Essays on Dynamic Supply Chains and Service Delivery Systems

By

James Edward Paine

B.S. Chemical Engineering
University of Florida, 2009

Master of Business Administration
Wake Forest University, 2014

M.S. Mechanical Engineering
Georgia Institute of Technology, 2012

S.M. Management Research
Massachusetts Institute of Technology, 2020

SUBMITTED TO THE DEPARTMENT OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN MANAGEMENT

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

Authored by: James Edward Paine
MIT Sloan School of Management
May 3, 2023

Certified by: Hazhir Rahmandad
Associate Professor of System Dynamics
MIT Sloan School of Management
Thesis Supervisor

Accepted by: Eric So
Sloan Distinguished Professor of Financial Economics
Professor, Accounting and Finance
Faculty Chair, MIT Sloan PhD Program

Essays on Dynamic Supply Chains and Service Delivery Systems

By

James Edward Paine

Submitted to the Department of Management on May 3, 2023 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Management

ABSTRACT

The field of Operations Management, and closely related fields of Operations Research and Industrial Engineering, focus intensely on addressing real-world problems associated with the design and management of product and service delivery systems in a human context. System Dynamics is a framework to understand, design for, and manage change emerging from both structural and behavioral features, and is uniquely suited to address policy questions in socio-technical supply chain contexts. Using System Dynamics, Operations Management, and Supply Chain Research methods this work expands on existing toolsets and theory and provides policy insights in dynamic supply chain and service delivery systems.

Chapter 1 presents a methodological contribution to the System Dynamics and Supply Chain Research communities by developing a novel framework for supply chain models by combining three classic methods: co-flow differential equation structures, spot price discovery, and multinomial logistic choice modeling. Chapter 2 applies this framework to build a structural theory explaining the simultaneous surge in food insecurity alongside surges in food surplus and purposeful disposal at the beginning of the COVID-19 pandemic in the United States. Utilizing this structural theory, this chapter further illustrates policies that could help mitigate these stresses. Chapter 3 continues the concepts of managing a behaviorally driven multi-echelon supply subject to shocks. Utilizing a simulated environment, different policy features implied by parallel streams of Operations Management and Supply Chain literature are directly tested. These include policies that range from myopic, limited information decision rules to more modern, but data-intensive machine learning methods.

Thesis supervisor: Hazhir Rahmandad

Title: Schussel Family Professor of Management Science
Associate Professor, System Dynamics

Acknowledgements

The last five years have been a time of immense growth, both professionally and personally. I am so very thankful to everyone I have met along the way, and to those who helped set me down this path. The research in this dissertation is only a small part of what I have learned and what I hope to contribute.

First, thank you to my partner Christina and our amazing kids, Amelia and Henry, who have supported me through the highs and lows of getting a PhD while raising a young family. We started this journey together nearly seven years ago, before Henry was even born. It is amazing that those first discussions in Winston-Salem about the future, and the multiple false starts along the way, led to a PhD from MIT.

I am thankful for all the incredible people I have met at MIT that have made me a better researcher and person. To my committee, Hazhir Rahmandad, John Sterman, and Rogelio Oliva, thank you for showing me what it means to be a rigorous researcher and dedicated teacher. You have invested so much time in me. I hope to one day live up to the standards that each of you have set for me.

Thank you to both David Keith and Zeynep Ton for giving me some of my first opportunities to collaborate on real research questions and fallowing me to learn through practical experience both as a researcher and teacher.

The PhD students who came before me, Sergey, Mahdi, James, Tanyi Li, Tse Yang, and Jad, set an amazing example. I am especially thankful for their guidance and support in the first years of the program. I cannot wait to see what the PhD students coming up behind me, Jose, Arya, Cathy, Alexander, and Jason will achieve in time ahead. I know they will all make my achievements look small, which is how our discipline will grow and improve. And thank you to Mohammed for putting up with my inconsistent presence in our shared office and cluttered mess of a desk. You will get the window seat next!

During my time at MIT, from 2018 through 2023, I got to see it before, during, and in the recovery from a global pandemic. I did my General Exam remotely (and with the one-question-a-day format!) and locked my keys in my office for nearly three months over Summer of 2020. The pandemic was a massive disruption and also an amazing teacher. It illustrated much of the supply chain phenomena that I came to MIT to study, put my family and friends through immense personal and professional pain, but also showed me new ways to be a good parent and partner.

Through adversity comes growth.

And of course, thank you to the MBTA for teaching me patience, problem-solving under duress, and fire safety.

Table of Contents

Table of Figures

Table of Tables

Chapter 1

Dynamic Supply Chains with Endogenous Dispositions

CHAPTER ABSTRACT

The movement of goods through a supply chain depends on both the physical flow of goods and on the economic decisions of each entity along the chain, including price discovery and inventory disposition decisions. This chapter presents a methodological contribution to the System Dynamics and Supply Chain Research communities by developing a novel framework for supply chain models by combining three classic modeling methods: co-flow differential equation structures, spot price discovery, and multinomial logistic choice modeling. The relative economic values of possible dispositions of goods, including outright disposal, are considered. For work-in-progress, development is considered in terms of the economic value that an additional unit of time will bring to the finished good, and the interplay of these considerations drive goods through, or out of, supply chains. Incorporating these mechanisms can produce materially different behavior modes and can be applied to multiple levels of aggregation within a production process.

Note:

Much of the content of this chapter was previously accepted for publication at the **System Dynamics Review** on November 21, 2022.

It was first published Online as Open Access under a Creative Commons Attribution-NonCommercial License at the System Dynamics Review on December 20, 2022 and is available at https://doi.org/10.1002/sdr.1725

The full citation for the published version of this chapter is:

Paine, J. (2023), Dynamic supply chains with endogenous dispositions. System Dynamics Review. https://doi.org/10.1002/sdr.1725

1.1 Introduction

The sudden emergence of the novel coronavirus commonly known as COVID-19 on the global stage in the early part of 2020 placed immense strain on supply chains and consumers alike. Articles from the middle part of that year describe “dumped milk, smashed eggs, plowed vegetables” as producers made the decision to outright terminate work-in-progress (WIP) rather than move goods into a finished state for sale (Bauer, 2020; Corkery et al., 2020; Corkery & Taffe-Bellany, 2020; Yaffe-Bellany & Corkery, 2020; Zhou, 2020). Stated bluntly in one article: “[Farmers] are being forced to destroy…fresh food that they can no longer sell” (Yaffe-Bellany & Corkery, 2020). This was not simply driven by an overall lack of demand, as simultaneously the number of food-insecure people in the United States was estimated to be rising in 2020 to 45 million from 2019’s 35.2 million after several years of steady decline. “Before the start of the pandemic, the overall food insecurity rate had reached its lowest point since it began to be measured in the 1990s, but those improvements were being upended by the pandemic” (Hake et al., 2021). Producers were responding to local economic forces, not to a global balance of supply and demand.

