This paper investigates two different yet related research questions about stock management in feedback environments. The first purpose is to analyse the effects of selected experimental factors on the performances of subjects (players) in a stock management simulation game. In light of these results, our second objective is to evaluate the adequacy of standard decision rules typically used in dynamic stock management models and to seek improvement formulations. To carry out the research, the generic stock management problem is chosen as the interactive gaming platform. In the first part, gaming experiments are designed to test the effects of three factors on decision-making behaviour: different patterns of customer demand, minimum possible order decision (‘review’) interval and, finally, the type of receiving delay. ANOVA results of these three-factor, two-level experiments show which factors have significant effects on 10 different measures of behaviour (such as max–min range of orders, inventory amplitudes, periods of oscillations and backlog durations). In the second phase of research, the performances of subjects are compared against some selected ordering heuristics (formulations). First, the patterns of ordering behaviour of subjects are classified into three basic types. Comparing these three pattern types with simulation results using different decision rules, we observe that the common linear ‘anchoring and adjustment rule’ can represent the smooth and gradually damping type of behaviour, but cannot generate the non-linear and/or discrete ordering dynamics. Thus, several alternative non-linear rules are formulated and tested against subjects’ behaviour patterns. Some standard discrete inventory control rules (such as \((s, S)\)) common in the inventory management literature are also formulated and tested. These non-linear and/or discrete rules, compared with the linear stock adjustment rule, are found to be more representative of the subjects’ ordering behaviour in many cases, in the sense that these rules can generate non-linear and/or
discrete ordering behaviours. Another major finding is the fact that the well-documented oscillatory dynamic behaviour of the inventory is a quite general result, not just an artifact of the linear anchor and adjust rule. When the supply line is ignored or underestimated, large inventory oscillations result also with the non-linear rules, as well as the standard inventory management rules. Furthermore, depending on parameter values, non-linear ordering rules are more prone to yield unstable oscillations—even if the supply line is taken into account. Further methodological and experimental research questions are suggested. Copyright © 2004 John Wiley & Sons, Ltd.

Keywords stock management; anchor and adjust heuristic; experimental testing of decision rules; non-linear decision heuristics; inventory control rules

1. STOCK MANAGEMENT GAME

For the purpose of experimental testing, the generic stock management problem, one of the most common dynamic decision problems, is chosen as the interactive gaming environment. The objective of the game is stated as ‘keeping the inventory level as low as possible while avoiding any backorders’. If there are not enough goods in the inventory at any time, customer orders are entered as backorders to be supplied later. ‘Order decisions’ are the only means of controlling the inventory level. The general structure of the stock management problem is illustrated in Figure 1 (see Sterman, 2000). The three boxes—Expectation Formation, Goal Formation and Decision Rule—are deliberately left blank in Figure 1, as they are unknown, since they take place in the ‘minds’ of the players. (Later, in the simulation version of the game, these three boxes will have to be specified. For instance, the expectation formation will be formulated by exponential smoothing; inventory goal will be set to inventory coverage times expected demand; and supply line goal will be order delay times expected demand. As for the decision rule, different formulations will be tried: linear stock adjustment rule, three different non-linear adjustment rules and finally various standard discrete inventory control rules.) This

![Image of stock management game](image-url)
notion of ‘gaming experimentation’ to analyse and test subjects’ decision heuristics has been successfully used in the system dynamics literature (Sterman, 1987, 1989), as well as in experimental psychology (Brehmer, 1989).

While playing the game, subjects can monitor the system from information displays/graphs showing their inventory, supply line levels and customer demand (Appendix A). Neither the costs associated with high inventories nor costs resulting from backorders are accounted for explicitly in the simulation game. However, the relation between keeping these costs as low as possible and the objective of the game is stated in the instruction given to subjects. In other words, subjects are instructed that keeping large safety stocks would result in high holding costs, but backlogs must also be avoided as they would incur large costs due to lost demand. Before beginning the game all subjects are given a written instruction presenting the problem and their task (Appendix B). The time available to accomplish the task is not limited. No explicit, tangible reward is used to motivate the subjects.

2. GAMING EXPERIMENTS

The first set of gaming experiments is designed to test the effects of three factors on the decision-making behaviour of subjects:

(a) length of order decision (review) interval;
(b) type of receiving delay;
(c) pattern of customer demand.

2.1. Length of Decision Interval

Subjects were allowed to order at ‘each time unit’ in the first group of experiments (Short Game), whereas they were allowed to order ‘once every five time units’ in the second group of experiments (Long Game). Short Games are simulated for 100 time units, whereas Long Games are simulated for 250 time units. The receiving delays are also shorter (four time units) in the short game, and longer (10 time units) in the long game. (Note that if one were to change the decision interval from one to five days but kept the receiving delay at four, then the nature of the game would have changed in an implicit and problematic way, in the sense that the receiving delay would be four times longer than the order interval in the short game but shorter than the order interval in the long one.) Thus the ‘length’ effect is really a ‘package’ (involving longer receiving delays and longer game length as well), but this package is summarized by the term length of decision interval effect, because, as will be seen below, this particular component will be the focus in interpreting the experimental results. It is hypothesized that subjects being free to make decisions at any point in time versus decisions allowed only every five time units would influence the difficulty of the game, and hence cause differences in the performances.

2.2. Type of Receiving Delay

The second independent experimental factor is the type of delay. (The length of the delay is not an independent factor, since it is changed as an integral part of the length of decision interval effect described above). We focus on the type of delay representation, as different delay types may be appropriate for different inventory acquisition systems, like continuous exponential delay or discrete delay representations. Since ‘receiving’ is the inflow to inventory, its transient behaviour may influence a decision-maker’s interpretation of the results of his/her own order decisions. The two extremes of the exponential delay family, namely the first-order exponential delay and infinite-order discrete delay, are chosen as the two levels of delay factor in the experimental design. This is further motivated by a common criticism of system dynamics games that their continuous delays are not realistic or intuitive; hence such games pose an artificial difficulty for players with no expertise in modelling. So, it would be interesting to see if subjects’ performances would actually deteriorate in cases involving continuous delays.

2.3. Patterns of Customer Demand

Until the fifth decision interval, average customer demand remains constant at 20. At the
beginning of the fifth decision interval (at time five in the Short Game and at time 25 in the Long Game) an unannounced, one-time increase of 20 units occurs in the customer demand patterns used in the experiments. Subjects react to the disequilibrium caused by this change. After the step up, demand remains constant at 40 in the first type of customer demand pattern which we call 'step-up in customer demand' (Figure 2).

In the second type of demand pattern, called 'step-up-and-down in customer demand', a second disturbance, a one-time decrease in demand, follows the first increase after some time interval, restoring the demand back to its original level of 20 (see Figure 3, where the step-down is set to occur at time 20). The time interval between the steps-up-and down in customer demand is chosen as roughly half of the natural periodicity of the model (about 25 days in Short Game and about 60 in Long Game). A common perception in stock management circles is that 'poor ordering performance is caused by complex demand patterns'. The purpose of this particular experimental factor is therefore to test this claim in a simplified context.

Before the demand patterns described above are used in games, 'Pink noise' (auto-correlated noise) is added to the average patterns to obtain more realistic demand dynamics. The standard deviation of the white noise is set to 15% of average customer demand. The delay constant of the exponential smoothing (the correlation time used to create pink noise) is taken as two time units (see Appendix C for equations).

