decisions similar to those of the newsvendors. We thus compare treatments across the two inventory settings to test the conjecture.

First, we observe increased standard deviations of the ordering decision in the dynamic treatments in contrast with their newsvendor counterparts (2.72 vs. 0.94, p-value < 0.001 for Qonly; and 3.04 vs.1.44, p-value < 0.001 for PQ). This is consistent with the theory prediction that a base stock policy in the dynamic setting should generate larger order variance than a stationary optimal order quantity of the newsvendor model. Second, we find the standard deviation of the selling price in Ponly to be larger under the dynamic setting than the single-period case (1.98 vs. 0.81, p-value < 0.001). This again is in accordance with theory that subjects should price according to the current inventory levels instead of a list price. Both observations reveal that subjects respond to the dynamic environment by some decision rule other than the one used under a newsvendor setting, providing evidence that subjects are not entirely myopic.

To summarize, we identify systematic behavioral deviations from predictions by normative theories, especially that more inventory control options does not necessarily enhance profit performance. One possible explanation is that, for a boundedly rational agent, more freedom in decision making means more room to err. The experimental results provide some insights of where the theory breaks down. In the next section, we propose a behavioral model to explain these behavioral departures.

IV. BEHAVIORAL MODEL

Our model is motivated by two general principles. First, people make mistakes. We incorporated this idea using the probabilistic choice framework (McFadden 1976, McKelvey and Palfrey 1995). Second, people do not look forward into the future perfectly when dealing with dynamics.

4.1. Probabilistic Choice with Partial Forward-Looking

The core idea behind the probabilistic choice framework is that individuals are subject to random errors while making decisions. We consider a representative agent who has the following utility function when evaluating an action i among possible alternatives:

$$U_i = \sum_{t=0}^{T} \beta_t (E\pi_t) + \varepsilon$$

where $E\pi_t$ is the expected profit at period t. $t = 0$ is set as the reference time period where an initial decision is made. ε is a random variable with zero mean, which can be interpreted as the error made in calculating the sum of the expected profit. β_t is a weight the agent assigned to future profits in period t. While it has a similar structure to the standard discounting factor, there is one key difference. β_t can decrease at a rate not constrained by the exponential decay. Thus, we do not restrict it to represent discounting based on the future value of money. In a laboratory experiment where subjects are expected to be paid within a short period of time, there is no true discounting. However, we still do not expect β_t =1. It is possible in this framework for the agent to place less weight on profits further away in the future because they may have less confidence of the validity of their estimation. Ultimately, the exact reason of this weighing is less important than the validation of its existence, as shown next in Section 5.

To make the model tractable, we make two further simplifying assumptions: 1) β_0 = 1 (without loss of generality), $\beta_1 = \beta$ and β_t = 0 for all t > 2. Thus, we reduce the model to two terms, one for the current period and the other for the next period. This is a strong assumption assuming that an agent does not consider what will happen beyond the immediate next period. The full dynamic optimization process is computationally intensive because of the combinatorial explosion of the possibilities. In addition, previous experimental studies suggest that human decision-makers would consider only few periods ahead. As we can also see in the later section, estimates of β from the two-period model using experimental data are significantly below 1, which helps justify this assumption. 2) ε is distributed with the extreme value distribution, which is a standard assumption used in choice models (McFadden 1976).

Under these assumptions, the choice probabilities are modeled by a multinomial logit distribution. Given specific parameter settings of our study, the action space is two-dimensional and consists of possible combinations of price and order quantity. Subjects were restricted to choose integer decisions within a certain reasonable range. Thus, modeling the action space as a finite set is adequate, however, the model can be generalized to continuous decision spaces (see Morgan 1992). Let i = 1... n be a finite set of possible actions. The probability that the decision maker chooses alternative i is given by:

$$p_i = \frac{e^{\gamma U_i}}{\sum_{i=1}^{n} e^{\gamma U_i}},$$

where Ui is the utility defined above and γ is a parameter related to the distribution characteristics of ε. At γ = 0, the agent randomly selects his/her decision with equal probability for all possible choices. As $\gamma \to \infty$, the agent picks the choice with the highest utility with probability 1. Hence, the traditional utility maximization model is a special case of the above probabilistic choice model. γ can be interpreted as the degree of rationality. An agent is completely irrational (random choices) at γ=0 and fully rational when $\gamma \to \infty$. In Figure 2, we plot three choice distributions given different levels of gamma for illustration. The probability distribution of the low gamma case is quite flat as the agent makes almost random decisions. As gamma increases, distributions become more concentrated on the utility-maximizing decision. Although the best alternative is no longer chosen with probability one, it is still the mode of the choice distribution.