This starkly illustrates how, in a supply chain context, starting a unit of production development does not mean that a producer will necessarily ultimately make a finished good available for sale. Rather, producers are continually assessing all the choices available to them during the production and development process. The movement of goods monotonically along from raw materials, to WIP, to finished goods available for sale occurs only because the individuals managing that process choose to move those units along each period. As another example consider recent research on policies to address ‘artificial shortages’ in which producers withhold goods strategically based on expected future earnings (Levi et al., 2021). In the first example, producers were choosing to end production early, and in the second choosing to withhold finished goods. In both cases, the implicit connection between *production starts* and *finished goods available for sale* seen in many of the supply chain modeling frameworks developed in prior System Dynamics, and even Operations Management, literature breaks down.

For many contexts, the act of starting production does *not* necessarily guarantee that those units of inventory will become available to ship to customers. Aside from line losses (such as those due to production errors, quality assurance sampling, or even natural losses such as crop failure or spoilage), each unit of inventory under development represents some measure of operational capacity that is reserved and thus prevented from other uses. Thus, the act of producing a unit of inventory has within it an opportunity cost in the form of captured production capacity or other resource utilization. Managers are continuously making decisions in which they weigh the expected value of each disposition of the units of production under their purview, both finished goods and WIP.

Inventory disposition decisions can include, but are not limited to, normal production activities (e.g. moving material along in production or transferring finished goods to a customer in a sale), withholding production activities (e.g. holding production to free resources for some other activity or holding finished goods from sale under expectation of better future earnings), modifying production (e.g. ending production early and moving into a finished goods state with less than typical development time, or conversely spending more time under development than typical prior to moving in to a finished goods state), or even outright disposal of goods (e.g. disposing of either WIP or finished goods, outright removing them from the production process and any inventory). While disposition routes that regularly move goods from raw materials, to WIP, to finished goods, to sold-and-shipped goods can, and perhaps even should, be regular and routine under long-run steady state conditions, such regularity is not guaranteed but rather an outcome of a producer considering the relative value underlying each possible route.

This chapter contributes to System Dynamics methodology by presenting a framework to close the methodological gap between traditional inventory management and development models that directly connect production starts to finished goods, and the observed reality that this progression is not guaranteed. This is achieved by extending traditional inventory management and supply chain models found in System Dynamics literature by allowing for the endogenous determination of dispositions of inventory and production in a supply chain via the application of multinomial logistic (MNL) choice modeling.

Furthermore, this framework embeds the MNL mechanisms within a wider set of economically motivated decision rules that track the relationship between the age, or development time, of goods under production and their corresponding market value. In doing so, the value of a unit of production started and placed under development in a WIP state to the producer is considered in the wider context of the interplay of supply and demand (and resultant price setting). The value of continuing development of a unit of production is considered versus alternative disposition routes, including even purposeful disposal if relevant to the production environment. The movement of WIP into a finished goods state is done not because a specific period has elapsed but rather because of the underlying economic value of that decision versus other disposition options.

The remainder of this chapter is organized as follows: First a literature review places this chapter in the context of prior System Dynamics and related modeling research, including key work laying the foundations of the MNL framework presented here. Second, this prior literature is built upon by presenting an explicit framework to extend traditional inventory management models by incorporating disposition choice formulations and dynamic valuation of WIP goods. Third, an example model is presented that illustrates how inclusion of these mechanistic features can yield fundamentally different short and long-run behavior modes in similarly parameterized systems, and furthermore that this methodological framework can be applied to multiple levels of disaggregation of a production or aging process. Finally, the discussion reiterates the assumptions and limitations of this framework, while also emphasizing how this framework differs from more traditional inventory and production management modeling methods and how this can generate insights of interest to modelers, production managers, and others.

1.2 Literature Review

System Dynamics has a long history with incorporating models of the physical flow of goods through supply chains with the human elements that interact with those supply chains, starting with Jay Forrester's original modeling of the interactions of labor scheduling with production planning at General Electric (Forrester, 1961, 1989). Further investigations have included inventory-workforce interactions (Mass, 1975), production scheduling and planning via material requirements planning (MRP) systems (Morecroft, 1983a), and consideration of the supply chain in larger settings that can yield business and capital equipment purchase cycles (Anderson & Fine, 1999; Sterman & Mosekilde, 1993).

Stability of production, inventory, and information signals within supply chains has been especially of interest in prior System Dynamics literature, with extensive research on the origins of instability, often characterized by the bullwhip effect (Lee et al., 1997), specifically arising from behavioral heuristics when used in ordering decisions (Sterman, 1989a, 1989b), and how these systems can be stabilized either via observations on the cognitive features of the people making decisions (Narayanan & Moritz, 2015), or specific modifications to the information structure of the system (Croson et al., 2014; Croson & Donohue, 2006).

Inventory management models appear in much of the above referenced literature, and are described in detail in multiple System Dynamics textbooks (for example, see chapters 18 and 19 of Sterman, 2000, chapter 5 of Morecroft, 2015, or other illustrative uses of similar model structures in articles such as in Kampmann & Oliva, 2009). These classic models use basic behavioral feedback, tied to a producer's desired inventory coverage level, to adjust a stock of inventory based on a perceived demand signal from a consumer. These core inventory management models can be readily extended by considering the time between the act of starting production of inventory and the availability of that inventory (see chapter 19 and chapter 20 of Sterman, 2000 for a detailed example of this extension), or by considering other endogeneity such as the influence of inventory availability on customer demand patterns (Gonçalves et al., 2005; Morecroft, 1983b). More generally, other work has described principles of dynamic systems such as adding a minor loop to oscillatory systems like those seen in these inventory management models (Graham, 1977).

A key and fundamental consideration of the inventory management structures described above, and of the System Dynamics modeling framework is the purposeful incorporation of the behavioral features and heuristics employed by the individuals interacting with physical and information systems. These are often captured via decision rules that attempt to capture how a model of a human decision maker in the larger system incorporates information and observations to make a choice or action. This concept of modeling choices is not unique to System Dynamics but is also used extensively in modern economics literature. Specifically, discrete choice models have emerged over the last few decades as a method of empirically modeling the probability of observing outcomes among a finite set (Greene, 2018). For the scenario where choices are collectively exhaustive, mutually exclusive, finite in number, and that have the feature of independence from irrelevant alternatives (IIA), then MNL can be used (McFadden, 1974).