To summarize, there are eight combinations of the above three factors across the two levels of each. So we have a $2^3$ factorial design. Each condition is played six times (random replications), yielding a total of 48 experiments (see Table 1). Since the demand pattern is discovered by the subjects once the game is played and because they can improve their performance by practice, in order to obtain unbiased results the same subject never played two Short Games or two Long Games. (However, due to limited number of subjects, a few subjects were asked to play one Short Game and one Long Game, since transferring experience between Short and Long Games is not easy).

2.4. Initial Conditions

All games start at equilibrium. The supply line level is initially set at 80 in Short Games and 200 in Long Games so that no backordering occurs even when the decision-maker does not order goods during the first four decision intervals (at the end of which the disturbance in customer demand causes disequilibrium). Inventory levels are initially set arbitrarily (at 40 and 200 respectively) so as to satisfy the average initial customer demand for the first two decision intervals (for two days in Short Game and 10 days in Long Game).

<table>
<thead>
<tr>
<th>Runs</th>
<th>Length of ordering interval</th>
<th>Type of receiving delay</th>
<th>Pattern of customer demand</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Every time unit (1 day)</td>
<td>Every five time units (5 days)</td>
<td>Exponential</td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>5</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>7</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>8</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 1. Design of experiments. 'X' indicates the selected level of factors for the corresponding runs

3. ANALYSIS OF EXPERIMENTS

The general, broad behaviour pattern of inventory in the majority of games is one of oscillations. (See Figures 4 and 5 and most other games illustrated in the figures throughout the article). This finding is consistent with overwhelming evidence on oscillating inventories in the system dynamics literature and elsewhere (for instance, Forrester, 1961; Sterman, 1989, 2000; Lee et al., 1997; Tvede, 1996). We will return later to this main ‘qualitative’ result. But first we take a more quantitative look at the effects of the three experimental factors.
Representative summary measures of orders and inventory levels computed from the experimental results are summarized in Table 2. Ten characteristics are tabulated for each of the 48 games. The averages of these 10 measures for each of the eight experiments are also displayed. From these averages it is possible to have some idea about the effects of the three factors on each of these measures. For instance, as one moves from Run 1 to Run 2, the only experimental factor that changes is ‘demand pattern’ (from step-up to step-up-and-down). In this case observe, for example, that the average max. order measure changes from 143.3 to 146.7, a minor change,
Table 2. Selected measures of game performances

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Min. order</th>
<th>Max. order</th>
<th>Range of orders</th>
<th>Min. inventory</th>
<th>Max. inventory</th>
<th>Range of inventory</th>
<th>Initial backorder time</th>
<th>Final backorder time</th>
<th>Duration of backorders</th>
<th>Inv. oscillation period</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>100</td>
<td>100 –125</td>
<td>50</td>
<td>175</td>
<td>4</td>
<td>38</td>
<td>34</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>100</td>
<td>100 –125</td>
<td>150</td>
<td>275</td>
<td>4</td>
<td>33</td>
<td>29</td>
<td>31</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>100</td>
<td>100 –175</td>
<td>150</td>
<td>325</td>
<td>4</td>
<td>18</td>
<td>14</td>
<td>29</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>200</td>
<td>200 –75</td>
<td>100</td>
<td>175</td>
<td>6</td>
<td>16</td>
<td>10</td>
<td>25</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>300</td>
<td>300 –225</td>
<td>100</td>
<td>325</td>
<td>5</td>
<td>43</td>
<td>38</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>60</td>
<td>60 –60</td>
<td>40</td>
<td>100</td>
<td>6</td>
<td>37</td>
<td>31</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Avg. of Run 1</td>
<td>3.3</td>
<td>143.3</td>
<td>140.0 –130.8</td>
<td>98.3</td>
<td>229.2</td>
<td>4.8</td>
<td>30.8</td>
<td>26.0</td>
<td>28.3</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>100</td>
<td>100 –60</td>
<td>115</td>
<td>175</td>
<td>10</td>
<td>21</td>
<td>11</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>100</td>
<td>100 –60</td>
<td>250</td>
<td>310</td>
<td>5</td>
<td>13</td>
<td>8</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>60</td>
<td>60 –60</td>
<td>60</td>
<td>120</td>
<td>7</td>
<td>20</td>
<td>13</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>400</td>
<td>400 –80</td>
<td>180</td>
<td>260</td>
<td>6</td>
<td>21</td>
<td>15</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>150</td>
<td>150 –100</td>
<td>250</td>
<td>350</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>40</td>
<td>70</td>
<td>30 –60</td>
<td>150</td>
<td>210</td>
<td>6</td>
<td>18</td>
<td>12</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Avg. of Run 2</td>
<td>6.7</td>
<td>146.7</td>
<td>140.0 –70.0</td>
<td>167.5</td>
<td>237.5</td>
<td>6.7</td>
<td>17.8</td>
<td>11.2</td>
<td>29.5</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>100</td>
<td>100 –225</td>
<td>300</td>
<td>525</td>
<td>5</td>
<td>34</td>
<td>29</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>220</td>
<td>220 –225</td>
<td>225</td>
<td>450</td>
<td>5</td>
<td>35</td>
<td>30</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>150</td>
<td>150 –75</td>
<td>160</td>
<td>235</td>
<td>14</td>
<td>25</td>
<td>11</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>60</td>
<td>60 –15</td>
<td>190</td>
<td>215</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>100</td>
<td>100 –125</td>
<td>250</td>
<td>375</td>
<td>5</td>
<td>38</td>
<td>33</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>300</td>
<td>300 –260</td>
<td>180</td>
<td>440</td>
<td>5</td>
<td>18</td>
<td>13</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Avg. of Run 3</td>
<td>0.0</td>
<td>155.0</td>
<td>155.0 –154.2</td>
<td>217.5</td>
<td>373.3</td>
<td>6.8</td>
<td>30.0</td>
<td>23.2</td>
<td>35.0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>150</td>
<td>150 –25</td>
<td>150</td>
<td>175</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>150</td>
<td>150 –100</td>
<td>290</td>
<td>390</td>
<td>6</td>
<td>21</td>
<td>15</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>80</td>
<td>80 –145</td>
<td>80</td>
<td>225</td>
<td>5</td>
<td>24</td>
<td>19</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>100</td>
<td>100 –40</td>
<td>350</td>
<td>390</td>
<td>12</td>
<td>18</td>
<td>4</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>100</td>
<td>100 –250</td>
<td>210</td>
<td>460</td>
<td>6</td>
<td>27</td>
<td>21</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>200</td>
<td>200 –150</td>
<td>250</td>
<td>400</td>
<td>5</td>
<td>22</td>
<td>17</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Avg. of Run 4</td>
<td>0.0</td>
<td>130.0</td>
<td>130.0 –118.3</td>
<td>221.7</td>
<td>340.0</td>
<td>6.8</td>
<td>22.4</td>
<td>15.2</td>
<td>35.7</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>600</td>
<td>600 –260</td>
<td>425</td>
<td>685</td>
<td>20</td>
<td>30</td>
<td>10</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>400</td>
<td>400 –350</td>
<td>200</td>
<td>550</td>
<td>25</td>
<td>85</td>
<td>60</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>300</td>
<td>300 –280</td>
<td>200</td>
<td>480</td>
<td>20</td>
<td>85</td>
<td>65</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>400</td>
<td>400 –300</td>
<td>200</td>
<td>500</td>
<td>25</td>
<td>70</td>
<td>45</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>560</td>
<td>560 –150</td>
<td>370</td>
<td>520</td>
<td>20</td>
<td>40</td>
<td>15</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>250</td>
<td>250 –20</td>
<td>425</td>
<td>445</td>
<td>35</td>
<td>45</td>
<td>10</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Avg. of Run 5</td>
<td>0.0</td>
<td>418.3</td>
<td>418.3 –226.7</td>
<td>303.3</td>
<td>530.0</td>
<td>24.2</td>
<td>59.2</td>
<td>34.2</td>
<td>54.8</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>40</td>
<td>375</td>
<td>335 –200</td>
<td>250</td>
<td>450</td>
<td>20</td>
<td>40</td>
<td>20</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>250</td>
<td>250 –100</td>
<td>200</td>
<td>300</td>
<td>25</td>
<td>55</td>
<td>30</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>0</td>
<td>400</td>
<td>400 –150</td>
<td>475</td>
<td>625</td>
<td>30</td>
<td>45</td>
<td>15</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>50</td>
<td>250</td>
<td>250 –60</td>
<td>200</td>
<td>260</td>
<td>30</td>
<td>55</td>
<td>25</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>50</td>
<td>300</td>
<td>300 –0</td>
<td>275</td>
<td>275</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>100</td>
<td>250</td>
<td>150 –160</td>
<td>250</td>
<td>410</td>
<td>25</td>
<td>55</td>
<td>30</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Avg. of Run 6</td>
<td>40.0</td>
<td>304.2</td>
<td>264.2 –111.7</td>
<td>275.0</td>
<td>386.7</td>
<td>26.0</td>
<td>50.0</td>
<td>24.0</td>
<td>70.7</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>100</td>
<td>270</td>
<td>170 –100</td>
<td>500</td>
<td>600</td>
<td>20</td>
<td>40</td>
<td>20</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>20</td>
<td>500</td>
<td>480 –500</td>
<td>300</td>
<td>800</td>
<td>25</td>
<td>90</td>
<td>65</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>80</td>
<td>250</td>
<td>170 –40</td>
<td>450</td>
<td>490</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

