



Instructor Guide to the Beer Game

Laboratory implementation, theoretical foundations,
experimental evidence, and variants

Behavioral and Experimental Economics Research Lab
Department of Economics



Revised edition | October 2026

<https://beer-lab.org/>

How to use this guide

This guide supports preparation, implementation, and discussion of the Beer Game family at the Behavioral and Experimental Economics Research Lab (BEER Lab) in the Department of Economics at Tecnológico de Monterrey. It is written for instructors in microeconomics, industrial organization, operations, behavioral economics, system dynamics, and introductory courses who want students to experience how locally plausible decisions can create inefficient aggregate outcomes.

It assumes no prior knowledge of oTree, inventory theory, or supply-chain analysis. The document covers three basic modes and eight extended conditions. Every mode uses a four-stage chain, delayed orders and shipments, inventories, backlogs, and a cooperative objective of minimizing cost.

The guide first explains access and classroom operation, then develops the mechanics, theory, and evidence, and finally provides variants, discussion prompts, worked exercises, and a feedback script. Screenshots come from the executable English interface. Payoffs use experimental currency units (ECU) and do not represent actual commercial transactions.

Author

This guide was prepared by Stanislao Maldonado, research professor in the Department of Economics at Tecnológico de Monterrey, Monterrey Campus. For questions or suggestions about its content, please contact the author at stanislao.maldonado@tec.mx.

For the instructor

The English and Spanish configurations use the same parameters, treatments, and formulas. Project aggregated results only. Do not associate participant codes, links, orders, inventories, or costs with names or student records.

Contents

How to use this guide	i
1 Quick implementation guide	1
1.1 Access through the updated laboratory website	1
1.2 Activity profile	4
1.3 Learning objectives	5
1.4 Operational sequence	5
1.5 A 90-minute plan	6
2 Access, isolation, and session creation	7
2.1 Institutional route	7
2.2 Four rooms	7
2.3 Creating and monitoring a session	7
3 Participant experience	9
3.1 Welcome and laboratory rules	9
3.2 Instructions and roles	9
3.3 Comprehension	12
3.4 Order and outcome	12
4 Economic mechanics and a worked path	19
4.1 Stocks, flows, and delay	19
4.2 Costs and the cooperative objective	19
4.3 One fully worked week	20
4.4 An eight-week path	20
4.5 A reference policy	20
5 Theory of the bullwhip effect	22
5.1 Definition and measurement	22
5.2 Operational causes	22
5.3 Behavioral causes	22
5.4 Rationality and system structure	23
5.5 Information, coordination, and physical speed	23
6 Reading the instructor report	24
6.1 Reveal control and denominators	24
6.2 A disciplined reading order	24
6.3 The order histogram	26
6.4 Feedback script	26
6.5 Multiple chains	26
7 Basic modes	27

7.1	Twelve weeks of constant demand	27
7.2	A twelve-week unexpected step	27
7.3	The twenty-four-week standard simulation	27
8	Extended variants	28
8.1	Decentralized baseline	28
8.2	Shared final demand	28
8.3	Full visibility	28
8.4	Structured communication	28
8.5	Shorter delay	28
8.6	Decision support	33
8.7	Order batching	33
8.8	Promotional pulse	35
8.9	Experimental comparison	35
9	Evidence and scope	38
9.1	Origin and classroom use	38
9.2	Sterman and feedback misperception	38
9.3	Information and mechanisms	38
9.4	Validity and limits	38
10	Teaching plan and discussion	40
10.1	Before class	40
10.2	During play	40
10.3	After play	40
10.4	Layered questions	40
11	Diagnosing trajectories and connecting them to economic analysis	42
11.1	Separate the stimulus from the response	42
11.2	Inventory, backlog, and the delivery pipeline	42
11.3	Decompose cost and locate the mechanism	43
11.4	Compare treatments without conflating mechanisms	43
11.5	A feedback routine grounded in evidence	44
11.6	Extensions across economics courses	44
12	Exercises with answer guidance	46
12.1	Exercise 1: weekly accounting	46
12.2	Exercise 2: shortage	46
12.3	Exercise 3: inventory position	46
12.4	Exercise 4: bullwhip ratio	46
12.5	Exercise 5: zero denominator	46
12.6	Exercise 6: chain cost	47
12.7	Exercise 7: shared demand	47
12.8	Exercise 8: shorter delay	47
12.9	Exercise 9: batches	47
12.10	Exercise 10: promotion	47
12.11	Exercise 11: causal design	47
12.12	Exercise 12: cautious inference	47

12.13	Exercise 13: policy rule	48
12.14	Exercise 14: attribution	48
12.15	Exercise 15: transfer	48
12.16	Integrated activity: diagnostic memo	48
13	Data, privacy, and troubleshooting	49
13.1	Exported data	49
13.2	Common problems	49
13.3	Final checklist	49

List of Figures

1.1	Current institutional website. Select CLASSROOM in the top navigation menu to open the teaching experiments. Captured on October 7, 2026.	1
1.2	Current catalog: search by title, topic, or JEL code. Then locate this family's card and select language and edition. Captured on October 7, 2026.	2
1.3	Demo contains only the configurations of this game family, in both languages and editions.	3
1.4	Rooms contains the four rooms for this family: basic and extended editions in Spanish and English.	3
1.5	Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.	4
2.1	Participant links and administrative tools for a basic session.	8
3.1	Institutional welcome screen and general laboratory rules. Continuation 1/2. . . .	9
3.3	Current basic-edition instructions, with initial resources and experimental units. Continuation 1/2.	10
3.5	Instructions for the factory. Continuation 1/2.	11
3.7	Current comprehension check with an incorrect answer, its specific explanation, and retained responses. Continuation 1/2.	13
3.9	The screens share a layout and use stable role colors. Retailer order decision. . . .	14
3.10	Factory production order.	15
3.11	The outcome separates physical flow, state, and cost. Retailer's weekly result. . . .	16
3.12	Factory's weekly result.	17
3.13	Explicit closure and cumulative individual summary.	18
6.1	The instructor controls when aggregates become visible. Report before session completion.	24
6.2	Report after reveal. Continuation 1/2.	25
8.1	Demand information and state information are distinct interventions. Shared final demand.	29
8.2	Full inventory and supply-line visibility. Continuation 1/2.	30
8.4	Communication reveals intentions without removing autonomy. Announcement of a planned order.	31
8.5	Shared plans before the final choice.	32
8.6	Changing physical structure differs from changing cognitive support. A one-week physical delay.	33
8.7	An inventory-position recommendation. Continuation 1/2.	34
8.9	Operational variants change the source of variability. Orders restricted to batches. .	35
8.10	A temporary demand pulse.	36

Quick implementation guide

1.1 Access through the updated laboratory website

Start at <https://beer-lab.org/en/>. Select **CLASSROOM** in the top navigation menu; it opens <https://teaching.beer-lab.org/en/>. In the **Classroom experiments** catalog, use the search box to locate the activity by name. You can also filter by **Topic** or **JEL code**. Clear any filters left from an earlier search before looking for another game.

Each family's card distinguishes **Español** and **English** access links. The **Show extended version** checkbox selects that edition before following a link. Catalog language and session language are separate choices: verify both. Students use the room specified by the instructor; the instructor prepares and monitors the session through this family's own administrative pages.

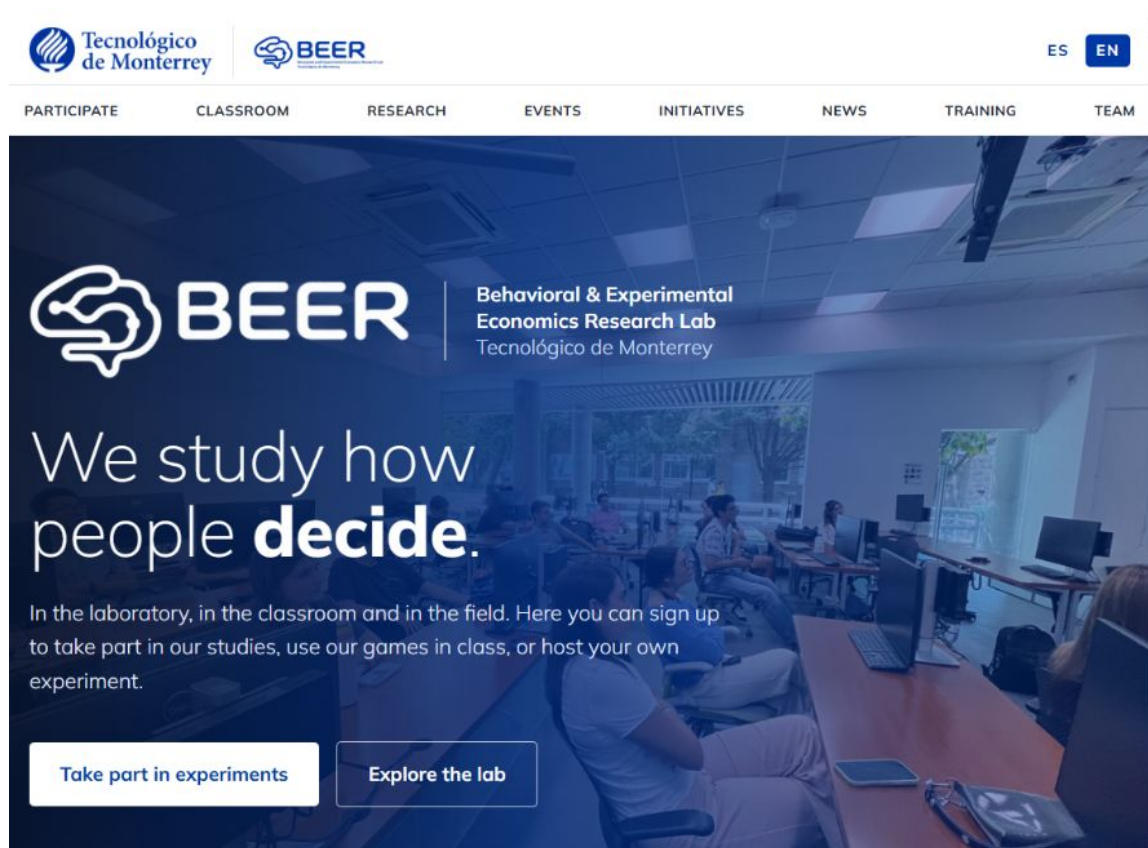


Figure 1.1: Current institutional website. Select CLASSROOM in the top navigation menu to open the teaching experiments. Captured on October 7, 2026.