The original article that popularized of this framework in marketing applications expressly described how MNL could be used to form models that resolve the, often invisible, heuristics employed by individual decision makers into population-level outcomes:

> *"The link between models of individual behavior and data on population choices is most critical when the decision-maker's alternatives are qualitative, or 'lumpy.' In conventional consumer analysis… one can often plausibly assume that all individuals in a population have a common behavior rule, except for purely random "optimization" errors… [MNL is a] general procedure for formulating econometric models of population choice behavior from distributions of individual decision rules." (McFadden, 1974)*

To be clear, the MNL framework does not claim to represent the actual decision process used by any one decision maker to come to a singular and discrete choice, but rather capture the aggregate behavior observed in a population of decision makers all making similar choices.

This MNL choice modeling framework has become the most widely used in econometric analysis primarily because the formula relating the utility of each choice to the probability of that choice is closed form and easily interpretable (Henserh et al., 2005; Train, 2009). While extensions of this choice model exist, such as the probit model which relax assumptions related to the independence of the choices (Train, 2009), in supply chain modeling the set of disposition choices for inventory after or during any one step of production often follow the assumption of being collectively exhaustive, mutually exclusive, and finite.

The usage of this modeling framework in a supply chain context has been relatively limited, with perhaps a notable exception for its use in models of transportation applications (Aloulou, 2018; de Bok et al., 2018), especially in comparison with its near ubiquitous use in other settings such as Marketing (Chandukala et al., 2007). However, this has begun to change with the emergence of Behavioral Operations Management as a distinct subfield, which has leveraged the MNL framework in other choice settings in an operations context (see chapter 2 and chapter 17 of Donohue et al., 2018 for recent examples).

Within the System Dynamics literature, the use of discrete choice modeling frameworks is similarly sparse, though much of the underlying mathematical theory overlaps with parameter estimation tools such as method of simulated moments (Hosseinichimeh et al., 2016; Jalali et al., 2015; Train, 2009). Explicit use of this choice framework in System Dynamics modeling has followed more closely to Marketing applications, determining expected market share of different options given perceived utility in specific consumer contexts (Keith et al., 2017; Rahmandad & Sibdari, 2012). In a single industrial context found by this author, seemingly more superficially similar to the supply chain context discussed in this chapter, the use of the multinomial logistic (MNL) choice model is still ultimately framed in terms of relative market share of fuel options for running electricity plants (Moxnes, 1990). Those examples from compartmental aggregate models in System Dynamics literature do highlight an important feature discussed in more detail below, namely that the probabilistic nature of the MNL choice model allows for discrete and mutually exclusive choices at an individual level to be expressed as expected outcomes at an aggregate level.

1.3 Adding Disposition Choice Formulations into Supply Chain Models

Consider the well-known inventory management model seen in Figure 1-1 (Sterman, 2000). This model captures the delay between starting production and having inventory on hand for shipment to a customer with a delay formulation between *Production Start Rate* and *Production Rate*, resulting in an accumulation of inventory in the form of work-in-progress (WIP). When creating a model of this system, this delay can be as simple as a fixed pipeline delay or is often represented as a more complex third-order delay to capture some sense of multi-stage production.

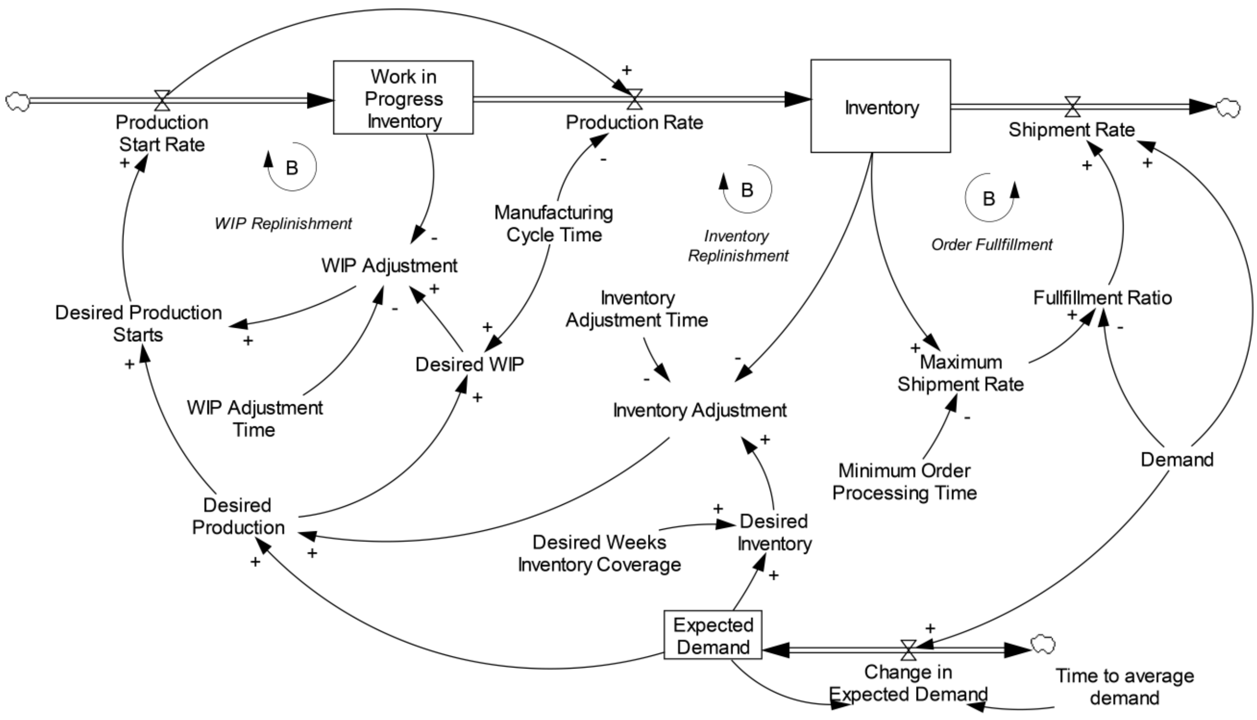


Figure 1-1. Inventory Management Model with Production Delays

Each unit of inventory under development represents some measure of operational capacity that is reserved and thus prevented from other uses. Thus, the act of producing a unit of inventory contains within it an opportunity cost in the form of captured production capacity. When applied to WIP inventory, the opportunity cost concept can be extended by noting that act of holding of inventory and value-added development processes that are assumed to occur in a WIP state are not without cost. Aside from the more obvious direct development costs and holding costs of literal work in process, holding inventory in such a state could prevent another unit of production starts from entering development (if production capacity is finite, fixed, and full), or at minimum contains some measure of opportunity cost by virtue of simply taking up physical space that could otherwise be utilized for a totally unrelated purpose.

