasing. A simple correlation measure is not prone to this problem and is recommended for use in future research.

Acknowledgments

This research was supported by the INSEAD Alumni Fund.

References

Ben-Zion U, Cohen Y, Peled R, Shavit T (2008) Decision-making and the newsvendor problem: An experimental study. *J. Oper. Res. Soc.* 59(9):1281–1287.

Bolton GE, Katok E (2008) Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing Service Oper. Management* 10(3):519–538.

Bostian AJA, Holt CA, Smith AM (2008) Newsvendor "pull-to-center" effect: Adaptive learning in a laboratory experiment. *Manufacturing Service Oper. Management* 10(4):590–608.

Cachon GP, Terwiesch C (2009) Betting on uncertain demand: The newsvendor model. *Matching Supply with Demand*, international ed. (McGraw-Hill, Singapore), 220–255.

Gavirneni S, Xia Y (2009) Anchor selection and group dynamics in newsvendor decisions—A note. *Decision Anal.* 6(2):87–97.

Kremer M, Minner S, Van Wassenhove LN (2010) Do random errors explain newsvendor behavior? *Manufacturing Service Oper. Management* 12(4):673–681.

Lurie NH, Swaminathan JM (2009) Is timely information always better? The effect of feedback frequency on decision making. *Organ. Behav. Human Decision Processes* 108(2):315–329.

Moritz BB, Hill AV, Donohue K (2009) Cognition and individual differences in the newsvendor problem: Behavior under dual process theory. Working paper, University of Minnesota, Minneapolis. Accessed November 11, 2012, http://faculty.chicagobooth.edu/workshops/omscience/archive/pdf/Moritz%20Hill%20Donohue%20Cognition%20and%20Newsvendor.pdf.

Newbold P, Carlson WL, Thorne B (2010) *Statistics for Business and Economics*, 7th ed. (Prentice Hall, Upper Saddle River, NJ).

Schweitzer ME, Cachon GP (2000) Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Sci.* 46(3):404–420.

Su X (2008) Bounded rationality in newsvendor models. *Manufacturing Service Oper. Management* 10(4):566–589.