Continues
whereas the average max. inventory changes from 98.3 to 167.5, clearly a more significant effect, at least at the base level of the other factors, ignoring any interactions. (a simple t-test between Run 1 and 2, ignoring other levels of the other two factors and possible interactions, would yield the same conclusion). But a more complete and definitive conclusion about the significance of each of the three factors on each of the 10 output measures can be obtained by a full factorial analysis of variance (ANOVA), considering the effects of each factor at all levels of other factors (and any possible interactions between them). A summary table derived from full ANOVA (using SPSS software) is shown in Table 3. These results are obtained from a full factorial ANOVA model involving seven effects (three main effects, three two-way interactions and one three-way interaction term). Since we have a total of 48 data sets, the degrees of freedom for residuals (errors) is $48 - 7 - 1 = 40$ and hence the $F$ statistic (=mean squared explained by regression/mean squared error) for each effect has $(1, 40)$ degrees of freedom for numerator and denominator respectively. So, if the $F$ value computed for any effect is ‘large enough’, we reject the hypothesis that the corresponding effect coefficient is zero, i.e. a significant effect is discovered. Typical significance levels are $alpha = 0.01$ (99% confidence), $alpha = 0.05$ (95% confidence) or $alpha = 0.10$ (90% confidence). In Table 3, we provide the $F$ values computed for each effect (for each output measure) and the ‘$P$ value’ at which the $F$ value would be found significant. To conclude, if a $P$ value is $\leq$ the chosen alpha level, we decide that the given factor has a significant effect on the selected output measure, at $(1 - alpha)$% confidence level. Although the ANOVA results come from a full factorial model, in Table 3 we show the main effects only for readability, because analysing individual interaction terms is beyond our research scope.

Before we examine the ANOVA results of Table 3, some statistical issues need to be discussed. First, the number of replications in each experimental cell (six in our case) must be large enough to meet the normality assumption and provide an acceptable test ‘power’ (i.e. a small enough variance). It is in general known that the larger the sample size, the better the normality assumption and the higher the power of the test. But since obtaining independent replications in each experimental cell can be quite impractical, the experimental design literature (e.g. Montgomery, 1991) provides approximate suggestions for the minimum sample size needed for normality and satisfactory test power. In multi-factor designs, minimum independent replications are typically found between three and six. So our six replications per cell seems to be a reasonable size, given practical difficulties of experimenting with human subjects. Yet this does not mean that the statistical requirements

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Min. order</th>
<th>Max. order</th>
<th>Range of orders</th>
<th>Min. inventory</th>
<th>Max. inventory</th>
<th>Range of inventory</th>
<th>Initial backorder time</th>
<th>Final backorder time</th>
<th>Duration of backorders</th>
<th>Inv. oscillation period</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>50</td>
<td>250</td>
<td>200</td>
<td>–200</td>
<td>425</td>
<td>625</td>
<td>30</td>
<td>40</td>
<td>10</td>
<td>75</td>
</tr>
<tr>
<td>41</td>
<td>0</td>
<td>1000</td>
<td>1000</td>
<td>–1000</td>
<td>2000</td>
<td>3000</td>
<td>20</td>
<td>100</td>
<td>80</td>
<td>133</td>
</tr>
<tr>
<td>42</td>
<td>0</td>
<td>450</td>
<td>450</td>
<td>–375</td>
<td>420</td>
<td>795</td>
<td>35</td>
<td>60</td>
<td>25</td>
<td>52</td>
</tr>
<tr>
<td>Avg. of Run 7</td>
<td>41.7</td>
<td>453.3</td>
<td>411.7</td>
<td>–369.2</td>
<td>682.5</td>
<td>1051.7</td>
<td>26.0</td>
<td>66.0</td>
<td>40.0</td>
<td>89.4</td>
</tr>
<tr>
<td>43</td>
<td>0</td>
<td>250</td>
<td>250</td>
<td>–125</td>
<td>400</td>
<td>525</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>75</td>
</tr>
<tr>
<td>44</td>
<td>0</td>
<td>450</td>
<td>450</td>
<td>–500</td>
<td>200</td>
<td>700</td>
<td>20</td>
<td>60</td>
<td>40</td>
<td>N/A</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
<td>375</td>
<td>375</td>
<td>–750</td>
<td>200</td>
<td>950</td>
<td>20</td>
<td>70</td>
<td>50</td>
<td>N/A</td>
</tr>
<tr>
<td>46</td>
<td>20</td>
<td>300</td>
<td>280</td>
<td>–330</td>
<td>330</td>
<td>660</td>
<td>25</td>
<td>55</td>
<td>30</td>
<td>55</td>
</tr>
<tr>
<td>47</td>
<td>50</td>
<td>250</td>
<td>200</td>
<td>–180</td>
<td>300</td>
<td>480</td>
<td>30</td>
<td>45</td>
<td>15</td>
<td>114</td>
</tr>
<tr>
<td>48</td>
<td>0</td>
<td>300</td>
<td>300</td>
<td>–375</td>
<td>350</td>
<td>725</td>
<td>30</td>
<td>60</td>
<td>30</td>
<td>N/A</td>
</tr>
<tr>
<td>Avg. of Run 8</td>
<td>11.7</td>
<td>320.8</td>
<td>309.2</td>
<td>–376.7</td>
<td>296.7</td>
<td>673.3</td>
<td>25.0</td>
<td>58.0</td>
<td>33.0</td>
<td>81.3</td>
</tr>
</tbody>
</table>
are automatically met. So we also ran some diagnostic tests at the end. The most typical diagnostic test is the ‘normal probability plot’ where the observed cumulative probabilities of residuals are plotted against the expected normal cumulative probabilities. The results of these diagnostic tests are somewhat mixed: in many cases the normal plots display acceptable deviations around the expected normal line, while in a few cases the deviations are too large. So the Normality assumption is not met for all 10 measures. There can be other causes than sample size of these non-normality instances. It seems that in many cases the non-normality is caused by the discrete nature of a given measure (like minimum order) or by a few extreme data points. In some other situations, the distribution is skewed, a deeper problem. In any case, as in most experiments involving human subjects, the normality assumption is not always met and must be further investigated. This can be done by leaving some extreme data points out, by trying out data transformations or by carrying out new experiments. Since this is an exploratory article, such further statistical analysis would be a significant distraction from our main focus, so we suggest them as a ‘further investigation’ topic. As can be seen in Table 3, our significant and non-significant results are mostly quite widely separated, so the non-normality instances are probably not strong enough to hurt our exploratory conclusions.