The model shown in Figure 1-1 has one clear inventory disposition route: movement into a finished goods state after some (average) manufacturing cycle time. However, given some unit of production starts, the time that unit is under development is ultimately a choice of the producer, not fixed a priori. As illustrated in the examples that opened this chapter, there exist environments where WIP can be terminated at any point and the unit under production either moved to a finished goods state, withheld from sale, or even outright disposed.

The farmers described in the introduction to this chapter made the difficult decision to destroy their crops because ultimately it made economic sense to do so. The costs of harvesting, processing, and transporting their goods exceeded the value they would get from selling finished goods, and even exceeded the opportunity cost of leaving the goods in the field, either in terms of holding up productive capacity or due to the loss of value from spoilage in the ground. The time that a unit of production is under development, and considered WIP inventory, may also have a meaningful economic impact on the final value of the product at hand. Consider a piece of software under development, where the value of the final product may increase with increased development time, but with decreasing marginal returns. Or consider a crop that is to be planted and harvested, with a specific window of maturation time in which it could be sold at full market value.

The ‘Manufacturing Cycle Time’ shown in Figure 1-1 is not fixed in this choice-centric view of production. At the level of an individual producer, it is not even an average of a distribution of times. Rather it is an explicit choice made by the producer based on the economic features of the landscape in which they operate.

From the point of view of a single producer, each possible disposition of a unit of production in a WIP state is likely to be mutually exclusive (e.g., in a single period a farm cannot simultaneously destroy, harvest, and continue to cultivate a single unit of food). Under a model of a single producer with fully known and fixed values (or costs) associated with each disposition decision, this economic decision becomes a straight-forward assessment of the expected value of each disposition route (for example weighing the costs of shipping and storing food versus the costs of destroying it, offset by the value that would come from selling if it were sold).

The multinomial logistic (MNL) framework discussed above does not capture the individual decision-making process for any one producer in these examples who must make a disposition decision for their goods under development in a given timeframe. However, when the unit of analysis is aggregated to a *system* of producers, each with structurally similar representations of the value of those choices, then the MNL choice model becomes an appropriate method of capturing how these individual level, and likely heuristic, decision processes resolve at a population level.

For an aggregate compartmental model of many producers, the MNL framework provides a probability that any unit of production will end up in any one of the disposition routes, which resolves to the total expected WIP inventory that is delegated to each of the possible disposition routes. As stated above, this requires that the individuals making these decisions have structurally similar decision rules, or more precisely individuals making these disposition decisions are fully informed with stationary costs (e.g. are fully informed about the cost structure of the system in which they are embedded), and that there is no correlation among choices (McFadden, 1974; Train, 2009). These assumptions can be relaxed in part by applying alternative methods that allow for correlation, like probit or mixed logit models (Revelt & Train, 1998; Train, 2009), or modifications to allow for stochasticity of observable data (Marcus, 1991).

The assumption that the individuals know the costs associated with each disposition in a supply chain context follows from the assumption that the model is capturing the aggregate decision processes of those actors that control the routing to those dispositions. However, the use of the MNL choice modeling framework is more generally consistent with classic discrete-choice foundations and contains within it a realistic assumption of human ordering behavior, namely that the proportion of goods relegated to any specific disposition route is proportional to the *relative* benefit of any one of those routes.

A model with the MNL formulation can still follow the principles of modeling decision making (Morrison & Oliva, 2018), even if the MNL framework itself does not reflect the individual-level details of the decision heuristics that result in specific inventory disposition choices. When embedded within a larger model that captures the delays and feedbacks that create the signals that form the perceived relative benefit of each choice, and that also acknowledges that the MNL framework results in an *expected* or *desired* outcome that may differ from the actual realized outcome, the resulting model is using the MNL formulation as a representation of the aggregate formation of a goal (or really series of goals, one for each disposition option) . The rest of the model still must translate that goal into action.

If these disposition routes follow the assumptions of MNL modeling (collectively exhaustive, mutually exclusive, finite, and IIA), the probability of choosing disposition route X_i of N available choices is given by the expression below, where v_i is the mean value of each disposition.

$$P(X_i) = \frac{e^{v_i}}{\sum_{I=1}^{N} e^{v_i}} \tag{1}$$

The components of what defines the 'value' of each disposition must be determined for each application of this framework and should be based on behaviorally realistic and grounded drivers under consideration by the decision makers being modeled. How a producer assigns the relative values of each of the disposition choices is a matter of modeling freedom and should be based on observations of how real producers value these choices. The advantage of the MNL model is that changing or updating the assumptions that form this value assessment only changes the relative value of each choice, and thus the relative proportion of the units under development delegated to each option, but not the underlying model.

For simplicity of presentation here and the examples elsewhere in this chapter, we assume that the sole driver of value is *economic* (e.g., proportional to expected profitability) for each disposition. This can be relaxed based on actual observations or assumptions of the system and problem being modeled to include additional drivers that determine the perceived relative value of each disposition route for the decision maker. For this simplified model, for some relative expected profitability π_i for choice X_i, expression (1) can be rewritten as follows:

$$P(X_i) = \frac{e^{\beta \pi_i}}{\sum_{I=1}^{N} e^{\beta \pi_I}} \tag{2}$$

In the above, β is the weight the producer places on the concept of expected profitability when making an inventory disposition decision. Given a wider definition of value, there would be corresponding parameters determining the weight the producer places on each driver. Under a full MNL model, these become free parameters used to help fit the model to observed data of disposition choices.

By inspection of (2), lower absolute values of β in this simplification have the effect of diminishing the relative influence of the economic value of each disposition. In the extreme case where $\beta = 0$, expression (2) reduces to allocating equally to all dispositions, independent of the economic value of those dispositions. For this extreme case, the model is stating that economic value is the sole driver of inventory disposition decisions, but that the influence of that driver is still null and thus the decision maker is indifferent among all options.

