**James Paine**
May 10, 2021

## To the Faculty of the System Dynamics Group at MIT:

The purpose of this document is to describe the existing work I have done on bullwhip mitigation research, and lays out my plans for next steps. The primary framing is in terms of the literature gaps that I fill in the research and publication potential. But I also address how this work fits into my overall development, and eventual dissertation.

I first present the current state of my research in this space by laying out my existing abstract, and then describe the gap in the literature I believe my current work already fills. Next, I highlight work I would like to continue to do in this space by developing a model-free approach to bullwhip mitigation (via a deep neural net or similar architecture). In that section, I highlight the interest being paid towards this area in OM and OR, and how I think that work not only makes the chances of publication higher, but also will help me in my future research in behavioral OM.

Finally, I emphasize the experimental gap in this work that I would like to fill via an experiment utilizing both the model-based and model free approaches. I think such an experiment will both increase the likelihood of future publication, and also fill a personal development gap. The end of this document contains a rough outline of how I envision this work fitting into my overall dissertation, along with some ideas for future work beyond my PhD related to this research.

## Abstract for existing work (emphasis added)

The ‘bullwhip effect’ is a classic, yet persisting, problem with reverberating consequences in inventory management and refers to how forecast errors and safety stock builds yield increasing amplitudes in both orders and on-hand inventory positions the further one moves away from a source of order variability. The bullwhip effect is responsible for both excessive strain on real world inventory management systems, stock outs, and unnecessary capital reservation though safety stock building. In this paper, the author develops algorithmic approaches to mitigating bullwhip using simulation modeling, including cost minimization and amplification minimization, and then **interprets the results in the context of existing models of human heuristics in ordering decisions.** The algorithmic approaches are utilized as one member within a model of a human decision makers operating within a multi-echelon supply chain with imperfect information sharing and information delays. Within the cost-minimization routine, human decision biases such as supply line under-weighting are compensated for by the developed methods via the control of the

flow of information and simulated physical goods both up and downstream. In all methods developed, inventory and ordering oscillations are minimized in the simulated environment. **The overall goal of this project is to develop useful, implementable, and (to the degree possible) understandable algorithms capable of mitigating bullwhip generated by real humans when placed into an actively evolving inventory management crisis in-progress.** To this end, **the parameters that emerge in the developed algorithm are mapped to previously observed modes of behavior that mitigate the effects of bullwhip.** The resulting algorithms act in a manner analogous to those exhibiting high levels of trust within the supply chain, coupled with a cautious approach to information signals outside of the supply chain. Desired stock levels of the resulting algorithms approach those found in optimal base-stock replenishment policies. Finally, it is observed that the algorithm does not fall prey to supply line under-weighting and can act to offset the ordering decisions that typically result in bullwhip in a simulated model of a multi-echelon supply chain.

## Gap addressed my work so far

Existing literature often either presents structural sources of instability that lead to bullwhip (e.g. variance amplification as seen in (Lee et al., 1997)), or behavioral influences (e.g. supply chain underweighting seen in (Sterman, 1989) or (Croson & Donohue, 2006) or cognitive reflection seen in (Narayanan & Moritz, 2015)), and associated interventions presented are often aimed managing either these structural or behavioral features. This work aims to present a feasible intervention that does not require either structural changes to the supply chain nor behavioral changes to the players.

By introducing an algorithmic entity in series in the supply chain (here by out right replacing one member), this work aims to mitigate bullwhip (which remains a persistent phenomenon despite 60+ years of research) by allowing people to still be people.

Moreover, the exiting work described in the abstract above is designed to be *interpretable* in the context of existing behavioral models of ordering decisions, with the parameters of resulting algorithmic intervention directly mapping to existing behavioral concepts in bullwhip research.

The above abstract also hints at a feature of the optimization that perhaps deserves more attention: the intervention reduces bullwhip when applied to *an actively evolving inventory management crisis in-progress*. In other words, replacing the human-like ordering rules *in the midst of bullwhip* still reduces costs and inventory builds.