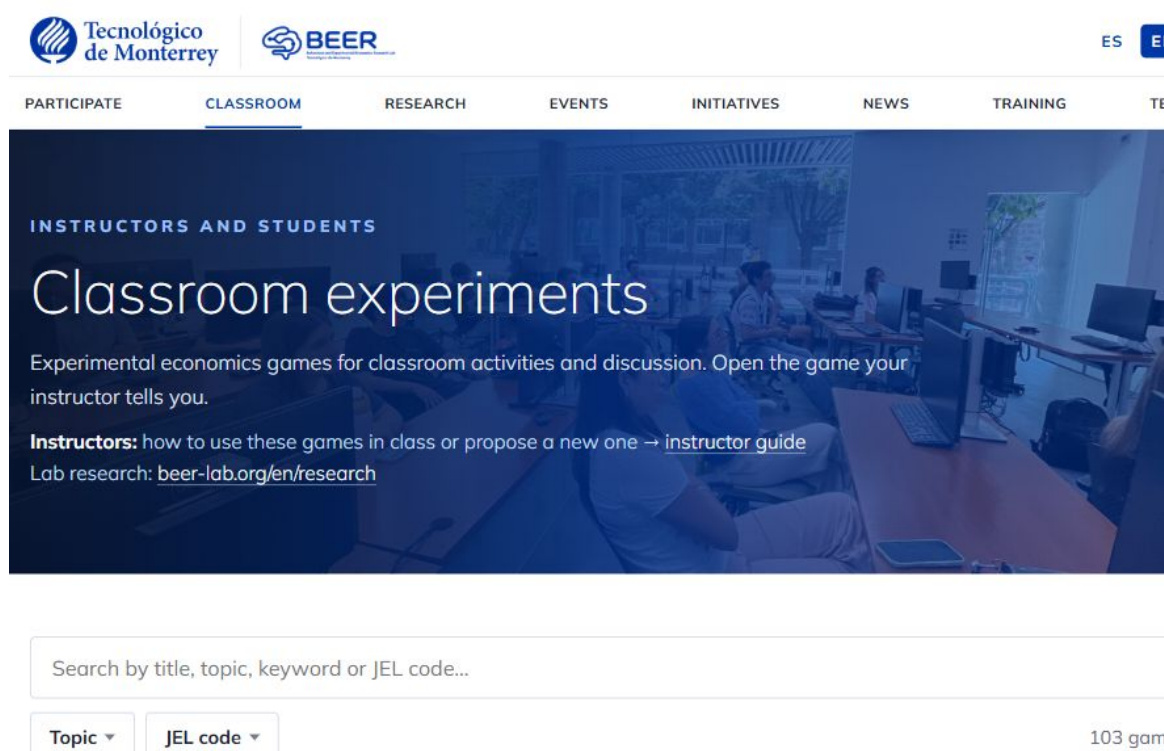


Figure 1.2: Current catalog: search by title, topic, or JEL code. Then locate this family's card and select language and edition. Captured on October 7, 2026.

Checking resources, comprehension, and participant requirements

Before opening the room, check the selected configuration and the number of students present. The session-creation form states the group size and minimum required by that configuration. An individual activity accepts any positive participant count; an interactive activity must satisfy its stated requirement: complete groups when size is fixed, or a minimum participant count when size is flexible. Requirements may change across treatments, for example when adding a third party or changing an institution's size. If you edit a grouping parameter, check the requirement for the edited configuration, rather than relying only on the demo's suggested count.

In the instructions, ensure that students can identify what each role receives, what they can decide, and how the outcome is calculated. An initial account, a distributable budget, and final earnings need not be the same amount. Refer to the resources panel and the example for the selected treatment. Do not carry the basic edition's endowment into a treatment that changes accounts or introduces additional positions.

ECU stands for *Experimental Currency Units*: the activity's unit of account. An outcome of 100 ECU does not automatically equal 100 pesos or 100 dollars. Conversion depends on the session configuration. Before announcing a payment, confirm with the laboratory whether money is paid, the conversion rate, and which decisions count for payment. In hypothetical modules, describe amounts as outcomes in the stated situation.

When a comprehension check is included, an incorrect response receives an explanation beside that question, identifying the submitted answer and the correct answer. Other answers remain in the form. Ask the student to return to the example, locate the relevant rule, and redo the calculation before correcting the response. Avoid revealing answers to the entire class for an individual error, and do not interpret a comprehension mistake as an economic choice in the game.

— 12 semanas, demanda constante	Beer Game: inventarios y efecto látigo / Beer Game: Inventory and Bullwhip Effect — BEER Lab Seleccione una variante de este juego. Select a version of this game.
— 12 semanas, salto inesperado	
— 24 semanas, simulación estándar	
— 12 weeks, constant demand	
— 12 weeks, unexpected step	
— 24 weeks, standard simulation	
extendido — cadena descentralizada	
extendido — demanda final compartida	
extendido — visibilidad completa	
extendido — comunicación estructurada	
extendido — menor retraso de entrega	
extendido — apoyo a la decisión	
extendido — pedidos por lotes	
extendido — pulso promocional	
Beer Game — decentralized chain	
Beer Game — shared final demand	
Beer Game — full visibility	
Beer Game — structured communication	
Beer Game — shorter delivery delay	
Beer Game — decision support	
Beer Game — order batching	
Beer Game — promotional pulse	

Figure 1.3: Demo contains only the configurations of this game family, in both languages and editions.

Rooms

Current rooms:

[ES] Beer Game
[EN] Beer Game
[ES] Beer Game: extendido
[EN] Beer Game: extended

Figure 1.4: Rooms contains the four rooms for this family: basic and extended editions in Spanish and English.

Room: [EN] Beer Game

Create a new session

Session Config

[EN] Beer Game — 12 weeks, constant demand ▼

Number of participants

Complete groups of 4 participants. Use a multiple of 4; minimum 4.

Create

[Configure session](#) ⚙️

App sequence

beer_game (24 rounds) A four-stage supply-chain game about inventory, delays, coordination, and the bullwhip effect.

Figure 1.5: Session-creation form for the basic English room. Check the selected configuration and participant requirement before creating the session.

Downloading a session with several rounds

For the instructor

Open **Data** within the relevant session and use that session's CSV or CSV-for-Excel download link. The general export has one row per participant and columns for the different rounds. Ten rounds do not necessarily mean ten rows per person. When the family provides its own long-format export, observations may instead be organized by participant and round or another unit defined by the design.

Save the original file together with the configuration name, session code, and active-round count. Check that columns for the last active round are present and that the data belong to the selected session. Empty cells can represent inactive rounds, a pending decision, or a field that does not apply to a role; do not automatically recode them as zeros. To compare learning, preserve round and group identifiers alongside participant identifiers. Share only anonymous aggregates in class, and record session incidents before constructing derived tables.

1.2 Activity profile

Purpose	Experience delays, feedback, coordination, and information in a supply chain.
Experimental unit	One chain of four participants.
Roles	Retailer, wholesaler, distributor, and factory.
Decision	An order from 0 to 40 units each week.
State	Inventory, backlog, shipment received, order received, and supply line.

Costs	0.5 ECU per unit of inventory and 1 ECU per backlogged unit per week.
Objective	Minimize cumulative cost across all four stages.
Duration	45–60 minutes for 12 weeks; 70–90 for 24 weeks and feedback.
Requirement	The participant count must be a multiple of four.
Participants	Basic English room .
Instructor	beergame.beer-lab.org/demo .
Basic edition	12 weeks, constant demand; 12 weeks, unexpected step; 24 weeks, standard simulation.
Extended edition	decentralized chain; shared final demand; full visibility; structured communication; shorter delivery delay; decision support; order batching; promotional pulse.
Outcome	Cumulative inventory and backlog cost, by stage and for the entire chain.
Topics	supply chains; inventory; bullwhip effect; coordination; information.
JEL	C92; D21; M11.

Attention

The canonical proposed host is <https://beergame.beer-lab.org/>. The verified local instance is <http://localhost:8054/demo>. The public host still requires laboratory DNS and deployment and should not be announced as live before it is checked.

1.3 Learning objectives

Students should distinguish a stock from a flow, explain why an order is not a delivery, reconstruct the sequence of a week, calculate holding and backlog costs, and recognize how a delay separates an action from its consequence. They should define the bullwhip effect as amplification of order variability upstream, recognize that oscillation can be endogenous, and evaluate demand sharing, full visibility, communication, shorter delays, batching, and decision support as distinct interventions.

A deeper objective is disciplined inference. A graph may show that factory orders varied more than retailer orders, but that pattern alone does not establish the mechanism. The instructor must connect the path to available information, delays, policy, and treatment. A classroom session describes one cohort under one institution. It does not estimate a general causal effect unless the comparison uses an appropriate assignment and repeated chains.