Conversely, if the value of β is large, then the decision maker is extremely sensitive to even small differences in the economic value of each disposition route, causing higher valued expected dispositions to be even more likely to be chosen. In the extreme case of $\beta \to \infty$ then the *only* disposition possibility for a given period is the one with the highest economic value, with a probability of 1 and all other dispositions being 0. A proof of these extreme conditions is provided in Appendix A.

To further simplify the modeling framework presented here, we can fix values of β to be the inverse of some reference price for the producer (e.g., the price at which a farm sells its goods under normal steady state conditions). We can do this in part because for a given weight β (that is neither null nor infinite) the output of the MNL choice model ultimately depends only on the relative difference in utility of each option (Greene, 2018), and this has the further advantage of allowing the relative values of each choice, π_i, to be expressed in terms of prices and monetary values, while allowing the expression above to properly reduce to a dimensionless probability.

When the unit of analysis is a single producer (or even a single unit of production under development), these probabilities collapse into one discrete outcome. In other words, a single unit of production cannot be simultaneously in multiple dispositions. However, when the unit of analysis is expanded to consider a *group* of producers or continuous flow of goods through a production system, then expressions (1) and (2) can be used to find the expected aggregate values of goods in each disposition at any period of time, given the expected value (or as simplified in (2), profitability) of each disposition.

As the value used in these expressions is an *expected value* with some variance, for any single unit of production the *realized value* may be higher or lower than that expected value. As the choices are assumed to be exhaustive and mutually exclusive, the expected amount of goods in each disposition will be nonzero at any given time, so long as the weight placed on economic value is not infinite as discussed above. This includes extreme cases such as a disposition route having an average expected zero value (often referred to as the null choice in Marketing a Behavioral Economics literature that uses this modeling framework), or even average expected negative values. Again, these disposition values are the expected values with some variance, and therefore there is assumed to always be some proportion of goods whose actual realized value for each possible route will result in it being assigned to any one of the dispositions under consideration. In general, we would expect higher average economic valued dispositions to receive the plurality of units, but not necessarily absolutely every unit due to internal variance between goods.

For example, consider a single unit of production that ends development with very low quality relative to the average of other units of production. Should the producer choose to move this unit of production through the rest of the supply chain, they will incur additional real costs of storage, transportation, and associated overhead, along with opportunity costs from locked up capital. If the quality is low enough, the good may have little to no value on the market, or even little to no secondary market or salvage value. The producer may *still* have to incur disposal costs at this later stage if no salvage is possible.

By choosing to hold on to this low-quality item through production through the subsequent, and costly, steps that move this good towards final sale the producer possibly incurs more costs versus if they had chosen to dispose of the product earlier. For this specific example, the producer has a choice among only negative outcomes, but can make the least costly decision by disposing of this unit of production early, and thus free up capacity and capital for production of an expected higher quality unit of production. On average, the producer expects that moving goods through the production process towards final sale is valuable, while purposeful disposal of goods under production is costly. As the average expected values of each disposition option widen from each other, then the likelihood of a unit of production falling into this most costly, on average, route diminishes. However, for this specific example realization, the choice of disposal makes the most economic sense for the producer. The power of this framework is the ability to capture this more behaviorally realistic variability in outcomes at an *aggregate* level of analysis without needing to track outcomes at the most granular unit of production level.

Figure 1-2 shows visually how the above MNL choice model can be incorporated into the basic inventory management model first presented in Figure 1-1. Note that the actual fraction of WIP that can continue development is not directly used in the model, rather it is implied via the use of the MNL choice model and the relative values of each disposition route.

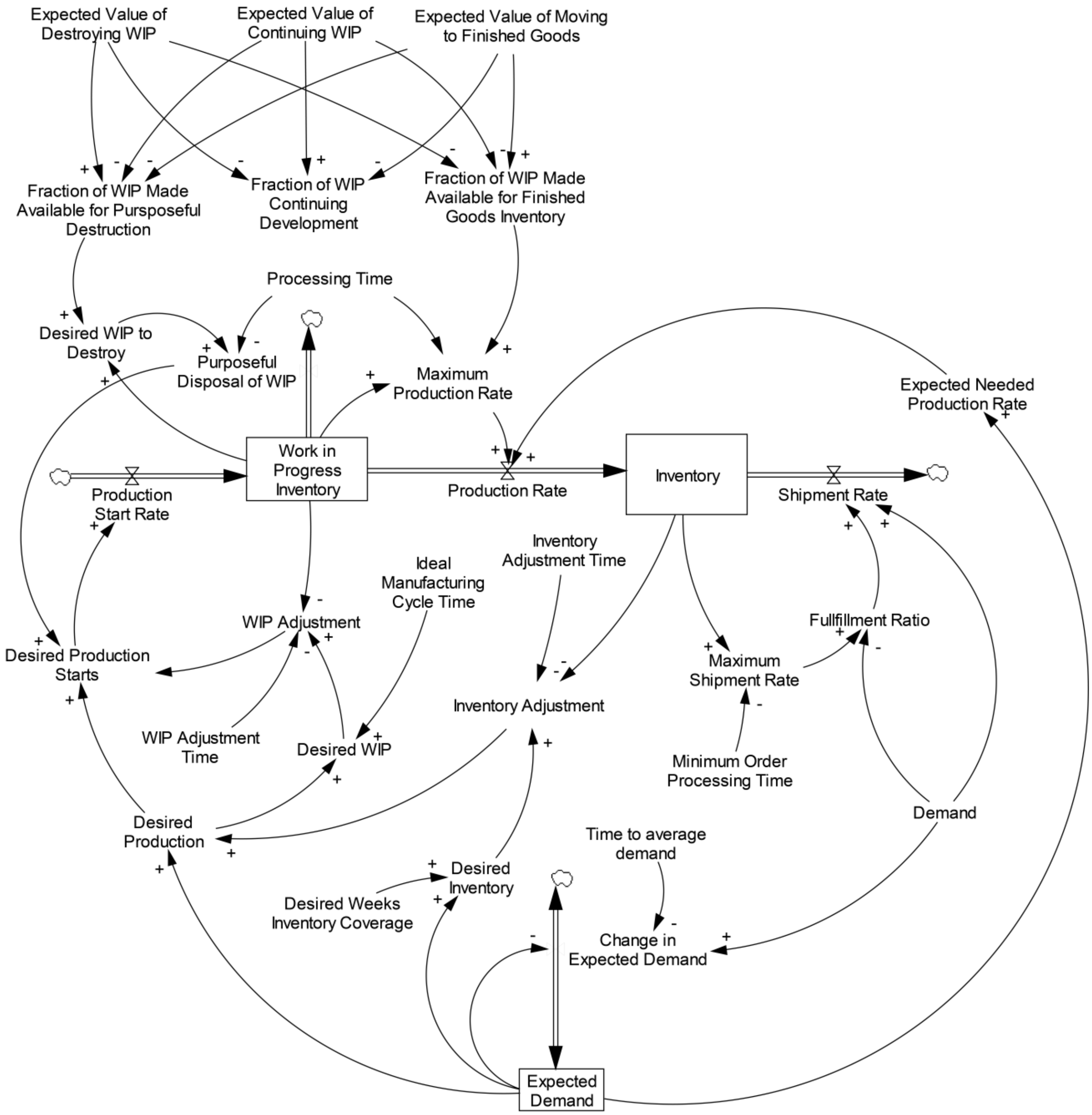


Figure 1-2. Multinomial Logistic Choice Model and the Inventory Management Structure

