Design, Implementation and Test of Collaborative Strategies in the Supply Chain

Thierry Moyaux[1], Brahim Chaib-draa[1], and Sophie D'Amours[2]

[1] Universit Laval, Dpt. d'Informatique et de Gnie Logiciel, DAMAS & FOR@C
Ville de Qubec G1K 7P4 (Qubec, Canada), {moyaux, chaib}@iad.ift.ulaval.ca,
[2] Universit Laval, Dpt. de Gnie Mcanique, FOR@C & CENTOR
Ville de Qubec G1K 7P4 (Qubec, Canada), sophie.damours@gmc.ulaval.ca,

Abstract. In general, game theory is used to analyze interactions formally described by an analytical model. In this paper, we describe a methodology to replace the analytical model by a simulation one in order to study more realistic situations. We use this methodology to study how the more-or-less selfishness of agents affects their behaviour. We illustrate our methodology with the case study of a wood supply chain, in which every company is seen as an agent which may use an ordering strategy designed to reduce a phenomenon called the bullwhip effect. To this end, we assume that every agent utility can be split in two parts, a first part representing the direct utility of agents (in practice, their inventory holding cost) and a second part representing agent social consciousness, i.e., their impact on the rest of the multi-agent system (in practice, their backorder cost). We find that company-agents often apply their collaborative strategy at whatever their same level of social consciousness. Our interpretation of this specific case study is that every company is so strongly related with one other, that all should collaborate in our supply chain model. Note that a previous paper outlined this methodology and detailed its application to supply chains; our focus is now on the presentation and the extension of the methodology, rather than on its application to supply chains.

1 Introduction

There is an increasing interest in game theory in the multi-agent community [1]. In particular, some methodologies have applied concepts from game theory to analyze multi-agent simulations [2, 3], even though, in general, game theory is used to analyze the behaviour of agents described with an analytical model. Our ultimate goal is to have a methodology suitable for the analysis of interactions in agent-based simulations that would profit from most concepts from game theory. To this end, we now enhance the methodology we proposed in a previous work [2] to study (i) how the obtained results vary with some change in the utility function of agents, and (ii) how robust the results are with regard to the stochasticity of some parameters.

Let us focus on these two enhancements. Concerning the first point, agents in our previous work had a common fixed way to calculate their utility that was based on their direct utility and on their impact on other agents. In contrast, the methodology proposed in the present paper, the way in which utility is calculated is *neither common* to all agents, *nor fixed*. In practice, we continue to assume that agent utility is the addition of two components: the direct utility that implements the individual utility of agents, and a utility that represents the social consciousness of agents in regard to other agents. In [2], the weight of the direct utility was half of the weight of the social utility for every agent; we now change this ratio to verify that agents still have incentives to use the same strategy, or conversely, if agents prefer to change their behaviour when they are more or less socially conscious. The second improvement to our methodology deals with taking *probability distribution* for some parameters into account in replacement of a fixed sequence of numbers, in order to study stochastic models. This means that we generate an instance of the considered distribution, and then, we apply our methodology with this fixed sequence of numbers. This is the same method as in our previous work, except that we repeat the process with other instances of the considered distribution in order to next use statistical tools. Our methodology is described in Section 2.

We illustrate this methodology in the context of the management of a wood supply chain. In other words, we present our methodology as generally as possible, while its application to a specific supply chain demonstrates it. We profit from the first enhancement provided by our methodology to study whether selfish as well as benevolent company-agents should collaborate by sharing information concerning market consumption. This shared information is of importance to companies when they place orders, because this information may be distorted by a phenomenon called the "bullwhip effect" [4] that causes an increase of their inventory level and a decrease in their customer service level, and information sharing is the most often proposed solution to the bullwhip effect [5].

We take the direct utility as agent profit, and social utility as the level at which company-agents take client satisfaction into account in their utility because the goal of the entire supply chain should be to satisfy end-customers. We profit from the second enhancement provided by our methodology to consider a stochastic market consumption pattern. More precisely, we assume that market demand is informally distributed, and we carry out our experiments with ten particular instances of this uniform distribution. We consider only ten instances of the distribution, because, for each of these ten instances, a great number of simulations is required by our methodology to simulate all the combinations of the possible strategies among the considered agents. However, all of these simulations are carried out with the same particular instance of the market consumption in order to build a game in the normal form. We next repeat all these simulations with the nine other considered instances of the market consumption in order to build nine other games. In the ten obtained games, we then change the ratio between the direct and the social utility of agents when we look for Nash equilibria. The experimental results show that the level of social consciousness does not have an impact on the choice of which level of collaboration is chosen by agents. In other words, selfish agents agree to collaborate as often as benevolent agents, which shows that every company-agent in our model is so strongly *dependent* on each other that they should all collaborate. This illustration of our methodology on an issue of supply chain management, namely the bullwhip effect, is detailed in Section 3.

Finally, we discuss how we have applied our methodology in practice to study supply chain management, and we identify some future enhancements on this methodology. Concerning our application to supply chains, we can note that we only consider scenarios in which all companies have the same level of social consciousness, while our methodology is more general and allows us to mix benevolent and selfish agents in a same supply chain, which would allow situations as the Prisoner's Dilemma to occur. We also note that we have only considered ten instances of the uniform distribution for the market consumption pattern, i.e., for the stochastic parameter. In both limitations of the application of our methodology to supply chains, we made these simplifications to save either analysis (first limitation) or simulation (second limitation) time, but we indicate how we expect to manage these two limitations in future work. Concerning the possible enhancements on our methodology, the main one deals with the addition of probability on the use of strategies, which corresponds to the concept of mixed strategies in game theory. Such an addition of probability in our context of supply chains is not as easy as in game theory, but it could be carried out in some contexts other than information sharing decisions in supply chains. The application of our methodology to supply chain management and the possible extensions of our methodology are discussed in Section 4.

2 Methodology

Before the presentation of our methodology, we need to introduce a few notations. For the sake of simplicity, we will continuously illustrate this section with the case study of a supply chain, even if this case study will only be detailed in Section 3.

2.1 Notations

We write $i \in I$ a particular agent, and $|.|$ the cardinal of a set, so that $|I|$ is the number of agents in the considered system. In the case study in Section 3, every agent i represents a company in a wood supply chain, and there are $|I| = 6$ agents. $s^i \in S$ denotes the strategy used by Agent i during an entire run of a simulation, and S is common to all agents, i.e., all agents have the same set of possible strategies, even if they do not use the same strategy during a particular simulation. Simulations are run over fifty weeks during which agents cannot change their strategy s^i. In the case study in Section 3, $S = \{\mu, \beta, \gamma\}$ for every company i to represent three levels of collaboration by information sharing strategies, where:

- μ implements no information sharing, so that placed orders are only based on inventory level. Notice that incoming demand (i.e., orders placed by the direct client) could also have been taken into account here.

– β represents little information sharing, because it allows companies to inform their direct suppliers of the market consumption, i.e., of what is demanded by end-customers. Specifically, every retailer informs its wholesalers of its sales when it places an order, then every wholesaler transmits this information to its suppliers, etc.
– γ implements full information sharing, when retailers broadcast the market consumption (approximated as their sales) to the rest of the supply chain. Such information sharing is referred as information centralization. γ shares the same information as β, i.e., the market consumption information, but instantaneously instead of as slowly as orders.

Agent i's utility is $u^i_{s^i}$ when it uses the strategy s^i, but for the purpose of simplicity we will always write u^i instead of $u^i_{s^i}$. To study social consciousness, we assume that every utility u^i is the convex combination of the direct utility u^i_{dir} and of the social utility u^i_{soc}, so that $u^i = ((1-\epsilon^i)u^i_{\text{dir}} + \epsilon^i u^i_{\text{soc}})/(1-\epsilon^i)$, with $\epsilon^i \in [0;1)$; we divide by $(1-\epsilon^i)$ because u^i_{dir} represents realistic costs in our case study and u^i thus remains a cost (notice that $\epsilon^i = 2/3$ for all i in [2]). This splitting of utility function as direct vs. social utility is similar to others, such as the representation of individual outcome vs. social consciousness [6], or of self vs. group interest [7].

