Clustering Human Decision-Making in Production and Logistic Systems

Christos Tsagkalidis, Rémy Glardon, Maryam Darvish

HAL Id: hal-01417508

https://hal.science/hal-01417508

Submitted on 15 Dec 2016

Clustering Human Decision-Making in Production & Logistic Systems

Christos Tsagkalidis[1], Rémy Glardon[1], Maryam Darvish[2]

[1]Laboratory for Production Management and Processes, Swiss Federal Institute of Technology at Lausanne (EPFL), CH-1015 Lausanne, Switzerland
christos.tsagkalidis@epfl.ch, remy.glardon@epfl.ch
[2]Faculté des Sciences de l'Administration, Université Laval, Québec, Canada
maryam.darvish.1@ulaval.ca

Abstract. Human decisions play an essential role in Operations and Supply Chain Management. However, these decisions are rarely integrated in simulation models of Production and Logisitc Systems. One main reason for this fact is the strong dispersion of human decisions among a population, as well as the variability of a single individual's decision over time. This work presents an experimental study of a human decision consisting in the dynamic selection of suppliers in a well-controlled laboratory environment. The analysis of the results obtained on a large population shows that individual decision behaviors can be grouped into representative clusters typifying different decision behaviors. The results obtained from this study opens up the prospect to significantly reduce the number of decision models required to simulate Production and Logistic Systems including human decisions and could also allow categorizing human decision behavior based on a set of known criteria.

Keywords: Decision-making, Behavioral Operations Management, Cluster Analysis

1 Introduction

A review of papers published in the field of operations management reveals that behavioral operations defined as "*the study of human behavior and cognition and their impacts on operating systems and processes*" (Gino and Pisano 2008) has gained considerable attention during the past few years. One stream of behavioral supply chain studies tries to identify the causes of instabilities within the supply chains using laboratory experiments and reveal some behavioral assumptions that are commonly held true in normative supply chain literature. Following Sterman (1989) a majority of the research on the behavioral causes of the bullwhip effect is using the Beer Distribution Game (BDG), which mimics a linear supply chain consisting of four echelons of retailer, wholesaler, distributor and factory. Studies show how the bullwhip effect could be reduced; shorter lead-times (Steckel, et al., 2004), sharing point of sale data (Croson and Donohue 2003) and inventory information (Croson and Donohue 2006) have shown to be important in reducing the overall

supply chain cost. The role human factors play in supply chain decision-making is also investigated; overreaction to backlogs (Oliva and Gonçalves, 2005), the role of trust on the inventory replenishment decision (Kaboli, et al., 2012), the role and impacts of power in distribution channels (Das Guru, et al., 2014) are very well studied.

To understand human behavior, the research in behavioral operations tends to study the aggregated behavior; Croson and Donohue (2006) pinpoint the need for more theoretical research analyzing the behavior at the individual level but much of the literature ignores the individual differences. Furthermore, the integration of human decision behavior in simulation models cannot efficiently be performed on an individual level, without identification and modeling of representative behaviors. Accordingly, the main goal of this study is to investigate if and how individual decision-behaviors can be adequatly grouped in representative clusters. To this end, a controlled experimental environment, a participatory simulation platform is implemented. More specifically, the human decision studied in this work consists in dynamically allocating the market demand to three competing supply chains in accordance with their delivery performances.

2 Experimental Procedure and Methodology

2.1 Experimental procedure

The use of participatory simulation as the experimental framework is chosen, as it is known to provide a stable and controlled environment (Katok, 2011) that allows reliable observations and eliminates most of the external perturbations (Das Guru et al., 2014). In this study, a participatory simulation platform developed at Université Laval, Québec, Canada, called XBG-platform, is used (Montreuil et al. 2008) which mimics the dynamics of inventory replenishment in a decentralized, linear, four echelon supply chain. The chosen experimental set-up is illustrated in Fig. 1. It is composed of three competing supply chains delivering products to a continental distributor that must satisfy the volatile market demand. The four echelons of each supply chain are operated by software agents. The human agent plays the role of the continental distributor; his/her decision consists of continuously allocating the market demand to the three supply chains. This decision is modeled by an algorithm called Demand Modulation Algorithm (DMA) that simulates the buyer's decision-making behavior by including the effects of perception, subjective judgment and memory, through a set of 11 parameters (Tsagkalidis, 2014). Thus, each individual behavior can be characterized by a vector of 11 parameters (Individual Decision Behavior Vector, IDBV) acquired by best fitting the algorithm to the obtained experimental results (human agent behavior). The inputs of the DMA are 3 quantitative parameters (Delivery performance, Order refusal, Order cancellation). The DMA ouputs are the shares of the demand allocated to each of the three supply chains. The computation takes into account behavioral aspects such as assymetry of the reaction to good or bad performances (i.e. gain or loss of trust in the supplier).