A cursory inspection of Figure 1-2 may imply that this example has two possible inventory dispositions: purposeful destruction of WIP or production (moving WIP into a finished goods state). However, this system represents *three* possible disposition outcomes in any unit of time by expression: Goods that the producer would like to continue to develop (i.e., stay in the WIP stock), goods the producer would like to dispose of, and goods the producer would like to move into a finished goods state. By expression (2) the relative size of each of these cohorts is based on the relative economic value of each disposition. In this aggregate model of WIP inventory, the producer is assumed to have an ideal manufacturing cycle time that yields maximum economic value and determines the expected needed production start time (this is relaxed in the sections below), and a single processing time for WIP inventory designated for movement to the finished goods state and that marked for disposal. The producer may have different times for these two dispositions, but for the sake of compact presentation they are combined here.

1.4 Dynamic Valuation of Work in progress

The model presented in Figure 1-2 assumes fixed economic values for each inventory disposition. However, as discussed earlier in this chapter, the time that a unit of production is under development may also have a meaningful economic impact on the final value of the product at hand, and thus the value of either holding or shipping inventory may change with a concept of time under development or age of the work-in-progress (WIP).

Adopting a co-flow structure typically used to keep track of continually time-accruing attributes (Hines, 2005; Sterman, 2000) can be used to keep track of a concept of average development time (or average age, or average maturation time, or any similar measure that is appropriate to the system under investigation). Figure 1-3 illustrates this extension, where a co-flow structure is used to produce a concept of the *Average Age of WIP Inventory*, which in turn is used to adjust the expected value of moving WIP into a state of finished goods for sale to a downstream wholesaler or other customer.

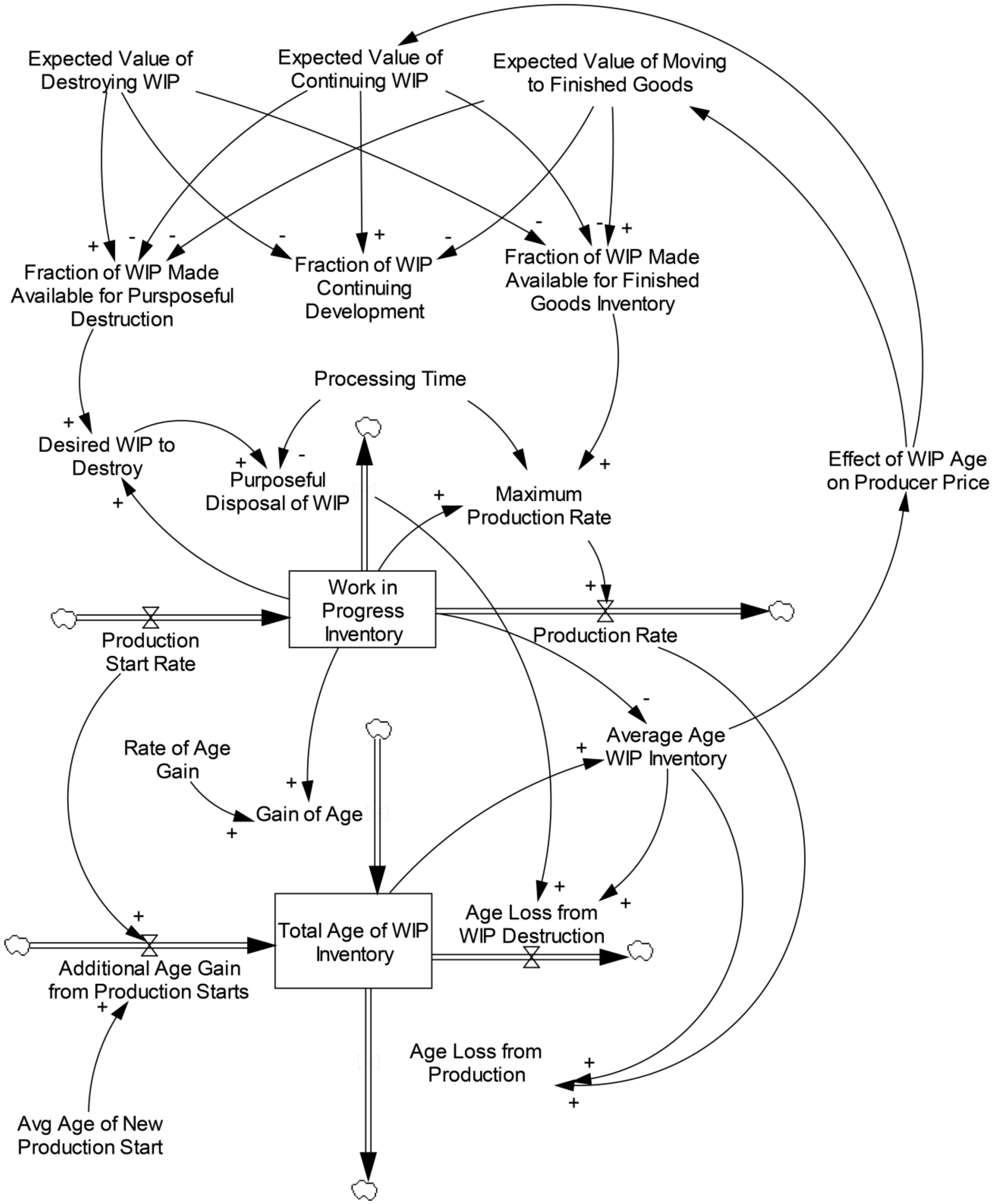


Figure 1-3. Extending the Model to Track Development Time

Note that in the above formulation, the ‘Average Age of New Production Starts’ may have a value of 0 units of time, or some other non-negative value. For example, when applied to employee experience in a firm, an employee may arrive with some pre-existing experience. However, in the context of a food producer planting crops it may be safe to assume that newly planted crops arrive with no pre-existing maturation. The model is flexible to allow for this assumption to be relaxed based on specific circumstance (for example, buying partially matured nut trees or fully matured sows rather than starting from seeds or piglets).