## Model-Free Extensions Based on Expressed interests of OM and OR Community

Algorithmic interventions in a real-world setting, utilizing the Beer Game as the empirical testing bed, are not new (see (Martin et al., 2004) for an early example) and more recently a paper was published in MSOM that received a notable about of attention that trained a DQN neural net to play in this space (Oroojlooyjadid et al., 2021).

I was approached by Professor Jordan Tang at Wisconsin to present at this year's INFORMS conference in the track *"Interface of Algorithms and Human Behavior in Operations"*, based on the work I've presented in the past, summarized in the abstract above, and the publication of the DQN paper in MSOM this year. The question came up about how a model-free method like the DQN used by Oroojlooyjadid compares to the more interpretable and model-based method like that I used. This caused me to write the abstract below for a planned presentation at INFORMS:

> Bullwhip is a classic, yet persisting, problem with reverberating consequences in inventory management. In this work, the author develops algorithmic approaches to mitigating bullwhip using simulation modeling based on classical behavioral models and compares the results to those obtained with a dueling network architecture reinforcement learning algorithm. The algorithmic approaches are utilized as one member within a model of a human decision makers operating within a multi-echelon supply chain with imperfect information sharing and information delays. The overall goal of this project is to develop useful, implementable, and (to the degree possible) understandable algorithms capable of mitigating bullwhip generated by real humans when placed into an actively evolving inventory management crisis in-progress.

Development of neural net architectures for application in supply chain problems is an ongoing area of interest and active publication. Part of my motivation to pursue this is to simply learn *how* to build and understand these model-free algorithms. We are a group that fundamentally believes in the value of models, and thus the rapid adoption of model-free approaches to problem solving is fascinating to me. I feel that in order to maintain relevancy in the field of OM, I must be familiar with these approaches, and be able to properly move from model-based to model-free frameworks in my future research.

To be blunt to those reading this, when talking to people at conferences it is the mention of my work on developing a reinforcement learning architecture to approach these problems that gets attention. **This gives me an opportunity to explore the intersection of model-based and model-free approaches to OM problems in a way that is still not fully explored in the existing literature.**

## Missing Pieces and Future Work

In my view, the paper as currently written is interesting to the OM community and could be worthy of publication in POMS and perhaps MSOM with the proper framing and with some editing down to the core concepts. It bridges two related, but often parallel, streams of research in the OM and OR communities that address structural and behavioral interventions separately. Furthermore, tolerance of bullwhip related research as greatly increased given the recent inventory stresses brought on by the coronavirus, as has been made clear to me in recent conference presentations.

By adding the pieces that compare model-free approaches to model-based approaches, I would then be able to add more richness to the resulting paper by being able to comment on both applicability and interpretability of both approaches. There is a minor contribution in that I plan on building a dueling DQN architecture, utilizing a combination of a greedy and annealing style optimization, to address short-comings I see in prior approaches like in the Oroojlooyjadid paper. Applying different network and optimization architectures in new places is an area of ongoing interest in some publications (but not outside of MSOM in the journals I'm interested in, an only there when combined with some other larger contribution).

**The real value, in my view, is contributing to the literature that explores the trade-offs that come from model-free and model-based approaches to problems.**

This exploration of the boundary of these two approaches is an area of active interest, especially as people are being to acknowledge the over-application of AI approaches and perhaps more need for model-centric approaches (see for example the recent discussion by Professor Ng (Anderson, 2021)). Looking forward, and within the System Dynamics community specifically, I think there is benefit to having these conversations rather than dismissing model-free approaches out of hand. Moreover, I feel the System Dynamics community will be uniquely positioned to provide expertise to other research groups that have championed the explosion of model-free approaches over the last decade. As those groups begin to incorporate more model-based features in the near future, System Dynamics possibly has an opportunity to build new bridges and expand its sphere of influence.

## Opportunity for Experiments

The cost minimization activity that I have already done and is described in the opening abstract takes place solely in my computer, utilizing a model of human ordering behavior. Similarly, the model-free approach still does rely on a model of human ordering in the other entities in the supply chain. This begs an entire series of questions about the applicability of these observations in real environments. Addressing these questions is not only interesting, but It also opens the possibility for publication in more empirically grounded journals like JOM (and greatly increases the chances of publication in POMS or MSOM when combined with the model-free architecture approaches).