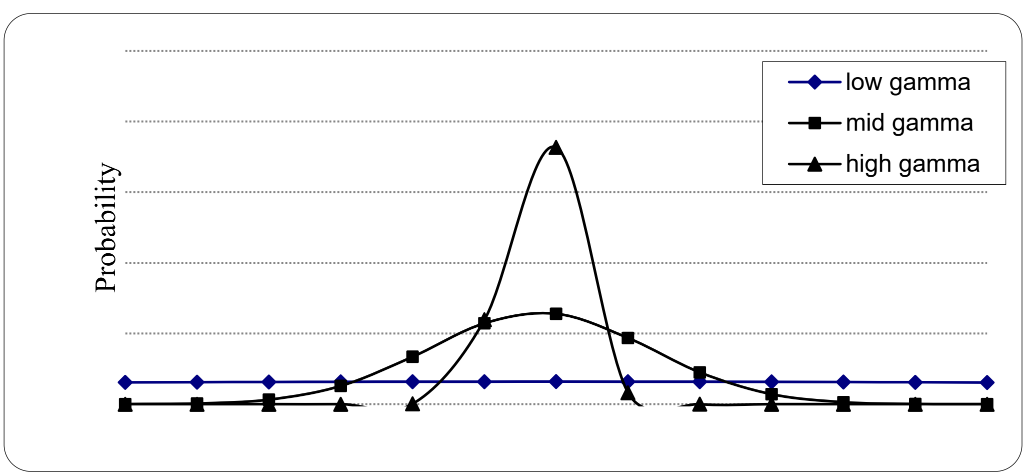

FIGURE 2: PROBABILITY DISTRIBUTIONS OF DECISIONS UNDER PROBABILISTIC CHOICE MODEL.

4.2. Modeling Implications

In this section, we summarize major implications from the behavioral model introduced above when applying to our game settings. Monte Carlo simulations are used to study the model under the newsvendor and the dynamic inventory settings. A "software" agent makes decisions in settings identical to our treatments (i.e., Qonly, Ponly and PQ). In every period, the agent calculates the expected profit for all potential decisions. All decisions were restricted to integers. Since prices have a finite range of [0, 15], an expected profit can be calculated for each possible price. Ordering decisions, on the other hand, are not bounded. Theoretically, as long as the decision space is not infinite, there is positive probability to choose even very large quantities. However, in practice, the probability of doing so diminishes quickly. We therefore placed an upper bound on the quantity decision in the simulation. The bound was chosen so that it would not impact the final results. Each simulation run lasts for 36 periods, and 500 simulations are used to calculate the average behavioral performance given different β and γ. For illustration purpose, we report simulation results from the behavioral model in which β is set to be zero. In this case, the agent is myopic who computes the expected profit of the current period only. While actual decision distributions will change when β is positive, the comparative static results shown below are robust.

Modeling Result 1: For boundedly rational decision makers, reduced options to control inventory can help improve empirical performance.

We first examine the profit performance predicted by the behavioral model. Figure 3 plots the average period profit per round under each treatment for $\gamma \in [0, 1]$ (the range was chosen in accordance with gamma estimations from the experiments reported in Section 5). We can see that profits increase with γ in all treatments. The behavioral model suggests that when γ increases, the deviations from the optimal policy reduce and thus profits increase. Note that differences in profit performance decrease as γ increases among all treatments, which is consistent with the idea that when agents become more rational, behaviors would converge to the optimal solutions by the normative theory.