Furthermore, while it may be safe to assume that the 'Rate of Age Gain' is constant and unitary (i.e., 1 weeks/week or 1 years/year or similar). The formulation itself does allow for some flexibility if, as an example, a fertilizer was applied to speed maturation, or a drought hit and slows maturation down.

1.5 Applying the Framework to a Model Supply Chain

The purpose of the formulation developed above is to provide a flexible framework for applying this concept of age-dependent economic features affecting the disposition choices of producers. As mentioned earlier in this chapter, this framework must be embedded into a larger model that captures the delays and feedbacks that allow these new structures to properly integrate into the principles modeling decision making (Morrison & Oliva, 2018). The section below illustrates this by combining the framework developed above with spot price discovery and capacity management in a simple example supply chain model to illustrate the additional insights adding this combination of multinomial logistic (MNL) choice modeling and price-value relationships can yield. The purpose here is to illustrate how the fundamental modes of simulated behavior differ when incorporating these structures, rather than providing a specific analysis of the example model itself. The interested reader can find more details on this example model beyond what is provided below in Appendix A and the accompanying model files.

While there may be multiple ways to construct the interplay of supply and demand that ultimately forms the spot price at each interface point of producers and consumers in a market, the example below utilizes inventory-sensitive spot pricing most often seen with commodity products (Chen et al., 2009; Sterman, 2000; Whelan & Forrester, 1996). At the core of this economic model are three balancing loops across each entity in the supply chain, with spot pricing driving either demand or supply. Producers offer a good at the price expected by the market modified by the current inventory levels. Given the current spot price in the market, the expected gross margin of the producer is affected, which in turn affects capacity utilization and production starts in a balancing loop. Additionally, the spot price is fundamentally anchored to what the market expects it to be, and this introduces a reinforcing loop around the spot price and the expected prices. These loops, in the context of a general n-entity supply chain, are visually summarized in Figure 1-4.

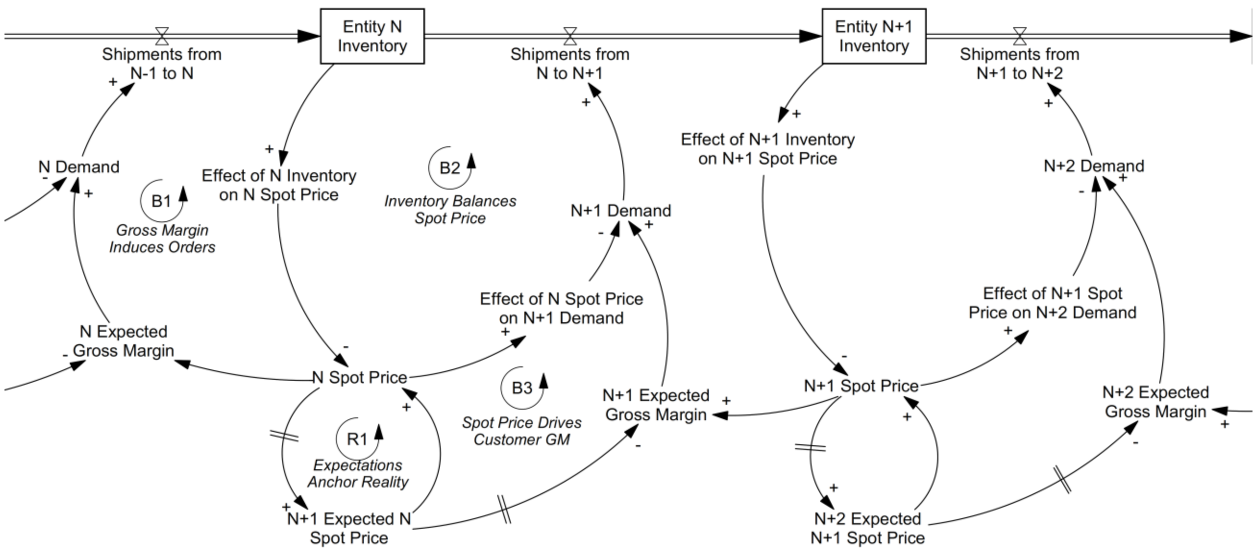


Figure 1-4. Ordering and Price Setting is Nested in Larger Interconnected Supply Chain

This price setting system seen in Figure 1-4 is then combined with the MNL choice modeling framework shown in Figure 1-3 to create the central contribution described in this chapter. The exact definition of the relationship between development time and the fraction of the full price that the producer can extract is context specific. The concept of 'Effect of WIP Age on Producer Price' seen in Figure 1-3 is left purposefully ill-defined. This relationship is ultimately context-specific and can vary depending on the product under development and how the market values that product as a function of the development or maturation time.

As discussed above, this relationship could be increasing with increased development time but with decreasing marginal returns, as in the case of software development. Or it could have an asymmetric gaussian shape with an ideal development time but with diminishing returns on either side of that time. This could be further simplified as an asymmetric trapezoidal shape if the window in which the full value of the good is not a single period but rather a window. An example of a product that could be conceptualized by a trapezoidal relationship between the time of development and the ability of the producer to extract the full market value of their goods would be a commodity food under cultivation. For such a product, there is an ideal window of maturation time, or work-in-progress (WIP) age utilizing the language of the frameworks presented above, at which the food can receive its full economic value. Outside of this window, the producer or farmer can expect less than full value or even no value at all. This trapezoidal relationship between development age and price can be operationalized via expression (3).

$$f\ (t) = \begin{cases} 0 & t \leq a \\ \left(\frac{1}{b-a}\right)t - \left(\frac{a}{b-a}\right) & a < t \leq b \\ 1 & b < t \leq c \\ \left(\frac{1}{c-d}\right)t - \left(\frac{d}{c-d}\right) & c < t \leq d \\ 0 & t > d \end{cases} \tag{3}$$

More general trapezoidal shapes are possible that do not necessarily have linear changes in value (for example see Dorp & Kotz, 2003) and may be more appropriate in specific contexts. Appendix A presents several other functional forms that this relationship could take on as examples for other modeling activities, including fixed-values, linear and saturating, and symmetric or asymmetric gaussian relationships.