1.4 Operational sequence

1. Confirm that attendance is a multiple of four and choose language, duration, demand pattern, and treatment.
2. Start at beer-lab.org, continue to teaching.beer-lab.org, and select **Beer Game**.
3. Open one of the four stable rooms and create a session. Keep **Monitor** and **Report** in separate tabs.
4. Share the room or current individual links. Never reuse participant links from an old session.

5. Require silence, allow participants to read, and wait until everyone passes comprehension.
6. Do not announce the demand step in a condition designed to make it unexpected.
7. Monitor progress without revealing results. Intervene only for technical problems.
8. At the end, collect predictions, reveal the report, and describe facts before proposing mechanisms.
9. Export data for assignments and record the configuration, date, and incidents.

Checkpoint

Before starting, verify six items: a multiple of four, language, edition, demand pattern, number of weeks, and treatment. The basic room must contain three options; the extended room must contain eight.

1.5 A 90-minute plan

Allow ten minutes for access and laboratory rules, fifteen for instructions and comprehension, thirty for play, five for written predictions, twenty-five for feedback, and five for closure. If you use 24 weeks, reduce commentary during play. Explaining each oscillation changes the learning process and turns the activity into coached inventory control.

Before play, ask: “If consumer demand rises from 4 to 8, should every stage simply order 8?” Record answers without correction. After play, ask: “What did each stage know when it acted? What was already in transit?” The gap between initial intuition and the observed path becomes the basis for analysis.

Access, isolation, and session creation

2.1 Institutional route

The route starts at beer-lab.org and continues through teaching.beer-lab.org. The game card leads to its dedicated instance. The **Demo**, **Rooms**, **Sessions**, and **Data** pages must contain only this family. Isolation prevents a participant who selected the Beer Game from opening a bank run, credit market, or unrelated experiment on the same host.

The deployment key is `GAME_FAMILY=beergame`. The catalog contains six basic configurations, three per language, and sixteen extended configurations, eight per language. If **Demo** shows more than 22 entries or another game appears, do not launch the class. The family filter is not active.

2.2 Four rooms

Room	Options	Use
juego_cerveza_es	3	Constant demand, short step, or standard simulation in Spanish.
beer_game_en	3	The same basic modes in English.
juego_cerveza_extendido	8	Eight institutional conditions in Spanish.
beer_game_extended	8	The same conditions in English.

The room address is stable and can be made into a QR code. oTree creates anonymous session-specific participant links. A room URL identifies no person; an individual link points to one position in one session. If the interface says that a user no longer exists, the link belongs to an old database state. Create a new session from the room and distribute its current links.

2.3 Creating and monitoring a session

Use four participants per chain. Twenty-eight students form seven chains; thirty students cannot form an eighth complete chain. Do not improvise a two- or five-person chain because positions and flows cease to match the design. If attendance is not divisible by four, reserve observers, let the instructor fill a position, or adjust participation before launch.

Use **Monitor** to identify progress. All four decisions are required before the platform computes

oTree Demo Sessions Rooms Data Server Check Logout

[EN] Beer Game — 12 weeks, unexpected step: session 'zlw3p0cn' (demo)

New Links Monitor Data Payments Report Description

Below are temporary links for testing and demonstration. To launch a real study, either create persistent links by setting up a [room](#), or create a session through the [sessions page](#).
Choose one of the following options to play the session:

Combined views

Split-screen view [Launch](#) Grid view [Launch](#)

Session-wide testing link
Open the below link in up to 4 browser tabs.
<http://127.0.0.1:8054/join/nepomogi>

Single-use links
Open each link in its own browser tab.

P1	http://127.0.0.1:8054/InitializeParticipant/1v1hps6p
P2	http://127.0.0.1:8054/InitializeParticipant/8xvoayjy
P3	http://127.0.0.1:8054/InitializeParticipant/4lt5qteb
P4	http://127.0.0.1:8054/InitializeParticipant/uec31fqy

Figure 2.1: Participant links and administrative tools for a basic session.

shipments, backlogs, and cost for a week. One delayed participant leaves the chain on a waiting page. Do not duplicate the link or open a second session. Resend the original link from Monitor.

Open **Report** before the activity begins. It displays participants, chains, expected decisions, and pending decisions while hiding outcomes. Only after all expected decisions are complete does the reveal button become active. This design preserves an opportunity to elicit predictions before showing aggregates.

Attention

Do not run `resetdb`, restart the server, or change the configuration during class. Use the family's local database for rehearsals. Production requires one process and one database per family host.

Participant experience

3.1 Welcome and laboratory rules

The first screen displays institutional logos, the purpose of the activity, and rules concerning silence, privacy, remaining on the page, and requesting assistance. The historical name *Beer Game* identifies the paradigm; the activity does not promote alcohol. The product can be described as cases of a generic beverage without changing the mechanism.



Welcome to the BEER Lab

The **Behavioral and Experimental Economics Research Lab** in the Department of Economics at Tecnológico de Monterrey is a space designed to learn about economic decision making through controlled activities.

Figure 3.1: Institutional welcome screen and general laboratory rules. Continuation 1/2.

You will manage one stage of a supply chain. The shared objective is to serve demand while keeping inventory and backlog costs low.

3.2 Instructions and roles

Each chain contains a retailer, wholesaler, distributor, and factory. Orders move from the consumer toward the factory; physical units move in the opposite direction. The retailer observes final demand. Every upstream stage receives an order from its immediate customer. Under decentralization, it cannot directly observe why that order changed.

The instructions separate five events: receiving a shipment, observing demand and backlog, shipping

- Remain silent and do not talk with other participants.
- Make your decisions individually and privately.

- Follow the instructions on screen.
- Do not refresh, close, or leave this page.
- If you need help, raise your hand and wait for laboratory staff.

Welcome to the BEER Lab

The **Behavioral and Experimental Economics Research Lab** in the Department of Economics at Tecnológico de Monterrey is a space designed to learn about economic decision making through controlled activities.



Beer Game: supply chain

You will manage one stage of a supply chain. The shared objective is to serve demand while keeping inventory and backlog costs low.

General laboratory rules

- Remain silent and do not talk with other participants.
- Make your decisions individually and privately.
- Follow the instructions on screen.
- Do not refresh, close, or leave this page.
- If you need help, raise your hand and wait for laboratory staff.

☐ I have read and accept the laboratory rules.

Start

Powered by [oTree](#)

Figure 3.2: Institutional welcome screen and general laboratory rules. Continuation 2/2.



Week 1 of 12

How the chain works

Your role: **Retailer**

Resources and unit of account

Each participant's initial inventory: 12 product units. The weekly score starts from 50 ECU and deducts costs; inventory is a separate physical resource.

ECU stands for Experimental Currency Units. It is the unit of account for this activity.

This session does not convert ECU into real money.

Decentralized chain

Each stage observes only its local operation and cannot communicate.

The four-stage chain

Consumer → Retailer → Wholesaler → Distributor → Factory

Each week, every stage receives an order from its customer, receives a delayed shipment from its supplier, and places a new order.

Weekly sequence

1. Receive the shipment already in transit.
2. Observe the new order and add any prior backlog.
3. Automatically ship as much as possible from inventory.

Costs

Ending inventory: 0.5 ECU per unit per week.

Backlog: 1 ECU per unit per week.

The goal is to minimize the total cost of all four

Continuation 1/2.

The four-stage chain

Consumer → Retailer → Wholesaler → Distributor → Factory

Each week, every stage receives an order from its customer, receives a delayed shipment from its supplier, and places a new order.

Weekly sequence

1. Receive the shipment already in transit.
2. Observe the new order and add any prior backlog.
3. Automatically ship as much as possible from inventory.
4. Place an order of 0 to 40 units with your supplier.
5. The order and shipment move with a delay.

Costs

Ending inventory: 0.5 ECU per unit per week.

Backlog: 1 ECU per unit per week.

The goal is to minimize the total cost of all four stages, not to maximize order size.

An order placed today does not arrive immediately. Under the current rule, the physical delay is 2 week(s).

Worked example

If you start with 10 units, receive an order for 7, and have no backlog, you ship 7 and end with 3. Your holding cost is 1.5 ECU. A new order affects future weeks, not this shipment.

Continue

Powered by oTree

Figure 3.4: Current basic-edition instructions, with initial resources and experimental units.
Continuation 2/2.



Week 1 of 12

How the chain works

Your role: **Factory**

Decentralized chain

Each stage observes only its local operation and cannot communicate.

The four-stage chain

Consumer → Retailer → Wholesaler → Distributor → Factory

Each week, every stage receives an order from its customer, receives a delayed shipment from its supplier, and places a new order.

Weekly sequence

1. Receive the shipment already in transit.
2. Observe the new order and add any prior backlog.
3. Automatically ship as much as possible from inventory.
4. Place an order of 0 to 40 units with your supplier.
5. The order and shipment move with a delay.

Costs

Ending inventory: 0.5 ECU per unit per week.

Backlog: 1 ECU per unit per week.

The goal is to minimize the total cost of all four stages, not to maximize order size.

Figure 3.5: Instructions for the factory. Continuation 1/2.

places a new order.

Weekly sequence

1. Receive the shipment already in transit.
2. Observe the new order and add any prior backlog.
3. Automatically ship as much as possible from inventory.
4. Place an order of 0 to 40 units with your supplier.
5. The order and shipment move with a delay.

Costs

Ending inventory: 0.5 ECU per unit per week.

Backlog: 1 ECU per unit per week.

The goal is to minimize the total cost of all four stages, not to maximize order size.

An order placed today does not arrive immediately. Under the current rule, the physical delay is 2 week(s).