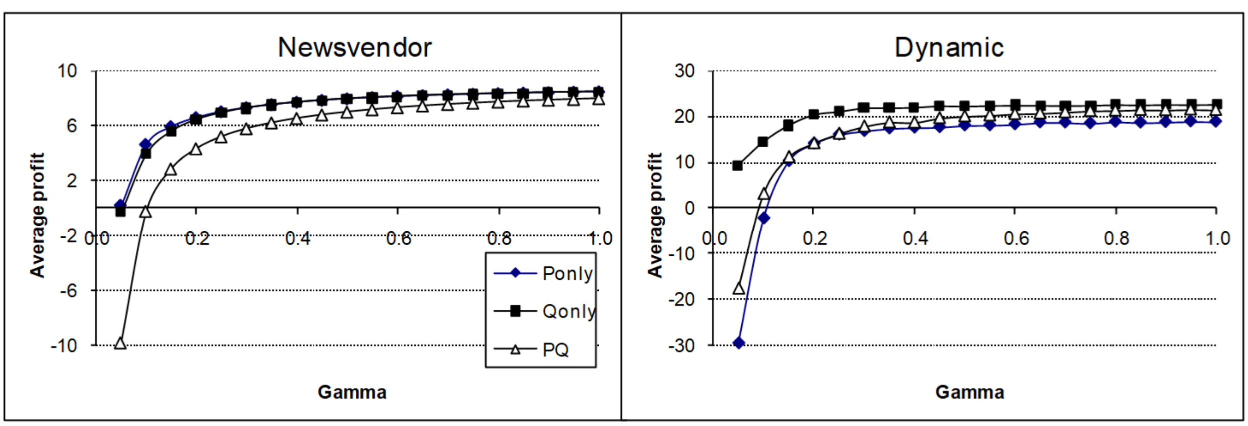

FIGURE 3: PROFIT PERFORMANCE AS A FUNCTION OF GAMMA PREDICTED BY THE BEHAVIORAL MODEL.

Consistent with Experimental Result 1, the behavioral model predicts that reduced options to control inventory can improve empirical profit performance. More specifically, in the newsvendor setting, our behavioral model predicts similar profit performance in Qonly and Ponly, and both of which strictly dominate the performance of PQ for all values of gamma. In the dynamic setting, profits in Qonly are always higher than that in PQ. The treatment of Ponly, however, is predicted to perform worse than PQ. This is driven by the fact that subjects face an ordering policy is set to be suboptimal with a constant order (of 6). The profit performance of PQ is therefore hampered by both system errors and human errors, as well as possible interactions between the two. The ordering of profit comparisons amongst various treatments predicted by the behavioral model is identical to what we observed in the experiments. The proposed model, which considers bounded rationalities of decision makers, is therefore capable to explain the "Optimizer's Paradox" from a behavioral aspect.

Modeling Result 2: Price and order comparisons amongst treatments predicted by the behavioral model are consistent with experimental observations.

We next look at the pricing decisions implied by the behavioral model with a focus on the dynamic setting. Figure 4 shows the simulation results on the average and standard deviation of the selling price as a function of γ for the dynamic treatments. Qonly, where the pricing policy is set exogenously at optimality, is also included in the graph for reference.

Consistent with Experimental Result 2, the behavioral model implies that average prices in Ponly and PQ to be lower than the optimal price of 13. The "underpricing" is in line with the "midpoint bias" suggested by Su 2008 under the newsvendor setting. The basic intuition is that the probabilistic choice model, with limited support in the decision dimension, predicts a shift to the midpoint (a price of 11.50 in our case) of the range of the support. This reasoning applies to the pricing decisions in our setting as well. The shift is more prominent when gamma is at its low end as illustrated by Figure 4. Furthermore, the behavioral model predicts the average price to be higher in Ponly than in PQ. Recall that orders are fixed at 6 in Ponly. Yet in PQ, orders can be in the range between 0 and 15 with a mid-point of 7.5 (assuming the subjects will not order beyond the highest possible demand under the lowest price of 8). If we assume the demand can go up to 23, the highest possible demand given a zero price, the order will be even higher which further strengthens the intuition we are trying to illustrate. Applying the same idea of the mid-point bias, the model predicts the order quantity to be higher than 6 in PQ with a further reduced selling price. Although the pricing differences observed in the experiments are not significant across the treatments but they are along the direction suggested by the behavioral model.

