

When Less Leads to More: Phantom Ordering in the Beer Game?

Gokhan Dogan

*MIT Sloan School of Management, 30 Wadsworth Street, E53-364, Cambridge,
Massachusetts 02142*

gdogan@mit.edu

John Sterman

*MIT Sloan School of Management, 30 Wadsworth Street, E53-351, Cambridge,
Massachusetts 02142*

jsterman@mit.edu

Abstract

We analyze experimental data from the Beer Game in which the customer orders are constant (4 cases/week) and all the subjects are informed about this fact before the game starts. Even though the experimental settings disfavor oscillation and amplification, we still observe them. To analyze the decisions made by the subjects, we first estimate the decision rule used by Sterman (1989). This analysis suggests that typically subjects do not understand the time delays and the stock and flow structure of the Beer Game. Next, we relax some assumptions of this decision rule and use more sophisticated alternatives. These alternative decision rules do not yield overall improvement in terms of fit to the real data. However, for some subjects, these decision rules lead to significant improvement. Our analysis reveals strong evidence that these subjects were caught up in a reinforcing phantom ordering loop even though the experimental conditions strongly disfavor such behavior.

Introduction

Previous studies have shown that subjects' performance at the Beer Game is far from optimal because most subjects do not understand the time delays and the implications of the stock and flow structure of the game (Sterman 1989). In the game, subjects manage a supply chain that consists of four positions: retailer (R), wholesaler (W), distributor (D) and factory (F). The retailer receives orders from an exogenous customer and the pattern of the exogenous customer orders is determined by the experimenter. Sterman (1989) uses a step function that is equal to 4 cases per week in the first four weeks and then increases to 8 cases per week at week 5 (See Figure 1).

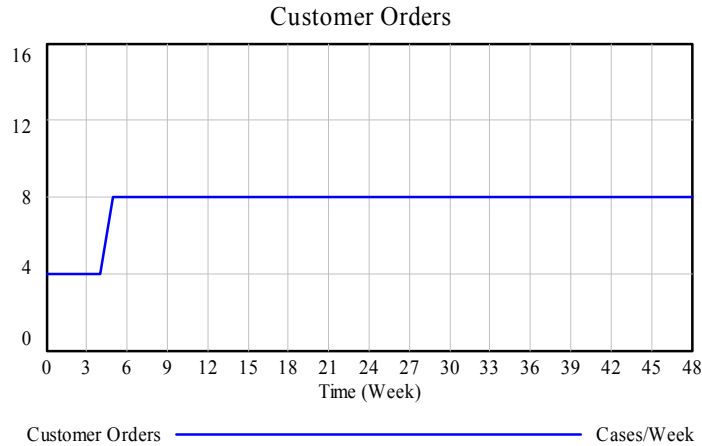


Figure 1: Exogenous customer orders used by Sterman (1989).

After receiving the customer orders, the retailer makes the decision of how much to order from the wholesaler. The wholesaler makes a similar decision and orders from the distributor, and so on. The goal of the subjects is to minimize total supply chain costs. In the game, costs are incurred for holding inventory or having backlog. Even though the customer orders follow a step function and they do not oscillate, subjects generate large fluctuations and these oscillations are amplified as one moves upstream from the retailer to factory (Sterman 1989). As a result of oscillation and amplification, inventory levels deviate from the optimal level of 0 and supply chain costs increase. Sterman (1989) estimated the decision rules of the subjects econometrically and showed that the main cause of oscillation and amplification is that the subjects typically underweight or ignore the time delays in the system, especially the supply line of beer on order.

In this paper, we use a different data set than the one used by Sterman (1989). In the new data set the customer demand was constant: equal to 4 cases per week and this information was announced to the subjects (See Figure 2).

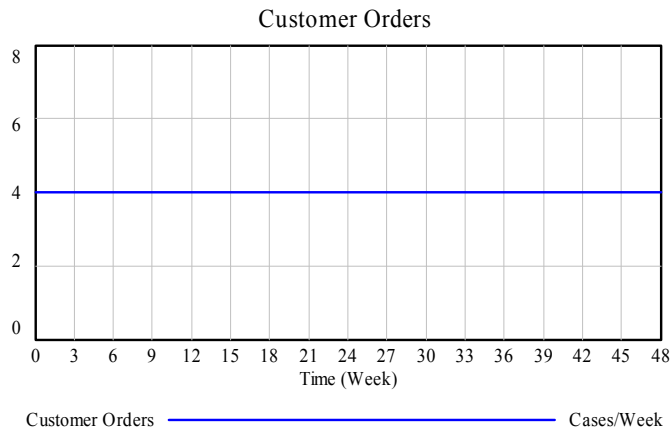


Figure 2: Exogenous customer orders used in this paper.

Before the start of the game, it was confirmed with a manipulation check that the subjects understood this fact. The supply chain was initialized at equilibrium and in some treatment groups initial inventory was equal to the optimal level of 0. Under these circumstances, the optimal strategy is clearly to order 4 cases per week throughout the game. However, we still observe oscillation and amplification for most subjects (See Figure 3 for a typical team).

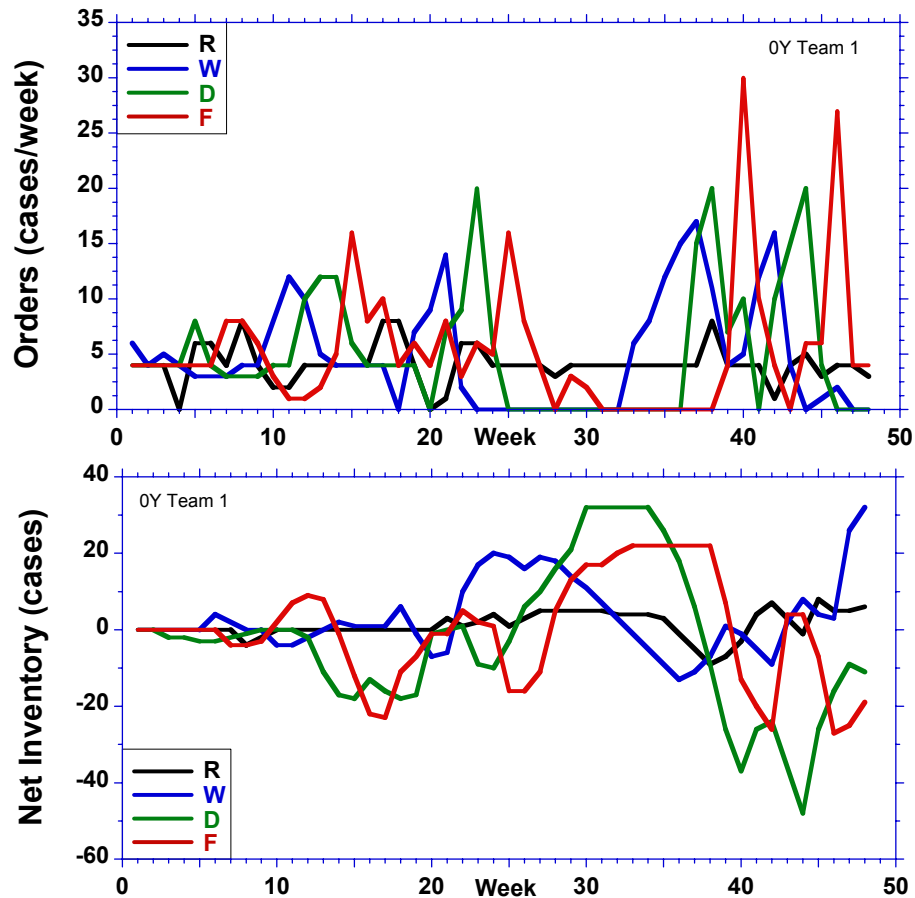


Figure 3: Orders placed and net inventory of a typical team.

Only 9 subjects out of the 240 subjects participated in the experiment ordered exactly 4 cases per week throughout the game. The maximum orders placed by a subject in one week were 40000 cases per week (See Figure 4). The standard deviation of all orders placed by all subjects was 761.86 cases per week. Croson et. al. find that subjects deviate from the optimum because they want to hold *coordination stock* against the risk that other subjects will not behave optimally (Croson et. al. 2004).

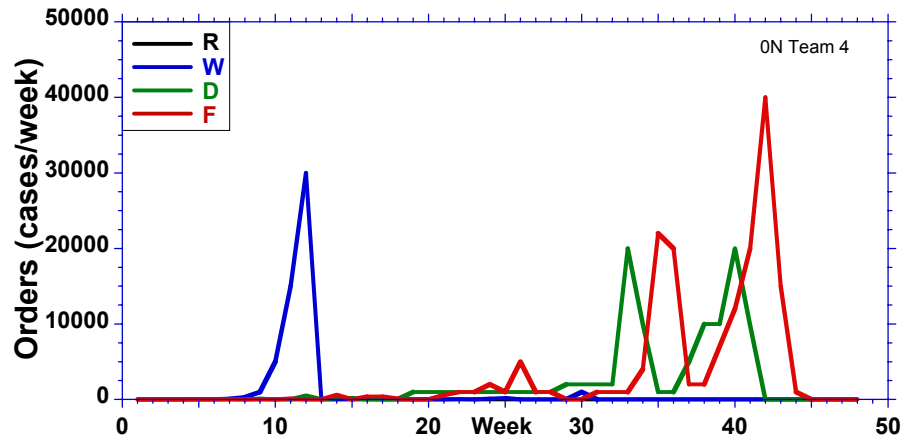


Figure 4: Orders placed by the team that includes the highest orders placed per week

The game was run for 48 weeks for each team during the experiment. A web-based simulator was used for conducting the experiment, so data series to be used in estimating the subjects' decision rules do not contain any accounting or measurement errors. This data set is of higher quality than most data collected in real organizations because typically the data collected at real organizations are subject to unknown measurement and reporting error. We will use this data set to estimate the subjects' decision rules in the next section.

Decision Rule Estimation

Following Sterman (1989), we base the decision rule on the anchoring and adjustment heuristic (Tversky and Kahneman 1974). The rule uses expected customer orders as the anchor and adjusts for the discrepancy between desired inventory and net inventory and the discrepancy between the desired supply line of beer on order and actual supply line. In the formulation, we use a non-negativity constraint since the orders cannot be negative:

$$O_t = \text{MAX}(0, CO_t^e + AS_t + ASL_t + \varepsilon_t) \quad (1)$$

where CO^e represents orders expected from the subject's customer next period. AS is the adjustment made to reduce the discrepancy between desired inventory and inventory and ASL is the adjustment made to reduce the discrepancy between desired supply line and actual supply line. ε is the additive disturbance term. We formulated the expected customer orders using exponential smoothing:

$$CO_t^e = \theta * IO_{t-1} + (1 - \theta) * CO_{t-1}^e \quad (2)$$

where IO is actual incoming orders. This equation represents the possibility that subjects do not forecast incoming orders optimally (4 cases per week) but instead respond to the actual orders they receive from their customers.

Inventory adjustment is linear in the discrepancy between desired inventory (S^*) and net inventory (S):

$$AS_t = \alpha_S (S_t^* - S_t) \quad (3)$$

where α_S is the fraction of inventory discrepancy ordered each period. Similarly, the supply line adjustment formulation is also linear in the discrepancy between the desired supply line (SL^*) and the actual supply line (SL):

$$ASL_t = \alpha_{SL} (SL_t^* - SL_t) \quad (4)$$

where α_{SL} is the fraction of supply line discrepancy ordered each period. So, orders placed equals:

$$O_t = \text{MAX}(0, CO_t^e + \alpha_S (S_t^* - S_t) + \alpha_{SL} (SL_t^* - SL_t) + \varepsilon_t) \quad (5)$$

This decision rule assumes that the desired inventory (S^*) and desired supply line (SL^*) are constant. Since customer demand is constant, subjects are informed about this fact and the chain is initialized at equilibrium, it seems reasonable to assume that desired inventory and supply line are fixed. Obviously, the optimum desired inventory is 0 cases since the customer demand is constant and known. Optimum level of desired supply line is customer demand (4 cases per week) times normal acquisition lag (4 weeks for R, W, D and 3 weeks for F), which is 16 cases for R, W, D and 12 cases for F. Of course the assumption that desired inventory and desired supply line are fixed might be relaxed by formulating them as endogenous variables. A subject might want to have enough beer in the supply line that will serve the customer for a time period that equals the expected time it takes the supplier to deliver the beer. In that case, higher expected acquisition lag or higher desired acquisition rate would lead to higher desired supply line. So, desired supply line formulation is:

$$SL_t^* = EAL_t * \text{DesiredAcquisitionRate}_t \quad (6)$$

where EAL is the expected acquisition lag. It equals the expected lag between the time subject places the order and receives the goods. If the supplier stocks out, the acquisition lag would exceed the normal delivery delay of 4 weeks. In that case, the expected acquisition lag would also increase. This desired supply line formulation is in line with the ones used by Forrester (1961, Chapter 15) and Sterman (2000, Chapter 17).