In particular, Glass and Grosz [6] note that computing a measure of social consciousness and combining this measure with individual outcome are two challenges. In our case study, we address these two challenges by defining u^i_{dir} as company-agent's inventory holding cost (thus, $u^i_{\text{dir}} < 0$ because the agents in our simulation cannot earn money), and u^i_{soc} as company-agent's backorder cost (again, $u^i_{\text{soc}} < 0$) to represent the more or less will to satisfy clients; backorders correspond to orders that were not filled in the past and that have to be shipped as soon as possible, or, in other words, there are no lost sales in the simulation model used in our case study because clients wait for product availability, and the backorder cost measures (industrial or end-) customer *dissatisfaction* due to this wait. Since dissatisfaction generates a bad reputation for agents, u^i_{soc} also measures the *reputation* that agents expect to have in the system. Since backorders are measured in our specific case as negative inventory levels, the calculation of u^i_{dir} and u^i_{soc} are similar, and thus, their aggregation into u^i represents a cost. Notice that, $\epsilon^i = 0$ means that Agent i only has a selfish behaviour because u^i_{soc} has no influence on u^i in the previous convex combination, while $\epsilon^i \to 1$ implements a fully benevolent agent because u^i_{dir} becomes negligible in comparison with u^i_{soc}. In the specific of our case study, we can reformulate the question addressed in this paper as:

> At what set of values $\{\epsilon^i\}_{i=1}^{i=|I|}$ do the Nash equilibria switch from an entire non-collaborating supply chain to an entire collaborating supply chain?

In this paper, the proposed methodology allows studying such a question, but the application of this methodology in the case study only addresses this

Algorithm 1 methodology

$\{\{Inv^{i,s}_{\text{init}}\}_{i=1}^{i=|I|}\}_{s\in S} = \textbf{optParam}(\{\epsilon^i\}_{i=1}^{i=|I|}, M)$
for each $m \in M$ **do**
  $\{\{u^i_{\text{dir}}, u^i_{\text{soc}}, s^i\}_{i=1}^{i=|I|}\}_{s^1\in S\times\ldots\times s^{|I|}\in S} = \textbf{simAllConfig}(\{\{Inv^{i,s}_{\text{init}}\}_{i=1}^{|I|}\}_{s\in S}, m)$
  sets of Nash equilibria $= \textbf{anaSimOut}(\{\{u^i_{\text{dir}}, u^i_{\text{soc}}, s^i\}_{i=1}^{|I|}\}_{s^1\times\ldots\times s^{|I|}}, \{\epsilon^i\}_{i=1}^{i=|I|})$
returns *all sets of the Nash equilibria corresponding to the considered* $\{\epsilon^i\}_{i=1}^{i=|I|}$ *and the* $|M|$ *instances of the stochastic parameter.*

question for sets $\{\epsilon^i\}_{i=1}^{i=|I|}$ in which $\epsilon^1 = \ldots = \epsilon^{|I|}$. This means that, in the case study, all agents have the same level of selfishness, but that this common level may be different between two of the considered situations. That is, we do not consider the case in which selfish agents cohabit with benevolent ones (i.e., Prisoner's Dilemma), although the presented methodology allows doing so. We make this assumption to simplify the analysis of the outputs of our methodology, but we will try to relax this assumption in future work, as discussed in Section 4. We call ϵ the common value of $\{\epsilon^i\}_{i=1}^{i=|I|}$, so that $\epsilon = \epsilon^i$ for every agent i.

In the previous reformulation of the question addressed in this paper, we specified that we focus on Nash equilibrium as the solution concept. Indeed, this equilibrium is the main solution concept in game theory [8], in particular, because it is the most general one and thus includes the other concepts of equilibrium. A Nash equilibrium is only a stable state of the multi-agent system, that is, no agent has incentives to unilaterally deviate from it. As a result, when we refer to such an equilibrium, we do not understand it as a description of what will always occur, but as what holds when it occurs. The issue of knowing what equilibria should be preferred by agents was addressed in [2]; now, we reconsider the same example of supply chains to study how equilibria change with ϵ.

Next, we write Inv^i_{init} the value of some parameter Inv^i at the beginning of a simulation. In the particular case of our study of a supply chain, Inv^i_{init} will be company-agent i's initial inventory level. Finally, m_A represents one particular instance of a probability distribution for a second parameter m. We write M the set of the considered instances of the probability distribution under study. In our case study, the distribution under study is the uniform distribution of 2×50 integer numbers over the interval $\{11, 12, \ldots, 17\}$, so that each of the 50 integers represents the consumption of one of the 2 markets during one of the 50 weeks of the simulation. We generate *once* 10 instances of this uniform distribution, so that $M = \{m_A, m_B \ldots m_J\}$.

2.2 Explanation of the Algorithms

With these notations, we can now describe our methodology, as outlined in Algorithms 1, 2, 3, and 4. Algorithm 1 outlines the methodology at the highest level:

(i) Parameters Inv are first optimized by **optParam** for a specific value of ϵ (Algorithm 1 is general, but as previously stated, we will only consider the sets

Algorithm 2 optParam

inputs: $\{\epsilon^i\}_{i=1}^{i=|I|}, M$
for each strategy $s \in S$ **do**
  for each considered instance $m \in M$ **do**
    for each agent $i \in I$ **do**
      set agent i to use $s^1 = s$
    optimize the initial value of Inv^i for every agent i, so that $\sum_i (u^i_{\text{dir}} + u^i_{\text{soc}})$ is maximized under m.
    save the set of optimal values $\{Inv^{i,s}_{\text{init}}\}_{i=1}^{|I|}$.
  save in a new set $\{Inv^{i,s}_{\text{init}}\}_{i=1}^{|I|}$ the average of the values in the $|M|$ sets $\{Inv^{i,s}_{\text{init}}\}_{i=1}^{|I|}$ saved in the previous loop.
returns $\{\{Inv^{I,s}_{\text{init}}\}_{i=1}^{|I|}\}_{s\in S}$.

$\{\epsilon^i\}_{i=1}^{i=|I|}$ in which $\epsilon^i = \epsilon$ for every company i, with $\epsilon = 1/6, 2/6, 3/6, 4/6$ or $5/6$). (ii) Next, the supply chain is simulated by **simAllConfig** for all combinations of s^i among the $|I|$ agents with the optimum values of Inv_{init}. (iii) Then, the simulation outcomes are used to build games in the normal form, in which Nash equilibria are next looked for. Depending on $\{\epsilon^i\}_{i=1}^{i=|I|}$, different games can be built here based on the output of **simAllConfig**. We now describe each step in this process.

Algorithm 2 presents how the optimum values of parameter Inv are found. This method returns the best intial value of Inv depending on the considered agent i and on what common strategy s is used by all agents (we only optimize for s common to all agents in order to save time by greatly reducing the number of optimizations) so that the social welfare of agents is maximized.

We assume here that the social welfare is the sum of all agent utilities $\sum_i (u^i_{\text{dir}} + u^i_{\text{soc}})$ [9] because this assumption is compatible with our case study in which this sum represents money, but, in general, social welfare can be much harder to define [10]. In brief, Algorithm 2 carries out $|M|$ simulations to find the best value of the parameter Inv and return the average of these $|M|$ values (we thus find values good on average in all instances in M), and repeat this process for every $s \in S$.

As a consequence, the word "optimize" in Algorithm 2 both refers to a simulation and to a parameter optimization.

In our case study, each invocation of **optParam** returns (i) a set of the six best values of Inv (one value per company-agent) when all agents use β, (ii) a set of the six best values of Inv when all agents use γ, (iii) and the set {0, 0, 0, 0, 0, 0} when all agents use μ. Concerning μ, we use the vector of zeros as optimal parameters, because we proved that $Inv^i_{\text{init}} = 0$ is locally optimal for every company i [2]. This local optimality is compatible with the fact that μ is a non-collaborative strategy.