Finally, note that the last four measures in Table 2 (initial backorder time, final backorder time, duration of backorders and oscillation period) show the symbol N/A in some of the cells. The meaning is that some of these measures were simply undefined in some of the game results. In a few instances, there are no backorders at all, so all three related measures are marked N/A. A more problematic and frequent situation arises when the dynamics of inventory has no clearly identifiable, computable constant period. This arises when the inventory is non-oscillatory, or when it exhibits very complex, highly noisy oscillations. In about two or three out of six runs in each cell we have this situation, so it is quite frequent. Since ANOVA requires ‘balanced’ input data, these N/A entries require some
pre-treatment. Prior to ANOVA, SPSS uses a mixture of two methods, depending on the suitability: ‘estimating the missing data from neighbouring points’ or ‘discarding data from other cells so as to balance’. In any case, it should be noted that the results about the last output measure, period of inventory oscillations, should be taken with more caution, because the effective number of replications for this measure is three to four.

Although there are 10 output measures in Tables 2 and 3, some of these can be seen as intermediate measures used to compute related end-measures. Minimum and maximum orders are measured to ultimately obtain a measure of ‘range of order fluctuation’ (or amplitude) and the same is true for minimum and maximum inventories. Lastly, initial backorder time and final backorder time (when applicable) are measured to compute the ‘duration of backorders’. So to conserve space, we focus on four end-measures only and leave more detailed examination to the interested reader. These four end-measures summarize four distinct performance characteristics: amplitude of order fluctuations, amplitude of inventory fluctuations, duration of backorder phase and finally the period of inventory oscillations (when it is defined).

3.1. Effect of Different Patterns of Customer Demand

ANOVA results in the bottom row block of Table 3 show the effects of ‘demand pattern’ (step-up only or step-up-and-down) on the output measures. As illustrative game dynamics, see Figures 4, 5, 7 for step-up-and-down and Figures 6, 8B and 10B for step-up demand. (For full results, compare pair-wise the results of Experiments 1 and 2; 3 and 4; 5 and 6; 7 and 8 in Table 2. Also see Özevin, 1999). F values in Table 3 show that the demand pattern does not have a strong enough effect on ‘range of orders’ measure (although it does have ‘some’ effect, since the significance just missed 90%). Similarly, demand pattern has no significant effect at all on the amplitude or period of inventory oscillations (or on any other ‘inventory’ measure). So a strong result is that when the order pattern is changed from step-up to step-up-and-down, this has no significant effect on the inventory dynamics. An interpretation is that this effect is negligible compared with the structural causes of oscillations, such as supply line delays and order batching. Lastly, we observe that the demand pattern does have a significant effect on

![Graph 1: Performance of the player in Game 14](image)

**Figure 6.** Performance of the player in Game 14 (Short Game with orders each period, step-up in customer demand, discrete delay)
the ‘duration of backorders’. The meaning of this last finding is interesting because the direction of the effect is that step-up-and-down demand, compared with step-up only, causes the backlog durations to become shorter. (See Table 2). So a seemingly more complex demand pattern actually happens to help the players’ compensate for their ordering weaknesses in terms of backlog durations.

3.2. Effect of Different Representations of Receiving Delays

Illustrative game dynamics with first order exponential delay and with discrete delay are shown in Figures 4, 5 and Figures 6, 7 respectively. In these and many other examples with continuous exponential delay, subjects are able to manage the inventories in a relatively more stable way. It seems that when the goods ordered arrive gradually over some period of time, it prevents the players from over-ordering or under-ordering excessively. In contrast, discrete delay representation affects their performance negatively, causing large fluctuations. In these experiments, subjects yield large magnitude, long-period oscillations in inventories and fail to bring these oscillations under control in most cases. (For complete results, compare pair-wise the results of Experiments 1 and 3; 2 and 4; 5 and 7; 6 and 8 in Table 2. Also see Özevin, 1999). Subjects seem to have difficulties in accounting for the effects of sudden receiving. The above results are consistent with research evidence on the effect of delays on dynamic decision-making performance in system dynamics literature (Sterman, 1989), as well as in experimental psychology research (Brehmer, 1989). The results are also consistent with the mathematical stability conditions of discrete-delay dynamical systems, which are much more difficult to obtain compared with continuous-delay systems (see Driver, 1977).

For further statistical analysis, ANOVA results about the ‘type of receiving delay effect’ are given in the middle row block of Table 3. Observe that this factor has highly significant effects on two critical output measures: range of inventory fluctuations and its period. This finding statistically confirms our qualitative assessment above; that the type of receiving delay has a significant effect on the stability of inventory fluctuations. Furthermore, observe in Table 2...
that the direction of the effect is that discrete delay, compared with the continuous one, causes the range and period of oscillations to be larger. In a nutshell, discrete delay representation makes the system more oscillatory, less stable, and thus harder to manage for the subjects (a result that is consistent with mathematical stability of dynamic systems as well as with experimental psychology research).

3.3. Effect of the Length of Decision Intervals

The last factor is the ‘length’ effect (involving the decision interval, as well as the length of the delay and the game duration as described earlier). As illustrative game dynamics with ‘order every time step’ and ‘order every five time steps’ experiments see Figures 4, 6 and Figures 5, 7 respectively. (For full results,
compare pair-wise the results of Experiments 1 and 5; 2 and 6; 3 and 7; 4 and 8 in Table 2. Also see Özevin, 1999). These table entries reveal that all output measures change significantly when the ‘length factor’ is changed. ANOVA results given in the first row block of Table 3 are also consistent with this observation: effects of ‘length’ on all measures are highly significant. There may be two different sources of this high level and sweeping significance. First, when the decision period is made longer (noting that the receiving delay is also made longer to be consistent), since the time constants of the system are larger, time-related output measures (like period of oscillations, backorder times and backorder duration) all naturally become larger in value—a technical result of the dynamics of the system. Second—and more interestingly—amplitude measures (like range of orders and range of inventory) also become statistically significantly larger. This is a more behavioural/decision-related result: when decisions can be made less frequently, orders must be larger in magnitude, feedback is less frequent, controlling of the inventory becomes harder and so inventory fluctuations are larger in magnitude as well. So, in sum, ‘order every five time steps’ and ‘longer delays’ together make the system more oscillatory, less stable, and thus harder to manage for the subjects. The way our experiments are designed means that it is impossible to separate the ‘decision interval length’ and the ‘delay length’ effects. In another experimental setting it would be interesting to test these two different ‘length effects’ separately and see if one of these two factors have strong effects on some of the above output measures but negligible effects on others. Such an experimental design can be incorporated in related further research.