Worked example

If you start with 10 units, receive an order for 7, and have no backlog, you ship 7 and end with 3. Your holding cost is 1.5 ECU. A new order affects future weeks, not this shipment.

Continue

Powered by oTree

Figure 3.6: Instructions for the factory. Continuation 2/2.

automatically, placing an order, and waiting through the delay. This ordering prevents three common errors. An order does not increase current inventory. The order received is demand; the order placed is a choice. A backlog does not disappear; it carries forward and costs one ECU per unit per week.

3.3 Comprehension

The check requires participants to recognize delayed replenishment, the higher unit cost of backlog, and the chain-wide objective. It is not a mathematics test. Its purpose is to prevent oscillations caused by misunderstanding. Incorrect answers return a common explanation without revealing an optimal policy.

3.4 Order and outcome

The order page shows the local state: incoming order, incoming shipment, available inventory, prior backlog, and units in transit. The retailer also sees consumer demand because it is the retailer's incoming order. The factory orders production, which enters the same delay structure as replenishment elsewhere.

After submission, the platform fulfills demand with available inventory. It reports units shipped, ending inventory, ending backlog, both cost components, chain cost, and a score. The score is a teaching transformation; cost remains the central economic outcome.

The last week ends on an explicit completion screen. It summarizes cost, mean order, order variance, and the participant's bullwhip ratio. The ratio is not computed under constant demand because its denominator is zero. The participant is told to wait for the instructor, avoiding the



Review the marked questions. Each shows your answer and an explanation; your other answers are retained.



Week 1 of 12

Comprehension check

Retailer

If you increase your order this week, when can it help replenish inventory?

- ☐ In a future week, after the delay
- ☒ Immediately, before serving the current order
- ☐ Never

Your answer: Immediately, before serving the current order. The correct answer is In a future week, after the delay. Review the answers. Orders arrive with a delay, backlogs cost more than inventory, and the objective is to minimize total supply-chain cost.

Which situation has the higher cost per unit per week?

- ☒ An unfilled order
- ☐ Placing an order
- ☐ Receiving a shipment

What is the economic objective of the activity?

- ☒ Minimize total supply-chain cost
- ☐ Always place the largest order
- ☐ Keep zero inventory regardless of backlogs

Continue

Powered by oTree

Figure 3.8: Current comprehension check with an incorrect answer, its specific explanation, and retained responses. Continuation 2/2.



Week 1 of 12

Order decision

Retailer

Decentralized chain

Each stage observes only its local operation and cannot communicate.

Your local state

Order received

4

Shipment received

4

Available inventory

12

Prior backlog

0

Units in transit

4

Consumer demand

4

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 3.9: The screens share a layout and use stable role colors. Retailer order decision.



Week 1 of 12

Order decision

Factory

Decentralized chain

Each stage observes only its local operation and cannot communicate.

Your local state

Order received

4

Shipment received

4

Available inventory

12

Prior backlog

0

Units in transit

4

How many units do you want to order from your supplier?

units

Place order

Powered by [gTree](#)

Figure 3.10: Factory production order.



Week 1 of 12

Weekly result

Retailer

Physical flow and order

Order received 4	Shipment received 4	Order 4	Units shipped 4	Ending inventory 8
Ending backlog 0				

Weekly costs

Holding cost 4.0 ECU	Backlog cost 0.0 ECU	Your cost 4.0 ECU	Supply-chain cost 16.0 ECU	Weekly score 46.0 ECU
--------------------------------	--------------------------------	-----------------------------	--------------------------------------	---------------------------------

Continue to the next week

Powered by [oTree](#)

Figure 3.11: The outcome separates physical flow, state, and cost. Retailer's weekly result.



Week 1 of 12

Weekly result

Factory

Physical flow and order

Order received 4	Shipment received 4	Order 4	Units shipped 4	Ending inventory 8
Ending backlog 0				

Weekly costs

Holding cost 4.0 ECU	Backlog cost 0.0 ECU	Your cost 4.0 ECU	Supply-chain cost 16.0 ECU	Weekly score 46.0 ECU
--------------------------------	--------------------------------	-----------------------------	--------------------------------------	---------------------------------

[Continue to the next week](#)

Powered by [oTree](#)

Figure 3.12: Factory's weekly result.

generic “No more pages left” message.



Activity completed

You have completed the Beer Game. Remain at your station and wait for the instructor's directions.

Weeks completed 12	Cumulative cost 46.0 ECU	Mean order 6.7	Order variance 16.9	Bullwhip ratio 4.75
Cumulative score 554.0 ECU				

Powered by [oTree](#)

Figure 3.13: Explicit closure and cumulative individual summary.

Economic mechanics and a worked path

4.1 Stocks, flows, and delay

Inventory is a stock measured at a point in time. Orders, shipments, and production are flows during a week. The accounting identity for stage i is

$$I_{i,t}^{\text{end}} = I_{i,t}^{\text{start}} - S_{i,t},$$

where $S_{i,t}$ is the shipment actually sent. Required service combines prior backlog $B_{i,t}^{\text{start}}$ and the new incoming order $D_{i,t}$:

$$S_{i,t} = \min\{I_{i,t}^{\text{start}}, B_{i,t}^{\text{start}} + D_{i,t}\}, \quad B_{i,t}^{\text{end}} = B_{i,t}^{\text{start}} + D_{i,t} - S_{i,t}.$$

These are accounting relations rather than a behavioral theory. Every feasible ordering policy must respect them.

The shipment received at t was dispatched earlier. The standard condition uses a two-week physical delay, while the short-delay treatment uses one week. The decision $Q_{i,t}$ enters the supply line and affects future availability. Delay creates a cognitive problem: a manager reacting to current backlog may raise orders without giving sufficient weight to units already ordered. When those units arrive, shortage can become excess, causing a sudden order reduction that travels upstream.

Theoretical foundation

The supply line consists of committed units that have not arrived. Ignoring it means responding only to visible inventory. Sterman (1989) identifies supply-line underweighting as a behavioral source of oscillation: people correct backlog forcefully but respond too little to replenishment already in transit.

4.2 Costs and the cooperative objective

The stage cost is

$$C_{i,t} = 0.5I_{i,t}^{\text{end}} + 1.0B_{i,t}^{\text{end}}.$$

After fulfillment, ending inventory and ending backlog are not both positive at the same stage. Holding and shortage are costly for different reasons. Backlog receives the larger coefficient to represent delayed service and lost goodwill, not as an estimate of a specific firm.

The chain cost sums all stages and weeks. The roles do not compete over a fixed surplus. A stage may lower its immediate cost while shifting instability elsewhere. A wholesaler that orders zero

avoids adding to its pipeline today but may create future shortages at the retailer. The feedback should therefore examine total cost as well as individual stages.

4.3 One fully worked week

Suppose a wholesaler begins with 10 units and backlog 2, receives a shipment of 4, and receives a new order of 7. Available inventory is 14 and required service is 9. The platform ships 9, ending with inventory 5 and backlog zero. Holding cost is $0.5 \times 5 = 2.5$ ECU. An order of 8 now does not change current fulfillment; it changes future availability after the delay.

If the incoming shipment were 3, available inventory would be 13, shipment 9, inventory 4, and cost 2 ECU. If available inventory were only 6, the stage would ship 6, end with backlog 3, and pay 3 ECU. The comparison shows why zero inventory is not automatically efficient and positive inventory is not automatically wasteful. The state must be evaluated relative to demand and lead time.

4.4 An eight-week path

Consider demand 4 for four weeks and 8 thereafter. With initial inventory 12 and flow 4, ordering 4 maintains a steady state. In week 5 the retailer observes 8 and may order 12 to cover the increase and rebuild a buffer. Because of information delay, the wholesaler still receives the earlier order of 4. Later it receives 12, interprets the change as persistent, and may order 16. The distributor and factory respond later still.

The rising order stream does not mean that final demand keeps rising. It combines a permanent level change, delayed information, and inventory correction. When earlier large orders arrive, stages accumulate product and cut new orders. The path produces amplification, phase lag, and oscillation. Amplification means a larger swing upstream; phase lag means the peak occurs later; oscillation means alternating shortage and excess around the required flow.

Feedback script

Ask in order: When did final demand change? When did retailer orders change? When did the factory peak? What was already in transit when each stage raised its order? Which change could reduce oscillation without increasing backlog?

4.5 A reference policy

A base-stock policy uses inventory position

$$IP_{i,t} = I_{i,t} - B_{i,t} + L_{i,t},$$

where $L_{i,t}$ is the supply line. A stylized rule orders expected flow plus a gradual correction toward a target inventory position. The decision-support treatment displays such a recommendation while leaving the participant free to choose. It is not presented as a universal optimum because the appropriate target and response depend on demand, delay, cost, and information.

An aggressive correction can amplify noise, whereas a weak correction can prolong backlog. The teaching value of the formula is to make three objects visible: expected flow, net stock, and units in transit. Basic play often focuses on the first two and neglects the third.

The rule also supports a discussion of policy stability. The goal is not to follow a recommendation mechanically, but to avoid making a full correction after every observation. Students can compare two criteria: react to the latest order received, or combine an estimate of demand with a gradual correction of inventory position. The second criterion uses more of the system state and reduces the chance of ordering the same replenishment twice. If the objective or cost parameters change, calibration should change as well; the decision structure still separates flow, imbalance, and pipeline.

Theory of the bullwhip effect

5.1 Definition and measurement

Lee, Padmanabhan, and Whang (1997) use the term bullwhip effect for the increase in order variability moving upstream. The report computes

$$BW_i = \frac{\text{Var}(Q_i)}{\text{Var}(D)},$$

where Q_i denotes orders at stage i and D denotes final demand. A ratio above one means that stage orders varied more than consumer demand in the same session. The pattern across roles also matters: a rising sequence from retailer to factory is descriptive evidence of upstream amplification.