Next, Algorithm 3 uses the values of Inv that are the best on average, and simulates with these average values all the possible combinations of the $|S|$ strate-

Algorithm 3 simAllConfig

inputs: $\{\{Inv_{\text{init}}^{i,s}\}_{i=1}^{|I|}\}_{s\in S}, m$
for each $s^1 \in S$ **do**
  set agent i to use s^1 with $Inv_{\text{init}}^{1,s^1}$.
  for each $s^2 \in S$ **do**
    set agent i to use s^2 with $Inv_{\text{init}}^{2,s^2}$.
    ⋮
    for each $s^{|I|} \in S$ **do**
      set agent $|I|$ to use $s^{|I|}$ with $Inv_{\text{init}}^{1,s^{|I|}}$.
      simulate the supply chain under instance m.
      save the simulated value of $\{u_{\text{dir}}^i, u_{\text{soc}}^i\}_{i=1}^{|I|}$ with the corresponding s^i.
returns $\{\{u_{\text{dir}}^i, u_{\text{soc}}^i, s^i\}_{i=1}^{|I|}\}_{s^1\in S\times\ldots\times s^{|I|}\in S}$

Algorithm 4 anaSimOut(m)

inputs: $\{\{u_{\text{dir}}^i, u_{\text{soc}}^i, s^i\}_{i=1}^{|I|}\}_{s^1\in S\times\ldots\times S^{|I|}\in S}, \{\epsilon^i\}_{i=1}^{i=|I|}$
for each considered set $\{\epsilon^i\}_{i=1}^{i=|I|}$ **do**
  for each $m \in M$ **do**
    copy the $|I|^{|S|}$ sets $\{s^i\}_{i=1}^{|I|}$ and $\{(1-\epsilon^i)u_{\text{dir}}^i + \epsilon^i u_{\text{self}}^i/(1-\epsilon^i)\}_{i=1}^{|I|}$ in a Gambit file.
    apply Gambit method 'EnumPureSolve' to find all Nash equilibria.
    save Nash equilibria in a file.
returns *all sets of the Nash equilibria corresponding to the considered* $\{\epsilon^i\}_{i=1}^{i=|I|}$ *and the* $|M|$ *instances of the stochastic parameter.*

gies among the $|I|$ agents, which is implemented by the imbricated 'for' loops. $|S|^{|I|}$ simulations are performed here. In our case study, we have $|I| = 6$ agents with $|S| = 3$ strategies per agent, and thus, $3^6 = 729$ simulations are performed for each invocation of Algorithm 3. In addition, Algorithm 3 is called $|M| = 10$ times by Algorithm 1.

Finally, Algorithm 4 builds games in the normal form based on the simulation outcomes provided by Algorithm 3.

More precisely, each of the $|S|^{|I|}$ simulations produces $|I|$ utilities u_{dir}^i (one utility per agent) and $|I|$ utilities u_{soc}^i. We apply here $u^i = ((1-\epsilon)u_{\text{dir}}^i + \epsilon^i u_{\text{soc}}^i)/(1-\epsilon)$, and put these $|I|$ utilities u^i in one entry of a $|I|$-dimension matrix. As a result, this matrix contains $|I| \times |S|^{|I|}$ utilities. Let us now illustrate Algorithm 4 with our case study, in which each of the $|S|^{|I|} = 3^6 = 729$ simulations produces $|I| = 6$ utilities u_{dir}^i and $|I| = 6$ utilities u_{soc}^i, that we put under the form u^i in one entry of a 6-dimension matrix. Specifically, this is a $3 \times 3 \times 3 \times 3 \times 3 \times 3$ matrix, which contains $6 \times 3^6 = 4,374$ utilities u^i (each of the 3^6 simulations produce one utility per agent). We obtain a matrix for each of the $|\{\epsilon^i\}_{i=1}^{i=|I|}| \times |M|$ simulations run by Algorithm 1. In our case study, $|\{\epsilon^i\}_{i=1}^{i=|I|}| \times |M| = 5 \times 10 = 50$ matrices.

Every matrix is a game in the normal form that we analyze with Gambit [11]. Gambit is a set of software tools used to analyze games. These tools may be used either with the graphical user interface, or with the Gambit Command Language, which is a powerful way to carry out complex or repetitive operations on games. In our work, we use the latter language for two reasons: it is able to save Nash equilibria in a file (while the graphical interface only displays them), and all operations are automatized by placing the set of instructions to perform in the initialization file of Gambit. In practice, we only have to implement our simulator so that it writes the simulation outcomes in the Gambit format. Roughly, this deals with making the simulator write $|S|^{|I|}$ lines, where each line starts with the $|I|$ individual strategies s^i followed by the $|I|$ individual utilities u^i, and to add the head and the tail to this file to respect the Gambit format, so that this file can be read as a game with $|I|$ and $|S|$ strategies per player by Gambit. Next, the initialization file of Gambit has to open all files/games, find their equilibria, and write these equilibria in a single output file.

3 Case Study: Do Companies in a Supply Chain Agree to Share Demand Information?

We now illustrate our methodology in the context of supply chain management to study when the companies in a supply chain have incentives to share information. This information sharing aims to reduce a phenomenon called the bullwhip effect, which makes the overall supply chain less efficient. We first present the bullwhip effect in Subsection 3.1, next we propose a solution to this effect which is based on information sharing in Subsection 3.2. Since this solution requires information sharing, companies may disagree to use it, and therefore, we use the methodology described in Section 2 to study when companies prefer to use an ordering strategy involving information sharing. The simulation model of the supply chain used by our methodology (cf. Algorithm 3) is presented in Subsection 3.3. In this simulation, companies can use one of the three ordering strategies presented in Subsection 3.4, in which two strategies involve information sharing. Finally, the obtained results are presented in Subsection 3.5, and discussed in Subsection 3.6.

3.1 The Bullwhip Effect: A Problem of Supply Chain Management

The goal of this case study is to illustrate the methodology by studying when companies agree to share demand information in order to reduce the bullwhip effect. To this end, we now present the bullwhip effect (this phenomenon was shortly introduced in Subsection 2.2 of Chapter 1 as one of the problems faced by companies). Figure 1 shows how this effect propagates in a simple supply chain with seven companies: two retailers, two wholesalers, a pulp mill, a sawmill and a forest.

In this figure, each retailer exclusively sells to its customer and buys from its wholesaler, each wholesaler sells to its retailer and buys from its supplier (either the sawmill, or the pulp mill), etc. The ordering patterns of the companies are

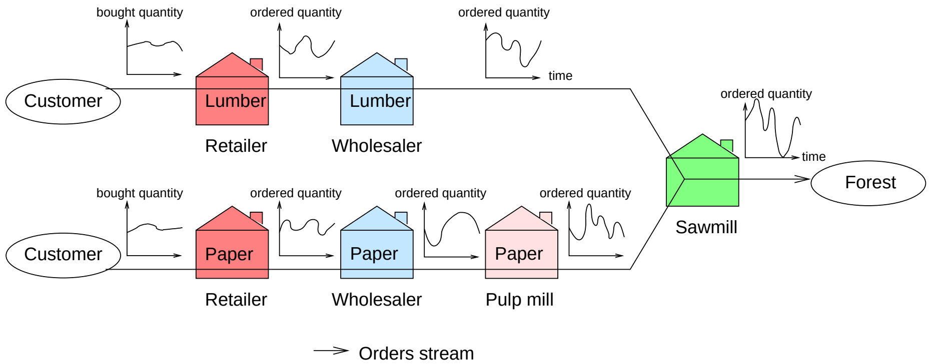


Fig. 1. The bullwhip effect [4, 12].

similar in the way that the variabilities of an upstream site are always greater than those of the downstream site [12]. As a variability, the bullwhip effect is measured by the standard deviation σ of orders.

This phenomenon is not harmful by itself, but because of its consequences. Here are the consequences reported by Carlsson and Fullér [13]:

– *Excessive inventory investments*: Since the bullwhip effect makes the demand more unpredictable, all companies need to safeguard themselves against the variations to avoid stockouts;
– *Poor customer service levels*: Despite the excessive inventory levels mentioned in the first consequence, demand unpredictability may cause stockouts anyway;
– *Lost revenues*: In addition to the poor customer service levels of the second consequence, stockouts may also cause lost revenues;
– *Reduced productivity*: Since revenues are lost (cf. third consequence), operations are less cost efficient;
– *More difficult decison-making*: Decisions-makers react to demand fluctuations and adapt (production and inventory) capacities to meet peak demands;
– *Sub-optimal transportation*: Transportation planning is made more difficult by demand uncertainties induced by the bullwhip effect;
– *Sub-optimal production*: As transportation, a greater demand unpredictability causes missed production schedules.

Carlsson and Fullér [13] noted that these consequences are not due to changes in the demand of end-customers, but only to inefficiencies in the supply chain. As an insight into the importance of these inefficiencies, these same authors claim that the bullwhip effect would cost the Finnish forest products industry 100-200 MFIM (17-34 million euros) per year, the industry having a total turnover of more than 100 BFIM (17 billion euros).

Causes	Proposed solutions	Authors
Demand forecast updating	Information sharing (e.g., VMI, CRP...), echelon-based inventory and leadtime reduction	[4, 12]
Order batching	EDI (Electronic Data Interchange) and Internet technologies	[4, 12] [15]
Price fluctuation	EDLP (Every Day Low Pricing)	[4, 12]
Rationing and shortage gaming	Allocation based on past sales	[4, 12]
Misperception of feedbacks	Giving a better understanding of the supply chain dynamics to managers	[16–20]
Local optimization without global vision	None	[5] [21–24]
Company processes	None	[15]

Table 1. Proposed causes and solutions of the bullwhip effect [14].

Next, some causes and solutions of the bullwhip effect have been identified. Table 1 summarizes them [14]. Lee and his colleagues [4, 12] proposed the first four causes and solutions presented in this table.