4. TESTING OF ALTERNATIVE DECISION FORMULATIONS

Section 3 completes the relatively more quantitative/statistical research objective of the paper. In light of the above results, the second objective is to evaluate the adequacy of standard decision rules typically used in dynamic stock management models and to seek improvement formulations. To this end, the performance patterns of subjects will be (qualitatively) compared with the dynamic patterns obtained using different simulated ordering formulations: as an initial step, the patterns of ordering behaviour of subjects are observed to fall into three basic classes: (i) smooth, continuous—oscillatory or non-oscillatory—damping orders (for example, see Figure 4 and 8); (ii) alternating large-and-zero discrete orders, like a high-frequency signal (for example, Figures 6, 22 and 24); and (iii) long periods of constant orders punctuated by a few sudden large ones (for example, Figures 10B and 12B). Using these three types of observed ordering patterns, the common linear ‘anchoring and adjustment’ rule, several ‘non-linear’ adjustment rules and some standard discrete inventory control rules (such as (s, Q)) found in the inventory management literature will be evaluated. The criterion of ‘adequacy’ for a given decision heuristic is: a formulation has acceptable adequacy if it is capable of generating an ordering pattern that qualitatively belongs to the pattern class of interest. Such an adequacy is necessary, but obviously not sufficient for a decision rule to be useful. For sufficiency, the parameters of the candidate decision rule must be further calibrated so as to obtain a good enough (or best?) fit to some given ordering data (see, for instance, Sterman, 1987, 1989). Such a calibration is beyond the scope of our article. Our purpose is to show that there are some types of ordering patterns that some rules (such as linear anchor-and-adjust heuristic) simply cannot generate, so proper formulations must be developed before any parameter calibration is attempted. We develop such rules and discuss their basic dynamical properties.

4.1. Linear Anchoring and Adjustment Rule

The linear anchoring and adjustment rule is often used to model decision-making behaviour in system dynamics models (Sterman, 2000, 1987). In this heuristic, one starts from an initial point, called the anchor, and then makes some adjustments to come up with the final decision. In the context of inventory management, a plausible
anchor point for order decisions is the expected customer demand. (If the inventory manager can order only once every five periods, the anchor should be the total of expected customer demand for five periods between subsequent decisions.) When there are discrepancies between desired and actual inventory levels and/or between desired and actual supply line, adjustments are made so as to bring the inventory and the supply line back to desired levels. Thus, the order equation based on linear anchoring and adjustment heuristic is formulated as:

\[ O_t = E_t + \alpha^*(I_t^* - I_t) + \beta^*(SL_t^* - SL_t) \]  (4.1)

When orders can be given once every five periods, then the anchor of the rule is modified as follows (the adjustment terms being as before):

\[ O_t = 5^*E_t + \alpha^*(I_t^* - I_t) + \beta^*(SL_t^* - SL_t) \]  (4.2)

where \( E_t \) represents expected customer demand per time period, \( I_t^* \) represents the desired inventory, \( I_t \) the inventory, \( SL_t^* \) the desired supply line and \( SL_t \) the supply line. \( \alpha \) and \( \beta \) are the adjustment fractions.

In real life, ‘safety stocks’ are determined by balancing the inventory holding and backordering costs. Although, an optimum inventory level minimizing these costs may be found mathematically, safety stocks are more often set approximately. The desired inventory \( I_t^* \) is thus modelled as proportional to customer demand to allow adjustments in safety stocks when changes in customer demand occur.

\[ I_t^* = k^*E_t \]  (4.3)

To maintain a receiving rate consistent with receiving delay \( \tau \) and customer demand, \( SL_t^* \) is formulated as a function of \( \tau \) and the expected customer demand \( E_t \).

\[ SL_t^* = \tau^*E_t \]  (4.4)

The linear anchor and adjust rule can mimic the subjects’ performances adequately in experiments where they tend to place smooth and continuously damping orders (See Figure 8 as an example and Özevin, 1999, for more). However, in certain games, many subjects tend to place non-linear or discontinuous orders (Figures 6, 10B, 12B, 22 and 24), especially when the receiving delay representation is discrete and/or the order interval is five. Such order patterns typically fall into class ii (alternating large-and-zero discrete orders, like a high-frequency signal—Figures 6, 22 and 24) or class iii (long periods of constant orders punctuated by a few sudden large ones—Figures 10B and 12B) described above. The linear ‘anchoring and adjustment rule’ by its nature can not yield such discrete-looking intermittent or non-linear order patterns. Non-linear rules may be able to represent better the subjects’ performances in such situations. Two different approaches will be discussed below in two separate sections: non-linear adjustment rules and standard inventory control rules.

4.2. Rules with Non-linear Adjustments

The linear adjustment rule is based on ‘adjustments’ to orders proportional to the discrepancy between the desired and observed stock levels. Orders are placed regularly, in proportion to this discrepancy. However, some players do not place such smooth orders. They completely cease ordering when the inventory seems to be ‘around a satisfactory’ level and place rather large orders if the discrepancy between the desired and actual inventory becomes larger. The resulting ordering behaviour is in class iii (long periods of constant orders punctuated by a few sudden large ones—Figures 10B and 12B) described above. In this section, three alternative non-linear decision rules will be developed and tested to address this particular class of non-linear ordering behaviour.

4.2.1. Cubic Adjustment Rules

Similar to the linear anchoring and adjustment rule, cubic adjustment rules also start with expected customer demand as an anchor point, but the adjustments are formulated as non-linear. One or both of the adjustment terms may be cubic in discrepancies in inventory and/or in supply line. Alternative order equations can be mathematically expressed as:

\[ O_t = E_t + \alpha^*(I_t^* - I_t)^3 \]  (4.5)
if only the inventory adjustment is taken into account, and:

\[ O_t = E_t + \alpha^*(I_t^* - I_t)^3 + \beta^*(SL_t^* - SL_t) \]  (4.6)

\[ O_t = E_t + \alpha^*(I_t^* - I_t) + \beta^*(SL_t^* - SL_t)^3 \]  (4.7)

where one of the adjustments is made cubic and the other is linear, and finally:

\[ O_t = E_t + \alpha^*(I_t^* - I_t)^3 + \beta^*(SL_t^* - SL_t)^3 \]  (4.8)

where both of the adjustments are formulated as cubic. In the equations above, \( E_t \) represents the expected customer demand; \( I_t^* \) and \( SL_t^* \) the desired inventory and supply line levels; \( I_t \) and \( SL_t \) the actual inventory and supply line, and \( \alpha \) and \( \beta \) are the fraction of the discrepancy corrected by the decision-maker at each period. The internal consistency of the rules can be shown mathematically (Özevin, 1999). When orders can be made once every five periods, the adjustments terms are as above; however, the anchor term is increased to five times the expected customer demand.

The general stability properties of the cubic adjustment rules are similar to the well-established linear adjustment rule. Thus, first the supply line must be taken into account (i.e., adjustment fraction \( \beta \) must be non-zero and preferably equal to \( \alpha \) for optimum stability), and secondly, the larger the values of \( \alpha \) and \( \beta \) the less stable the system tends to be. For example, Figure 9 compares two behaviours of the cubic rule with non-zero and zero \( \beta \) values. Observe that the behaviour becomes quite unstable when the supply line adjustment term is zero. Although the behaviour in Figure 9A illustrates a very stable case, we should note that choosing stable values for the adjustment fractions is not easy with cubic adjustment rules (whereas the linear rule guarantees stability for any \( \alpha = \beta < 1 \)). In other words, the performance of the cubic rule is too sensitive to the adjustment parameter values. In particular, when orders are given once every five periods, the cubic rule most often fails to generate stable dynamics. On the other hand, the primary advantage of the cubic rule is that it is possible to generate ordering patterns that can mimic the non-linear ordering behaviour of subjects characterized by long periods of constant orders punctuated by a few sudden large ones (described as class iii, above). Figure 10 provides a comparative evaluation of such dynamics. The disadvantage is that the parameter range in which cubic adjustment formulation can yield stable dynamics is too narrow to be useful (Özevin, 1999). This is the motivation for the discussion of other non-linear rules in the following sections.