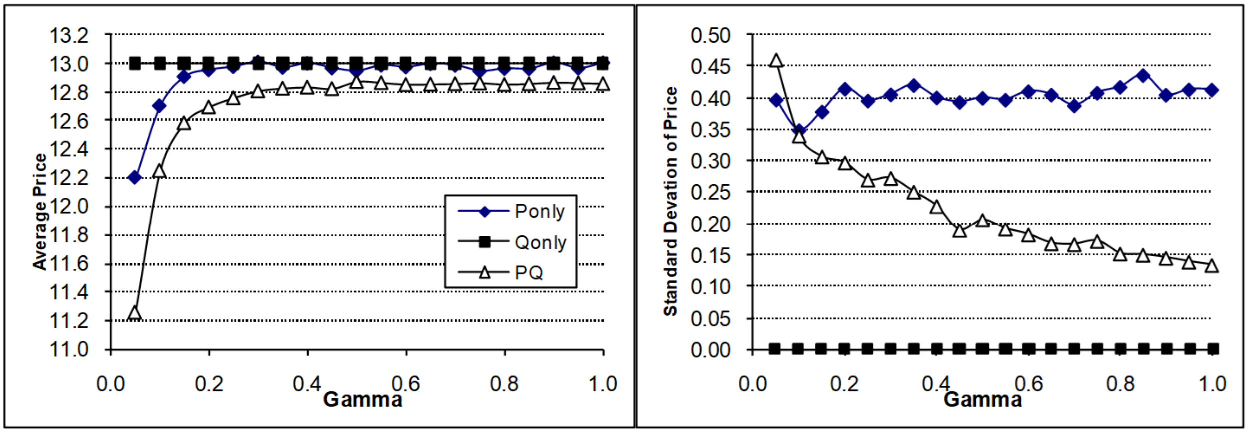

FIGURE 4: PRICING DECISIONS UNDER DYNAMIC SETTING AS A FUNCTION OF GAMMA PREDICTED BY THE BEHAVIORAL MODEL.

As for the decision variance, the model predicts it to be lower in PQ than in Ponly. When gamma increases, the pricing policy under PQ converges to the list price with smaller variations, whereas the price in Ponly requires dynamic adjustments. When β is positive, standard deviations of the pricing decision enlarge in general under these two treatments, but the relative comparisons remain the same. In the newsvendor setting, the behavioral model suggests the pricing standard deviation to be larger in PQ for all gamma values, and the difference is more seemingly when gamma is less than 1. The standard theory implies the exact same pricing strategy to be used in Ponly and PQ under the newsvendor model. The larger decision variance predicted in PQ by the model therefore can be attributed to the fact of increased decision freedom.

Lastly, we examine the stocking decisions implied by the behavioral model. Figure 5 displays simulation results on averages and standard deviations of the order quantity for the three treatments under the dynamic setting. In Qonly, our model predicts the average order quantity to be close to 6, which indeed was observed in the experiment. The average order in PQ, however, is predicted to be above 6; and the treatment difference is more significant at lower levels of gamma. The intuitive explanation for the above results can again trace back to the mid-point bias. In Qonly, when the price is set to be 13, the resulting critical fractile is 50%. We do not see much variation from the mean demand of 6. In PQ, since the mid-point bias can cause the price to be below 13, the corresponding stocking decision is then higher than 6 due to the demand function. The larger decision variance in PQ can again be explained by the increased decision complexity. When β is positive, decision variances go up under both Qonly and PQ. The ordering strategy converges to the base stock policy as decision makers become less myopic. Yet at any given value of beta, the standard deviation of orders in PQ is no less than that in Ponly.