The experiment lasts approximately 40 minutes, which is equivalent to a 3.5-month business horizon. Through an experimental campaign, 278 individual results; i.e 278 IDBVs have been acquired and stored, providing a statistically significant sample. Furthermore, information about possibly relevant individual characteristics of each human agent is also stored for later use.
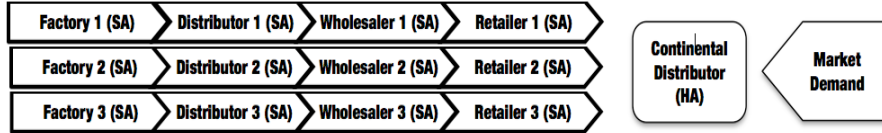

Fig. 1. Structure of the experimental set-up (SA, software agent; HA, human agent)

The set of 278 IDBVs is used for the clustering process that aims at building groups (clusters) of individuals having similar decision behaviors within a given cluster and dissimilar between different clusters. The number of possible clusters is not defined a priori but results from the clustering process.

2.2 Methodology

The methodology consists of 4 main steps: 1) experiment with 278 individual human agents, 2) determination of the 278 IDBV through fitting of the DMA, 3) clustering of the 278 IDBV into k clusters, 4) determination of the representative behavior of each cluster (centroid coordinates) that leads to definition of k Cluster Decision Behavior Vectors (CDBV), 5) study of possible relations between clusters and individual human agent characteristics. It must be noted that the optimal number of clusters k is not known in advance.

2.3 Clustering procedure

The clustering is conducted using SPSS (version 16) from IBM Company. The initial step is the definition of the clustering procedure. Generally two different approaches are available: hierarchical and non-hierarchical clustering. In the current case, as the number of clusters is not known beforehand and the data size is relatively small, a hierarchical, agglomerative clustering procedure is chosen. This means that the process starts by considering each individual as a single cluster and proceeds with a progressive merging process. The main steps are as follows:

- **Selection of distance measurement and clustering algorithm.** The Square Euclidian Distance criterion (Hair et al. 1995, Grimm et al. 2000) is chosen for similarity measurement between cases and is implemented for all variables. The Ward's method (Ward. 1993) is selected as the clustering algorithm. This criterion creates clusters characterized by variables having as much similarity as possible (smallest variation) within clusters and as much dissimilarity as possible (biggest variation) between clusters. Ward's criterion is chosen as it works well with

Squared Euclidean Metrics, attempts to reduce each cluster's overlay by minimizing the Sum of Squares Error, produces clusters with low variability and is considered to be computationally intensive, while it is statistically simple (Murtagh and Legendre, 2014).

- **Determination of the number of clusters.** The optimal number of clusters represents a critical issue in a hierarchical, agglomerative clustering process. Two indicators are used to support the determination of the cluster number: the fussion coefficient from the agglomeration schedule and the dendrogram.
- **Output of clustering analysis.** In each cluster, each human agent is defined by the 11 parameters of the IDBV and a set of individual characteristics. The following results are finally obtained as an output of the clustering process: 1) Representative behavior of each cluster (11 centroid coordinates, CDBV), 2) Standard deviation of the 11 parameters within the cluster, 3) percentage of each individual characteristics within the cluster.

3 Results of the Clustering analysis

As stated above, the decision behavior is characterized by the IDBV. Not all 11 parameters have the same significance concerning the clustering process; two of the parameters are constant in the DMA and four show non-significant variations among human agents. Consequently, the clustering process is based on 5 relevant parameters (see 3.1) of the IDBV.

The number of clusters that leads to a maximum value of the fussion coefficient is found to be $k = 3$ or 4. Additionally, the use of dendrogram indicates the choice of a value $k = 3$. The main characteristics of the main 3 clusters are indicated in Table 1. The CDBVs and the standard deviations of the parameters within clusters are not provided here, as their single values do not have much significance.

Table 1. Final cluster size and characteristics

Cluster	Cluster population	Population percentage
1	62	22.3 %
2	29	10.4 %
3	187	67.3 %
Total	278	100 %

3.1 Cluster Population Characteristics

Based on previous experience and the typology of the experiment participants, the population is characterized by the following criteria:

- **Master Study Field.** Table 2 indicates the frequency of the Master study field per cluster. In all 3 clusters the ratio of population not holding a master degree is still significant (13.8 to 21%). Furthermore, it can be noticed that almost half of the

population in cluster 1 owns a master in supply chain management, while this ratio is much lower in clusters 2 and 3. Finally, cluster 3 is singled out by a large population ratio (total 28.3%) with Master in Business or Basic sciences, this ratio is much lower in clusters 1 and 2 (4.8%, 13.8%).