**4.2.2. Variable Adjustment Fraction Rule**

Analogous to the anchoring and adjustment rule and the cubic adjustment rules, we define a ‘variable adjustment fraction rule’ that anchors at expectations about customer demand. However, the adjustments are increased sharply (non-linearly) when the discrepancy in inventory increases. The simplest version of the order equation of the rule can be mathematically expressed as:

\[ O_t = E_t + \alpha^*(I_t^* - I_t) \]  (4.9)

The variable fraction \( \alpha \) is a function of the discrepancy in inventory, normalized by the desired inventory. The shape of the function yields increased adjustments when the discrepancy in inventory is increased. Normalized inventory discrepancy \( \delta \) is defined as:

\[ \delta = \frac{I^* - I}{I^*} \]  (4.10)

According to the function in Figure 11, the rule is not mathematically unbiased in the ideal known demand case; there will be some small, deliberate steady-state discrepancy between the inventory and its desired level. But this may well be a ‘realistic’ bias in order to be able to obtain a non-linear ordering behaviour similar to some subjects. The rule performs quite realistically in the ‘noisy’ demand case, where the steady-state bias is negligible anyway and may be irrelevant in real life. The more general (and more stable) version of the variable adjustment fraction rule will be of the form \( E_t + \alpha^*(I_t^* - I_t) + \beta^*(SL_t^* - SL_t) \), where \( \beta \) is defined by a function exactly equivalent to the one in Figure 11, except that the input would be ‘normalized supply line discrepancy’. Alternatively, the supply line
adjustment may be linear. (We omit this discussion further in this article to conserve space.) In any case, inclusion of the supply line term will increase the stability of the system.

The variable adjustment fraction rule is able to generate long periods of constant orders punctuated by a few sudden large ones (pattern class iii). Therefore they may be used to represent subjects’ behaviour when orders are characterized by this type of non-linearity, where linear adjustment rules would fail (See Figure 12 and Özevin, 1999, for more illustrations).
4.2.3. Non-linear Expectation Adjustment Rule

The order equation of another non-linear rule that we call ‘expectation adjustment rule’ can be mathematically expressed by:

\[ O_t = \alpha E_t \quad (4.11) \]

where the variable adjustment coefficient \( \alpha \) is a function of discrepancy in inventory normalized by desired inventory and \( E_t \) represents the expected customer demand. \( \alpha \) is equal to one when the inventory is at the desired level, since adjustments are not necessary when the system is in equilibrium (See Figure 13). The shape of the \( \alpha \) function causes increasing upward adjustments in orders when the inventory level is below the desired level and it causes reductions in orders when the inventory is above its desired level.

Figure 10. Comparison of (A) performance of ‘linear supply line and cubic inventory adjustment rule’ (\( \alpha = 1/500, \beta = 1 \)) with (B) the performance of the player in Game 1 (Short Game with orders each period, step-up in customer demand, exponential delay)
Like the previous variable adjustment rule, non-linear expectations rule can generate sudden large orders, separated by no-action intervals. Therefore, it may provide an alternative to the linear anchoring and adjustment rule when subjects' behaviours exhibit such patterns (see Figure 14 for example).

In addition to the specifics of three alternative non-linear formulations, this section yields a general result: the well-documented oscillatory dynamic behaviour of the inventory when supply line is underestimated is true not only for the linear anchor and adjust rule but also for the non-linear rules seen above. Furthermore, in non-linear rules, stability is achieved only for a rather narrow range of parameter/function values.

5. STANDARD INVENTORY CONTROL RULES

For class ii type of ordering behaviour (alternating large-and-zero discrete orders like a high-frequency signal—Figures 6, 22, 24) neither the linear anchor and adjust rule nor are the non-linear ones developed above are suitable. These discrete orders and the resulting zigzagging inventory patterns suggest that classical discrete inventory control rules used in inventory management may be suitable. The following four policies are most frequently used in inventory management literature:

- order point–order quantity (s, Q) rule;
- order point–order up to level (s, S) rule;
- review period–order up to level (R, S) rule;
- (R, s, S) rule.

Two fundamental questions to be answered by any inventory control system are 'how many' and 'when' (or 'how often') to order. 'Order Point' systems determine how many to order, whereas 'Periodic Review' systems determine how often to order as well (Silver and Peterson, 1985; Tersine, 1994). When subjects can order every time unit, they are free to order at any time they desire. Therefore, order point systems, rather than periodic review systems, are more appropriate to represent the ordering behaviour in these situations. In contrast, when subjects can order only, say, once every five periods, periodic review systems with five as the review period may be more appropriate as decision rules. These inventory control rules assume that time flow and changes are discrete. Therefore, these rules will be tested only with discrete delays. As will be seen below, these rules are non-linear in the sense that they consist of piecewise, discontinuous-in-derivative functions.

5.1. Order Point–Order Quantity (s, Q) Rule

The order point–order quantity (s, Q) rule can be mathematically expressed as:

\[ O_t = Q, \quad \text{if } EI_t \leq s \]
\[ \quad 0, \quad \text{otherwise} \]  \hspace{1cm} (5.1)

where EI, represents the effective inventory and s the 'order point'. Effective inventory and order point are calculated as follows:

\[ EI_t = I_t + SL_t \]  \hspace{1cm} (5.2)
\[ s = DAVGSL_t + DMINI_t \]  \hspace{1cm} (5.3)
$I_t$ and $SL_t$ represent goods in inventory and in supply line respectively. $DAVGSL_t$ refers to the desired average supply line and $DMINI_t$ to desired minimum inventory. To be consistent with the previous continuous adjustment rules, desired average supply line and desired minimum inventory can be defined as:

$$DAVGSL_t = \tau E_t$$  \hspace{1cm} (5.4) \\
$$DMINI_t = E_t + SS$$  \hspace{1cm} (5.5)
in terms of receiving delay \( t \), expected demand \( E_t \) and safety stocks \( S_S \). Order quantity \( Q \) and safety stock \( SS \) are fixed arbitrarily as constants in this research. Desired minimum inventory \( D_{MINI} \) is defined as the sum of a constant safety stock and demand expectation. With such a definition, desired minimum inventory can be adapted to variations in customer demand. The performance of this rule with deterministic demand is seen in Figure 15. Orders exhibit alternating zeros and \( Q \)'s and inventory zigzags around a constant level. Observe that the \( (s, Q) \) rule cannot prevent the inventory from falling below the desired minimum inventory, even when no noise exists, primarily because the order quantity \( Q \) is constant. This particular rule is therefore not suitable for our purpose (i.e., for comparative evaluation against the continuous stock adjustment rules). The \( (s, Q) \) rule will not be further evaluated; it is defined only as a preliminary for the other more realistic rules to follow.