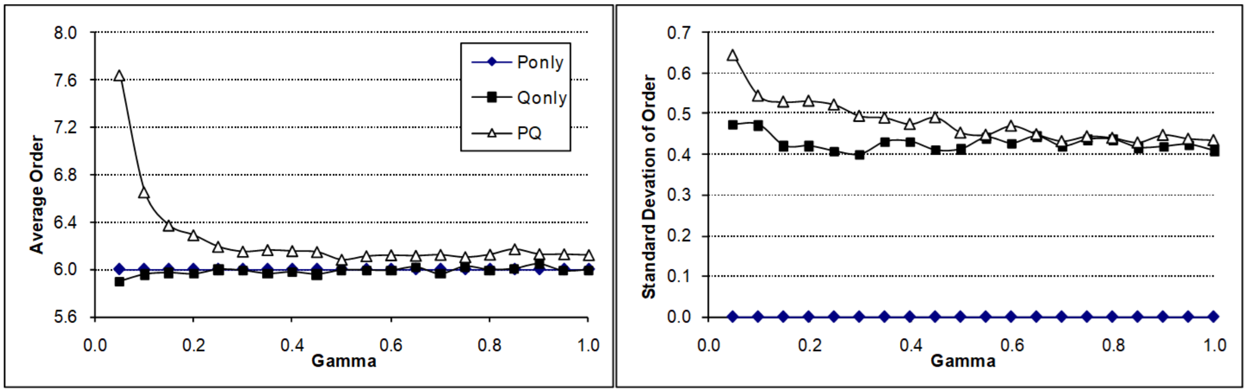

FIGURE 5: ORDERING DECISIONS AS A FUNCTION OF GAMMA BY THE BEHAVIORAL MODEL.

It is worth noting that results in this section are obtained under the assumption that γ remains at the same level across various treatments. From model estimations results reported in Section 5, we find that γ actually decreases as the "complexity" of the decision task increases (for example, from Qonly to PQ). The smaller γ estimated from PQ does not change the comparative results discussed above. In fact, it makes the gap even larger. The basic comparative statics of the model predictions still hold even if γ changes across treatments, as long as γ decreases with increased decision complexity. To conclude, directional predictions from the proposed behavioral model are consistent with lab observations in general. In the next section, we apply the model to estimate behavioral parameters of the subjects for their tendency to make random errors and their ability to look forward into the future.

V. MODEL ESTIMATION RESULTS

The method of maximum likelihood estimation is used to estimate γ and β, with individual differences being independent across subjects. Estimations of both γ and β are performed at an aggregate level, using observed decisions from all subjects under a treatment. For treatments under the newsvendor setting, since the initial stocking level is always 0, decisions are simulated to be independent across periods. For treatment under the dynamic setting, we assume that, conditional on the starting inventory in a period, decisions are independent. The total number of data points used = (the number of subjects) x (the number of periods) x (number of decisions in a period). Table 3 reports the corresponding behavioral estimates.

TABLE 3: ESTIMATION RESULTS FROM THE BEHAVIORAL MODEL.

Treatment		Gamma	Beta	Log Likelihood
Newsvendor	Ponly	0.29	n/a	528.51
	Qonly	0.27	n/a	553.24
	PQ	0.13	n/a	1463.56
Dynamic	Ponly	0.11	0.14	1644.43
	Qonly	0.14	0.33	1938.82
	PQ	0.10	0.35	2528.67

In all treatments, log-likelihood ratio tests reject the null hypotheses that γ=0, or $\gamma=\infty$ (p-values < 0.001). We can conclude that subjects are neither perfectly rational ($\gamma=\infty$) nor did they make totally random decisions (γ=0). In other words, while subjects respond to profit incentives in making decisions, they are not capable of full optimization. The behavioral model proposed can capture the level of decision variations due to bounded rationalities. Furthermore, we note that estimates in PQ are significantly *lower* than those in Ponly and Qonly under each of the two inventory settings (all-values < 0.05). This again demonstrates that subjects have a higher propensity to make mistakes (lower gamma) when they can manage inventory through both pricing and ordering decisions (i.e., more complexity in decision making). Note that it is not possible to compare gamma estimates between the newsvendor and dynamic settings due to different profit scale under the respective models.

In treatments under the dynamic setting, log likelihood ratio tests reject the hypothesis that $\beta = 0$ with p-values < 0.001. This is strong evidence that subjects are not entirely myopic and do consider the future payoff while deciding for the current period. Moreover, we can reject the hypothesis that $\beta = 1$ in all three treatments with p-values < 0.001. This result cannot be explained by the discounting behavior. In our experiments, subjects are paid in cash right after they finish the game. Hence, we expect subjects to treat a dollar made in the last period the same as a dollar earned in the first period. An alternative explanation is that, due to limited cognitive ability, decision makers have to reduce the weight placed on periods further into the future while calculating the expected payoffs. This behavior is consistent with what was found in the beer game: subjects underweight the supply line, which are shipments to be received in the future (Sterman 1989).

Figure 6 provides an illustration for the goodness-of-fit of our behavioral model. The graph is generated using data and estimates from the dynamic Qonly treatment (it is picked for convenience. Similar graphs can be obtained from other treatments). In particular, we plot the observed frequencies of the order-up-to level"freq"), versus predictions from a myopic model with β=0 ("myopic"), and the limited-forward looking model given the estimated β("beta"). As we can see, the limited-forward looking model is an improvement over the myopic model, which is statistically confirmed by log likelihood ratio tests as well.