Table 2. Master study field

Cluster		Engineering	Life Sciences	Business & Economics	Basic Sciences	Supply Chain	None	Total
1	Count	12	4	3	0	28	13	62
	% cluster population	19.4%	6.5%	4.8%	0.0%	45.2%	21.0%	100.0%
2	Count	9	0	2	2	3	4	29
	% cluster population	31.0%	0.0%	6.9%	6.9%	10.3%	13.8%	100.0%
3	Count	51	10	24	29	9	29	187
	% cluster population	27.3%	5.3%	12.8%	15.5%	4.8%	15.5%	100.0%
Total	Count	72	14	29	31	40	46	278
	% cluster population	25.9%	5.0%	10.4%	11.2%	14.4%	16.5%	100.0%

- **Familiarity with simulation platforms.** Cluster 1 is characterized by a large population ratio having previous experience with simulation platforms in general (58.1%) and with the XBG-platform in particular (24.2%) as shown in Table 3. On the contrary, the ratio of the population with general simulation platform experience drops too much lower values in clusters 2 and 3 (17.2%, 20.9%) and the ratio with XBG-experience is even lower (3.4%, 5.9%).

Table 3. Familiarity with simulation platforms in general and XBG in particular

Cluster		Familiarity with simulation platform in general		Familiarity withXBG-simulation platform		Total
		Yes	No	Yes	No	
1	Count	36	26	15	13	62
	% cluster population	58.1%	41.9%	24.2%	21.0%	100.0%
2	Count	5	24	1	11	29
	% cluster population	17.2%	82.8%	3.4%	37.9%	100.0%
3	Count	39	148	11	88	187
	% cluster population	20.9%	79.1%	5.9%	47.1%	100.0%
Total	Count	80	198	27	112	278
	% cluster population	28.8%	71.2%	9.7%	40.3%	100.0%

- **Acquaintance with supply chain management.** Participants' prior knowledge of SCM is considered important, as due to the experimental set-up, phenomena typical of suply chains are expected. The results presented in Table 4 reveal that the highest population ratio with SCM experience is to be found in cluster 1 (75.8%) and the lowest in cluster 3 (50.8%), cluster 3 being close to cluster 2 (55.2%).

Table 4. Acquaintance with supply chain management

Cluster		Familiarity with Supply Chain Management		Total
		Yes	No	
1	Count	47	15	62
	% cluster population	75.8%	24.2%	100.0%
2	Count	16	13	29
	% cluster population	55.2%	44.8%	100.0%
3	Count	95	92	187
	% cluster population	50.8%	49.2%	100.0%
Total	Count	158	120	278
	% cluster population	56.8%	43.2%	100.0%

- **Working experience.** As shown in table 5, roughly 25% of the population in clusters 1 and 3 has no working experience, while this ratio is lower for cluster 2 (17.2%). However, cluster 1 is also characteried by the lowest population ratio (43.5%) without any specific working experience in logistics and procurement; this ratio is the highest for cluster 3 (76.5%). About specific working experience in logistics or procurement, the difference between the three clusters becomes more evident. Cluster 1 presents the highest population ratio with experience (total 56.4%), cluster 2 a medium value (37.9%), and cluster 3 the lowest ratio (total 23.6%).

Table 5. Working experience in general and in logistics or procurement in particular

Cluster		Working experience in general (years)			Working experience in logistics & procurement			Total
		none	1-5	>5	none	1-5	>5	
1	Count	15	30	17	27	26	9	62
	% cluster population	24.2%	48.4%	27.4%	43.5%	41.9%	14.5%	100.0%
2	Count	5	15	9	18	6	5	29
	% cluster population	17.2%	51.7%	31.0%	62.1%	20.7%	17.2%	100.0%
3	Count	48	107	32	143	22	22	187
	% cluster population	25.7%	57.2%	17.1%	76.5%	11.8%	11.8%	100.0%
Total	Count	68	152	58	188	54	36	278
	% cluster population	24.5%	54.7%	20.9%	67.6%	19.4%	12.9%	100.0%

3.2 Cluster Population, Synthesis

The main features of the three clusters are summuraized in Table 6, according to the most significant population characteristics described above. These results indicate that cluster 1 contains a meduim size population (62 members) with the strongest background in SCM, logisitcs and simulation. Thus, it groups the participants with the closest and stongest knowledge about decision-making issues in Supply Chain. Cluster 2, with the smallest population (29 members) is characterized by the largest population ratio with background in engineering and general working experience. Finally, cluster 3 is by far the largest with 67.3% of the total population (187 members). It is characterized by the least familiarity with SCM professional experience and the largest population ratio with a business and economics education.