Note that this desired supply line formulation implies the existence of a reinforcing loop: as desired supply line goes up, orders placed increases and supply line goes up. Due to the increasing supply line of beer on order, supplier stocks out and it takes longer for the supplier to ship the orders. As a result, expected acquisition lag increases and desired supply line increases even further (See Figure 5). This reinforcing loop has the potential to destabilize the system very rapidly by increasing orders to very high amounts. This is particularly the case for subjects that try to reduce the inventory discrepancy aggressively.

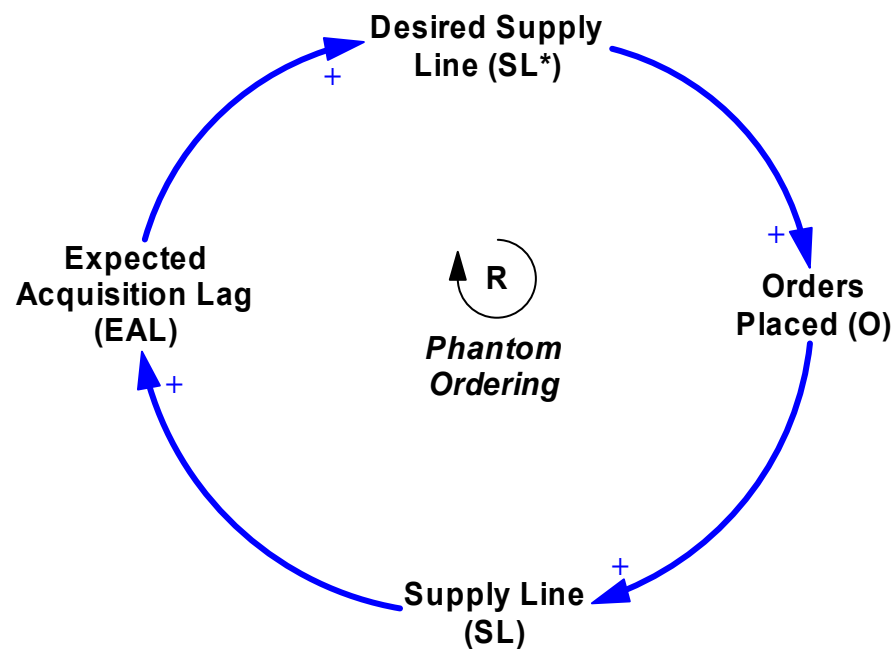


Figure 5: The Phantom Ordering loop.

As mentioned above, the conditions of the experiment strongly disfavor this reinforcing loop: Demand is constant, all the subjects are informed about it and the game is initialized at equilibrium. Furthermore, the subjects know that they will receive their orders eventually since they do not compete against other parties as in real life. In real life, typically a supplier serves more than one customer and if the supplier can only ship less than the total of customer orders, the customers cannot get all the orders they place. So, they might order more than what they need because they know that they will receive less than what they order. This is called *phantom ordering*. The fact that each supplier serves only one customer in our experimental setting disfavors phantom orders even further. Hence, we will assume that desired supply line and desired inventory are constant and estimate the decision rule accordingly.

Assuming S^* and SL^* are constant and defining $\beta = \alpha_{SL}/\alpha_S$ and $S' = S^* + \beta SL^*$, the equations to be estimated become:

$$O_t = \text{MAX}(0, CO_t^e + \alpha_s(S' - S_t - \beta SL_t) + \varepsilon_t) \quad (7)$$

$$CO_t^e = \theta * IO_{t-1} + (1 - \theta) * CO_{t-1}^e$$

So, we need to estimate four parameter values: θ (for CO_t^e), α_s , β and S' . The parameter β is the fraction of supply line the subjects take into account while placing orders. Since the subjects should take into account the supply line as much as their on-hand inventory, the optimum value of β is 1. The optimum value of α_s is also 1 because they should try to reduce the entire inventory discrepancy each period.

We minimize the sum of squared errors between actual orders, AO_t , and model orders, O_t , to estimate the parameters.

$$\text{Min}_{\alpha_s, \beta, \theta, S'} \sum_{t=1}^{48} (AO_t - O_t)^2 \quad (8)$$

subject to

$$0 \leq \theta \leq 1$$

$$0 \leq \alpha_s \leq 1$$

$$0 \leq \beta \leq 1$$

$$0 \leq S'$$

Estimation results are presented in Table 1. Median fraction of inventory discrepancy corrected at each time period (α_s) is 0.27 and median fraction of supply line subjects take into account (β) is 0.16. Both values are far below the optimal value of 1. So, even if the customer demand is constant and this is known by the subjects, most of them ignore the supply line of goods on-order. β is significantly smaller than 1 for 83% of the subjects and α_s is significantly smaller than 1 for 87% of the subjects. β is not significantly bigger than 0 for 45% of the subjects.

Overall, the decision rule captures the orders placed by subjects successfully. The median R^2 is 0.51 and the median Root Mean Square Error (RMSE) is 2.78 cases per week (Croson et. al. 2004). Since the standard deviation of orders placed is 761.86 cases per week, the median RMSE value signals a good fit.

	θ	α_s	β	S'	R^2	RMSE
Median Estimate	0.24	0.27	0.16	5.59	0.51	2.78
Median Width of 95% Confidence Interval ¹	0.84	0.24	0.48	10.10		
N	231	231	212	212		

Table 1: Median of the parameter estimates and widths of 95% confidence intervals².

¹ Confidence intervals are estimated using the bootstrapping method (Dogan 2004).

² As mentioned above, 9 subjects out of 240 ordered 4 cases per week throughout the game, so the parameter estimates for these subjects are not identified. On the other hand, β and S' are not identified when α_s is zero. This was the case for 19 subjects.

As expressed above, the decision rule fits the data well. However, for some subjects the fit is very poor. See Appendix 1 for the plots of these subjects' actual orders and model output. So, as the next step we estimated the alternative decision rule with the endogenous desired supply line (SL*) formulation. In this alternative decision rule, desired supply line is formulated according to equation (6). Desired inventory is fixed as in the original decision rule because we wanted to change only the desired supply line formulation to see its impact. If variable desired supply line formulation has considerable impact on the fit of the model to the data, this will signal that the reinforcing Phantom Ordering loop in Figure 5 is active. After testing the improvement of this alternative decision rule, we will use endogenous formulations for desired inventory as well and assess their impact on the fit.

In equation (6), desired supply line is formulated as the amount of goods that will cover the desired acquisition rate for a time period equal to the expected acquisition lag. We used two formulations for expected acquisition lag (EAL_t) and two formulations for desired acquisition rate (DAR_t). As mentioned above, expected acquisition lag might change over time because, if the subjects receive fewer shipments than what they expect, they would figure out that the supplier has stocked out and it will take longer to receive the orders than the normal acquisition lag. Our first alternative formulation for expected acquisition lag is based on Little's law. According to Little's law, the acquisition lag is equal to the ratio of the supply line (SL) to shipments received at steady state. Also, acquisition lag cannot be smaller than the normal acquisition lag ($NAL = 4$ for R, W, D). So, perceived acquisition lag (PAL) is:

$$PAL = \text{MAX}(NAL, SL/\text{Shipments Received}) \quad (9)$$

The second alternative formulation for expected acquisition lag is based on the fraction of expected deliveries received. If the supplier has not stocked out, shipments received by the subject this period should be equal to what the subject ordered 4 weeks ago (for R, W and D). If they receive less than what they ordered 4 weeks ago, it means that the supplier has stocked out. In that case, perceived acquisition lag would go up. Again, perceived acquisition lag (PAL) cannot be smaller than normal acquisition lag of 4 weeks. So, PAL is equal to³:

$$PAL = NAL * \text{Effect of Deliveries on PAL} \quad (10)$$

$$\text{Effect of Deliveries on PAL} = \text{MAX}(1, \text{Expected Deliveries} / \text{Shipments Received})$$

$$\text{Expected Deliveries} = \text{Orders Placed}_{t-NAL}$$

We also considered the possibility that the subjects might anchor their expectations about the acquisition lag on the normal acquisition lag. So, we used a weighted average of normal acquisition lag and perceived acquisition lag for the expected acquisition lag formulation.

$$EAL = w*NAL + (1-w)*PAL \quad (11)$$

where w is the weight on normal acquisition lag and it is a parameter to be estimated.

³ In both PAL formulations, acquisition lag gets very big as shipments received approaches zero. When it is exactly zero, the ratio is not determined. In this case, we used the value of 1 for shipments received since it is the next smallest integer value for shipments received and makes the ratio as big as possible.

As mentioned above, we used two alternatives for Desired Acquisition Rate (DAR) as well. First one assumes that the subjects adjust the supply line according to the expected loss rate. In this case, desired acquisition rate is equal to expected customer orders (CO^e).

$$DAR = CO^e \quad (12)$$

The second alternative is more sophisticated and it assumes that in addition to the expected loss rate, subjects also account for the temporary gaps between desired inventory and actual inventory while adjusting the supply line (Sterman 2000, Chapter 17). In that case, the desired acquisition rate is:

$$DAR = CO^e + \alpha_s(S^* - S) \quad (13)$$

In brief, for the alternative decision rule with variable desired supply line (SL^*) and fixed desired inventory (S^*), we used four alternative formulations for SL^* since we have two alternative formulations for the expected acquisition lag (EAL) and two for desired acquisition rate (DAR). Note that the factory does not have a supplier and the delivery delay is always three weeks for the factory. So, EAL is 3 weeks for the factory all the time. Thus, we have two alternative formulations for the factory.

The means and medians of parameter estimates and summary statistics for the alternative decision rule are presented in Appendix 2. Most subjects underweight the time delays and ignore the supply line according to the alternative decision rule as well. Mean and median β values are substantially smaller than 1. So, essentially the alternative decision rule reaches the same conclusion as the original decision rule in terms of the extent to which the subjects understand the time delays.

Appendix 2 also reveals that the alternative decision rule's performance is very close to that of the original decision rule in terms of R^2 and RMSE. In addition, we formally tested the hypotheses that the alternative decision rule's R^2 and RMSE values have the same mean and median with the original decision rule and they come from the same distribution as the original decision rule. We used nonparametric tests for testing the equality of the medians and distribution and we used the t-test for the equality of the means. We were not able to reject any of these hypotheses with 95% significance level for R, W and D, so the alternative decision rule does not lead to overall improvement. For F, we would expect the alternative decision rule to lead to more improvement because standard deviation of the incoming orders to the factories is higher than R, W and D due to the amplification of orders. Hence, F might benefit more from the variable desired supply line formulation since it accounts for changes in incoming orders. Nevertheless, we were not able to reject the hypotheses that the means and medians of R^2 and RMSE of the original decision rule are equal to those of the R^2 and RMSE of the alternative rule (95% significance level). We only rejected the hypotheses that R^2 and RMSE of the original rule and the alternative rule are from the same distribution for F, which is an expected result.

So, the endogenous desired supply line formulation does not lead to overall improvement. However, for the 16 subjects that the original decision rule performed very poorly, the new decision rule improves the fit considerably (See Appendix 3). 11 of these subjects are R, W and D and 5 of them are F. The variable supply line formulation improves mean value of R^2 from 0.24 to 0.56 and the median from 0.28 to 0.59 for R, W and D. We reject the hypotheses that the means and medians of the R^2 of the original decision rule are equal to the R^2 of the alternative rule (99% significance level). Also, their distributions are not identical (99% significance level). Median RMSE values are also not equal and their distributions are not identical (99% significance level). We cannot reject the hypothesis that the mean RMSE values are equal ($p=0.09$) but we have only 11 data points and t-test is not very reliable for such a small sample. Thus, these results present strong evidence that the alternative decision rule leads to significant improvement for the R, W and D subjects for which the original decision rule performs poorly.

On the other hand, we have weaker evidence of improvement for the F subjects for which the original decision rule performs poorly. We cannot reject the hypothesis that the medians of R^2 are equal for the original decision rule and the alternative decision rule (95% significance level). For RMSE, we cannot reject the hypothesis that their means and medians equal (95% significance level).

Hence, especially for the R, W and D subjects for which the original decision rule performs poorly, alternative decision rule leads to significant improvement. These findings signal that the reinforcing Phantom Ordering loop in Figure 5 might be operating for these subjects. Given that the customer demand is constant and the subjects are informed about it, the presence of the Phantom Ordering loop is striking because the experimental conditions disfavor such a reinforcing loop. Since desired inventory is fixed in the alternative decision rule, the improvement comes from the variable desired supply line formulation. This formulation consists of two factors: expected acquisition lag and desired acquisition rate. The Phantom Ordering loop is active if the expected acquisition lag formulation has a significant influence on the improvement. To test the influence of the expected acquisition lag on the improvement, we estimated a slightly modified version of the decision rule for these subjects. We treated the expected acquisition lag as fixed (4 weeks) in this modified version of the decision rule. If there is significant difference between the performance of the alternative decision rule and its slightly modified version, we would conclude that the variable expected acquisition lag formulation contributes to the performance improvement and hence the Phantom Ordering loop is active. Note that this test is meaningful for R, W and D only because for the factory the acquisition lag is constant and equal to 3 in all cases. For R^2 , we reject the hypotheses that the means and medians of the alternative decision rule and its slightly modified version are equivalent and the distributions are identical (95% significance level). For RMSE, we reject the hypotheses that the medians are equivalent and the distributions are identical. We cannot reject the hypothesis that the means are equal but given that we have only 11 observations and the distribution is not normal, t-test is not a very accurate indicator. So, for these 11 subjects, we have strong evidence that the

variable expected acquisition lag formulation contributes to the performance improvement and the Phantom ordering loop is active.