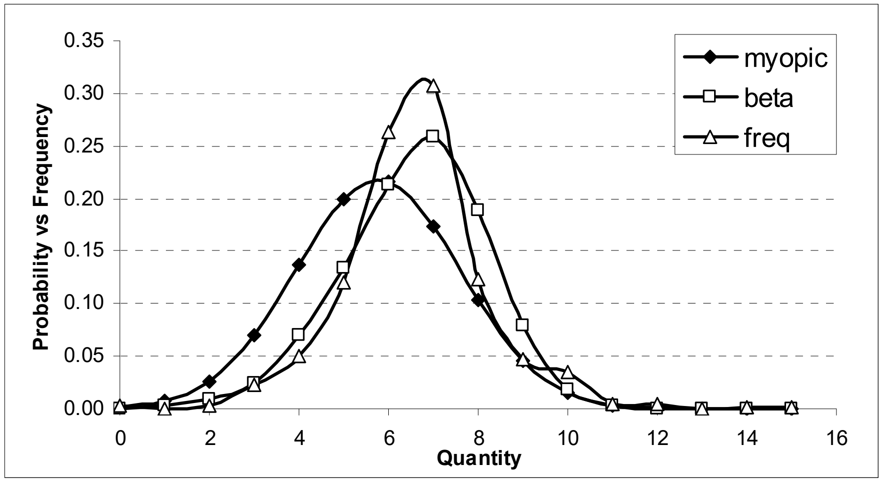

FIGURE 6: GOODNESS-OF-FIT OF THE BEHAVIORAL MODEL UNDER QONLY IN THE DYNAMIC SETTING.

Lastly, we check if there is any behavioral adjustment over time with the help from the behavioral model. Behavioral parameters are estimated for the first half and the last half of the game (each with 18 periods). Table 4 summarizes the results. Under the dynamic setting, we note that gamma increases while beta decreases from the first to the second half of the game significantly in *all* three treatments as shown by the likelihood ratio tests. The increase in gamma can be interpreted as subjects "learned" to reduce errors with more experience. On the other hand, the decrease in beta seems to suggest that subjects tend to be more myopic over time. We speculate that, since the calculation of future payoffs becomes more challenging due to the convolution of probabilities, subjects may choose to focus more on improving the evaluation of the current period, and thus pay less attention to the future periods. Under the newsvendor setting, we observe similar increasing trend in the gamma estimates. However, none of them is statistically significant. This observation is line with Bolton and Katok (2008), which showed learning from experience is difficult and slow under the newsvendor problem.

TABLE 4: TIME TREND IN BEHAVIORAL ESTIMATES.

Treatment			Gamma	Beta	LRT p-value
Dynamic	Ponly	1-st Half	0.09	0.15	0.0004
		2-nd Half	0.14	0.13	
	Qonly	1-st Half	0.13	0.38	0.0150
		2-nd Half	0.16	0.28	
	PQ	1-st Half	0.07	0.43	0.0000
		2-nd Half	0.14	0.28	
Newsvendor	Ponly	1-st Half	0.26	n/a	0.3527
		2-nd Half	0.33	n/a	
	Qonly	1-st Half	0.25	n/a	0.8229
		2-nd Half	0.28	n/a	
	PQ	1-st Half	0.11	n/a	0.0810
		2-nd Half	0.14	n/a	

VI. CONCLUSIONS AND FURTURE WORK

We present a behavioral study that systematically compares empirical decision making under two classical inventory management models (single-period versus multi-period) in the operations management literature. Our research differs from previous studies in that agents are provided with different options to control inventory, which affects the complexity of their decision making. Using controlled experiments, we identify biases of underpricing and understocking in human subjects. Their profit performance does not necessarily improve with more inventory control options. This contradicts optimization theories and we coin this phenomenon the "Optimizer's Paradox". A behavioral intuition behind this result is that more decision freedom or options may allow more room for human errors. We propose a behavioral model to capture behaviors observed in the laboratory. The model is developed based on two principles: agents are prone to make mistakes in utility evaluations, and they have limited capability to look forward under a dynamic setting. The behavioral model is successful in prediction decision biases and profit comparisons from the experiments. Estimations from the model further help understand adjustment in the empirical behaviors. The principles used for modeling are not limited to the specific experimental settings of this study.