Table 6. Synthesis of the cluster population characteristics

Cluster Population		Cluster 1 62	Cluster 2 29	Cluster 3 187
a) Master study field	Engineering	L	H	
	Business & Economics	L		H
	Supply Chain	H		L
	None	H	L	
b) & c) Familiarity with simulation platforms	In general	H	L	
	XBG platform	H	L	
d) Familiarity with SCM	yes	H		L
e) & f) Working experience	In general		H	L
	Logistics & procurement		H	L
Lowest population ratio		L		
Highest population ratio		H		

4 Discussions and Conclusion

The objective of this work is to study the possibility of grouping human decision-behavior in clusters to simplify the integration of decision-making processes in Supply Chain simulation models. The human decision investigated in this work continuously allocates the market demand to 3 competing supply chains. An experimental campaign is performed using a participatory simulation platform with continuous time scale. The experiment is run with a statistically significant sample of 278 human participants.

Each human decision-maker (agent) is characterized using a vector (IDBV) of 11 parameters obtained by best fitting a decision algotihm developed in a previous study (Tsagkalidis, 2014). These experimental results (3058 single values) constitute the database used for a clustering process that aims at identifying typical and representative decision-behaviors. A hierarchical clustering procedure that uses the Square Euclidian Distance criterion as similarity measurement is chosen. Clustering is performed according to Ward's clustering algorithm.

A solution with three clusters arrises from the clustering process. They are characterrized by their polulation size and their centroid, an 11-parameter vector called Cluster Behavior Decision Vector (CDBV) that represents the average decision-behavior of the cluster population.

Each human agent is further described by a set of individual characteristics. A first attempt at identifying links between clusters, i.e. typical decision-behavior, and individual characteritics leads to some preliminary conlusions. They show in particular that previous knowledge and experience in Supply Chain Management appears to have a significant influence on the decision-making behavior. Similarly, an engineering education with no professional experience seems to lead to another specific decision behavior. According to the current results, these two categories appear to build specific groups that single out from the majority of participants characterized by low professional experience and no or weak previous knowledge in SCM.

A further extension of this work is to study how experience influences the decision process as previous limited observations seem to indicate a more random and

nervous reactions from less experienced people. The relations between cluster average decision behavior (CDBV) and financial performances in the experiment would also bring interesting conclusions concerning a possible economically optimal decision behavior. Finally finding possible links between individual characteristics and decision behavior constitutes the ultimate goal of this general field of research. Once these relations are identified, they could be validated by reversing the experiment; i.e. by selecting participants belonging a priori to a cluster and verifying if they do behave according to the expectations.

5   References

1. Croson, R., and Donohue, K. (2003). "Impact of POS Data Sharing on Supply Chain Management: An Experimental Study." Production and Operations Management 12:1–11.
2. Croson, R., and Donohue, K. (2006). "Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information." Management Science 52:323–36.
3. Das Guru, R., Kaboli A. and Glardon R. (2014). "The Effect of Coercive Power on Supply Chain Inventory Replenishment Decisions." Pp. 230–37 in Advances in Production Management Systems. Innovative and Knowledge-Based Production Management in a Global-Local World,. Springer Berlin Heidelberg, 2014.
4. Hair J., Joseph, F., Anderson E., Tatham R., Ronald L., and Black. W. C. (1995). Multivariate Data Analysis (4th Ed.): With Readings. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
5. Gino, F., and G. Pisano. (2008). "Toward a Theory of Behavioral Operations." Manufacturing & Service Operations Management 10:676–91.
6. Grimm, L.G. and Yarnold, P.R. 2000. Reading and Understanding More Multivariate Statis- tics. American Psychological Association.
7. Kaboli, A., Cheikhrouhou, N. Darvish M. and Glardon R.. (2012). "An Experimental Study of the Relationship between Trust and Inventory Replenishment in Triadic Supply Chain." in Proceedings of the POMS world conference.
8. Katok, E. (2011). "Using laboratory experiments to build better operations management models". Foundations and Trends in Technology, Information and Operations Management, 5(1):1–86.
9. Murtagh, F., Legendre, (2014). P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? Journal of Classification, 31(3), 274-29.
10. Oliva, R., & Gonçalves, P. (2005). Behavioral Causes of Demand Amplification in Sypply Chains: "Satisficing" Policies with Limited Information Cues. In *Proceedings of the 2005 System Dynamics Conference* (pp. 118–119).
11. Steckel, J. H., Gupta S. and Banerji, A. (2004). "Supply Chain Decision Making: Will Shorter Cycle Times and Shared Point-of-Sale Information Necessarily Help?" Management Science 50:458–64.
12. Sterman, J. D. (1989). "Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment." Management Science 35:321–39.
13. Tsagkalidis, C. (2014). "Study of Human Decision Behavior for the Operational Selection of Suppliers in a Compatitive Framework using a Participatory Simulation Platform". Master Thesis, EPFL, March 2014.
14. Ward, J. H. (1963). "Hierarchical grouping to optimize an objective function". Journal of the American Statistical Association, 58(301): 236–244.