- *Demand Forecast Updating:* Companies base their orders on forecasts, which are themselves based on their incoming orders while such forecasts are not perfectly accurate. Therefore, companies order more or less than what they really require to fulfill their demand. In other words, forecasting errors amplifies the variability of orders.

  A solution proposed for this cause is information sharing: each client provides more complete information to its supplier in order to allow the supplier to improve its forecasting. Information sharing is already part of industry practices, such as VMI (Vendor-Managed Inventory), CRP (Continuous Replenishment Program), etc. Two of the three ordering strategies available to the company-agents in our simulation employ such information sharing.
- *Order Batching* (lot sizing in a more general way): Companies discretize orders in order to profit from economies of scales (e.g. transportation with full truck-loads), and therefore, place orders for more or less products than what they actually need.

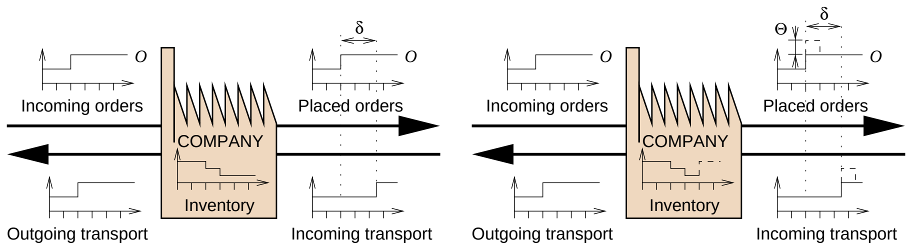
  The proposed solution for this cause is electronic transactions (e-commerce, EDI...) to reduce transaction costs and thus make companies' orders more frequent and for smaller quantities. Similarly, the size of production batches may be reduced with SMED (Single Minute Exchange of Die), which may next reduce the quantities ordered.

- *Price Fluctuation:* Every client (company or end-customer) profits from promotions by buying more products than what it really requires, and next, buying nothing when the promotion stops because it has enough products in inventory.
  The proposed solution is the EDLP (Every Day Low Pricing) policy, where price is set at the promotion level. However, EDLP also has some drawbacks, e.g., always looking for the lowest price may put a stress on the supply chain that may eventually reduce profits [25].
- *Rationing and Shortage Gaming:* Since every client behaves opportunistically, it overorders when its supplier cannot fulfill its entire demand, e.g., in the case where the supplier has a machine breakdown. Through such behaviour, this client does not hope to receive the quantity that it has ordered, but a lower quantity that matches its actual need.
  Since this behaviour occurs when the supplier allocates shippings in proportion to the ordered amount, it is preferable to allocate the few available products in proportion with the history of past orders.

Other authors have extended Lee and his colleagues' causes to the bullwhip effect:

- *Misperception of Feedback:* Sterman [20] has noted that players in the Beer Game place orders in a non-optimal way because they do not understand the whole dynamics in their supply chain. For example, they do not correctly interpret their incoming orders, and in consequence, smooth their orders when they should order more, because they do not understand that market consumption has increased.
- *Local Optimization without Global Vision:* Several authors [22–24] have noted that companies maximize their own profit without taking into account the effect of their decisions on the rest of the supply chain. In particular, some companies use an ordering scheme, such as the (s, S) policy (in which the company orders for $S - I$ products when inventory level I falls below s) that is the operationalization of this local optimization. It has been formally proven that some of these policies induce the bullwhip effect [5, 16, 21].
- *Company Processes:* Taylor and his colleagues [15] propose two causes to the bullwhip effect: variability in machine reliability and output, and variability in process capability and subsequent product quality. In these two causes, which are summarized as "Company processes" in Table 1, production problems at each workstation are amplified from one workstation to another. This cause recalls that intracompany problems and uncertainties may affect each company's behavior, which in turn may make them change the way they place orders.

In this context, our problem is first to find the way to place orders that reduces the bullwhip effect. Such ordering strategies should be the most efficient for the whole supply chain, which has been checked in a previous paper [14], and we next have to check whether companies agree to use it, which is the topic of this case study. The next subsection presents such an ordering strategy.



(a) Lot-for-lot ordering policy (b) (O, Θ) orders

Fig. 2. Lot-for-lot ordering policy and (O, Θ) orders.

3.2 One Cause and its Solution to the Bullwhip Effect

In this subsection, we present one of the possible refinements of the cause "misperception of feedbacks" in Table 1, then we propose a solution based on information sharing which addresses this specific cause of the bullwhip effect.

To avoid the bullwhip effect, the basic idea is that if companies' orders *follow* their clients' demand with a *lot-for-lot* ordering policy, there is *no bullwhip effect*, but inventories fluctuate greatly. In other words, either there is a bullwhip effect or inventories fluctuate greatly. This assertion is illustrated in Figures 2(a) and 2(b) that represent a company travelled by an ordering and a product stream. In these figures, we assume the company places orders strictly equal to its demand, following a strict lot-for-lot ordering policy. We now detail Figure 2(a) in four points, to show why companies prefer the bullwhip effect, rather than use the lot-for-lot policy.

1. The lot-for-lot ordering policy eliminates the bullwhip effect, because each company has the same ordering pattern as its client and thus, as the market consumption. Therefore, the two curves Incoming orders and Placed orders are identical. Since the bullwhip effect is measured as the standard-deviation of placed orders, we can see that the standard-deviation of each company's orders is exactly the same as the standard-deviation of its client's orders, and therefore, as the standard-deviation of the market consumption.

This first point explains why a lot-for-lot ordering policy eliminates the bullwhip effect, but not why companies do not use this ordering policy. As we shall see, the goal of companies is not to avoid the bullwhip effect, but to have products to carry out their activity:

2. The considered company tries to fulfill its entire demand, and thus, the two curves Incoming orders and Outgoing transport are the same, that is, as many products are shipped as ordered. The curve Outgoing transport in Figure 2(a) is valid as long as no stockouts occur by the considered company.

In short, the first two points say that the three curves Incoming orders, Placed orders and Outgoing transport are similar. The third point below says that the fourth curve also has the same pattern, but with a temporal shift δ:

3. The fourth curve Incoming transport has the same pattern as the three other curves, except that it is delayed by δ in comparison with the three other curves. This curve represents the reception of products by the company. The temporal shift δ corresponds to the ordering and shipping lead times, because items ordered by the company are not immediately received. The problem with lot-for-lot orders is that the inventory is not managed, because the temporal shift δ makes inventory decrease (respectively increase, when we inverse the pattern of "Incoming Orders"). In fact, the company ships more (respectively less) products than it receives during the ordering and shipping lead times δ.

   Note that Incoming transport has the same pattern as the three other curves only when the supplier has no stockouts, because the supplier is assumed to want to fulfill its entire demand, like the considered company.
4. Since every company wants to avoid stockouts (respectively huge inventory), rather than eliminate the bullwhip effect, it does not use the lot-for-lot ordering policy. Instead, it overorders (respectively underorders) to stabilize its inventory, which amplifies the demand variabilities, because the company overorders (respectively underorders) when the demand increases (respectively decreases). This shows that the *bullwhip effect always appears* each time the market consumption has an infinitesimal change, if companies want to *keep a steady inventory*.

Notice that "lead times" refer to the cause of the bullwhip effect that has just been presented. One should also note, that some of the other causes of the bullwhip effect presented in Table 1 induce the bullwhip effect even with a steady demand, while lead times amplify order fluctuations, but do not induce fluctuations when the demand is steady. Specifically, the bullwhip effect amplifies, because if a retailer overorders to stabilize its inventory, a worse phenomenon takes place with its suppliers: since the demand variation is now bigger, their inventories decrease much more, and thus, they must overorder more.

As we can see, the problem is not only to reduce the bullwhip effect, because companies just have to apply the lot-for-lot ordering policy to eliminate this effect, but inventories also have to be managed so that stockouts and high inventory levels are avoided. The solution we proposed in [2] is illustrated in Figures 2(b) and 3. The idea is to have an ordering strategy with a unique order amplification for each change in market consumption. Since companies have to know the market consumption to make this unique order amplification, this solution requires the sharing of demand information. As a result, our solution is the same as the one proposed by Lee and his colleagues to improve demand forecasting updating (cf. "information sharing" in Table 1), because companies have to share their incoming orders information with their suppliers. In fact, companies signal to their suppliers when they over- or underorder. This information sharing is presented in Figures 2(b) and 3, in which each company uses a vector (O, Θ) of two orders, where the actual order is the sum $(O + \Theta)$:

1. orders O follow the lot-for-lot policy to avoid the bullwhip effect;

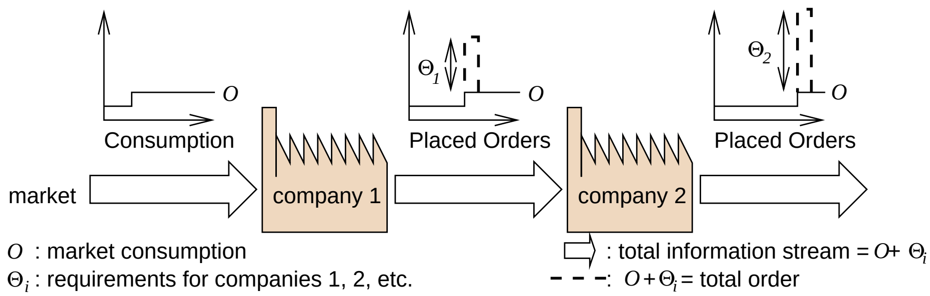


Fig. 3. Information streams cut into two parts.