The ratio requires demand variation. Under constant demand, $\text{Var}(D) = 0$, so the report displays “Not applicable.” That does not imply zero oscillation. Positive order variance with constant demand is meaningful, but it must not be expressed as a ratio with a zero denominator.

5.2 Operational causes

Lee et al. organize four sources: demand-signal processing, order batching, price variation, and rationing games. The Beer Game strips away several sources to show that delay and decentralized response can suffice. The batching treatment adds indivisibility. The promotional pulse represents temporary demand movement without explicitly modeling price. Shared demand and shorter delay represent different remedies.

Chen et al. (2000) quantify how forecasting and lead times amplify orders and show that centralized demand information can reduce but need not eliminate bullwhip. The qualification is central: information does not determine policy. A participant may know final demand and still overreact to backlog or neglect the supply line.

5.3 Behavioral causes

Sterman (1989) studies decisions in a multi-actor system with feedback and delay. An anchoring-and-adjustment model captures ordering around a reference flow, correcting inventory and backlog, and underweighting the supply line. The policy is understandable from a local screen yet unstable in aggregate.

Croson and Donohue (2006) remove operational explanations such as batching and price changes and find that bullwhip persists. Shared inventory information helps, especially upstream, but does

not eliminate underweighting. This evidence motivates separate treatments for final demand, full state visibility, and decision support. They change different parts of the decision problem.

5.4 Rationality and system structure

The game does not establish that participants are irrational. Each observes only part of the chain, faces delay, and coordinates without a binding contract. A choice can be reasonable under local beliefs and collectively costly. A good explanation therefore combines behavior and structure. Saying that the factory “caused” the problem mistakes a position in the system for an exogenous source.

The cooperative objective removes deliberate distributive conflict. If inefficiency appears, attention turns to information, mental models, policy, and delay. Real firms add incentives, prices, capacity, contracts, and bargaining power. Those features are omitted to isolate dynamic feedback.

Theoretical foundation

Orders are both actions and signals. Each stage reacts to the system and simultaneously changes the information seen upstream. That dual role makes the outcome endogenous and distinguishes dynamic feedback from a collection of independent decisions.

5.5 Information, coordination, and physical speed

Cachon and Fisher (2000) compare shared information with flow improvements. In their setting, shorter lead times and smaller batches can create larger gains than information sharing alone. The lesson is conditional: the value of information depends on whether it changes policy and whether physical delay remains.

Full visibility reveals stocks and pipelines. Structured communication reveals intended orders, which may be inconsistent even with a shared objective. Shorter delay changes physical dynamics. Decision support changes the translation of state into action. Comparing these variants prevents the shallow conclusion that “more information solves everything.”

Reading the instructor report

6.1 Reveal control and denominators

The report initially hides outcomes. It displays participants, chains, expected stage-week decisions, and pending decisions. A chain of four over 12 weeks has 48 expected decisions. This denominator prevents a chain-week from being confused with a participant or a complete chain.

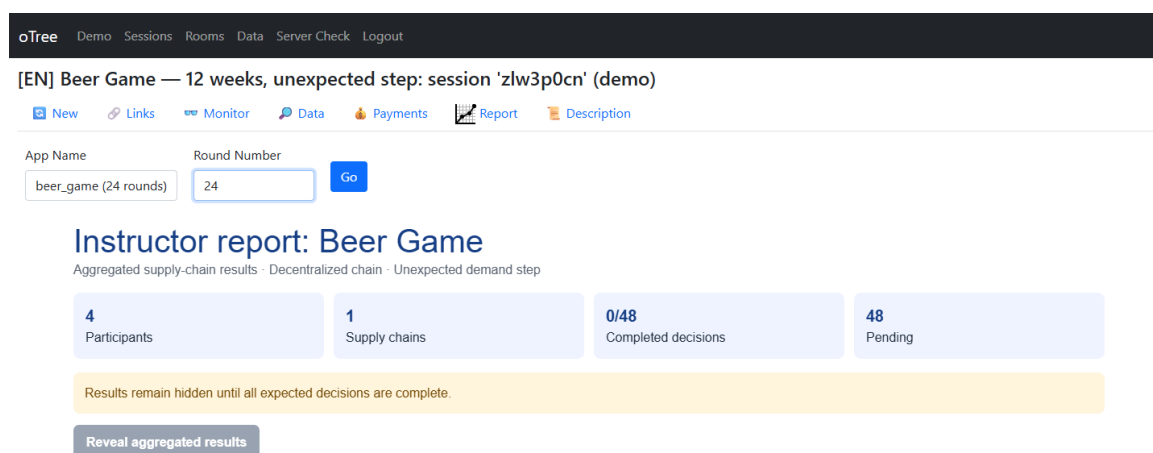


Figure 6.1: The instructor controls when aggregates become visible. Report before session completion.

The revealed panel starts with total cost, mean cost per chain-week, cumulative inventory, and cumulative backlog. It then separates roles and reports observations, mean order, variance, bullwhip ratio, mean inventory, mean backlog, and cost. The line chart places final demand and all four order series on a common scale; the table preserves exact values.

6.2 A disciplined reading order

Describe final demand first: initial level, change week, and persistence. Examine retailer orders second. Move upstream through wholesaler, distributor, and factory. Compare peak timing and magnitude. Connect oscillations to inventory, backlog, and cost. Only then discuss mechanisms.

This order limits hindsight. Starting with “the class ignored the supply line” turns the graph into confirmation. Starting with dates and magnitudes allows competing accounts: backlog reaction, forecasting, copying the incoming order, safety seeking, or supply-line underweighting.

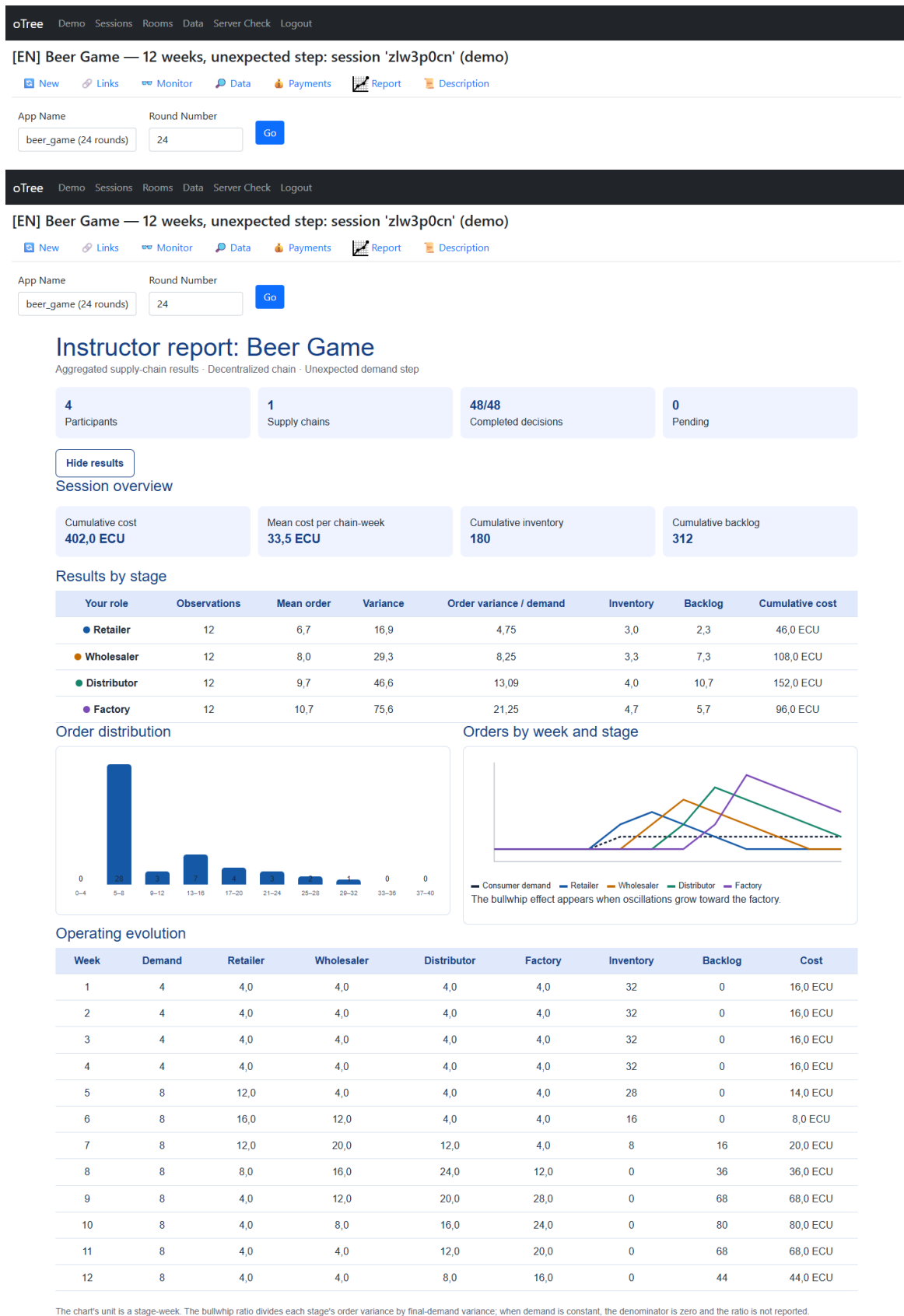


Figure 6.3: Report after reveal. Continuation 2/2.

6.3 The order histogram

The histogram uses four-unit intervals and counts stage-weeks. Mass near 4–8 shows orders close to the stable flow; a tail toward 24–40 reveals overordering episodes. Because the histogram omits timing and role, it must be paired with the line chart and role table.