This study offers several managerial implications. If more decision freedoms can lead to worse-off empirical performance, reducing inventory control options may become a necessary strategy in practice. Feng and Gallego (1995) provide field evidence that even in industries that adopt dynamic pricing policies, many companies still restrict the number of price changes during the selling horizon, and such restrictions become an important management strategy in the real world. Results from this study also challenge the popular notion of "employee empowerment" in the organizational behavior literature, which encourages autonomy and independent decision makings. It is important to point out that decision restrictions do not necessarily guarantee better empirical performance. As evidenced by observations in Ponly under the dynamic setting, when the inventory management policy is structured suboptimally, simply reducing the control options may not help improve decision makings. Therefore, it is important to tailor inventory control options based upon specific scenarios. The model we propose in the study can provide such a framework to understand how restrictions should be optimized behaviorally.

There are several limitations of the study, which future research can build upon. First, we only consider a dynamic setting without any delays. In practice, the replenishment decisions are often subject to lead times whereas pricing adjustment is relatively immediate. A natural research extension is to study the effect of inventory control options that differ in feedback delays. It is also interesting to apply our model to such dynamic scenarios to see if its predictions are robust. The current behavioral model is static and does not include behavioral preferences such as risk or loss aversion. Future research may consider the modeling framework of Experience Weighted Attraction (EWA) by Camerer and Ho (1998) to incorporate more behavioral effects into the discussion. Another direction for future work is to incorporate strategic interaction into the theory. For example, consider the inclusion of an upstream player who can modify the transfer price dynamically while determining his own production level. This requires a series of new experiments using a modified version of the beer game. It also demands to extend the current choice model to an equilibrium one to capture behavioral interactions. We consider this research offers an important first step to bridge the gaps.

REFERENCES

Anderson, S.P., De Palma, A., and Thisse, J.F. (1992), *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge, MA.

Bearden, J.N., Murphy, R. O. and Rapoport, A. (2008), Decision biases in revenue management: some behavioral evidence. *Manufacturing & Service Operations Management*, Vol. 10, Issue 4, pp. 625 - 636.

Bendoly, E., Donohue, K. and Schultz, K.L. (2006), Behavior in operations management: assessing recent findings and revisiting old assumptions. *Journal of Operations Management*, Vol. 24, Issue 6, pp. 747-752.

Bertsekas, D.P. (1976), *Dynamic programming and stochastic control*, New York: Academic Press, pp. 395.

Block, H.D. and Marschak J. (1960), Random ordering and stochastic theories of response. In I. Olkin, ed. *Contributions to Probability and Statistics*, Stanford University Press, Stanford, CA, pp. 97–132.

Bloomfield, R.J. and Kulp, S.L. (2013), Durability, transit lags, and optimality of inventory management decisions. *Production and Operations Management*, Vol. 22, No. 4, pp. 826-842.

Bolton, G.E. and Katok, E. (2008), Learning-by-doing in the newsvendor problem: a laboratory investigation of the role of the experience. *Manufacturing and Service Operations Management*, Vol. 10, pp. 519-538.

Busemeyer, J.R. (1999), International Encyclopedia of the Social and Behavioral Sciences: Methodology. *Mathematics and Computer Science,* Amsterdam: Pergamon.

Camerer, C. and Ho, T. (1999), Experience-weighted attraction learning in normal form games. *Econometrica*, Vol. 67, No. 4, pp. 827-874.

Chen, K.Y. and Wu, Y. (2020). Learning under the Economic Order Quantity problem. SJSU Working paper.

Diehl, E. and Sterman, J.D. (1995), Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Process*, Vol. 62, Issue 2, pp. 198-215.

Edwards, W. (1962), Dynamic decision theory and probabilistic information processing. *Human Factors.* Vol. 4, pp. 59-73.

Federgruen, A. (1993), Cebtralized planning models for multi-echelon inventory systems under uncertainty. *Logistics of Production and Inventory*, pp.133-173.

Federgruen, A. and Heching, A. (1999), Combined price and inventory control under uncertainty. *Operations Research*, Vol. 43, Issue 3, pp. 454-475.