Furthermore, we have some subtle and indirect evidence from the factories in this subject group of 16 subjects as well. As mentioned above, we have stronger evidence for improvement with the alternative decision rule for the R, W and D subjects than the F subjects. Since the acquisition lag is always fixed for F but not for R, W and D, the stronger improvement evidence for R, W and D signals that the subjects treated acquisition lag as a variable and hence the Phantom Ordering loop was active for these subjects.

After analyzing the results of the first alternative decision rule, we relaxed the assumption that the desired inventory is fixed. We used two alternative formulations for Desired Inventory (S^*):

2nd alternative decision rule:

$$S_t^* = EAL_t * DesiredAcquisitionRate_t \quad (14)$$

$$SL_t^* = EAL_t * DesiredAcquisitionRate_t$$

and

3rd alternative decision rule:

$$S_t^* = DesiredInventoryCoverage * DesiredAcquisitionRate_t \quad (15)$$

$$SL_t^* = EAL_t * DesiredAcquisitionRate_t$$

In both alternative decision rules, we used two alternatives for expected acquisition lag (EAL) and two alternatives for desired acquisition rate (DAR) as we did for the first alternative decision rule. The results are presented in Appendix 4. Again, most subjects ignore the time delays and underweight the supply line. Mean and median values of β are substantially smaller than 1. Also, mean and median R^2 and RMSE values are close to the original decision rule and first alternative decision rule.

In terms of R^2 , we cannot reject the claim that the means or medians of the decision rule with fixed desired inventory (first alternative) and the third alternative are equal. We also can't reject the hypothesis that they have identical distribution for all positions (R, W, D and F). For RMSE, test results are similar for R, W and D. Only for the factory, we reject the claims that the medians are equal and the distributions are identical. So, the overall performances of the first alternative and third alternative are very similar and the third alternative does not perform significantly better than the first alternative. However, both of them consistently perform better than the second alternative.

None of the alternative decision rules leads to overall improvement when compared to the original decision rule. In fact, the second alternative decision rule performs worse than the original decision rule. On the other hand, the alternative decision rules perform

significantly better than the original rule for the subjects for which the original decision rule performs poorly. See Appendix 5 for a comparison of the performances of the original decision rule and the alternative decision rules for these subjects. Appendix 6 plots the actual orders and model outputs of the best fitting alternative decision rule and compares them with the results of the original decision rule.

Discussion

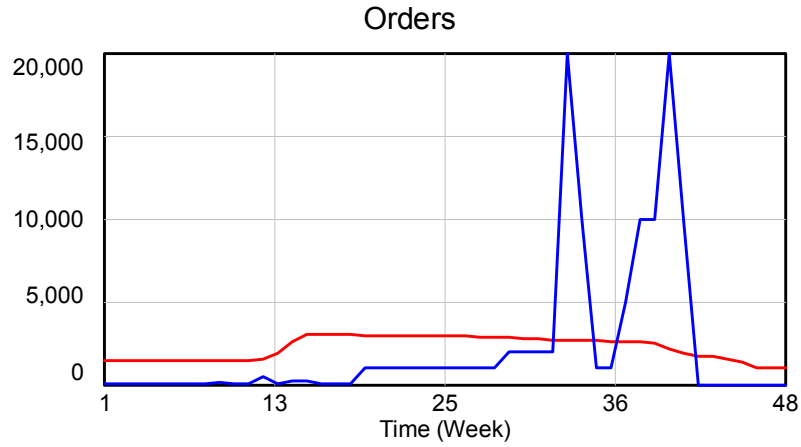
The conditions of the experiment explained in this paper disfavor oscillation and order amplification. Yet, we still observe oscillation and amplification and find strong evidence that most subjects do not understand the time delays and the implications of the stock flow structure of the game. Typically, subjects significantly underweight the supply line of unfilled orders.

Experimental conditions of this paper also disfavor more sophisticated decision rules than the one used by Sterman (1989). As expected, our test results show that more sophisticated decision rules that use variable desired inventory or variable supply line do not yield overall improvement. In fact, they can make the overall performance even worse (second alternative decision rule). So, the assumptions of fixed desired inventory and desired supply line are typically reasonable for this setting. However, even under these conditions, the alternative decision rules yield significant improvement for some subjects. Furthermore, we have evidence that these subjects were caught up in a reinforcing Phantom Ordering loop. Given the experimental conditions, this is a striking finding. Since the subjects do not compete against other individuals for getting their orders from their supplier, phantom ordering is not necessary at all. Yet, we still find evidence that they place phantom orders. This might be a transfer from the real world since in the real world it might be tempting to order more than the necessary amount when the supplier stocks out and puts the orders on allocation. The fact that we find evidence for the existence of Phantom Orders in such an experimental setting suggests that we need further investigation to understand the behavioral causes of phantom ordering.

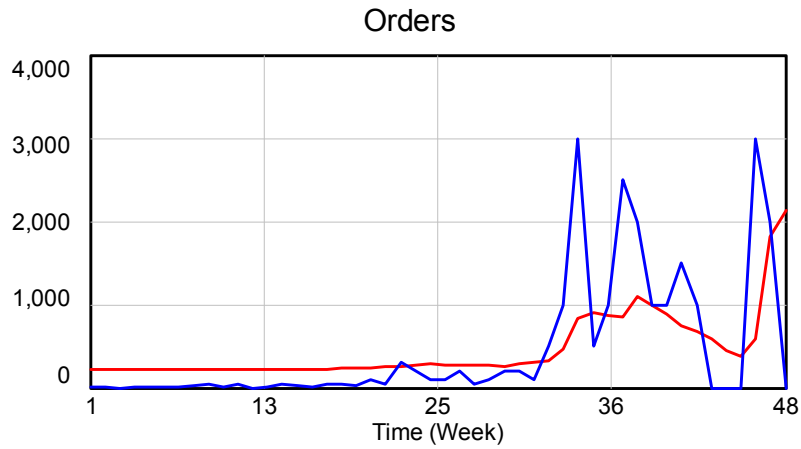
Appendix 1: Plots of actual orders and model output for the subjects for which the original decision rule performs poorly.