Therefore, high on my list of immediate next steps is to do an experiment in which this algorithmic agent is placed in a run of the beer game with real people.

My envisioned future abstract is below (with totally hypothesized parts italicized):

> The ‘bullwhip effect’ is a classic problem with reverberating consequences in inventory management and refers to how forecast errors and safety stock builds yield increasing amplitudes in both orders and on-hand inventory positions the further one moves away from a source of order variability. As made starkly evident during the Coronavirus Pandemic, the bullwhip effect is responsible for both excessive strain on real world inventory management systems, stock outs, and unnecessary capital reservation though safety stock building. In this paper, the author develops algorithmic approaches to mitigating bullwhip using simulation modeling based on classical behavioral models and compares the results to those obtained with a dueling network architecture reinforcement learning algorithm. Furthermore, both methods are empirically tested, and it is observed that *both methods are capable of reducing costs along a supply chain in which real humans are making ordering decisions.* In comparing these model-based and model-free approaches, differences in both interpretability and flexibility are highlighted. In the model-based approach towards cost-minimization, the parameters that emerge in the developed algorithm are mapped to previously observed modes of behavior that mitigate the effects of bullwhip, allowing for direct interpretation of this intervention in a manner that is more readily ported to other non-algorithmic contexts. *The model-free approach is often able to better reduce costs when applied to a real-world setting, but not consistently so*. Ultimately, this work contributes to exploring the boundary between model-based and model-free approaches to interventions in behavioral operations management.

## Context in my Dissertation

The sections above discuss where I think my research on bullwhip mitigation has gone and could go, with an emphasis on personal development and publication in top OM journals. However, I would also like to incorporate this work as a major piece of my dissertation in partial fulfillment of the requirements for my PhD.

I plan on having conversation with the faculty over the next few months about the structure of my dissertation, and thus the discussion here is still somewhat half-formed. However, I would like to incorporate my work on food security and supply chain modeling, which contains a large compartmental differential equation model, with this work on behavioral ordering in a supply chain, which explores the intersection of model-based and model free approaches, with a so far unidentified third project that would explore supply chains or product development in more of an empirical context. The connective tissue through all three is human decision making in supply chains, with an inadvertent subframe of the influence of the coronavirus on those supply chains.

For the third project, I am working on developing a project with either the Toyota Production System Support Center (TSSC), or with Moderna Pharmaceuticals. Both would, hopefully, result in a smaller conceptual model about either organizational change (for TSSC) or new product development (for Moderna) under stress. I am also reaching out to contacts from my former employer HanesBrands to help identify any other projects that fit in this larger framing. Any project or connection developed here but not used for my dissertation will provide an opening for future research in my career ahead.

## Setting Up for Future Research Beyond the PhD

The experimental piece proposed in the main document above is somewhat underdeveloped and is more of an 'demonstration' of the core ideas than a true 'experiment'. That leaves space open for a series of future research ideas which I've been keeping track of as they become apparent to me. Below are my notes on two of these projects.

### *Future Proposal 1: Bits and Bias: Does the Presence of a Machine Change Ordering Decisions?*

- Do people act differently when they are communicating with a machine?
    - Prior work (Shechtman & Horowitz, 2003) imply they might
- Two questions to answer:
    - Does the presence of an algorithmic intervention mitigate bullwhip?
        - *Work proposed above as part of PhD will largely address this question.*
    - Does knowledge of the presence of such a machine modify human ordering behavior?
- Treatments to test both influence of algorithmic intervention in general and if knowledge of presence of the algorithm changes ordering behavior.
- Variables and Controls
    - Follow example of Narayanan to quantify 'bullwhip' and costs.
    - Control for experience in supply chain, gender, age, and also CRT scores
- Study Structure
    - Demonstration test for Machine-Machine quadrant for initial feasibility study and to investigate mapping of algorithm to previous behavioral work (reinforce existing PhD work)
    - Use online platform like Emperica.ly to run full 2x2 experiment.