Means can hide cycles. Alternating 0 and 16 has the same mean as ordering 8 every week but a very different operational effect. Variance and trajectories are essential. Total cost can likewise combine excess inventory in some weeks and backlog in others, so both components remain visible.

6.4 Feedback script

Feedback script

1. “Final demand rose from 4 to 8 in week 5.”
2. “The retailer’s mean order was X and its variance was Y.”
3. “Amplitude grew upstream and the factory peak arrived later.”
4. “The pattern is compatible with local correction and underweighted supply lines.”
5. “We could distinguish mechanisms by comparing visibility, delay, or decision support under the same demand.”
6. “This session illustrates a mechanism; it does not establish a universal response.”

6.5 Multiple chains

With several chains, do not display only the most dramatic path. Compare distributions of cost and role-specific ratios. Ask whether amplification appears in every chain, which chains stabilize sooner, and which visible policies differ. Heterogeneity is informative: common structure constrains outcomes without determining a single path.

Do not rank people as good or bad managers. The relevant unit is a stage policy embedded in a chain. A participant may inherit a difficult state because downstream stages already generated backlog. Interpret role cost with states and denominators.

Basic modes

7.1 Twelve weeks of constant demand

This mode teaches accounting without an external shock. If every stage orders 4, the chain remains near steady state. Any oscillation arises from exploration, mistakes, beliefs, or endogenous response. It is useful before introducing the term bullwhip because it poses a direct question: how can a stable environment become unstable?

The variance ratio is undefined. Use absolute variance, range, cost, and inventory. Do not say there is no amplification merely because a ratio cannot be computed. Describe whether upstream order variability grows.

7.2 A twelve-week unexpected step

Demand equals 4 in weeks 1–4 and 8 thereafter. This is the recommended 90-minute option. The horizon captures propagation and cost, although the chain may not fully recover. Do not announce the step in advance; advance notice creates a different information treatment.

Before week 5, ask without giving the date: “If demand changes once and stays at the new level, what will factory orders look like?” Compare the written prediction with the final line chart.

7.3 The twenty-four-week standard simulation

This mode preserves the step and extends the horizon. It can show arrival of excessive orders, inventory accumulation, order collapse, and possible recovery. It fits operations and system-dynamics courses or a two-hour session. Participants need time to read states; rushing weakens the exercise.

Divide feedback into steady state, response, amplification, and correction. Ask students to locate the week when enough product was already in transit but orders continued to grow. That observation connects directly to Sterman’s mechanism.

Extended variants

8.1 Decentralized baseline

Each stage sees only its local operation and cannot communicate. Demand steps from 4 to 8 and shipments take two weeks. Behavioral evidence predicts amplification and phase lag, although disciplined policies can reduce both. Look for rising variance and later peaks upstream.

8.2 Shared final demand

Every stage sees consumer demand and its history. This reduces uncertainty about whether an incoming order reflects consumption or local inventory correction. Chen et al. predict reduced, but not necessarily eliminated, amplification. Compare distributor and factory variance with baseline. Ask: “If the factory knows consumer demand is 8, why might it rationally produce more than 8?” A complete answer mentions backlog, targets, and the supply line.

8.3 Full visibility

Every stage sees inventory, backlog, and supply line across the chain as well as final demand. The treatment approximates an integrated dashboard. Croson and Donohue find that inventory information can help upstream without automatically removing supply-line underweighting. Determine whether smoother orders reduce backlog, inventory, or both.

8.4 Structured communication

Participants first announce a planned order. All four plans appear before final choices. The design avoids free text and identifying information. Communication can support coordination, but plans are not binding. Compare planned and actual orders in exported data. Ask what makes a plan credible and whether agreement requires a common demand forecast, inventory target, or both.

8.5 Shorter delay

Shipments and production take one week. Local information remains unchanged, isolating physical speed. A shorter lead time limits the period over which forecast errors accumulate and makes



Order decision

Week 1 of 24
Retailer

Shared final demand

Every stage observes consumer demand and its history.

Your local state

Order received 4	Shipment received 4	Available inventory 12	Prior backlog 0	Units in transit 4
Consumer demand 4				

Final-demand history

1
4

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 8.1: Demand information and state information are distinct interventions. Shared final demand.



Week 1 of 24

Order decision

Retailer

Full visibility

Every stage observes inventory, backlog, and supply line throughout the chain.

Your local state

Order received 4	Shipment received 4	Available inventory 12	Prior backlog 0	Units in transit 4
Consumer demand 4				

Shared supply-chain state

Your role	Inventory	Backlog	Units in transit
● Retailer	12	0	4
● Wholesaler	12	0	4
● Distributor	12	0	4
● Factory	12	0	4

Final-demand history

1
4

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 8.3: Full inventory and supply-line visibility. Continuation 2/2.



Week 1 of 24

Announce your plan

Retailer

Structured communication

Each stage announces a planned order before recording its final order.

Enter the number of units you plan to order. All four intentions will be shared before the final decision.

Order received

4

Available inventory

12

Prior backlog

0

Planned order

units

Share plan

Powered by [oTree](#)

Figure 8.4: Communication reveals intentions without removing autonomy. Announcement of a planned order.



Week 1 of 24

Order decision

Retailer

Structured communication

Each stage announces a planned order before recording its final order.

Your local state

Order received
4

Shipment received
4

Available inventory
12

Prior backlog
0

Units in transit
4

Consumer demand
4

Announced plans

Retailer
4

Wholesaler
4

Distributor
4

Factory
4

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 8.5: Shared plans before the final choice.

correction faster. Cachon and Fisher show that reducing lead time can be more valuable than adding information. Look for shorter phase lag and faster recovery.

8.6 Decision support

The screen computes a recommendation from incoming demand, net inventory, and the supply line. Participants remain free to disagree. The treatment makes an inventory-position rule salient without automating the chain. Compare adherence, cost, and variability. Improvement suggests that translating state into action was difficult; no improvement may reflect distrust, a poorly matched target, or response to other stages.





Week 1 of 24

Order decision

Retailer

Shorter delivery delay Shipments and production take one week instead of two.

Your local state

Order received 4	Shipment received 4	Available inventory 12	Prior backlog 0	Units in transit 0
----------------------------	-------------------------------	----------------------------------	---------------------------	------------------------------

Consumer demand
4

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 8.6: Changing physical structure differs from changing cognitive support. A one-week physical delay.

8.7 Order batching

Orders must be multiples of four. A desired order of 6 becomes 4 or 8, and rounding errors propagate. Lee et al. identify batching as an operational cause. Examine the histogram, zeros,



Order decision

Retailer

Decision support

The screen gives an inventory-position recommendation; the participant remains free to choose.

Your local state

Order received

4

Shipment received

4

Available inventory

12

Prior backlog

0

Units in transit

4

Consumer demand

4

Final-demand history

1
4

Decision-support recommendation

4 units

The recommendation uses the incoming order, net inventory, and units in transit. You may choose a different value.

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 8.8: An inventory-position recommendation. Continuation 2/2.

and jumps of eight or more. Do not attribute all variability to behavioral bias when the institution imposes discontinuity.

8.8 Promotional pulse

Final demand rises to 12 for weeks 5 and 6 and returns to 4. The shock is temporary. If stages extrapolate the peak, factory production may crest after consumers have returned to 4. This treatment supports discussion of promotions, orders versus sales, and reactive destocking. Distinguish necessary response during the pulse from surplus inventory afterward.





Week 1 of 24

Order decision

Retailer

Order batching Orders must be multiples of four units.

Your local state

Order received 4	Shipment received 4	Available inventory 12	Prior backlog 0	Units in transit 4
----------------------------	-------------------------------	----------------------------------	---------------------------	------------------------------

Consumer demand
4

In this condition the order must be 0, 4, 8, ..., 40.

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 8.9: Operational variants change the source of variability. Orders restricted to batches.

8.9 Experimental comparison

To compare conditions, hold language, weeks, demand, and chain size constant. Assign complete chains rather than individuals because decisions interact within chains. Compare cost per chain-



Order decision

Week 1 of 24
Retailer

Promotional pulse

Final demand temporarily rises to 12 and then returns to 4 units.

Your local state

Order received

4

Shipment received

4

Available inventory

12

Prior backlog

0

Units in transit

4

Consumer demand

4

How many units do you want to order from your supplier?

units

Place order

Powered by [oTree](#)

Figure 8.10: A temporary demand pulse.

week, role variance, and backlog. With only one chain per condition, describe differences without presenting them as precise treatment effects.

A stronger comparison records three hypotheses before play: which series will respond first, which role will have the greatest variance, and which cost component will dominate. It then applies the same time window and formulas to all chains. An incomplete chain should not be silently pooled with complete chains. If students learn across sessions, treatment order must also be recorded. These simple choices teach that experimental design begins before the graph is observed and that selecting outcomes afterward can exaggerate chance differences.

Evidence and scope

9.1 Origin and classroom use

The Beer Distribution Game was developed at MIT within system dynamics. The manual activity uses tokens, order slips, and four positions. The digital implementation preserves the main stock-flow structure and automates accounting, delays, and visualization. This classroom adaptation uses 12 or 24 weeks and caps orders at 40 for interface stability; classical implementations often run longer.

9.2 Sterman and feedback misperception

Sterman (1989) analyzes game data and estimates a stock-management heuristic. Participants respond to orders, inventory gaps, and supply lines but give too little weight to units in transit. The same shortage is corrected repeatedly before prior orders arrive. The paper is the central reference for connecting individual decision rules to aggregate dynamics.