Actual Orders: Blue
Model Output: Red

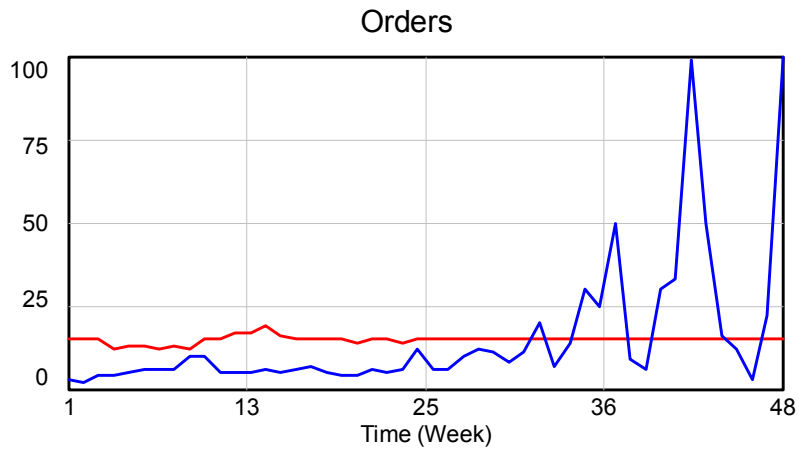
- $R^2= 0.04$ and $RMSE=4512.07$



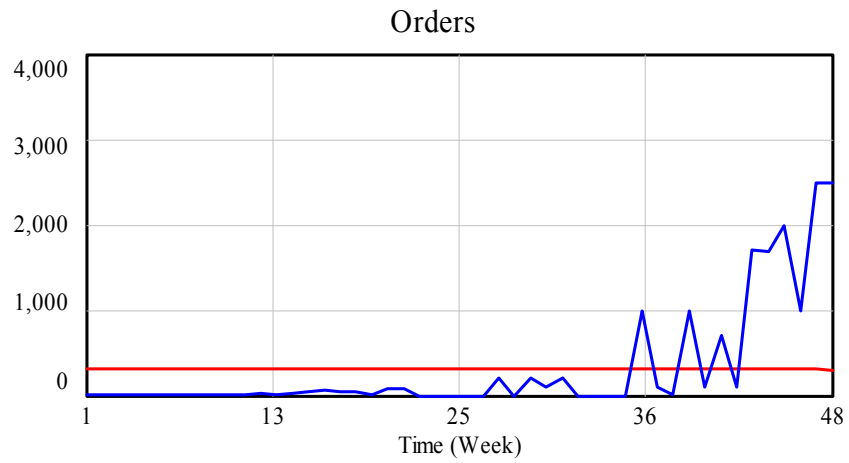
- $R^2= 0.29$ and $RMSE=668.67$



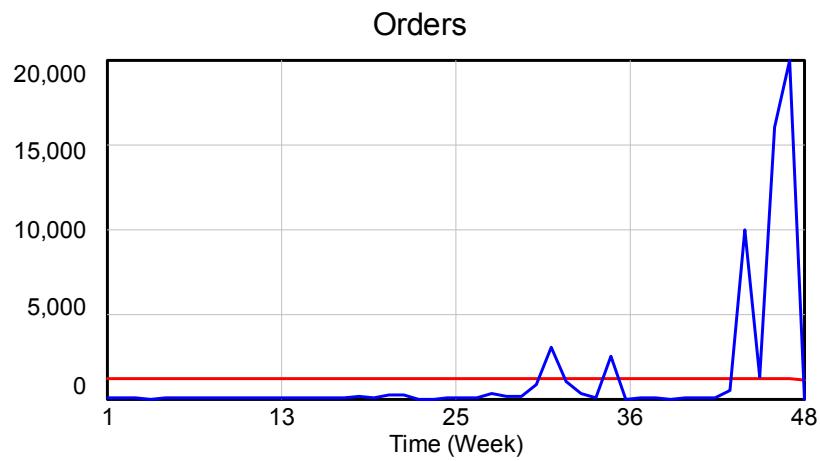
- $R^2= 0.00$ and $RMSE=20.62$



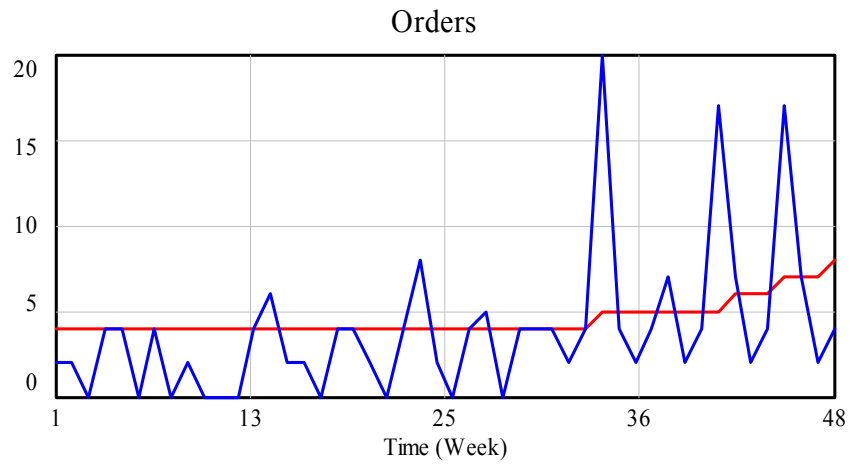
- $R^2= 0.68$ and $RMSE=662.76$



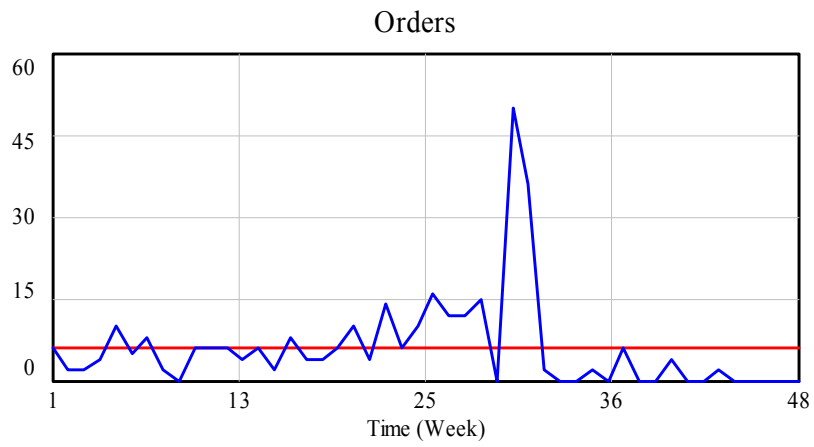
- $R^2= 0.28$ and $RMSE=3845.28$



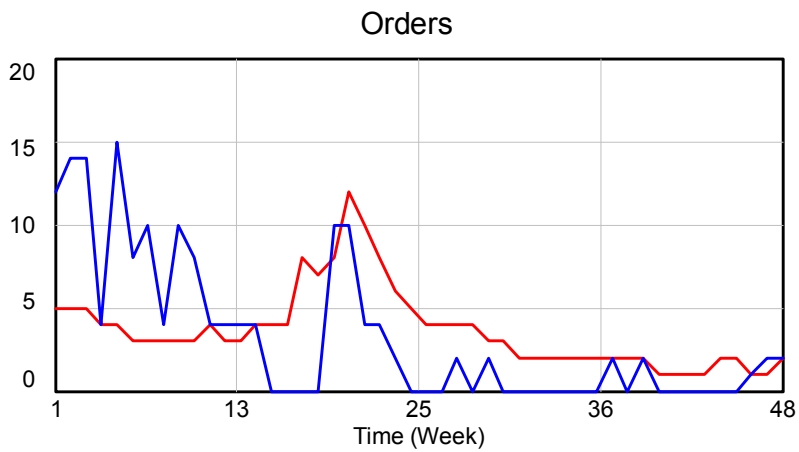
- $R^2= 0.14$ and $RMSE=4.00$



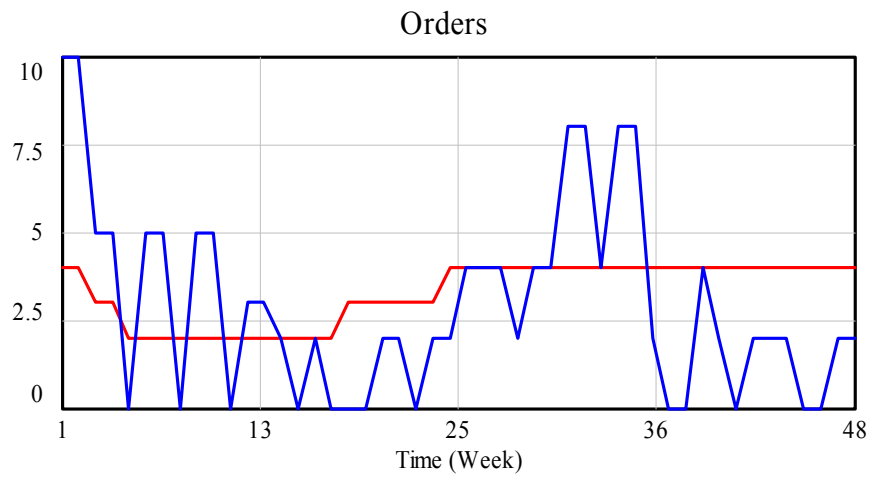
- $R^2= 0.00$ and $RMSE=8.96$



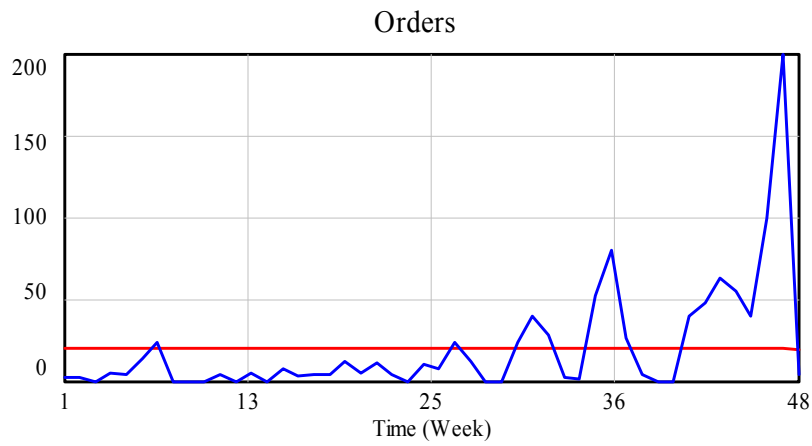
- $R^2= 0.14$ and $RMSE=4.16$



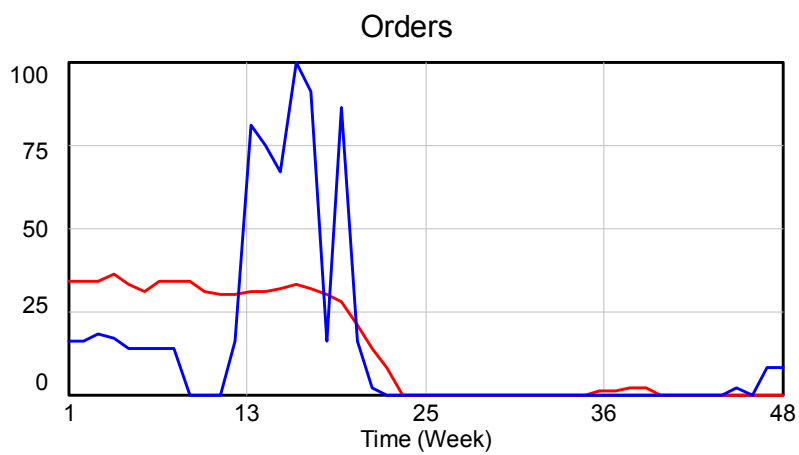
- $R^2= 0.05$ and $RMSE=2.69$



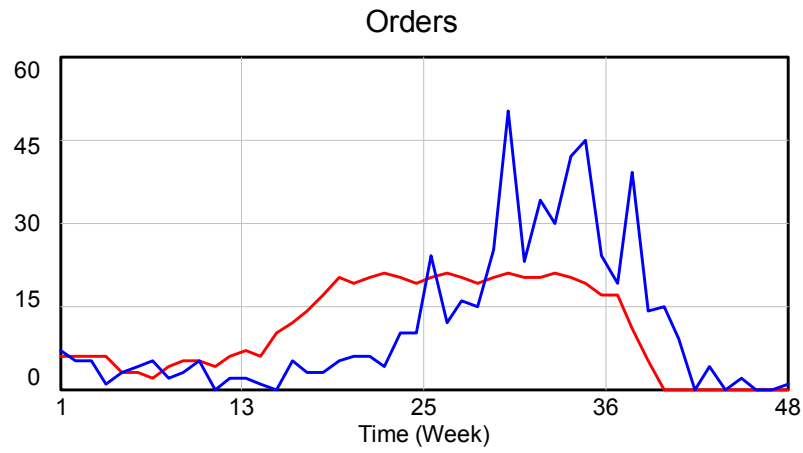
- $R^2= 0.00$ and $RMSE=34.68$



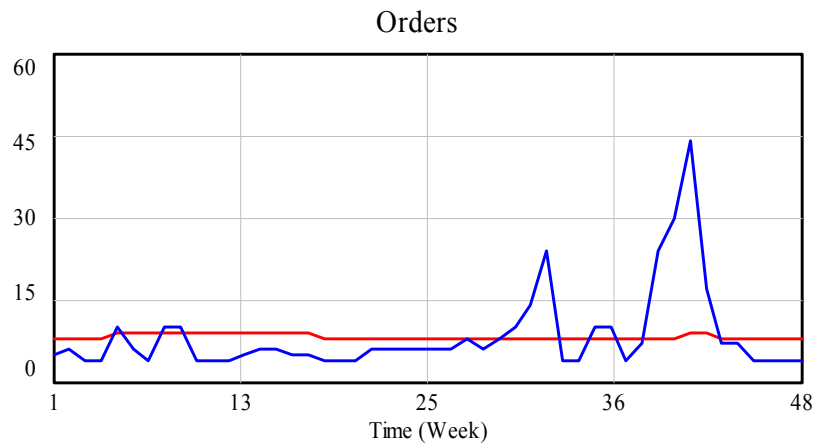
- $R^2= 0.33$ and $RMSE=22.12$



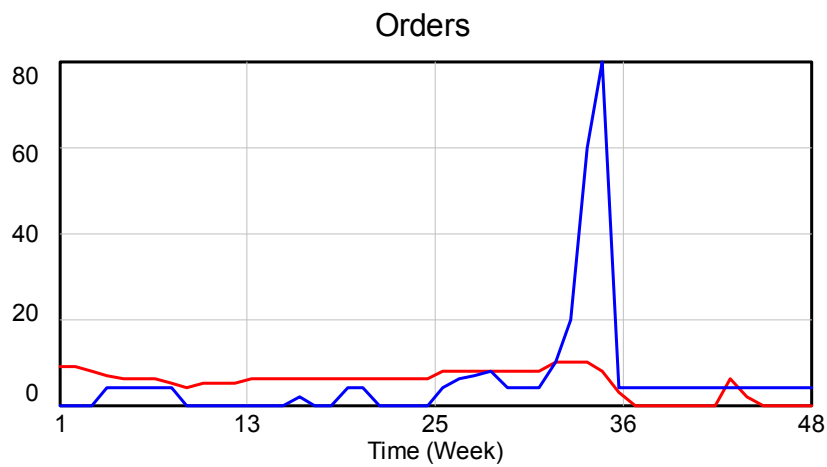
- $R^2=0.35$ and $RMSE=10.55$



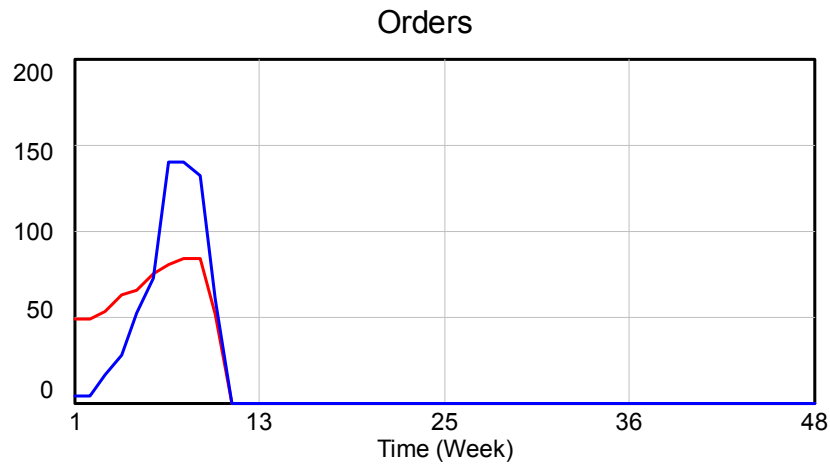
- $R^2=0.01$ and $RMSE=7.56$



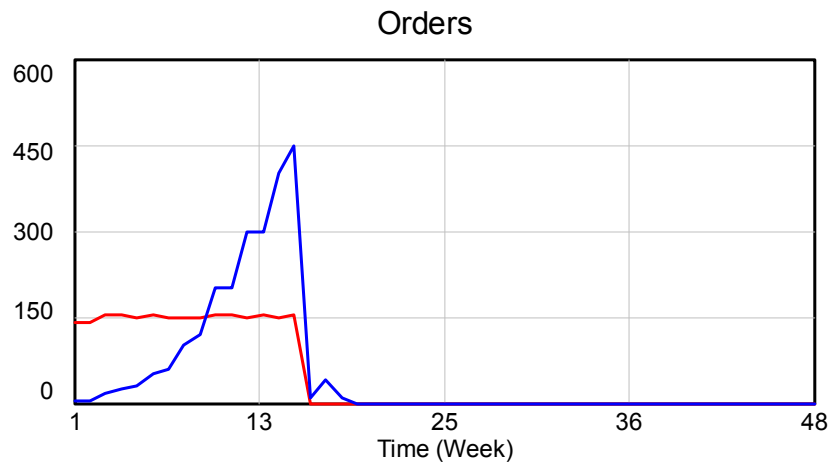
- $R^2=0.07$ and $RMSE=13.47$



- $R^2= 0.75$ and $RMSE=18.22$



- $R^2= 0.43$ and $RMSE=80.42$



Appendix 2:

a) Results of the alternative decision rules with variable desired supply line (SL*) and fixed desired inventory (S*) for all subjects. Table shows the means of parameter estimates and summary statistics for R, W and D.

Desired Inventory	Expected Acquisition Lag Based on:	Desired Acquisition Rate	w	θ	α_s	β	S*	Desired Inv Coverage	S'	R ²	RMSE
<i>Fixed</i>	Little's Law	CO ^e	0.58	0.41	0.27	0.25	43.67	-	-	0.47	84.26
		CO ^e + $\alpha_s(S^*-S)$	0.59	0.40	0.32	0.21	94.01	-	-	0.48	76.25
	Fraction of Expected Deliveries Received	CO ^e	0.59	0.35	0.26	0.26	85.27	-	-	0.46	86.07
		CO ^e + $\alpha_s(S^*-S)$	0.58	0.32	0.31	0.20	9895.47	-	-	0.47	78.13
Original Decision Rule											
<i>Fixed</i>	Fixed	Fixed	-	0.36	0.36	0.28	-	-	42.83	0.49	90.42

b) Means of parameter estimates and summary statistics for F.

Desired Inventory	Expected Acquisition Lag	Desired Acquisition Rate	w	θ	α_s	β	S*	Desired Inv Coverage	S'	R ²	RMSE
<i>Fixed</i>	Fixed (3)	CO ^e	-	0.19	0.31	0.27	28.61	-	-	0.48	119.24
		CO ^e + $\alpha_s(S^*-S)$	-	0.29	0.40	0.20	30.50	-	-	0.50	119.42
Original Decision Rule											
<i>Fixed</i>	Fixed	Fixed	-	0.32	0.45	0.44	-	-	32.34	0.55	112.03

c) Medians of parameter estimates and summary statistics for R, W and D.