### 5.2. Order Point–Order up to Level \( (s, S) \) Rule

Order point–order up to level \( (s, S) \) rule can be mathematically expressed as:

\[
O_t = S - EI_t, \quad \text{if } EI_t \leq s \\
0, \quad \text{otherwise}
\] 

(5.6)

where \( EI_t \) represents the effective inventory, \( s \) the order point and \( S \) the upper level of inventory. Effective inventory, the order point and the upper level \( S \) of inventory can be defined as follows:

\[
EI_t = I_t + SL_t \quad (5.7)
\]

\[
s = DAVGSL_t + DMINI_t \quad (5.8)
\]

\[
S = s + Q \quad (5.9)
\]

\( I_t \) and \( SL_t \) represent goods in inventory and in supply line respectively. \( DAVGSL_t \) refers to the desired average supply line and \( DMINI_t \) to desired minimum inventory. Desired average supply line and desired minimum inventory are defined as before:

\[
DAVGSL_t = \tau E_t \quad (5.10)
\]

\[
DMINI_t = E_t + SS \quad (5.11)
\]

in terms of receiving delay \( \tau \), expected demand \( E_t \) and safety stocks \( SS \). Order size \( Q \) and safety stock \( SS \) are initially set as constants. But note that the actual order quantity \( O_t \) (Equation 5.6.) is a variable in this rule. Desired minimum inventory \( DMINI_t \) is defined as the sum of expectations and safety stocks. As such, desired minimum inventory \( DMINI_t \) may be adapted to variations in customer demand. This rule can be shown to be unbiased and mathematically...
consistent in the deterministic case (Özevin, 1999). The inventory reaches equilibrium at the desired minimum inventory level, but in the ‘noisy’ case it may move up and down around the desired minimum, due to differences between the ‘expected’ and ‘actual’ customer demands (Figure 16).

The $(s, S)$ rule takes fully into account the supply line in the sense that the decision is based on the effective inventory $EI = I + SL$. Thus, in

Figure 14. Comparison of (A) non-linear expectation rule with (B) the performance of the player in Game 11 (Short game with orders each period, step-up-and-down in customer demand, exponential delay)
accordance with the fundamental result for the linear anchor and adjust rule, the resulting inventory dynamics is non-oscillatory. The zig-zagging behaviour of the inventory seen in Figure 16 is caused by the time-discrete, piece-wise ordering rule yielding discrete, alternating zero and non-zero orders. (There are also some minor wave-like dynamics caused simply by the autocorrelated random demand.) To explore this point further a modified version of the \((s, S)\) rule

![Figure 15. Performance of order point–order quantity rule \((Q = 4^tD_{r}, DAVGSL_t = 4^tD_{r}, DMINI = 0)\) in Short Game with orders each time unit, known step-up in customer demand, discrete delay](image1)

![Figure 16. Performance of order point–order up to level \((s, S)\) rule \((DAVGSL_t = 4^tE_t, DMINI = E_t, EI_t = I_t + SL_t)\) in Short Game with orders each time unit, Step-up in customer demand, discrete delay](image2)
is run by redefining $EI = I + kSL$, where $k$ is the supply line inclusion coefficient. Normally, for full inclusion of supply line $k$ is one. To demonstrate partial inclusion of supply line, the dynamics of the $(s, S)$ model is illustrated with $k = 0.70$ in Figure 17. Observe that the inventory now does exhibit an oscillatory pattern. So the hypothesis ‘Players that generate oscillatory inventories ignore/underestimate the supply line’ is compelling in general, whether the ordering heuristic is linear, non-linear or piece-wise/discontinuous.

5.3. Review Period, Order up to Level $(R, S)$ Rule

Order point–order up to level rule can be mathematically expressed as:

$$O_t = S - EI_t \quad \text{if} \quad t = Rk$$
$$O_t = 0, \quad \text{otherwise} \quad (5.12)$$

where $EI_t$ represents the effective inventory, $t$ the time, $S$ the upper level of inventory; $k$ is an integer and $R$ is the review period. Effective inventory, the order point and the upper level $S$ of inventory are defined as follows:

$$EI_t = I_t + kSL_t$$

$$S = DAVGSL_t + DMINI_t + R^*E_t$$

$I_t$ and $SL_t$ represent goods in inventory and in supply line respectively. Review period $R$ is five. $DAVGSL_t$ refers to the desired average supply line and $DMINI_t$ to desired minimum inventory. Desired average supply line is defined as:

$$DAVGSL_t = \tau^*E_t$$

in terms of receiving delay $\tau$, expected demand $E_t$. Desired minimum inventory $DMINI_t$ corresponds to the safety stock. $DMINI_t$ is arbitrarily fixed as constant.

This rule can be shown to be unbiased and mathematically consistent in the deterministic case (Özevin, 1999). The inventory reaches equilibrium consistent with the desired minimum inventory level, but in the ‘noisy’ case it may fall below the desired minimum due to noise effects (Figures 18 and 19 provide two illustrations). The basic behaviours are again alternating large discrete and then zero orders and zigzagging inventory patterns.
5.4. \((R, s, S)\) Rule

Finally, the \((R, s, S)\) rule can be mathematically expressed as:

\[
O_t = S - EI_t, \quad \text{if } t = kR \text{ and } EI \leq s
\]
\[
0, \quad \text{otherwise}
\]

where \(S\) represents the upper level of inventory, \(EI_t\) the effective inventory, \(R\) the review period, \(t\) the time and \(s\) the order point. Review period \(R\) is five. Effective inventory, the order point, the upper level \(S\) of inventory and safety stock \(SS\) are:

\[
EI_t = I_t + SL_t
\]
\[
SS = R^*E_t + DMINI_t
\]
\[
s = DAVGSL_t + SS_t
\]
\[
S = s + R^*E_t
\]

\(I_t\) and \(SL_t\) represent goods in inventory and in supply line respectively. \(DAVGSL_t\) refers to the desired average supply line, \(DMINI_t\) to desired minimum inventory, \(SS\) to safety stock, \(E_t\) to expected demand. Desired average supply line is defined as:

\[
DAVGSL_t = \tau^*E_t
\]

in terms of receiving delay \(\tau\) and expected demand \(E_t\). \(DMINI_t\) is determined as a constant. This rule can be shown to be unbiased and mathematically consistent in the deterministic case (Özevin, 1999). Due to the difference between the expected and actual customer demand in the noisy case, the \((R, s, S)\) rule may not result in ordering each time the effective inventory falls to or below the order point; orders are sometimes delayed until the following period. Consistent with the desired minimum inventory level, the inventory reaches equilibrium, but in the ‘noisy’ case it may occasionally fall below the desired minimum (see Figure 20). The fundamental patterns of behaviour are similar to those seen with the previous discrete inventory rules.

5.5. Comparison of the Standard Inventory Rules with Game Results

As discussed above, the order point–order quantity \((s, Q)\) rule is not a plausible decision rule...
formulation when demand is not constant. On the other hand, the order point–order up to level \((s, S)\) rule may provide an adequate representation of subjects’ performance in cases where the continuous anchoring and adjustment rules are inadequate. One such case is depicted in Figures 21 and 22. Observe that the ordering behaviour of subjects is of class ii, consisting of alternating large-and-zero discrete orders, reasonably well represented by the \((s, S)\) orders of Figure 21. The zigzagging...
inventory patterns of Figure 21 and 22 are also consistent. The linear or non-linear continuous adjustment rules, on the other hand, cannot produce such behaviour patterns, especially if orders are allowed in each time step.

The ordering patterns generated by the review period–order up to level \((R, S)\) rule, on the other hand, are generally quite similar to the ones produced by the anchoring and adjustment rules when the value of \(R\) is matched with minimum
ordering interval (five days in our experiments). Therefore, the \((R, S)\) rule does not provide novel behaviour patterns that anchoring and adjustment rules fail to represent.