2. orders Θ are used to order more or less products than O to stabilize inventory level.

We now present the two principles ruling the use of O and Θ. As presented in Figures 2(b) and 3, each company using (O, Θ) orders places a vector of orders (O, Θ), instead of a unique number O that englobes and hides these two numbers[3]. In (O, Θ), O is the market consumption transmitted from each company to its supplier(s) with the lot-for-lot policy, and Θ is chosen such that $O + \Theta$ represents what the company needs. As O follows a lot-for-lot policy, the bullwhip effect cannot occur in it. Unfortunately, it may occur in Θ, that is, non-zero Θs may be emitted anytime and anyhow. This is the reason why our (O, Θ) ordering strategies rest on the following two principles, and not only one of them.

First Principle: *The lot-for-lot ordering policy eliminates the bullwhip effect, but does not manage inventories.*
This first principle rules the way of choosing O in (O, Θ). Lot-for-lot orders means that each company orders what is demanded from it; if its client wants 10 products, the company places an order for 10 products. As previously stated in Figure 2(a), the problem is that the bullwhip effect is eliminated with the lot-for-lot policy, but inventory level is not managed. Therefore, we keep lot-for-lot orders for ruling O, but we add another piece of information Θ to manage inventory level. Note that lot-for-lot orders allow O to share the market consumption information, as illustrated in Figure 3.

Second Principle: *Companies should react only once to each market consumption change.*
This second principle rules the way of choosing Θ in (O, Θ). Θ is equal to zero all the time, except when the market consumption changes, in which case companies react to this change by sending non-zero Θ, in order to stabilize their inventory to the initial level. The purpose of Θ is to trigger a product

[3] O like *O*rder and Θ like *T*oken, as these two pieces of information were called in our previous papers. Moreover, O and Θ have the advantage of looking similar; while they have a very similar meaning: they are both *O*rders.

wave from the most upstream company when this company receives this Θ. This product wave will increase, or decrease when $\Theta < 0$, each company's inventory as it travels the supply chain down to the retailers.

We now illustrate these two principles with two ordering strategies, namely β and γ. In strategy β, we use the following trick to satisfy the second principle: since the market consumption is transmitted in O by the lot-for-lot policy (cf. first principle), a strategy in which $\Theta = 0$ when O is steady, and $\Theta \neq 0$ when O fluctuates, would satisfy the second principle. If we write i is the considered company, $i-1$ its unique client, and if we assume time is continuous (which is not the case in the simulation in this case study), then Equation 1 describes how the company i would place orders (O_t^i, Θ_t^i) with strategy β at instant t:

$$(O_t^i, \Theta_t^i) = (O_t^{i-1}, \Theta_t^{i-1} + \lambda \frac{dO_t^{i-1}}{dt}) \tag{1}$$

Note that $\frac{dO_t^{i-1}}{dt} \geq 0$ (respectively $\frac{dO_t^{i-1}}{dt} \leq 0$) represents the forecasted inventory decrease (respectively increase) during the ordering and shipping lead times. This quantity has to be overordered (respectively underordered) in order to keep a steady inventory. The constant λ depends on the duration of ordering and shipping lead times. We shall see how we adapt strategy β to our simulation in Subsection 3.4, i.e., after the presentation of the model of this simulation.

Before the presentation of the simulation, let us present another way of implementing our two principles, i.e., strategy γ. We assume in γ that information centralization is used, that is, retailers multicast the market consumption along the whole supply chain. Information sharing with information centralization is much quicker than information sharing with (O, Θ) orders, because the market consumption transmitted in O is as slow as orders, while information centralization is assumed to be instantaneous, reflecting thus the actual market consumption in real-time.

As a result, γ should be more efficient for the entire supply chain than β: as soon as the market consumption changes, non-zero Θs are sent by all companies to satisfy the second principle as quickly as possible. To do so, Θ is proportional to the variation of the market consumption transmitted by retailers. Moreover, companies also base O on the market consumption transmitted by retailers, instead of on incoming O, again in order to react quicker to the market consumption change. If we note again i the considered company, O_t^{retailer} the market consumption (approximated as the demand received by a retailer), and if we assume time is continuous, Equation 2 presents how company i places its orders (O_t^i, Θ_t^i) when there is information centralization. Again, λ is a constant depending on lead times:

$$(O_t^i, \Theta_t^i) = (O_t^{\text{retailer}}, \Theta_t^{i-1} + \lambda \frac{dO_t^{\text{retailer}}}{dt}) \tag{2}$$

We shall see an adaptation of β and γ to discrete time in Subsection 3.4. But before that, we present the simulation model of a supply chain that we use in this case study.

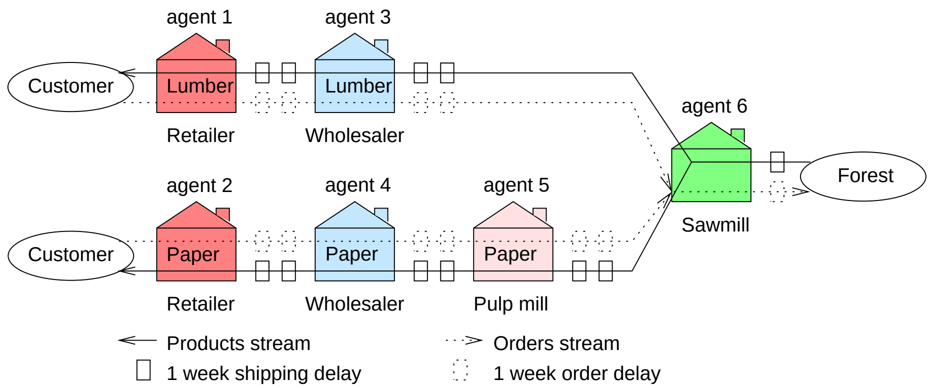


Fig. 4. Structure of the supply chain in our simulation [2].

3.3 Simulation Model

In Section 2, we assumed that the simulator was a source of data without giving additional details about it. Actually, the simulator used by our methodology is the *only link* with the application field. Of course, the results that our methodology will produce and their interpretation are only specific to this field.

In this context, each agent represents a company that places orders with its supplier in order to be able to fulfill its client's orders. Specifically, we use the same simulator as in our previous work [2, 14]. This simulation model relies on the supply chain structure and on the company model in the Quebec Wood Supply Game. In this model, agent direct utility is the opposite of their inventory holding cost, since agents cannot earn money and only incur costs, and thus, only have a negative direct utility. Similarly, agents' social consciousness is measured as backorders, that are equal to the duration and to the amplitude of stockouts. It is important to understand that we use backorders as the measure for social consciousness, because agents are *not obliged* to consider them when they make decisions (while they always have to take their inventory holding costs into account, which is the reason why inventory levels measure the direct utility), but the *goal of the entire supply chain* should be to satisfy end-customers. More technically, u^i_{dir} is the sum of the inventory level over the 50 weeks of a simulation, and u^i_{soc} is the sum of the backorder level over these 50 weeks.

Next, we detail the supply chain structure illustrated in Figure 4.

There are $|I|$=6 company-agents, where the first one represents the LumberRetailer ($i = 1$), which receives orders from the lumber market and places orders with the LumberWholesaler ($i = 3$). To fulfill these orders, this latter company places orders to the Sawmill ($i = 6$). Then, the Sawmill places orders to the Forest (which is not an agent because it does not place orders), but the Sawmill has to base its orders on two incoming demands in order to fulfill, as best as possible, the needs of the LumberWholesaler and the PulpMill. In fact, the Sawmill is the only company in our model that has to manage two types of products (lumber

Ordering strategy	Level of collaboration	Type of communication
μ	No collaboration	No communication
β	Information sharing with direct neighbours	Point to point
γ	Information centralization	Bulletin board

Table 2. The three ordering strategies.

and paper). Finally, there is a two-week delay to transmit an order and another two-week delay to ship the corresponding products.