Desired Inventory	Expected Acquisition Lag Based on:	Desired Acquisition Rate	w	θ	α_s	β	S*	Desired Inv Coverage	S'	R ²	RMSE
<i>Fixed</i>	Little's Law	CO ^e	0.67	0.39	0.19	0.08	2.28	-	-	0.47	2.57
		CO ^e + $\alpha_s(S^*-S)$	0.70	0.38	0.24	0.07	2.32	-	-	0.48	2.50
	Fraction of expected deliveries received	CO ^e	0.68	0.27	0.18	0.08	2.12	-	-	0.46	2.65
		CO ^e + $\alpha_s(S^*-S)$	0.68	0.19	0.22	0.07	2.08	-	-	0.47	2.64
Original Decision Rule											
<i>Fixed</i>	Fixed	Fixed	-	0.28	0.26	0.11	-	-	4.79	0.48	2.63

d) Medians of parameter estimates and summary statistics for F.

Desired Inventory	Expected Acquisition Lag	Desired Acquisition Rate	w	θ	α_s	β	S*	Desired Inv Coverage	S'	R ²	RMSE
<i>Fixed</i>	Fixed (3)	CO ^e	-	0.19	0.31	0.27	28.61	-	-	0.48	4.26
		CO ^e + $\alpha_s(S^*-S)$	-	0.29	0.40	0.20	30.50	-	-	0.50	4.27
Original Decision Rule											
<i>Fixed</i>	Fixed	Fixed	-	0.16	0.41	0.43	-	-	7.98	0.57	3.01

Appendix 3: Results for the subjects for which decision rules with variable desired supply line (SL*) and fixed desired inventory (S*) improves the fit the most when compared to the original decision rule.

a) Retailer, wholesaler and distributor.

Subject	Desired Supply Line	Desired Inventory	R ² of Best Fitting Alternative Decision Rule	R ² of Original Decision Rule (Sterman 1989)	RMSE of Best Fitting Alternative Decision Rule	RMSE of Original Decision Rule (Sterman 1989)
1	Variable	Fixed	0.23	0.05	2.46	2.69
2	Variable	Fixed	0.34	0.04	3804.00	4512.07
3	Variable	Fixed	0.59	0.28	2501.00	3845.28
4	Variable	Fixed	0.65	0.29	480.83	668.67
5	Variable	Fixed	0.53	0.33	18.65	22.12
6	Variable	Fixed	0.37	0.00	27.96	34.68
7	Variable	Fixed	0.77	0.43	53.51	80.42
8	Variable	Fixed	0.66	0.00	12.91	20.62
9	Variable	Fixed	0.83	0.68	274.03	662.76
10	Variable	Fixed	0.48	0.14	3.33	4.00
11	Variable	Fixed	0.70	0.35	7.15	10.55
		Mean	0.56	0.24	653.26	896.72
		Median	0.59	0.28	27.96	34.68

b) Factory.

Subject	Desired Supply Line	Desired Inventory	R ² of Best Fitting Alternative Decision Rule	R ² of Original Decision Rule (Sternan 1989)	RMSE of Best Fitting Alternative Decision Rule	RMSE of Original Decision Rule (Sternan 1989)
12	Variable	Fixed	0.34	0.07	11.40	13.47
13	Variable	Fixed	0.15	0.00	8.51	8.96
14	Variable	Fixed	0.97	0.75	6.26	18.22
15	Variable	Fixed	0.56	0.14	3.01	4.16
16	Variable	Fixed	0.25	0.01	6.64	7.56
		Mean	0.45	0.19	7.17	10.48
		Median	0.34	0.07	6.64	8.96

Appendix 4:

a) Results of the alternative decision rules with variable desired inventory (S^*) are added to the tables at Appendix 2. Table shows the means of parameter estimates and summary statistics for R, W and D.

Desired Inventory	Expected Acquisition Lag Based on:	Desired Acquisition Rate	w	θ	α_s	β	S^*	Desired Inv Coverage	S'	R^2	RMSE
Fixed	Little's Law	CO^e	0.58	0.41	0.27	0.25	43.67	-	-	0.47	84.26
		$CO^e + \alpha_s(S^*-S)$	0.59	0.40	0.32	0.21	94.01	-	-	0.48	76.25
	Fraction of Expected Deliveries Received	CO^e	0.59	0.35	0.26	0.26	85.27	-	-	0.46	86.07
		$CO^e + \alpha_s(S^*-S)$	0.58	0.32	0.31	0.20	9895.47	-	-	0.47	78.13
EAL*DAR	Little's Law	CO^e	0.73	0.29	0.12	0.28	-	-	-	0.40	88.34
		$CO^e + \alpha_s(S^*-S)$	0.60	0.34	0.19	0.36	-	-	-	0.44	65.96
	Fraction of Expected Deliveries Received	CO^e	0.75	0.30	0.12	0.31	-	-	-	0.40	88.27
		$CO^e + \alpha_s(S^*-S)$	0.63	0.30	0.19	0.35	-	-	-	0.42	84.42
Desired Inv Coverage*DAR	Little's Law	CO^e	0.51	0.35	0.27	0.23	-	10.80	-	0.47	82.52
		$CO^e + \alpha_s(S^*-S)$	0.56	0.36	0.33	0.20	-	24.23	-	0.49	71.12
	Fraction of Expected Deliveries Received	CO^e	0.53	0.34	0.25	0.24	-	19.85	-	0.46	84.39
		$CO^e + \alpha_s(S^*-S)$	0.54	0.32	0.30	0.20	-	1983.99	-	0.47	76.26
Original Decision Rule											
Fixed	Fixed	Fixed	-	0.36	0.36	0.28	-	-	42.83	0.49	90.42

b) Means of parameter estimates and summary statistics for F.

Desired Inventory	Expected Acquisition Lag	Desired Acquisition Rate	w	θ	α_s	β	S*	Desired Inv Coverage	S'	R ²	RMSE
Fixed	Fixed (3)	CO ^e	-	0.19	0.31	0.27	28.61	-	-	0.48	119.24
		CO ^e + $\alpha_s(S^*-S)$	-	0.29	0.40	0.20	30.50	-	-	0.50	119.42
EAL*DAR	Fixed (3)	CO ^e	-	0.23	0.26	0.44	-	-	-	0.46	117.58
		CO ^e + $\alpha_s(S^*-S)$	-	0.26	0.38	0.39	-	-	-	0.48	117.64
Desired Inv Coverage*DAR	Fixed (3)	CO ^e	-	0.18	0.30	0.26	-	7.12	-	0.49	116.26
		CO ^e + $\alpha_s(S^*-S)$	-	0.20	0.40	0.19	-	8.21	-	0.50	116.29
Original Decision Rule											
Fixed	Fixed	Fixed	-	0.32	0.45	0.44	-	-	32.34	0.55	112.03

c) Medians of parameter estimates and summary statistics for R, W and D.

Desired Inventory	Expected Acquisition Lag Based on:	Desired Acquisition Rate	w	θ	α_s	β	S*	Desired Inv Coverage	S'	R ²	RMSE
Fixed	Little's Law	CO ^e	0.67	0.39	0.19	0.08	2.28	-	-	0.47	2.57
		CO ^e + $\alpha_s(S^*-S)$	0.70	0.38	0.24	0.07	2.32	-	-	0.48	2.50
	Fraction of expected deliveries received	CO ^e	0.68	0.27	0.18	0.08	2.12	-	-	0.46	2.65
		CO ^e + $\alpha_s(S^*-S)$	0.68	0.19	0.22	0.07	2.08	-	-	0.47	2.64
EAL*DAR	Little's Law	CO ^e	0.95	0.20	0.06	0.09	-	-	-	0.40	2.86
		CO ^e + $\alpha_s(S^*-S)$	0.72	0.24	0.10	0.20	-	-	-	0.44	2.82
	Fraction of expected deliveries received	CO ^e	0.95	0.29	0.06	0.07	-	-	-	0.40	2.81
		CO ^e + $\alpha_s(S^*-S)$	0.80	0.20	0.10	0.14	-	-	-	0.40	2.83
Desired Inv Coverage*DAR	Little's Law	CO ^e	0.58	0.26	0.19	0.09	-	0.67	-	0.47	2.58
		CO ^e + $\alpha_s(S^*-S)$	0.60	0.28	0.23	0.06	-	0.78	-	0.49	2.48
	Fraction of expected deliveries received	CO ^e	0.60	0.26	0.18	0.07	-	0.62	-	0.46	2.66
		CO ^e + $\alpha_s(S^*-S)$	0.60	0.23	0.22	0.07	-	0.62	-	0.47	2.67
Original Decision Rule											
Fixed	Fixed	Fixed	-	0.28	0.26	0.11	-	-	4.79	0.48	2.63

d) Medians of parameter estimates and summary statistics for F.

Desired Inventory	Expected Acquisition Lag Based on:	Desired Acquisition Rate	w	θ	α_s	β	S*	Desired Inv Coverage	S'	R ²	RMSE
Fixed	Fixed (3)	CO ^e	-	0.19	0.31	0.27	28.61	-	-	0.48	4.26
		CO ^e + $\alpha_s(S^*-S)$	-	0.29	0.40	0.20	30.50	-	-	0.50	4.27
EAL*DAR	Fixed (3)	CO ^e	-	0.23	0.26	0.44	-	-	-	0.46	4.33
		CO ^e + $\alpha_s(S^*-S)$	-	0.26	0.38	0.39	-	-	-	0.48	4.32
Desired Inv Coverage*DAR	Fixed (3)	CO ^e	-	0.18	0.30	0.26	-	7.12	-	0.49	4.26
		CO ^e + $\alpha_s(S^*-S)$	-	0.20	0.40	0.19	-	8.21	-	0.50	4.26
Original Decision Rule											
Fixed	Fixed	Fixed	-	0.16	0.41	0.43	-	-	7.98	0.57	3.01

Appendix 5: Results of all alternative decision rules for subjects for which the alternative decision rule improves the fit the most compared to the original decision rule. The results of the decision rules with variable desired inventory are added to the table in Appendix 3.

a) Retailer, wholesaler and distributor.

Subject	Desired Supply Line	Desired Inventory	R ² of Best Fitting Alternative Decision Rule	RMSE of Best Fitting Alternative Decision Rule	Desired Supply Line	Desired Inventory	R ² of Best Fitting Alternative Decision Rule	RMSE of Best Fitting Alternative Decision Rule	R ² of Original Decision Rule (Sterman 1989)	RMSE of Original Decision Rule (Sterman 1989)
1	Variable	Fixed	0.23	2.46	Variable	Variable	0.27	2.42	0.05	2.69
2	Variable	Fixed	0.34	3804.00	Variable	Variable	0.52	3256.10	0.04	4512.07
3	Variable	Fixed	0.59	2501.00	Variable	Variable	0.59	2473.10	0.28	3845.28
4	Variable	Fixed	0.65	480.83	Variable	Variable	0.65	480.83	0.29	668.67
5	Variable	Fixed	0.53	18.65	Variable	Variable	0.63	16.43	0.33	22.12
6	Variable	Fixed	0.37	27.96	Variable	Variable	0.52	24.26	0.00	34.68
7	Variable	Fixed	0.77	53.51	Variable	Variable	0.82	48.24	0.43	80.42
8	Variable	Fixed	0.66	12.91	Variable	Variable	0.66	12.60	0.00	20.62
9	Variable	Fixed	0.83	274.03	Variable	Variable	0.83	274.02	0.68	662.76
10	Variable	Fixed	0.48	3.33	Variable	Variable	0.50	3.32	0.14	4.00
11	Variable	Fixed	0.70	7.15	Variable	Variable	0.77	6.42	0.35	10.55
		Mean	0.56	653.26			0.62	599.79	0.24	896.72
		Median	0.59	27.96			0.63	24.26	0.28	34.68

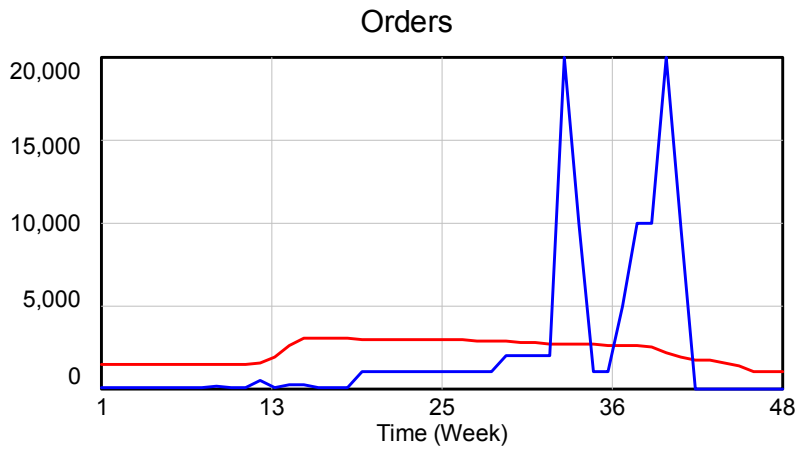
b) Factory.