Do not claim that every student follows the estimated equation. The report does not elicit subjective forecasts or response weights. A path consistent with underweighting may also reflect changing targets, coordination, or errors. Decision support creates a useful comparison, but one class is not a structural estimate.

9.3 Information and mechanisms

Lee et al. (1997) provide the four-cause framework. Chen et al. (2000) formalize forecasting, lead times, and information. Cachon and Fisher (2000) compare information with flow improvements. Croson and Donohue (2006) document persistent bullwhip after several operational causes are removed and study inventory sharing. Together these studies argue for separating visibility, policy, and physical flow.

9.4 Validity and limits

Internal clarity requires simplification. There are no prices, finite production capacity, competition, bankruptcy, contracts, or rich stochastic demand. The four roles share an objective and cost parameters. Choices occur on computers rather than a physical board. The activity supports reasoning about feedback and delay; it does not estimate the cost of enterprise software or prove how a particular firm should redesign its chain.

Responsible transfer identifies a shared mechanism and institutional differences. A pharmaceutical chain adds expiration and critical service; electricity cannot be stored in the same way; server capacity has delay without cases of product. The stock-flow framework remains useful only after those differences are stated.

External validity can be framed as a correspondence test. Is there a stock that accumulates past decisions? Is there a delay between action and consequence? Do actors observe only part of the system? Does a local decision change the state faced by someone else? The more often the answer is yes, the more useful the game mechanism becomes as a starting point. Prices, capacity, substitution, and contracts can still change the magnitude or even the direction of a recommendation. The guide transfers questions and mechanisms rather than an automatic policy.

Teaching plan and discussion

10.1 Before class

Rehearse with four tabs. Confirm colors, roles, comprehension, waiting pages, and closure. Open the report and verify that it remains hidden. Prepare three definitions: inventory, backlog, and supply line. Do not reveal an optimal rule. Learning comes from experiencing tension between local correction and delayed consequence.

10.2 During play

Read only common rules. If asked how much to order, refer the participant to available information and the cost objective. Do not suggest matching demand or multiplying backlog. Record confusion without commenting on performance. In communication, allow only the structured field; oral conversation would create an unplanned treatment.

10.3 After play

Before reveal, ask each chain to predict the role with the largest variance, the factory peak week, and the dominant cost component. Display demand and order paths first, then the role table, inventory, backlog, and histogram. End by presenting one intervention and eliciting a revised prediction.

Feedback script

Close with four statements: “We observed amplification and delay.” “The path is compatible with local correction and underweighted pipelines.” “Information and shorter lead time work through different channels.” “A causal conclusion requires comparable chains assigned to treatments.”

10.4 Layered questions

Start descriptively: When did demand change? What was each role’s maximum? Which cost grew first? Move to mechanism: What did each role observe? Which orders were already committed? Why does an order of 20 not imply consumer demand of 20? End with design: Which information should be shared? Which delay should be reduced? Which outcome should be measured? What comparison separates explanations?

For introductory courses, use ordinary language and one path. In microeconomics, emphasize information, internal externalities, and coordination. In behavioral economics, discuss mental models and delay. In operations, derive inventory position, base-stock policy, and cost trade-offs.

A second round of questions can turn the conversation into a test of explanations. Ask each group to identify one observation that supports its proposed mechanism and another that could contradict it. Then request one minimal intervention: share a variable, remove one week of delay, or modify an ordering rule. The answer must state which report panel should change and in what direction. Finally, compare proposals by implementation cost, information requirements, and the possibility of shifting costs to another role. This sequence moves students from narrative to a testable prediction.

Diagnosing trajectories and connecting them to economic analysis

11.1 Separate the stimulus from the response

The first diagnostic step is to distinguish final customer demand, which is the external stimulus, from the orders placed by each role, which are endogenous responses. This distinction prevents a common but consequential mistake: attributing every oscillation to supposedly unstable demand. The instructor can temporarily hide the order series and ask the class to describe final demand alone. The four order series can then be revealed. If demand changes once while orders display several peaks, troughs, and reversals, the additional movement must be explained by the internal dynamics of the chain. This comparison turns a complicated graph into a clear causal question: which decisions and delays transformed a simple signal into a volatile response?

The timing of the first adjustment is also informative. The retailer observes the demand change first, whereas upstream roles receive it through orders that travel along the chain. The roles should therefore not be expected to react in the same week. A delayed factory response does not by itself indicate inattention; it may reflect the information structure. A very large order after the signal becomes visible, in contrast, may indicate overcorrection. Ask each role to reconstruct what information was available, when it arrived, what shipment was expected, and which demand was anticipated. This reconstruction protects the discussion from hindsight bias.

The same principle helps distinguish amplification from a simple change in scale. An order series can have a larger mean than demand because the chain is clearing a backlog, yet its week-to-week variance can remain modest. Conversely, a series can share the same mean as demand while alternating between zero and very large orders. The report therefore combines levels, trajectories, and the variance ratio. None of these objects alone is a complete description. Students should learn to state which feature they are interpreting and why it matters for the question at hand.

11.2 Inventory, backlog, and the delivery pipeline

A complete economic interpretation cannot rely on on-hand inventory alone. The participant decides while units are in transit and customer orders remain unfilled. Two players with the same physical inventory can face very different positions if one expects a large delivery and the other carries a substantial backlog. Inventory position summarizes these dimensions: on-hand inventory plus outstanding deliveries minus backlog. When a participant ignores the pipeline, units that are already coming may be ordered again. When all those units finally arrive, excess inventory prompts a sharp order reduction. The resulting oscillation can continue after demand has stabilized.

To teach this logic, select one week from the history and freeze the display. First ask how many units are physically available. Then add the backlog, and finally recover the delivery pipeline. This sequence demonstrates why a single state variable yields an incomplete recommendation. It also clarifies that backlog represents demand that remains due, rather than a sale that has permanently disappeared. Future shipments must first satisfy accumulated obligations. A participant who treats backlog as lost demand will underestimate the units required to restore service; a participant who treats every backlog unit as a new permanent increase in demand will overshoot.

The instructor can use a balance identity as a check. Beginning inventory plus the arriving shipment equals the units available before service. Shipments to the customer cannot exceed available units, while unmet orders add to backlog. Ending inventory and ending backlog should not both be positive for the same role in the same week. If students produce an interpretation that violates one of these relationships, return to the accounting before discussing strategy. The discipline is useful beyond supply chains because it teaches how stocks accumulate from flows.

11.3 Decompose cost and locate the mechanism

Total chain cost combines two distinct losses. Holding cost reflects storage and capital tied up in inventory; backlog cost reflects delayed service. A high total can arise from persistent excess inventory, persistent shortage, or alternation between the two. An aggregate number does not identify the underlying mechanism. Read the report in three steps: locate the weeks with the largest cost, determine whether inventory or backlog dominates, and connect that pattern to orders placed in earlier weeks. The lag between an action and its consequence is a central part of the learning objective.

The allocation of cost across roles also deserves attention. A treatment can reduce total chain cost while shifting inventory toward one stage. This outcome opens a discussion about incentives and coordination: a policy that is efficient for the chain may be rejected when one participant bears most of its private cost. In an industrial organization or contracts course, ask which transfer, cost-sharing formula, service requirement, or contingent contract could support the improvement. Students can then distinguish technological efficiency from the distribution of surplus.

Cost comparisons require a consistent denominator. A longer session naturally generates more total cost, and a chain with additional active weeks has more opportunities to accumulate backlog. Compare average weekly cost when durations differ, and compare total cost only when the horizon is held fixed. The same caution applies across demand treatments: a promotional pulse creates a different service burden from constant demand. The report facilitates comparison, but it does not make unlike treatments automatically commensurable.

11.4 Compare treatments without conflating mechanisms

The extended variants modify different parts of the environment. Shared final demand improves the signal available to upstream players. Full visibility adds information about the state of the chain. Structured communication allows expectations and intentions to be coordinated. A shorter delay changes the physical technology. Decision support changes the available rule. Order batching imposes a constraint. A promotional pulse changes the demand process. A reduction in volatility should not be interpreted as evidence for the same mechanism in every case. A disciplined comparison holds the horizon and demand path constant and changes one feature.

Before revealing outcomes, record a directional prediction for each treatment. The prediction should cover chain cost, order variance, and the timing of reactions, together with a mechanism. The class can then compare prediction and evidence. A mismatch may reflect noise across chains,

learning, group composition, a treatment that was not binding, or a mistaken theory. With only a few chains, results describe the session and help evaluate mechanisms; they do not precisely estimate a population average treatment effect.

Full visibility and shared final demand provide a useful example. Both increase information, but they answer different questions. Final demand tells upstream roles about the original signal. Full visibility also reveals inventories, backlogs, and recent orders across stages. If the second treatment performs better, the difference may come from state information or coordination through observed actions. If the treatments perform similarly, final demand may have contained most of the relevant information in that session. The conclusion must follow the implemented contrast rather than a generic claim that transparency always works.

11.5 A feedback routine grounded in evidence

A debrief can be organized in five moves. First, ask for a neutral description of final demand. Second, identify the timing and size of each role's first response. Third, overlay inventory and backlog to reveal delayed consequences. Fourth, compare costs and bullwhip ratios across roles or treatments. Fifth, return to individual decisions and ask which feasible rule might have produced a more stable path using only the information that was available at the time.

This routine separates description, mechanism, and recommendation. It also prevents the discussion from becoming a moral assessment of participants. Statements such as “the factory caused the problem” should be replaced with testable descriptions such as “the variance of factory orders was larger and the peak followed the accumulation of backlog.” The second statement admits alternative explanations and directs attention to timing and structure. The instructor should recognize decisions that were reasonable under limited information even when their aggregate outcome was poor.