3.4 Agent Strategies

To decide how much to order, every company i applies its ordering strategy s^i that has to be chosen in set S. In the case of the Sawmill, we operate as if there were a lumber Sawmill and a paper Sawmill that apply the same ordering strategy s^6, and the orders that are eventually placed to the Forest are the average of the orders that would have been placed by each sub-Sawmill.

For all companies, we have already introduced the fact that we assume $|S| = 3$ possible ordering strategies s^i for each of the six companies (including the Sawmill) and $S = \{\mu, \beta, \gamma\}$ in our case study, but we have not yet detailed μ, β and γ. Every strategy μ, β and γ implements a different level of collaboration in the supply chain, as outlined in Table 2. In comparison with our previous work [2, 14], β and γ are identical, while μ is an enhancement on α to take backorders more efficiently into account by considering inventory position instead of on-hand inventory[4]. Here are more details about μ, β and γ:

– With ordering strategy μ, company-agents make their decision on their own and can only rely on their incoming orders and shippings to decide what order to place. The only information transmitted with μ are orders and no additional information is shared to help other agents place their own orders. In practice, μ is an (s, S) policy, which is classic in Inventory Management, and in which a company orders $(S - InvP)$ items when the level of its inventory position $InvP$ falls below s. The inventory position $InvP$ is the sum of on-hand inventory Inv and inventory on order minus backorders.We show in [2] that taking $s = 0$ and S equals to the current incoming order is optimal.
– There is little collaboration with strategy β, because agents now transmit the market consumption information to their supplier with (O, Θ) orders, which

[4] The on-hand inventory Inv represents the products that are actually in inventory, while the inventory position $InvP$ represents what would be in inventory if there were no lead times due to transportation lead times, suppliers' stockouts, etc.

allows every agent to place orders in a more accurate way.As explained in Subsection 3.2, every placed order is a two-dimension vector $(Op, \theta p)$, and the quantity ordered is equal to the sum $(Op + \theta p)$ of its two elements: the first element corresponds to the market consumption transmitted by each company to its direct supplier, and the second element represents the products to be ordered in more or in less than the market consumption.
In our simulation, agents place orders $(Op, \theta p) = (Oi, \theta i + 4 * \Delta Oi)$, where $(Op, \theta p)$ is the two-dimension placed order, $(Oi, \theta i)$ is the incoming order, 4 is the value of a parameter depending on the ordering and shipping lead times between two companies (i.e., 2 week ordering lead time plus 2 week shipping lead time), and ΔOi is the variation of Oi since the previous week.

– The highest level of collaboration is implemented in strategy γ, in which retailer-agents write the market consumption information on a bulletin board, and the rest of the agents can read this information (this is information centralization). In practice, γ works as β, except that companies take the first element O of their two-dimension orders equal to the current market consumption, which approximated as retailer's incoming demand Oi^{retailer}. Consequently, agents place orders $(Op, \theta p) = (Oi^{\text{retailer}}, \theta i + 2 * \Delta Oi^{\text{retailer}})$.Note that the value of the parameter is now 2 rather than 4, because information centralization eliminates the ordering lead time by transmitting information instantaneously.

See [14] for details about these three strategies, about the parameters 2 and 4 depending on lead times in β and γ, and about the efficiency of these strategies for the overall supply chain when all companies use the same strategy.

3.5 Results and Analysis

In this section, we first introduce Table 3 to show the raw data obtained with our methodology for the particular instance m_A of the market demand. Next, we show how we aggregate this data in order to present results for ten instances of the market demand in Table 4.

As just stated, we first introduce the raw data obtained with our methodology. To this end, Table 3 presents the results obtained with the particular instance m_A of a uniform distribution of market demand. The distribution is the same for the ten considered instances m_A to m_J, and it is made of integer numbers uniformally spread over $\{11, 12, \ldots, 17\}$. In particular, the instance m_A in Table 3 corresponds to the two series of 50 numbers $D^{\text{lumber}} = \{16, 13, 16, 11, 15, \ldots, 15, 15, 12, 16, 15\}$ and $D^{\text{paper}} = \{15, 15, 14, 15, 14, \ldots, 13, 14, 14, 11, 13\}$, where the w^{th} number represents the market demand in Week w, e.g., $D_2^{\text{lumber}} = 13$ means that 13 units are demanded by the lumber market in Week 2. We approximate $D^{\text{lumber}} = O_i^1$ and $D^{\text{paper}} = O_i^1$.

Next, Table 3 enumerates all the Nash equilibria obtained with different values of ϵ.

For example, for $\epsilon = 5/6$, there are 6 Nash equilibria, and the first one is the *strategy profile* $(s^1, s^2, s^3, s^4, s^5, s^6) = (\beta, \mu, \beta, \mu, \mu, \beta)$; in this equilibrium, the

Value of ϵ	Nash equilibria
$\epsilon = 1/6$	$(\gamma, \mu, \gamma, \mu, \gamma, \gamma)$, $(\gamma, \gamma, \gamma, \mu, \mu, \gamma)$
$\epsilon = 2/6$	$(\gamma, \mu, \gamma, \mu, \mu, \beta)$, $(\gamma, \gamma, \gamma, \mu, \mu, \gamma)$
$\epsilon = 3/6$	$(\gamma, \beta, \gamma, \mu, \mu, \beta)$, $(\gamma, \gamma, \gamma, \mu, \gamma, \gamma)$
$\epsilon = 4/6$	$(\gamma, \gamma, \gamma, \gamma, \gamma, \gamma)$
$\epsilon = 5/6$	$(\beta, \mu, \beta, \mu, \mu, \beta)$, $(\beta, \beta, \beta, \beta, \beta, \beta)$ $(\beta, \gamma, \beta, \gamma, \beta, \gamma)$, $(\beta, \gamma, \beta, \gamma, \gamma, \beta)$ $(\gamma, \beta, \gamma, \beta, \beta, \beta)$, $(\gamma, \gamma, \gamma, \gamma, \beta, \beta)$

Table 3. Results for Instance m_A.

LumberRetailer ($i = 1$), the LumberWholesaler ($i = 3$) and the Sawmill ($i = 6$) use Strategy β, while all other agents use γ. The interpretation of all the equilibria in Table 3 is that no company has an incentive to unilaterally deviate from these equilibria. This does not mean that these strategy profiles are the best either for the whole supply chain or for every agent. This only means that these profiles are stable for the multi-agent system under this particular value of ϵ and under the particular instance m_A of the market demand. Note that μ occurs 15 times, β 29 times and γ 40 times in overall Table 3, which corresponds respectively to frequencies of 20%, 29% and 51%. Therefore, the highest level of collaboration occurs in half of the found Nash equilibria.

Next, the question addressed in this paper is to determine at which value of ϵ the Nash equilibria switch from a full-collaborating system (that is, most companies collaborate at the highest level, e.g., $(\gamma, \gamma, \gamma, \beta, \gamma, \gamma)$) to a full non-collaborating one (that is, most companies disagree to collaborate, e.g., $(\mu, \beta, \mu, \mu, \mu, \mu)$). To answer this question, we check if the Nash equilibria with low ϵ (selfish agents) often apply μ (i.e., the non-collaborative strategy), and/or if the Nash equilibria with high ϵ (benevolent agents) frequently use γ (i.e., the highly collaborative strategy). To see this, we calculate the frequency at which every strategy is used depending on the value of ϵ. That is, we calculate the central column 'Instance m_A' in Table 4 based on the numbers in Table 3. This calculation is achieved in the following way. In line '$\epsilon = 5/6$' in Table 3, there are 6 equilibria that count a total of 3 μ, 20 β and 13 γ. These three numbers are reported in line $\epsilon = 5/6$ in Table 4 as '$3/36 = 8\% \to \mu$', which means that 3 of the 36 (=3+20+13) strategies used in a Nash equilibrium are μ.

When we look at the central column in Table 4, we cannot make a clear distinction between the equilibria obtained for the five considered value of ϵ, because there is no regularity in the distribution of percentages: we were waiting for μ with a maximum frequency when $\epsilon = 1/6$ (while γ is more frequent than μ in this entry), and for γ with a maximum value when $\epsilon = 5/6$ (which is again not true since β has a greater frequency). On the other hand, the collaborative strategy γ in general occurs the most frequently over most values of ϵ, which indicates that all agents are strongly related in the system whatever the level of selfishness is: if one behaves in a harmful way for the rest of the agents, this agent will also suffer from its incorrect behaviour.