## Existing Research

| Perceived Retailer Player \ Actual Retailer Player | Human | Machine |
|---|---|---|
| Human | **Well defined and researched (Sterman, Donohue, etc)**<br>High bullwhip and high costs | **Research on Computer/Human interaction gives differing outcomes**<br><br>Higher relationship building, more negativity? |
| Machine | **Research on Computer/Human interaction gives differing outcomes**<br><br>Unresearched space<br>***(PhD work lives here)*** | **Limited research**<br>Oroojlooyjadid et al imply performance boost.<br>Reeve and Nass imply no behavior shift but Shechtman does!<br>***(PhD work lives here)*** |

## Hypothesis

| Perceived Retailer Player \ Actual Retailer Player | Human | Machine |
|---|---|---|
| Human | **Existing Baseline**<br><br>High Costs, High Bullwhip<br><br>**$$$$$** | **Test of new Effects** **$$$**<br><br>Outcome unclear but hypothesize that cost will be reduced versus baseline, but less so than with knowledge of machine |
| Machine | **Test of new Effects**<br><br>Outcome unclear but hypothesize that cost will be slightly lower than baseline<br>**$$$$** | **Test of Validity of Algorithmic Intervention**<br>Hypothesize lower costs due to higher supply chain weighting **$** |

- Part 1 (which is largely what I am proposing as part of my PhD above):
    - Create optimizable functional form of the Beer Game.
        - Self-contained simulation of the system over a given time horizon.
        - Callable function that takes a given state-action pair and returns an updated state.
        - Create dynamic environment to train a model based on actual performance and feedback in the Beer Game.
    - Qualitatively examine the ‘optimized’ ordering choices
        - How do choices vary based on position in the supply chain?

        - How do these choices map to existing models of human behavior in multi-echelon supply chains?
    - Empirically Test
        - Real-world tests of optimization techniques missing from existing literature.
        - Opportunity to see how applicable different classes of optimization are.
- Part 2
- Extend and Test for Human-Machine Interactions
    - Design version of empirical test that varies if the players know that the algorithmic intervention is present or not.
    - Specifically, if it is a 'human' or 'machine' player.
- Scale up Study 1 and Study 2 to increase study power.
    - Extend beyond demonstration mode.
    - Allow more robust testing of DNN and other similar methods.

### *Future Proposal 2: Influence of Regulatory Focus on Decision Making in Multi-Echelon Supply Chains*

- Hypothesis: Regulatory Focus (Crowe & Higgins, 1997) of participants may influence ordering behavior in the Beer Game
- Previous studies, and current MIT online version of Beer Game, are largely framed as prevention focused.
    - Reduction of costs, not maximization of profits
    - Focus on individual costs, though reward based on team performance.
    - 'Winner' is the team that does the least bad.
- Framing matters in behavior
    - Crowe and Higgins show that framing of the problem influences performance and behavior.
    - Even subtle changes in instructor prompt can change outcomes.
    - Are we rewarding A (individual prevention focus), while hoping for B (global bullwhip reduction)? Influenced in part by (Kerr, 1975)
- Part 1: Measure of influence of Regulatory Focus in Multi-Echelon Supply Chain
    - Is there a prevention focus during the ordering decisions in the beer game?

- Propose to use the “Regulatory Focus Questionnaire” as baseline
- Run the game using groups of promotion-focused and prevention-focused individuals and observe bullwhip and supply chain underweighting effects.
- Use same measures introduced in (Croson & Donohue, 2006)
- Concern about if it is possible measure this during or after the game.

- Part 2: Manipulation of Regulatory Focus via framing in Multi-Echelon Supply Chain
  - Can *promotion framing* reduce bullwhip and underweighting?
  - Change team scoring system to be based on *profit* generated each round, not cost incurred.
  - Compare existing game framing (which emphasizes the cost of backlog) to neutral or positive framing.
  - Control for regulatory focus of participants in each treatment
- Part 3: Manipulation of Regulatory Focus via framing in Multi-Echelon Supply Chain
  - Can promotion framing reduce bullwhip and underweighting?
  - Change team scoring system to be based on profit generated each round, not cost incurred.
  - Compare existing game framing (which emphasizes the cost of backlog) to neutral or positive framing.
  - Control for regulatory focus of participants in each treatment
- Note that Parts 2 and 3 presuppose that Part 1 shows difference in Regulatory Focus on ordering decisions.