But the \((R, s, S)\) rule can represent subjects’ decisions where anchoring and adjustment rules fail, especially in cases where intervals between subjects’ discrete non-zero orders are not constant. One such comparison is depicted in Figures 23 and 24. Note that continuous anchor and adjust rules, even if applied under ‘orders are given every five time steps’ conditions,

Figure 23. Performance of \((R, s, S)\) rule (DAVGSLL = 10\*Et, DMINI = 0) in Long Game with orders once every five time units, step-up-and-down in customer demand, discrete delay

Figure 24. Performance of the player in Game 45 (Long Game with orders once every five time units, step-up-and-down in customer demand, discrete delay)
would fail to generate variable time intervals in between non-zero orders.

6. CONCLUSION

This paper has two different research objectives. The first is to analyse the effects of selected experimental factors on the performances of subjects (players) in a stock management simulation game. To this end, the generic stock management problem is chosen as the interactive gaming platform. Gaming experiments are designed to test the effects of three factors on decision-making behaviour: different patterns of customer demand, minimum possible order decision (‘review’) interval and finally the type of receiving delay. ANOVA results of these three-factor, two-level experiments reveal which factors have significant effects on 10 different measures of behaviour (such as max–min range of orders, inventory amplitudes, periods of oscillations and backlog durations). In most cases, the significant results are highly significant and they can be logically explained. Discrete delays have significant destabilizing effects on inventory oscillations compared with continuous delays. On the other hand, the step–up–and–down demand pattern has no obvious effect on inventory dynamics, when compared with step-up only demand pattern. The length of order interval has strong destabilizing effects on all performance measures. In a few cases the results are mixed and/or confounded and further experimental research is suggested.

The second research objective is to evaluate the adequacy of standard decision rules typically used in dynamic stock management models and to seek improvement formulations. The performances of subjects are compared against some selected ordering heuristics (formulations). First, a classification of the patterns of ordering behaviour of subjects reveals three basic types: (i) smooth damping (oscillatory or non-oscillatory) orders, (ii) alternating large-and-zero orders, like a high-frequency discrete signal, (iii) long periods of constant orders punctuated by a few sudden large ones. Comparing these pattern types with simulation results using different decision heuristics, we observe that the common linear ‘anchoring and adjustment rule’ can represent the first (i) type of behaviour, but cannot, due to its linear nature, replicate the other two types. Several alternative non-linear rules are formulated and tested against subjects’ behaviours. Some standard discrete inventory control rules—such as \((s, Q)\)—common in the inventory management literature, are also formulated and tested. The non-linear adjustment rules are found to be more representative of subjects’ ordering patterns in cases where subjects’ ordering patterns fall in class iii above (long periods of constant orders punctuated by a few sudden large ones). The standard discrete inventory rules are more representative of the subjects’ ordering behaviour in cases where decision patterns consist of alternating large then zero discrete orders. Another finding is the fact that the well-documented oscillatory dynamic behaviour of the inventory is a quite general result, not just an artifact of the linear anchor and adjust rule. When the supply line is ignored or underestimated, large inventory oscillations result, not just with the linear anchor-and-adjust rule but also with the non-linear rules, as well as the standard inventory management rules. Furthermore, depending on parameter values, non-linear ordering rules are more prone to instability—even if the supply line is taken into account.

More research is needed on both the experimental and methodological aspects of our work. Additional experiments can be run so as to increase the sample size, increase the test power and test the normality assumption. The way our experiments are designed, it is impossible to separate the ‘decision interval length’ and ‘delay length’ effects. In further experimental research, these two different ‘length effects’ can be separately analysed with respect to each output measure. More research is needed to formulate and test other non-linear formulations. There is also a need to test these rules in more complex and realistic game environments (such as more stocks, delays and multi-player supply chains). Currently we are examining the effects of information delays in the order decisions (in addition to supply line delays). Initial results indicate that the oscillations become more
unstable and the basic anchor and adjust heuristic needs to be modified properly in order to reduce the instability. Finally, a long-term research question is this: if individuals tend to order intermittently and discrete delays further exacerbate the situation, are mostly continuous system dynamics decision structures invalid? Or is it true that even if individuals decide in discrete and intermittent fashion, the macro decision-making behaviour of the aggregate system (that we are typically interested in) can/should be modelled continuously? There is a fundamental research question here that should bridge the micro behaviour of actors to the macro behaviour of the aggregate system.

ACKNOWLEDGEMENT

This work was supported by Boğaziçi University Research Fund no. 02R102. The authors would also like to thank the guest editors and anonymous referees for useful suggestions.

REFERENCES


Forrester JW. 1961. Industrial Dynamics. Pegasus Communications: Waltham, MA.


APPENDIX A: INTERACTIVE INVENTORY
MANAGEMENT GAME INTERFACE

APPENDIX B: GAME INSTRUCTION SHEET

Inventory Management Game (Short Version)

1. Objective
In this game, as an inventory manager, you will control the inventory level of a certain good such that your company is not required to backlog too many orders from your customers. You can achieve this by keeping a large safety stock. However, large safety stocks result in high inventory costs. Therefore, you should keep your inventory level as low as possible while trying not to backlog.
2. How the Inventory is Controlled

You will control your inventory by ordering new goods. You can order new goods once every
day. While ordering new goods you should consider the following three variables: inventory,
demand and supply line. Inventory is the quantity of goods you have in hand. Demand is the
quantity of goods requested within a time unit (a day). If there are not enough goods in your
inventory at any time, you will take the order as a backlog and supply the goods later. In this case,
your inventory will be negative until you receive enough goods from your suppliers. Supply line
corresponds to the goods you have ordered previously but have not received yet. Remember that
there are time delays between your placing of orders and receiving them.

3. Remarks

- There is no cost associated with ordering. Therefore you may order as frequently as every time
  period.
- By inspecting the graph of inventory and your orders over time, you may have some idea
  about the supply line delay and the future demand.
- You can enter your order either by typing it or by sliding the input device. After entering your
  order, press the play button.
- The game will last 100 days. At the end, please save your game under C:\STELLA5\ as
  yourname.stm

APPENDIX C: MODEL EQUATIONS (ILLUSTRATED WITH LINEAR ANCHOR AND
ADJUST RULE)

Demand Sector

demand = step_input + noise
noise = SMTH1(whitenoise, 2, 0)
step_input = 20 + STEP(20, 4) - STEP(20, 19)
whitenoise = NORMAL(0, 0.15* step_input, 67779)

Goal Sector

desired_inventory = inventory_coefficient* step_input
desired_onorder = step_input* receiving_delay
inventory_coefficient = 2

Order Sector

adjtime1 = 5
adjtime2 = 3
decision_rule = demand + inventory_adjustment + supplyline_adjustment  
inventory_adjustment = inventory_discrepancy / adjtime2  
supplyline_adjustment = supplyline_discrepancy / adjtime1  

**Production Line Sector**

inventory(t) = inventory(t - dt) + (receiving - delivery) * dt  
INIT inventory = inventory_coefficient * step_input  

INFLOWS:  
receiving = supplyline / receiving_delay  

OUTFLOWS:  
delivery = demand  
supplyline(t) = supplyline(t - dt) + (order - receiving) * dt  
INIT supplyline = 80  

INFLOWS:  
order = decision_rule  

OUTFLOWS:  
receiving = supplyline / receiving_delay  
inventory_discrepancy = desired_inventory - inventory  
receiving_delay = 4  
supplyline_discrepancy = desired_onorder - supplyline