Subject	Desired Supply Line	Desired Inventory	R ² of Best Fitting Alternative Decision Rule	RMSE of Best Fitting Alternative Decision Rule	Desired Supply Line	Desired Inventory	R ² of Best Fitting Alternative Decision Rule	RMSE of Best Fitting Alternative Decision Rule	R ² of Original Decision Rule (Serman 1989)	RMSE of Original Decision Rule (Serman 1989)
12	Variable	Fixed	0.34	11.40	Variable	Variable	0.45	10.46	0.07	13.47
13	Variable	Fixed	0.15	8.51	Variable	Variable	0.15	8.48	0.00	8.96
14	Variable	Fixed	0.97	6.26	Variable	Variable	0.97	6.26	0.75	18.22
15	Variable	Fixed	0.56	3.01	Variable	Variable	0.56	2.98	0.14	4.16
16	Variable	Fixed	0.25	6.64	Variable	Variable	0.26	6.59	0.01	7.56
		Mean	0.45	7.17			0.48	6.95	0.19	10.48
		Median	0.34	6.64			0.45	6.59	0.07	8.96

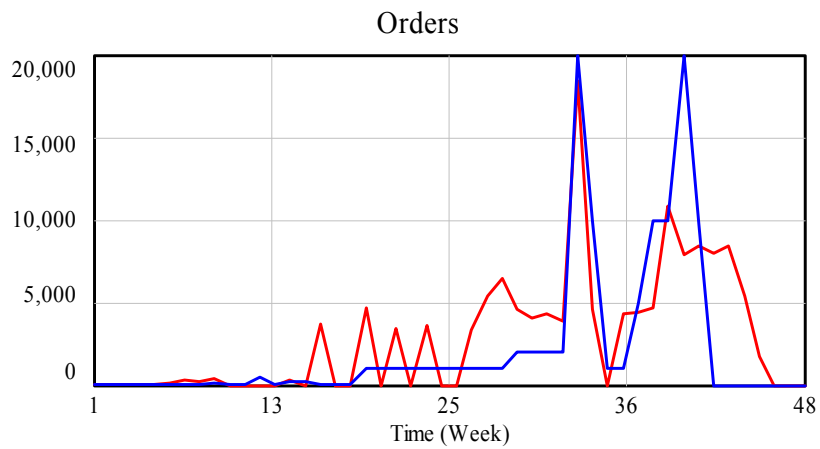
Appendix 6: Plots of actual orders and model output for the subjects for which the original decision rule performs poorly. Model output is the result of the best fitting alternative decision rule.

Actual Orders: Blue
Model Output: Red

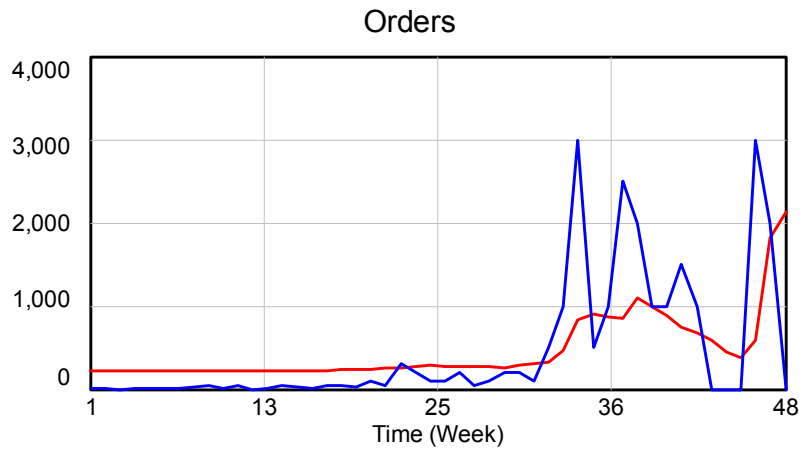
- **Original Decision Rule:** $R^2=0.04$ and $RMSE=4512.07$



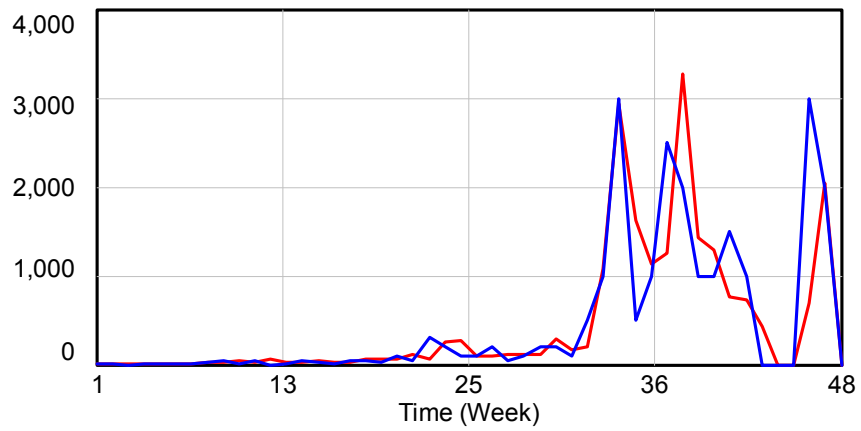
Best Fitting Alternative Rule: $R^2=0.52$ and $RMSE=3256.10$



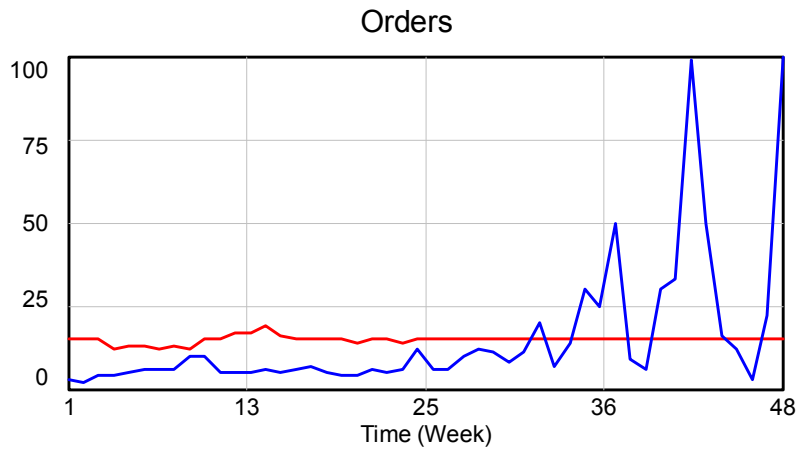
- **Original Decision Rule:** $R^2=0.29$ and $RMSE=668.67$



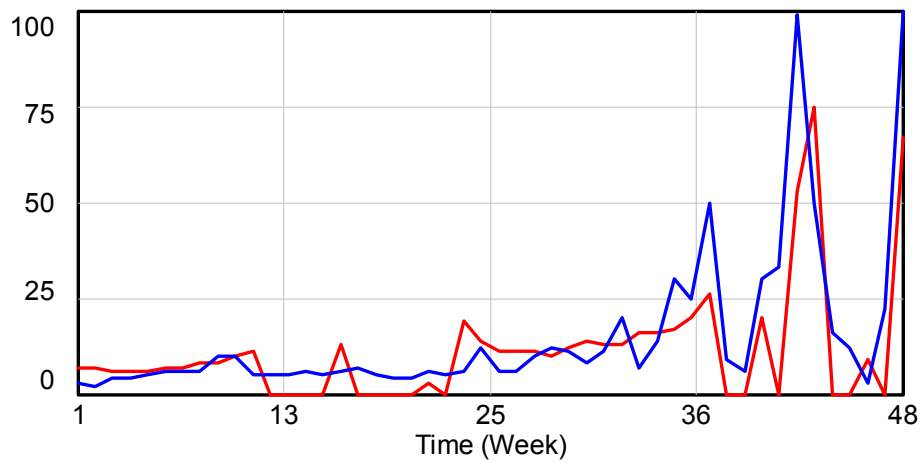
- **Best Fitting Alternative Rule:** $R^2=0.65$ and $RMSE=480.83$



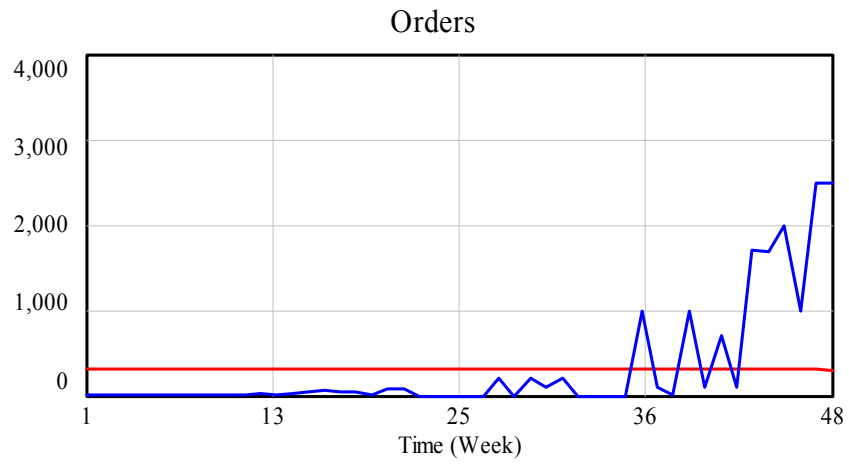
- **Original Decision Rule:** $R^2 = 0.00$ and $RMSE = 20.62$



- **Best Fitting Alternative Rule:** $R^2 = 0.66$ and $RMSE = 12.60$



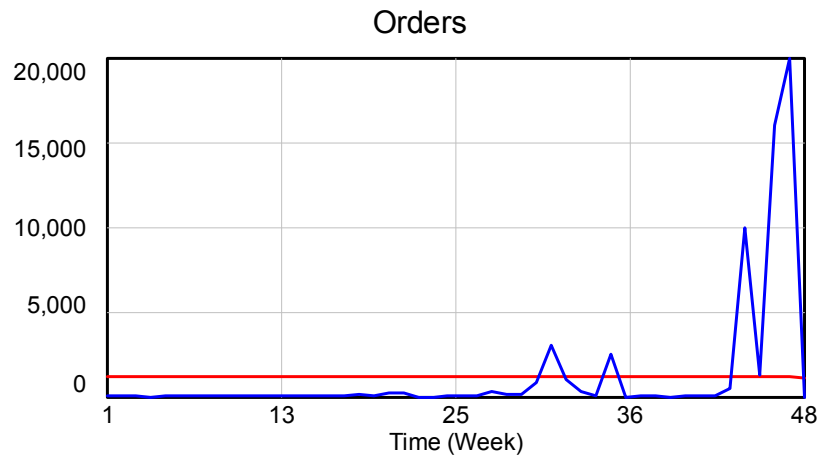
- **Original Decision Rule:** $R^2=0.68$ and $RMSE=662.76$



- **Best Fitting Alternative Rule:** $R^2=0.83$ and $RMSE=274.02$



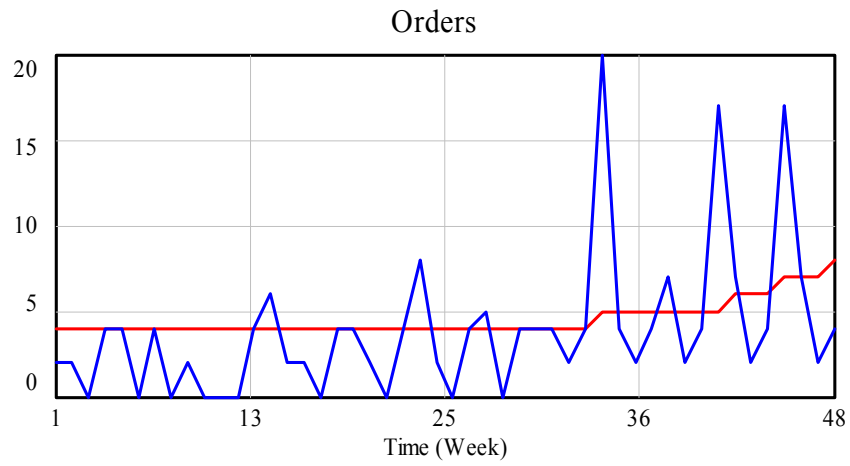
- **Original Decision Rule:** $R^2= 0.28$ and $RMSE=3845.28$