Value of ϵ	Frequency of μ, β and γ under Instance m_A	Average frequency of μ, β and γ over Instances m_A to m_J
$\epsilon = \frac{1}{6}$	$4/12 = 33\% \rightarrow \mu$ $0/12 = 0\% \rightarrow \beta$ $8/12 = 67\% \rightarrow \gamma$	$27\% \rightarrow \mu$ $9\% \rightarrow \beta$ $64\% \rightarrow \gamma$
$\epsilon = \frac{2}{6}$	$5/12 = 42\% \rightarrow \mu$ $1/12 = 8\% \rightarrow \beta$ $6/12 = 50\% \rightarrow \gamma$	$22\% \rightarrow \mu$ $6\% \rightarrow \beta$ $72\% \rightarrow \gamma$
$\epsilon = \frac{3}{6}$	$3/12 = 25\% \rightarrow \mu$ $2/12 = 17\% \rightarrow \beta$ $7/12 = 58\% \rightarrow \gamma$	$17\% \rightarrow \mu$ $14\% \rightarrow \beta$ $69\% \rightarrow \gamma$
$\epsilon = \frac{4}{6}$	$0/6 = 0\% \rightarrow \mu$ $0/6 = 0\% \rightarrow \beta$ $6/6 = 100\% \rightarrow \gamma$	$9\% \rightarrow \mu$ $16\% \rightarrow \beta$ $75\% \rightarrow \gamma$
$\epsilon = \frac{5}{6}$	$3/36 = 8\% \rightarrow \mu$ $20/36 = 56\% \rightarrow \beta$ $13/36 = 36\% \rightarrow \gamma$	$7\% \rightarrow \mu$ $43\% \rightarrow \beta$ $50\% \rightarrow \gamma$

Table 4. Relative frequency of the occurence of μ, β and γ in the incurred Nash equilibria.

This conclusion only holds for the particular instance m_A of the uniformly distributed market consumption. We have repeated for the nine other instances m_B to m_J of market consumption. Their results are in the right column in Table 4. To calculate every frequency in the right column, we take the average of the frequencies obtained with the ten market demands, i.e., these frequencies are not weighted by the number of equilibria: to obtain '$7\% \rightarrow \mu$' with $\epsilon = 5/6$ in the right column of Table 4, the 8% obtained with m_A has the same importance as the percentages obtained with $m_B, m_B, \ldots, m_J$, even though they do not represent the same quantity of equilibria. We can draw the same conclusions with the right column as with the center column in Table 4, i.e., μ only has low occurrence frequencies. We now detail our findings in the context of the reduction of the bullwhip effect.

3.6 Discussion on the Case Study

We have just seen that the collaborative strategy γ occurs the most frequently, and that its frequency does not depend on the value of ϵ. From the viewpoint of supply chain management, this is an important result, because it indicates (as for m_A) that every agent is strongly related to each other, and therefore, all agents prefer collaboration with β and γ, whatever their value of selfishness is (i.e., for all values of ϵ). In addition, the frequency of μ (i.e., the only considered non-collaborating strategy) depends on the value of ϵ: the frequency of μ regularly decreases from 27% to 7% when the social consciousness ϵ increases.

Consequently, the frequency of μ is *intuitive*, since companies in the selfish supply chain ($\epsilon = 1/6$) do not collaborate more frequently than companies in the socially conscious supply chain ($\epsilon = 5/6$). These two points mean that, if the supply chain stabilizes on a strategy profile (i.e., on a Nash equilibrium), this strategy profile will probably be made of many γ. In this result, we do not detail the utility incurred by every company, but we noticed in [2] that some equilibria will be avoided because they are Pareto-dominated by another equilibrium, which means that all agents are worse off in the Pareto-dominated equilibria than in another equilibrium.

From the viewpoint of supply chain management, this is an excellent result, because it was possible that each company-agent would have liked the rest of the supply chain to collaborate except itself, because this company-agent would have profited from the effort made by the rest of the supply chain (problem of the free rider). Fortunately, such a scenario does not occur in our simulations, but this is only true when all agents have the same level of self consciousness because we do not consider mixes of selfish and benevolent agents.

The answer to the question addressed in this paper is thus that collaboration holds in the entire supply chain whatever the level of social consciousness is, i.e., for all considered value of ϵ. Our interpretation is that company-agents are strongly dependent on each other. However, the non-collaborating strategy is *more frequently chosen* when the level of social consciousness is low. What we call dependence is that, if a company-agent disturbs the operation of the supply chain, it may earn a little from this in its direct utility (which represents what is incurred by the agent and which cannot be disregarded), but this agent will also lose a lot because of the supply chain dynamics that will disturb this agent as a feedback of its bad behaviour. As a consequence, benevolent agents ($\epsilon \to 1$), as selfish agents ($\epsilon = 0$), will collaborate by sharing information, that is, in general, all agents prefer to use Strategy γ (or at least β) rather than μ, whatever their selfishness level is. This is a high level result that should lead the design of lower level interaction mechanisms for agents in supply chains [26].

4 Discussion on the Methodology

The previous section has illustrated the methodology proposed in this paper by showing how to study social consciousness in supply chains. We now discuss our methodology and possible improvements. First of all, it would be interesting to be able to run exactly the presented methodology. This means that we should use different values of ϵ^i for our agents, instead of a ϵ common to all agents. As previously stated, we would be able to study if the free rider problem occurs. Unfortunately, this will greatly increase the bulk of outputs produced by our methodology.

To see this, let us consider that ϵ^i may take E values for every agent i. Thus, $E^{|I|}$ combinations of the E possible values of ϵ^i among the $|I|$ agents should be considered. For example, if $E = 3$ (e.g., $\epsilon^i \in \{1/4, 2/4, 3/4\}$) and $|I| = 6$, $3^6 = 729$ cases have to be considered *per instance m of M*. To analyze this great number of outputs, we could, for example, define some *classes of joint strategies*, and observe the distribution of the equilibria among these classes depending on

$\{\epsilon^i\}_{i=1}^{i=|I|}$. Such classes of joint strategies may be something like "all the $|I|$ agents use one specific strategy, or $|I|-1$ agents use one specific strategy except one agent that uses any other strategy", or any other definition in which the classes do not overlap. Finally, a few manipulations are not yet automatized in the implementation of our methodology, which obliges us to reduce the number of instances of $|M|$ to 10, while it should be much greater to get significant results.

Moreover, it would be interesting to include the concept of mixed strategies, which comes from game theory [10]. In fact, we only consider pure strategies in this paper, that is, agents use the same strategy during the entire run (i.e., over 50 weeks) of each of the $|S|^{|I|}$ simulations. If mixed strategies were used, we would use probabilities for all these strategies, e.g., Agent i would use s_1^i for 75% of the time (i.e., over 38 weeks), and s_2^i during the rest of the run of the simulation (i.e., over 12 weeks). Unfortunately, we cannot use mixed strategies for the following three reasons:

1. *Reason from supply chain management:* Companies cannot switch between collaboration and non-collaboration all the time, because contracts have to be signed to insure the secrecy of the information shared with β and γ, and collaborative devices have to be bought by companies.

The two other reasons can be found in any application field:

2. *Dynamics in the simulated system may create some "inertia":* Since this phenomenon occurs in our case study, we use it to illustrate what we mean by inertia; when a company switches from Strategy μ to γ, the products previously ordered with μ will still arrive, which has an impact on the supplier's inventory level, and therefore on its future orders. Such a transition does *not exist* in traditional game theory, and is hard to handle.
   As a result, we are not allowed to calculate the expected outcome of a simulation with mixed strategies based on two simulations with pure strategies because the result would have no meaning in reality. For instance, if the PaperRetailer uses μ 75% of the time and γ during the remaining time, the expected PaperRetailer's outcome is different from 75% of its outcome in simulation in which only μ is used, plus 25% of its outcome in a second simulation in which only γ is applied, e.g., if the PaperRetailer uses μ first, γ next, the outcome will be different than if the PaperRetailer starts with γ, even if the ratio of 75%/25% is the same.
   Therefore, the method used in traditional game theory to calculate expected utilities is not appropriate in the context of our case study, because of the transition time between the two ordering rules that we have just described. Fortunately, this issue of inertia does not occur in all contexts, e.g., in all traditional games studied with game theory [10], such as the "Prisoners' Dilemma", etc. When this problem does not happen, the methodology presented in Section 2 is the same, except that Gambit has to be asked for the Nash equilibria in mixed strategies instead of in pure strategies in Algorithm 4.