Feng, Y. and Gallego, G. (1995), Optimal starting times for end-of-season sales and optimal stopping times for promotional fares. *Management Science*, Vol. 41, No. 8, pp. 1371-1391

Ho, T.H. and Weigelt, K. (1996), Task complexity, equilibrium selection, and learning: an experimental study. *Management Science*, Vol. 42, No. 5, pp. 659-679.

Johnson, E. and Payne, J. (1985), Effort and accuracy in choice. *Management Science*, Vol. 31, No. 4, pp. 395-414.

Katok, E., Donohue, K. and Leider, S. (2018), The Handbook of Behavioral Operations. Wiley.

Kremer, M., Mantin, B. and Ovchinnikov, A. (2017), Dynamic Pricing in the Presence of Myopic and Strategic Consumers: Theory and Experiment. *Production and Operations Management*, Vol. 26, No. 1, pp. 116-133.

Lim, N. and Ho, T. (2007), Designing price contracts for boundedly rational customers: does the number of blocks matter? *Marketing Science*, Vol.2 6, No. 3, pp. 312-326.

Luce, R. D. (1959), *Individual Choice Behavior: A Theoretical Analysis,* Wiley, New York.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior. In P. Zarembka, ed. *Frontiers in Econometrics*, Academic Press, New York.

McFadden, D. (1976), Quantal choice analysis: a survey. *Annals of Economic and Social Measurement*, Vol. 5, No. 4, pp. 363-390.

Morgan, B.J.T. (1992), Analysis of Quantal Response Data. *Monographs on Statistics and Applied Probability*, Chapman & Hall: London.

Palfrey, T.R. and McKelvey, R.D (1995), Quantal response equilibria for normal form games. *Games and Economic Behavior*, Vol. 10, Issue 1, pp. 6-38.

Perera, H.N., Fahimnia, B. and Tokar, T. (2020), Inventory and ordering decisions: a systematic review on research driven through behavioral experiments, *International Journal of Operations & Production Management*, Vol. 40, No. 7/8, pp. 997-1039.

Petruzzi, N.C. and Dada, M. (1999), Pricing and the newsvendor problem: a review with extensions. *Operations Research*, Vol. 47, Issue 2, pp. 183-194.

Porteus, E. L. (2002), *Foundations of Stochastic Inventory Theory*, Stanford University Press, Stanford, CA.

Rapoport, A. (1966), A study of human control in a stochastic multistage decision task. *Behavioral Science,* Vol. 11, pp. 18-32.

Rapoport, A. (1967), Variables affecting decisions in a multistage inventory task. *Behavioral Science,* Vol. 12, Issue 3, pp. 194-204.

Rapoport, A. (1975), Research paradigms for studying dynamic decision behavior. In D. Wendt & C. Vlek, eds. *Utility, Probability, and Human Decision Making,* Dordrecht-Holland: Reidel, pp. 347-369.

Rubinstein, A. (1998), *Modeling Bounded Rationality*, MIT Press, Cambridge, MA.

Schweitzer, M.E. and Cachon, G.P. (2000), Decision bias in the newsvendor problem with known demand distribution: experimental evidence. *Management Science*, Vol. 46, pp. 404-420.

Simon, H. A. (1955), A behavioral model of rational choice. *Quarterly Journal of Economics,* Vol. 69, Issue 1, pp. 99–118.

Simon, H. A. (1982), *Models of Bounded Rationality*, MIT Press, Cambridge, MA.

Sterman, J. (1989), Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. *Management Science*, Vol. 35, No. 3, pp. 321-339.

Su, X. (2008), Bounded Rationality in Newsvendor Models. *Manufacturing & Service Operations Management*, Vol. 10, No. 4, pp. 566-589.

Thowsen, G.T. (1975), A dynamic, nonstationary inventory problem for a price/quantity setting firm. *Naval Research,* pp. 305-321.

Tversky, A. and Kahneman, D. (1971), Belief in the law of small numbers. *Psychological Bulletin*, Vol. 76, pp.105-110.

Veinott, A.F. Jr. (1965), Optional policy for a multi-product, dynamic, nonstationary inventory problem. *Management Science*, Vol. 12, Issue 3, pp. 206-222.

Wu,Y. and E. Katok. (2006), Learning, communication, and the bullwhip effect, *Journal of Operations Management*, Vol. 24, Issue 6, pp. 839-850