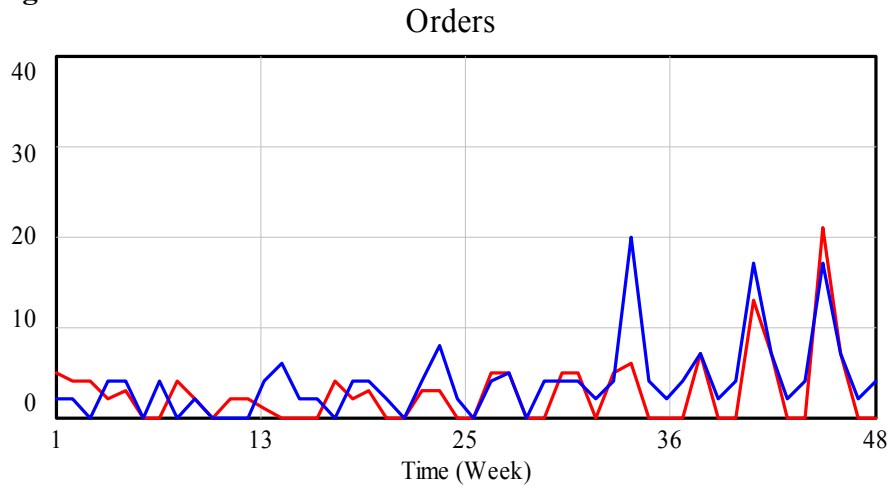
- **Best Fitting Alternative Rule:** $R^2= 0.59$ and $RMSE= 2473.10$



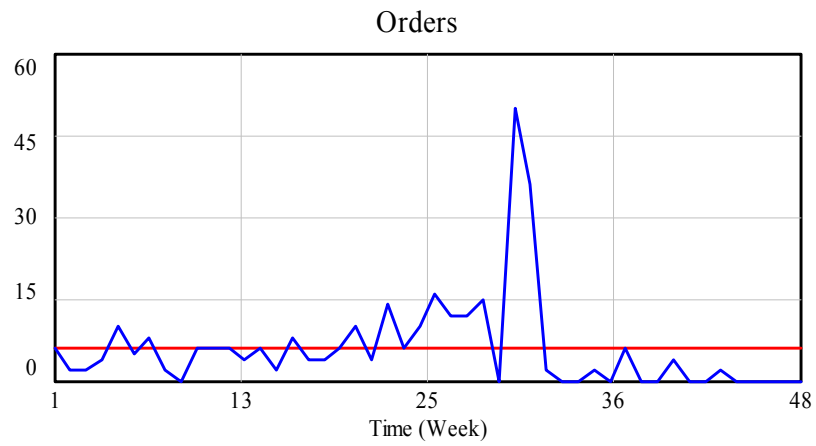
- **Original Decision Rule:** $R^2=0.14$ and $RMSE=4.00$



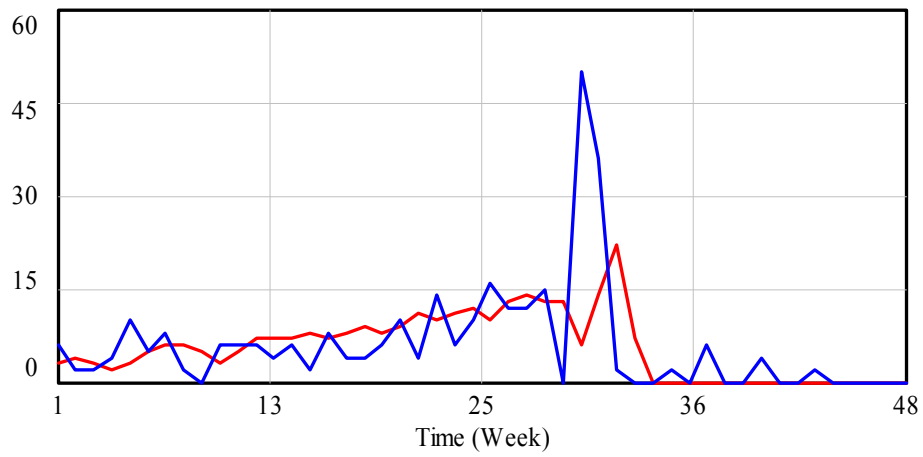
- **Best Fitting Alternative Rule:** $R^2=0.50$ and $RMSE=3.32$



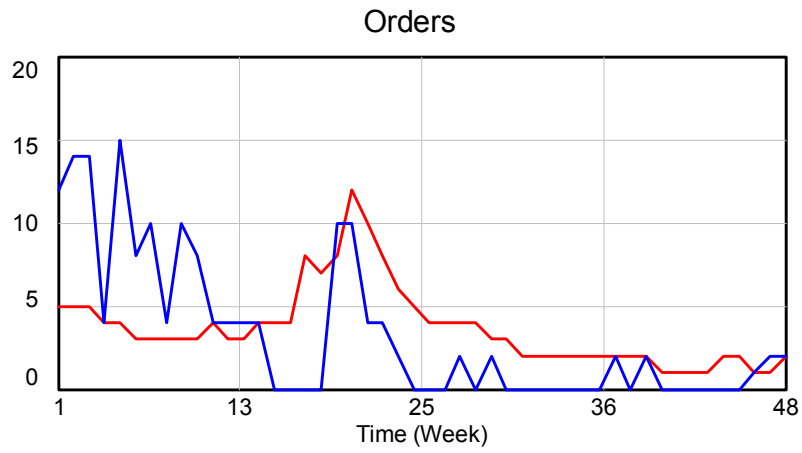
- **Original Decision Rule:** $R^2=0.00$ and $RMSE=8.96$



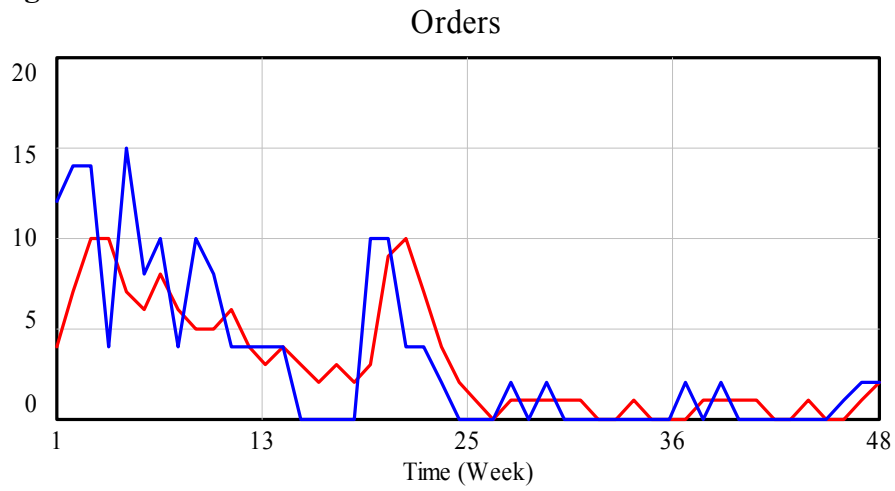
- **Best Fitting Alternative Rule:** $R^2=0.14$ and $RMSE=8.48$



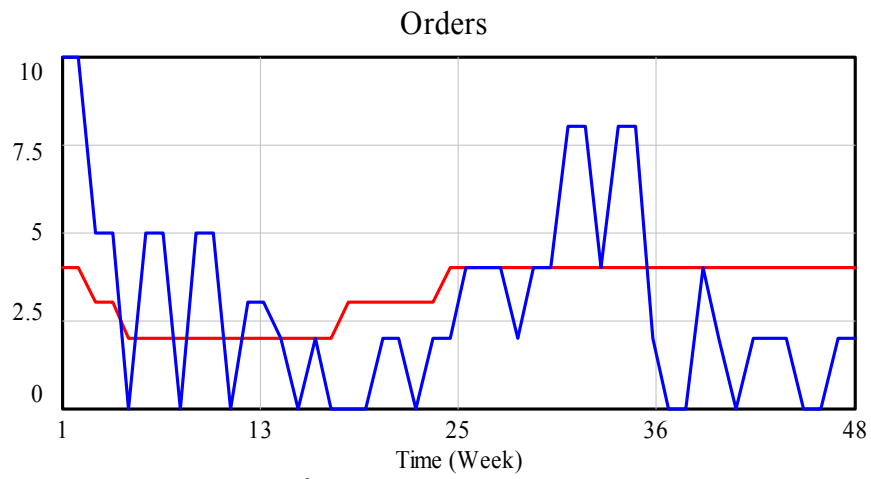
- **Original Decision Rule:** $R^2=0.14$ and $RMSE=4.16$



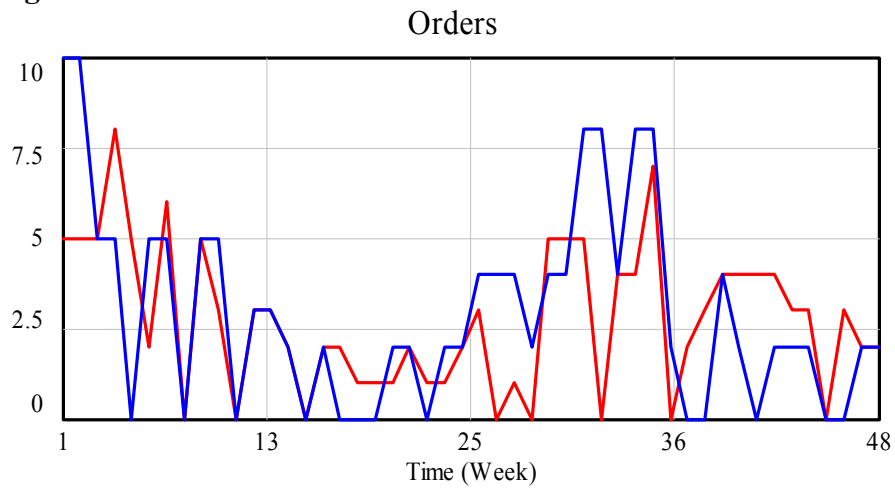
- **Best Fitting Alternative Rule:** $R^2=0.56$ and $RMSE=2.98$



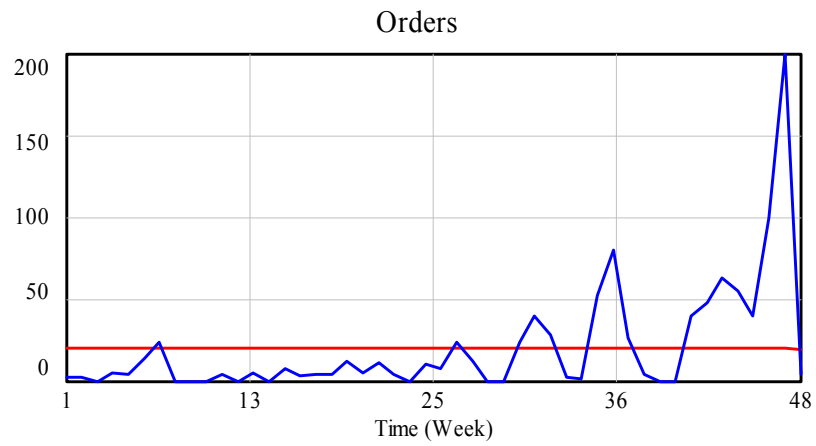
- **Original Decision Rule:** $R^2=0.05$ and $RMSE=2.69$



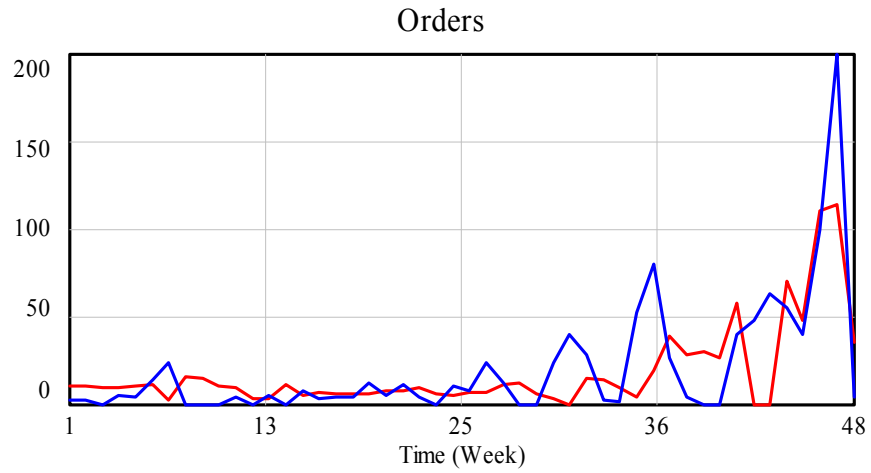
- **Best Fitting Alternative Rule:** $R^2=0.27$ and $RMSE=2.42$



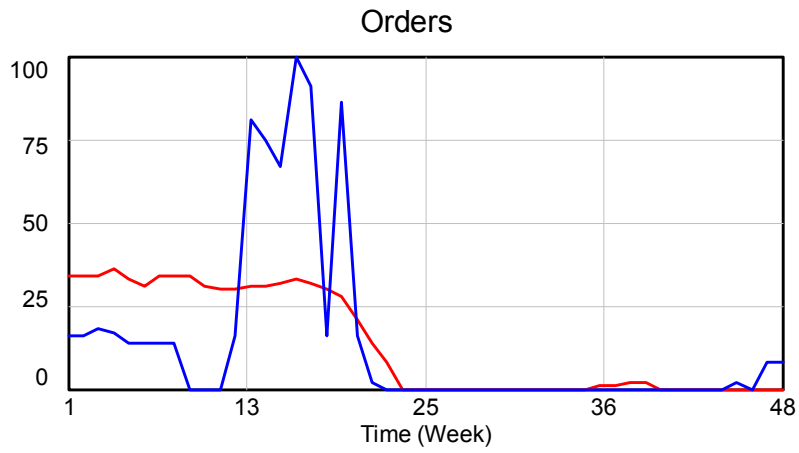
- **Original Decision Rule:** $R^2 = 0.00$ and $RMSE = 34.68$