3. *Algorithmic complexity:* We do not know of any algorithms used to determine Nash equilibria in mixed strategies requiring a reasonable time for any game. The determination of Nash equilibria in pure strategies requires the comparison of individual outcomes, while in mixed strategies, it requires the resolution of linear equations, a more complex task. From a practical viewpoint, computing (i.e., finding in a given game) Nash equilibria may also be a problem in pure strategies. For instance, it took around six minutes per game for Gambit to obtain our results, but it is much lower than what was required for mixed strategies because we stopped the computation of some of these same games after a week on a 2.7 GHz PC while the computation was not yet finished! From a theoretical viewpoint, Papadimitriou and Roughgarden [27] presented the finding of a polynomial-time algorithm for computing a Nash equilibrium as the "holy grail". In fact, Papadimitriou [28] thinks that this problem is harder than P and easier than NP-hard.

Let us now compare our methodology with another one that also studied interactions in a supply chain with simulations. This other work was carried out by Wellman and his colleagues [3] who refer to their methodology as "empirical game theory". They used the TAC/SCM (Trading Agent Competition/Supply Chain Management) [29] as data input to analyze the strategic behaviour of agents applying strategies inspired by the ones performed by the opponents in the TAC/SCM 2003 competition (note that Chapter 5 presents RedAgent, which is the winner of TAC/SCM 2003).

The two main differences between the approach of Wellman's team and ours are (i) that their data relies on the more detailed and complex simulation model of the TAC/SCM, in which our methodology would be too long to run (they claim that each simulation instance of the TAC/SCM takes 55 minutes, while our simulator runs within a second; as a result, the TAC/SCM simulator applied to our methodology would require $3^6 * 55$ minutes, i.e., almost a month, to build the game corresponding to one particular market consumption pattern), and (ii) they can consider symmetric games because all agents have the same role in TAC/SCM, while our agents have different roles in the supply chain. Considering symmetric games allows reducing the number of strategy profiles to consider from 729 to 28 when each of 6 agents can choose between 3 strategies.

Finally, readers interested in game theory applied to supply chains may be interested in several other chapters of this book. First, Chapter 4 defines the optimal production quantity and transactions of manufacturers as a Nash equilibrium. Then, Chapter 8 uses marketing science and cooperative game theory to propose a coordination and negotiation mechanism amongst the members of a virtual enterprise. The chapters related to auctions also deal with game theory, and in particular, Chapter 12 relies on mechanism design, which is the contrary of game theory (that is, game theory studies the properties of a system in which agents are given some strategies, while mechanism design searches for the strategies agents should be given so that the system has some properties).

5 Conclusion

In this paper, we presented a methodology to analyze stochastic simulations with game theory, when every utility function can be split in an individual utility and a utility representing the social consciousness of the agent. More precisely, we studied the impact on agents' choice of strategy when these agents consider more or less the part of their utility that represents social consciousness. We illustrated this methodology with a case study of the reduction of the bullwhip effect (i.e., the reduction of the amplification of order variability) in a supply chain. Specifically, we studied whether selfish company-agents would use non-collaborative strategies more frequently than benevolent company-agents. In this context, we assumed that every agent's individual utility is the inventory holding cost, and that agents take their impact on the rest of the system into account as the service delivered to their clients (which is measured as their backorder level, i.e., as the amplitude and duration of their stockouts).

We found that companies collaborate frequently, whatever the level of social consciousness is. We think that this result is due to the fact that companies are strongly dependent on each other in our model. However, non-collaboration is more frequent when the level of social consciousness is low, even if non-collaboration is never the most frequently used strategy.

Finally, we discussed our methodology and pointed out some possible improvements on it. In particular, we noted that the concept of mixed strategies may be difficult to introduce in some contexts. In particular, in our case study, there is some inertia in the supply chain that results in a transition time that forbids the calculation of expected utilities as in traditional game theory. Fortunately, this problem does not occur in all contexts.

6 Acknowledgment

We would like to thank FOR@C, the Research Consortium in E-business in the forest products industry (Université Laval, Québec, Canada), for supporting this research. This work was also partially supported by the National Sciences and Engineering Research Council of Canada (NSERC).

References

1. Parsons, S., Wooldridge, M.: Game theory and decision theory in multi-agent systems. Autonomous Agents and Multi-Agent Systems **5** (2002) 243–254
2. Moyaux, T., Chaib-draa, B., D'Amours, S.: Experimental study of incentives for collaboration in the Québec wood supply game. IEEE Transactions on Engineering Management (2005) (submitted).
3. Wellman, M.P., Estelle, J., Singh, S., Vorobeychik, Y., Kiekintveld, C., Soni, V.: Strategic interactions in a supply chain game. Computational Intelligence **21** (2005) 1–26
4. Lee, H.L., Padmanabhan, V., Whang, S.: Information distortion in a supply chain: The bullwhip effect. Management Science **43** (1997) 546–558

5. Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E.: Designing and managing the supply chain. McGraw-Hill Higher Education (2000)
6. Glass, A., Grosz, B.J.: Socially conscious decision-making. Autonomous Agents and Multi-Agent Systems **6** (2003) 317–339
7. Stirling, W.C., Goodrich, M.A., Packard, D.J.: Satisficing equilibria: A nonclassical theory of games and decisions. Autonomous Agents and Multi-Agent Systems **5** (2002) 305–328
8. Lipton, R.J., Markakis, E.: Nash equilibria via polynomial equations. In: Proc. LATIN 2004: Theoretical Informatics, 6th Latin American Symposium. (2004) 413–422
9. Sandholm, T.W.: Distributed rational decision making. In Weiss, G., ed.: Multiagent Systems - A Modern Approach to Distributed Artificial Intelligence. MIT Press, Cambridge (Massachusetts, USA) (1999) 201–258
10. Jehle, G.A., Reny, P.J.: Advanced Microeconomic Theory. Addison Wesley (2000) 2d Edition.
11. McKelvey, R.D., McLennan, A.M., Turocy, T.L.: Gambit: Software tools for game theory, version 0.97.0.4 (2004) `http://econweb.tamu.edu/gambit/` (accessed 16 October 2004).
12. Lee, H.L., Padmanabhan, V., Whang, S.: The bullwhip effect in supply chain. Sloan Management Review **38** (1997) 93–102
13. Carlsson, C., Fullér, R.: Soft computing and the bullwhip effect. Economics and Complexity **2** (2000)
14. Moyaux, T., Chaib-draa, B., D'Amours, S.: Information sharing as a coordination mechanism for reducing the bullwhip effect in a supply chain. IEEE Trans. on Systems, Man, and Cybernetics (2005) (to appear).
15. Taylor, D.: Measurement and analysis of demand amplification across the supply chain. The International Journal of Logistics Management **10** (1999) 55–70
16. Daganzo, C.: A theory of supply chains. Springer (2003)
17. Dejonckheere, J., Disney, S.M., Lambrecht, M.R., Towill, D.R.: Transfer function analysis of forecasting induced bullwhip in supply chains. International Journal of Production Economics **78** (2002) 133–144
18. Forrester, J.W.: Industrial dynamics - A major breakthrough for decision-makers. Harvard Business Review **36** (1958) 37–66
19. Forrester, J.W.: Industrial Dynamics. The M.I.T. Press - Massachusetts Institute of Technology - Cambridge (Massachusetts, USA) (1961)
20. Sterman, J.D.: Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. Management Science **35** (1989) 321–339
21. Chen, F., Drezner, Z., Ryan, J.K., Simchi-Levi, D.: Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. Management Science **46** (2000) 436–443
22. Kahn, J.A.: Inventories and the volatility of production. The American Economic Review **77** (1987) 667–679
23. Naish, H.F.: Production smoothing in the linear quadratic inventory model. The Economic Journal **104** (1994) 864–875
24. Shen, Y.: Impact of asymmetric information on inventory policy (2001) Term paper to Dr. Miller for the partial fulfillment of "Production Control".
25. Butman, J.: A pain in the (supply) chain. Harvard Business Review **80** (2002) 31
26. Rosenschein, J.S., Zlotkin, G.: Rules of Encounter: Designing Conventions for Automated Negotiation Among Computers. MIT Press, Cambridge (Massachusetts, USA) (1994)

27. Papadimitriou, C.H., Roughgarden, T.: Computing equilibria in multiplayer games. In: Proc. SODA. (2005) Algorithmic results for the special case of symmetric and other succinctly representable games: correlated equilibria can be found in polynomial time for a broad class that includes all previously known cases.
28. Papadimitriou, C.H.: Algorithms, games, and the internet. In: Proc. STOC 2001 (extended abstract in ICALP 2001). (2001)
29. Arunachalam, R., Sadeh, N.M.: The supply chain trading agent competition. Electronic Commerce Research and Applications **4** (2005) 63–81