After the first account, ask the class for a counterfactual. What would the graph have looked like if every role had kept ordering four units for two more weeks? What would shorter delivery time change immediately, and what would remain a behavioral choice? What could communication solve that visibility alone could not? A useful counterfactual changes one element and names the variables expected to move. Broad claims such as “more coordination fixes everything” should be converted into a precise prediction.

Instructor key: what students should be able to explain

By the end, students should distinguish external demand from endogenous orders; reconstruct on-hand inventory, backlog, and the delivery pipeline; explain how delayed feedback generates correction and overcorrection; interpret cost and the bullwhip ratio without conflating them; and propose an intervention tied to a specified mechanism. If one of these links is missing, return to a concrete week in the history before adding further theory.

11.6 Extensions across economics courses

In microeconomics, the chain illustrates internal externalities: an order changes the information and costs faced by other stages. In industrial organization, it supports discussion of contracts, bargaining power, vertical coordination, and inventory allocation. In behavioral economics, it provides evidence on anchoring, insufficient adjustment, selective attention, and misperception of feedback. In quantitative methods, it generates time series for calculating variances, covariances, lags, and sensitivity to measurement windows. In public policy or management, it permits separate

evaluation of information, coordination, and technological change as policy instruments.

The activity can be adapted without losing its causal core. In a short class, run constant demand first and then the unexpected step. With two meetings, use the first to establish the phenomenon and the second to compare a treatment against the decentralized baseline. For assessment, require a short memo containing one graph, one reproducible calculation, one mechanism, and one recommendation with a possible cost or limitation. The task then moves beyond a visual impression and becomes an evidence-based economic argument.

An advanced class can analyze the participant-week export. Students can compute the bullwhip ratio for alternative windows, check whether results depend on including the initial adjustment, and compare role-specific autocorrelation. They can plot cross-correlations between a downstream order and an upstream response to estimate the observed transmission lag. They can also separate within-chain from between-chain variation. These exercises reveal that a summary statistic is a modeling choice: definitions, windows, and units of observation must be stated before conclusions are compared.

Finally, connect the simulation to institutions outside the classroom. Retail promotions, minimum shipment sizes, production lead times, service-level agreements, and shared point-of-sale information each map to a specific variant. The comparison should remain cautious because real supply chains include capacity constraints, strategic suppliers, product substitution, and demand estimation. The laboratory environment is valuable because it isolates mechanisms. Its simplicity supports explanation; it does not imply that one classroom result mechanically predicts every industry.

Exercises with answer guidance

12.1 Exercise 1: weekly accounting

A stage begins with inventory 7 and backlog 3, receives 4, and gets a new order for 6. Find shipment, ending inventory, ending backlog, and cost. **Answer:** available inventory is 11 and required service is 9. It ships 9, ends with inventory 2 and no backlog, and pays 1 ECU. A new replenishment order does not alter this calculation.

12.2 Exercise 2: shortage

Repeat with no incoming shipment. **Answer:** available inventory is 7, required service is 9, shipment is 7, backlog is 2, and cost is 2 ECU. Ordering 20 now changes a future week, not the current cost.

12.3 Exercise 3: inventory position

Inventory is 6, backlog 5, and supply line 12. **Answer:** $IP = 6 - 5 + 12 = 13$. Looking only at inventory suggests severe shortage; including the pipeline reveals committed resources. A forecast and target are still required.

12.4 Exercise 4: bullwhip ratio

Demand is [4, 4, 8, 8] and orders are [4, 8, 12, 8]. Population demand variance is 4. Mean orders equal 8 and order variance equals 8. The ratio is 2. It describes amplification but does not establish its cause.

12.5 Exercise 5: zero denominator

Demand is [4, 4, 4, 4] and orders are [4, 8, 4, 8]. Demand variance is zero. **Answer:** report the ratio as not applicable and report order variance 4. Do not substitute infinity or zero.

12.6 Exercise 6: chain cost

The four stages incur 2, 5, 9, and 4 ECU. **Answer:** chain cost is 20. If the lowest-cost stage reduces inventory and creates six more backlog units elsewhere, its local improvement may raise total cost.

12.7 Exercise 7: shared demand

The factory sees final demand 8 but receives an order for 20. Must it produce 8? **Guidance:** not necessarily. It must account for backlog, inventory target, supply line, and downstream correction. Demand information helps interpret 20 but does not uniquely determine an order.

12.8 Exercise 8: shorter delay

Explain why reducing a two-week delay to one may dampen oscillation. **Guidance:** actions correct shortages sooner, less product remains invisible, and forecast errors persist for fewer periods. A very aggressive response can still destabilize the chain.

12.9 Exercise 9: batches

A policy would order 6, but only multiples of 4 are feasible. Compare alternating 4 and 8 with always ordering 8. **Answer:** alternating preserves mean 6 but creates variance; always ordering 8 biases the mean and builds inventory. Batching creates an operational trade-off.

12.10 Exercise 10: promotion

Demand rises to 12 for two weeks and returns to 4. The factory produces 16 four weeks later. Give two mechanisms. **Guidance:** delayed signals, accumulated backlog, extrapolation, underweighted pipeline, or late inventory rebuilding are compatible. Orders alone do not distinguish them all.

12.11 Exercise 11: causal design

Design a visibility comparison. **Answer:** assign complete chains to baseline or visibility, hold demand, cost, and delay fixed, measure chain-week cost and role variance, and analyze chains as the experimental units.

12.12 Exercise 12: cautious inference

One visible chain costs less than one baseline chain. What follows? **Answer:** a descriptive difference. With one chain per condition, treatment is confounded with group composition and behavior. Replication and assignment are needed.

12.13 Exercise 13: policy rule

Propose a simple ordering rule. **Guidance:** expected demand plus a partial correction for net inventory and the pipeline. Define all terms and explain why a response weight below one may prevent overcorrection.

12.14 Exercise 14: attribution

The factory has the highest cost. Is the factory responsible? **Answer:** the conclusion does not follow. The factory receives endogenous amplified orders and faces delay. Analyze the chain rather than assigning blame to a structural position.

12.15 Exercise 15: transfer

Apply the mechanism to hospital beds, server capacity, or school supplies. **Guidance:** identify stock, demand flow, delay, and cost, then name at least two institutional differences that limit the analogy.

12.16 Integrated activity: diagnostic memo

Give each team the anonymous export from one chain and request a one-page memo. It must contain a graph separating demand from orders, a bullwhip-ratio calculation or an explanation of why the ratio is not applicable, the week when the main inventory or backlog accumulation began, and an intervention tied to the observed mechanism. **Guidance:** score accounting accuracy, use of evidence, causal clarity, and recognition of limits separately. A strong memo distinguishes description from explanation and does not infer causality from an unassigned comparison.

Data, privacy, and troubleshooting

13.1 Exported data

The export includes role, week, order, shipment, inventory, backlog, supply line, costs, treatment, and language. A row is a participant-week. Aggregate by chain and session without publishing participant codes. Record incomplete decisions rather than imputing fabricated orders.

The report excludes pending decisions from realized metrics and displays the expected denominator. A partial session supports monitoring, not a final treatment comparison. If a participant leaves, document the incident and explain why that chain is not comparable.

13.2 Common problems

The screen says “Please wait.” Another stage has not submitted; check Monitor before restarting.

A participant link fails. It belongs to an old session; create a new session from the room.

Other games appear. Stop and verify `GAME_FAMILY=beergame`.

The basic room defaults to an extended treatment. The room-prefix filter is wrong; do not proceed.

The bullwhip ratio is blank. Final demand is constant, so the denominator is zero.

One chain is delayed. Identify the pending position and resend its original link.

The chart looks flat. Check the scale, completed weeks, and whether demand was constant.

Attention

Do not resolve a waiting page by inventing another participant’s decision. Synchronization is part of the mechanism. Document an incomplete chain and exclude it from final comparisons.

13.3 Final checklist

- Does the instance show only 22 configurations and four rooms?
- Is attendance a multiple of four?
- Are language, duration, demand, and treatment correct?
- Are Monitor and Report open before links are shared?
- Is communication prohibited except in the intended treatment?

- Does the report remain hidden until completion?
- Does feedback distinguish facts, mechanisms, and limits?
- Does the export preserve anonymity and the participant-week unit?

Session closure

Save the configuration name, number of complete chains, any incident, and the anonymous export. Close or archive the session under the laboratory protocol and remove individual links from the projected screen. Record which definition or page caused confusion. That evidence can improve the next class without altering data from a session that has already taken place.

Bibliography

- Cachon, G. P., and M. Fisher. 2000. Supply Chain Inventory Management and the Value of Shared Information. *Management Science* 46(8):1032–1048. doi:10.1287/mnsc.46.8.1032.12029.
- Chen, F., Z. Drezner, J. K. Ryan, and D. Simchi-Levi. 2000. Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. *Management Science* 46(3):436–443. doi:10.1287/mnsc.46.3.436.12069.
- Croson, R., and K. Donohue. 2006. Behavioral Causes of the Bullwhip Effect and the Observed Value of Inventory Information. *Management Science* 52(3):323–336. doi:10.1287/mnsc.1050.0436.
- Forrester, J. W. 1961. *Industrial Dynamics*. Cambridge, MA: MIT Press.
- Lee, H. L., V. Padmanabhan, and S. Whang. 1997. Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science* 43(4):546–558. doi:10.1287/mnsc.43.4.546.
- MIT Sloan School of Management. 2022. MIT Sloan Beer Game Online. <https://mitsloan.mit.edu/teaching-resources-library/mit-sloan-beer-game-online>.
- Sterman, J. D. 1989. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science* 35(3):321–339. doi:10.1287/mnsc.35.3.321.