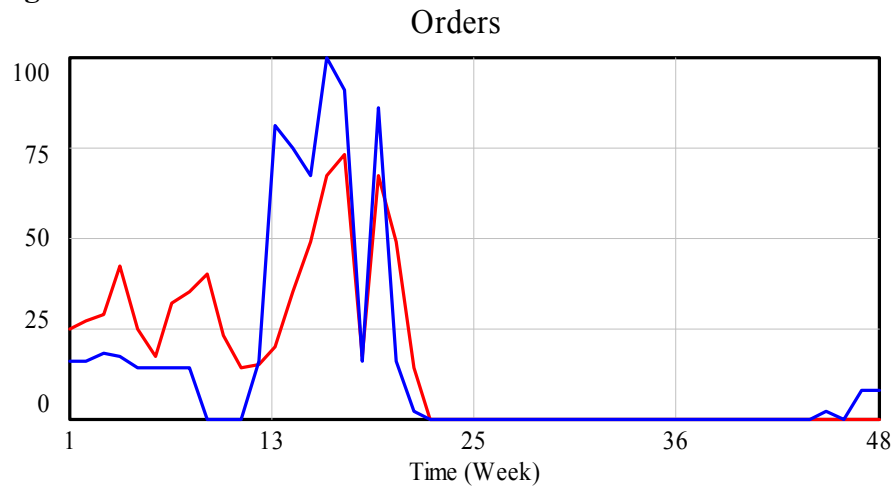
- **Best Fitting Alternative Rule:** $R^2 = 0.52$ and $RMSE = 24.26$



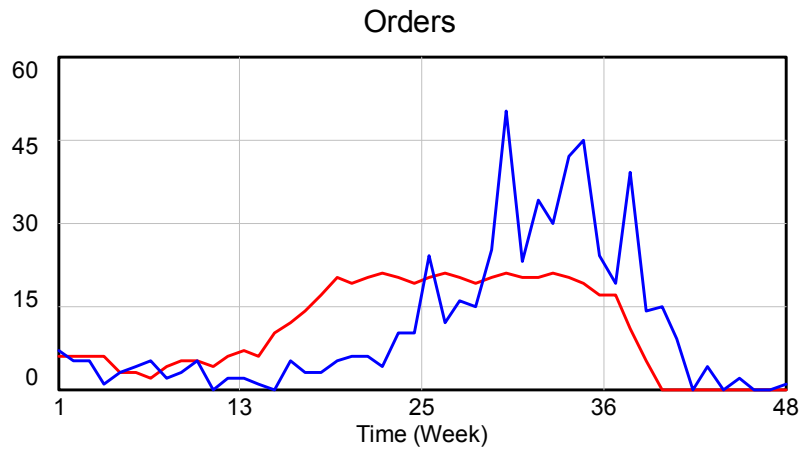
- **Original Decision Rule:** $R^2=0.33$ and $RMSE=22.12$



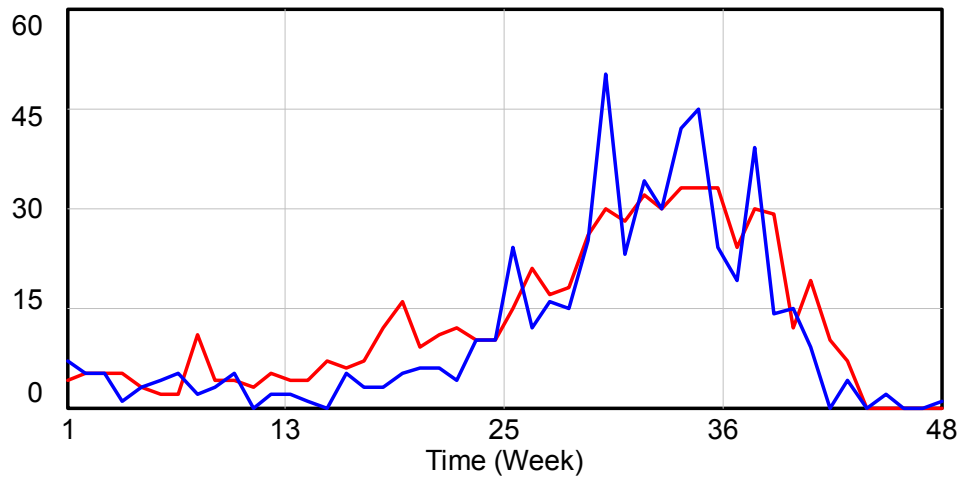
- **Best Fitting Alternative Rule:** $R^2=0.63$ and $RMSE=16.43$



- **Original Decision Rule:** $R^2=0.35$ and $RMSE=10.55$



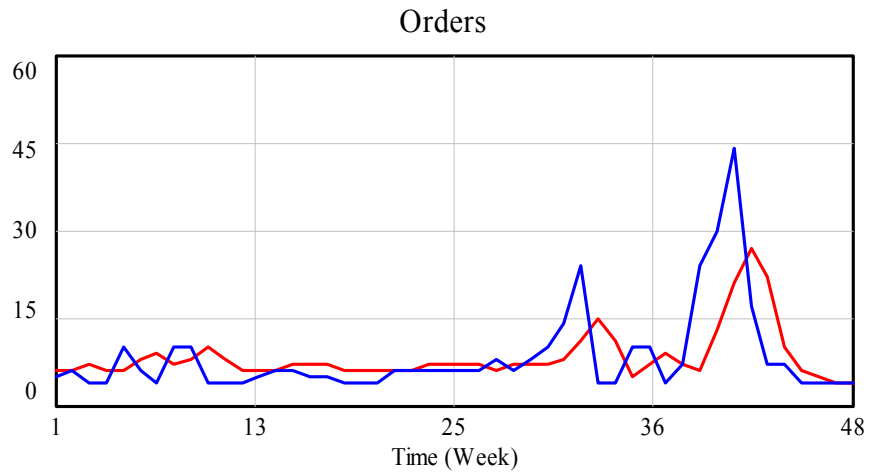
- **Best Fitting Alternative Rule:** $R^2=0.77$ and $RMSE=6.42$



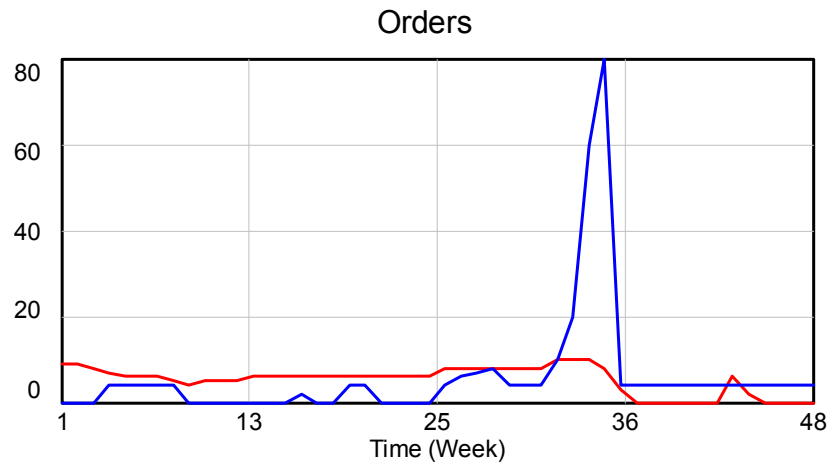
- **Original Decision Rule:** $R^2=0.01$ and $RMSE=7.56$



- **Best Fitting Alternative Rule:** $R^2=0.25$ and $RMSE=6.59$



- **Original Decision Rule:** $R^2=0.07$ and $RMSE=13.47$



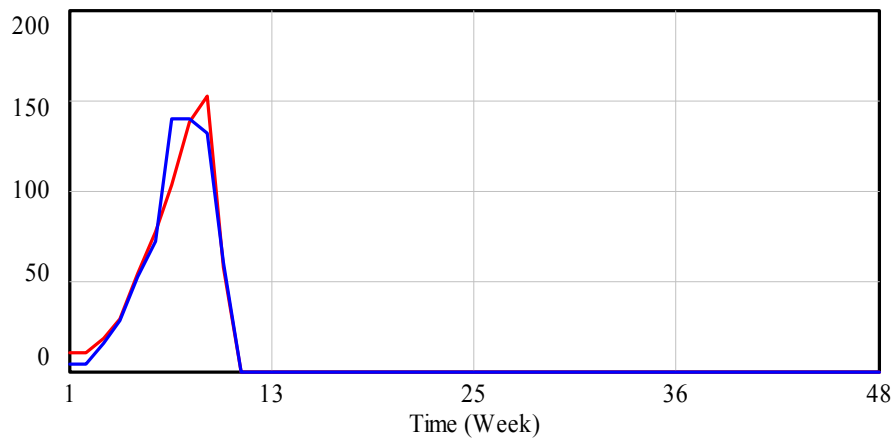
- **Best Fitting Alternative Rule:** $R^2=0.45$ and $RMSE=10.46$



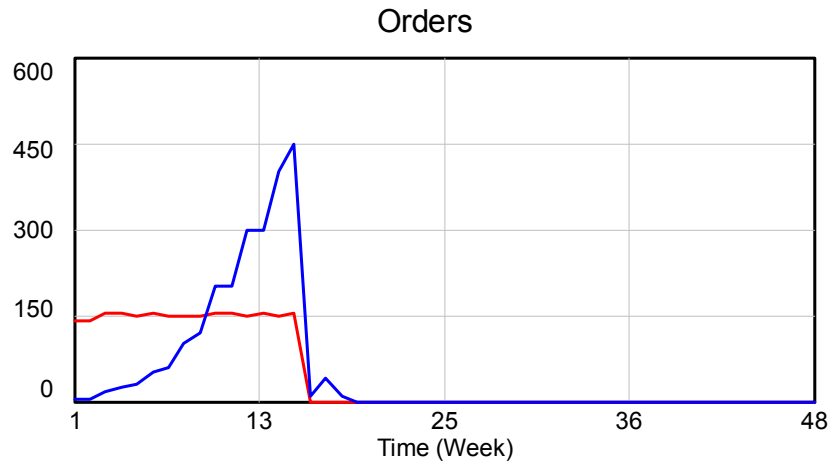
- **Original Decision Rule:** $R^2=0.75$ and $RMSE=18.22$



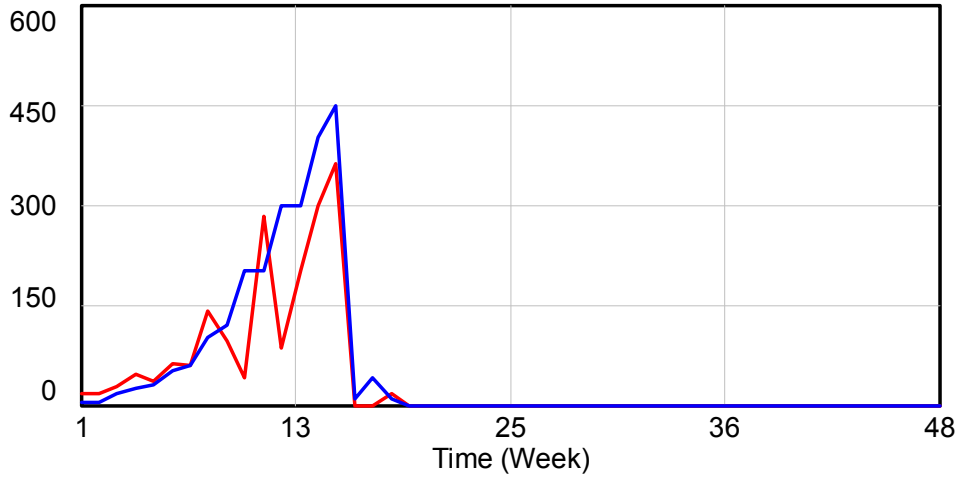
- **Best Fitting Alternative Rule:** $R^2=0.97$ and $RMSE=6.26$



- **Original Decision Rule:** $R^2=0.43$ and $RMSE=80.42$



- **Best Fitting Alternative Rule:** $R^2=0.82$ and $RMSE=48.24$



References

Croson R., Donohue K., Katok E., Sterman J. D. (2004). "Order Stability in Supply Chains: Coordination Risk and the Role of Coordination Stock." Working Paper, MIT Sloan School of Management.

Dogan, G. (2004), "Confidence Interval Estimation in System Dynamics Models: Bootstrapping vs. Likelihood Ratio Method." Proceedings of the 22nd International System Dynamics Conference, Oxford, England.

Forrester, J.W. (1961). Industrial Dynamics. MIT Press, Cambridge.

Sterman, J. D. (1989). "Modeling Managerial Behavior - Misperceptions of Feedback in a Dynamic Decision-Making Experiment." *Management Science* 35(3): 321-339.

Sterman, J.D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill, New York.

Tversky A. and Kahneman D. (1974). "Judgment Under Uncertainty: Heuristics and Biases." *Science* 185: 1124-1131.