The example supply chain model was started at steady state using the parameterization shown in Appendix A and model documentation that accompanies this chapter utilizing the trapezoidal age-value relationship presented immediately above. The exact parameters of the model were chosen to be semi-realistic, but ultimately, they are illustrative only and not the focus of the discussion below or the illustration of the new modeling structures combining co-flow differential equation structures, spot price discovery, and MNL choice modeling. The model was then exposed to an exogenous shock in the underlying consumer demand for goods, increasing 50% over the baseline value for a total of 40 weeks (from week 10 to week 50 in simulated time) before returning to baseline demand.

First, Figure 1-5 shows the demand and production pattern of the system *without* the MNL choice model nor the age-value relationship. In this baseline scenario, the supply chain is modeled in a manner like prior work, with production starts increasing to match increased demand, and all WIP inventory eventually moved into a finished goods state. Most importantly, this system as parameterized is relatively insensitive to the rectangular pulse in consumer demand, quickly returning to the prior level of production once the demand surge subsides. Note that this baseline model still allows for losses. Here, it is assumed that some finished goods are lost due to spoilage or obsolescence, resulting in more needed production starts than shipments in steady state. However, these losses occur as a result of the structure of the system, not due to explicit decisions made by the producer.

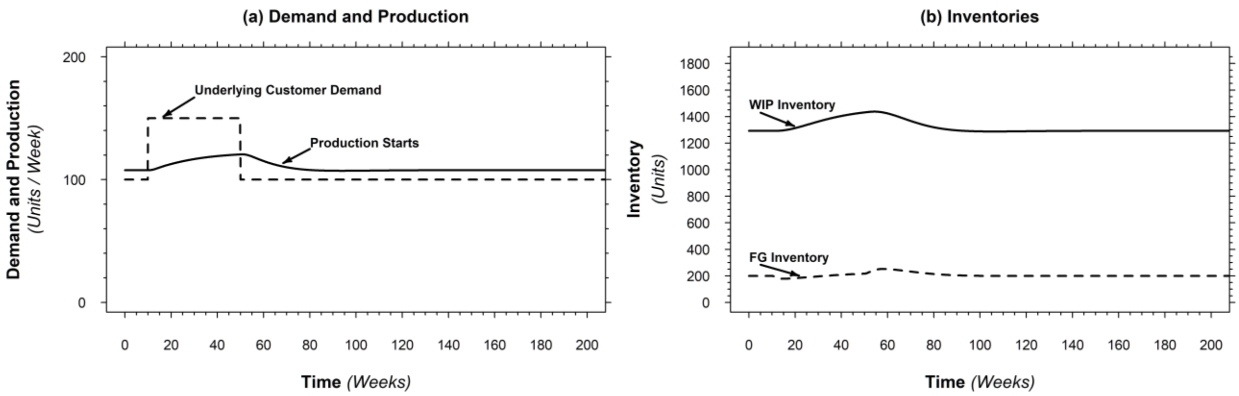


Figure 1-5. Baseline Response to Rectangular Pulse in Consumer Demand

Figure 1-6 shows the same system with the same parameterization but with WIP inventory advancing through the system subject to the MNL choice model introduced in expression (2) and as illustrated in Figure 1-2.

Incorporation of the MNL choice model means that prior to the shock in demand the producers will be disposing of some fraction of their WIP every period. In other words, the initial steady state value of disposal is not zero in Figure 1-6. Even though the expected value of disposal is significantly lower than other options, even taking on a negative value as parameterized in this example, some goods are still disposed of each period. As discussed in the sections above introducing the MNL model, the values of each disposition route are assumed to be expected values with some variance, and thus for any individual unit of production, the actual realized values of each disposition option may result in purposeful disposal being the most economically advantageous choice.

A key change to the behavior of the system introduced by the MNL choice model is the introduction of non-zero flows towards purposeful disposal in an aggregated representation of a supply chain[1]. As described in the example above, there is expected to be some amount of

[1] Note that non-zero flows towards purposeful disposal will occur to at least some degree for all parameter values unless under the extreme case of zero-valued reference prices. As built here, the β in expression (2) is the inverse of the reference price in the system, so zero-valued reference prices would drive β to infinity and, as discussed in the introduction to the MNL framework, route goods towards the highest valued disposition route at each period with probability 1. For infinite reference prices, the probability of disposal would be the same as any other disposition option, and for any other value of reference prices between 0 and infinity, the disposal probability would follow expression (2).

disposal in real supply chains that produce goods with some degree of variability, either in terms of quality of the goods or in terms of the production process itself, and use of a framework like MNL helps capture this expected behavioral outcome that is often outright excluded in the more traditional supply chain models illustrated in Figure 1-1.

As discussed in the Literature Review section, the MNL framework does not claim to represent the actual decision process used by decision makers to come to a singular and discrete choice, but rather capture the expected aggregate outcome of many similar choices. There may be known boundedly rational heuristics that capture the specific process being used by decision makers that are applicable to specific production scenarios. These explicit decision rules with discrete outcomes would then require keeping track of individual disposition outcomes. However, at an aggregate level of modeling, wherein the specific disposition of an individual unit of production is inconsequential relative to an aggregated response or outcome, the aggregated framework is preferred.

> *"Discrete choice models operate at the level of individual decision makers. However, the researcher is usually interested in some aggregate measure, such as the average probability within a population or the average response to a change in some factor."* (Train, 2009)

As discussed above, the MNL framework is used in this chapter because supply chain inventory disposition decisions can often be reasonably assumed to follow the underlying assumptions of this framework (e.g., collectively exhaustive, mutually exclusive, finite in number, and IIA). Should the specific decision-making process being captured in a model not follow these assumptions, then either an alternative framework could be used, or an agent-level modeling process adopted which aggregates individual disposition outcomes towards system level observations.

Independent of the framework utilized, it is expected that there will be some non-zero flow of goods towards disposal. This non-zero flow each period under steady state means that the producer must carry more WIP than they would otherwise to meet the same level of demand. With the rectangular pulse in consumer demand comes increasing spot prices, and the relative value of moving goods into a finished goods state versus continuing development or disposing of those goods also increases. This results in the immediate drop in the disposal of units and a small but still present drop in the average age of goods under development. As production starts begin to catch up with this increase in demand, the demand surge subsides. The value relationships shift again, with the relative value of holding goods in the WIP state or outright disposing of goods becoming more attractive versus finishing production and holding finished goods. This results in both an increase in maturation times and surge in disposal of goods.

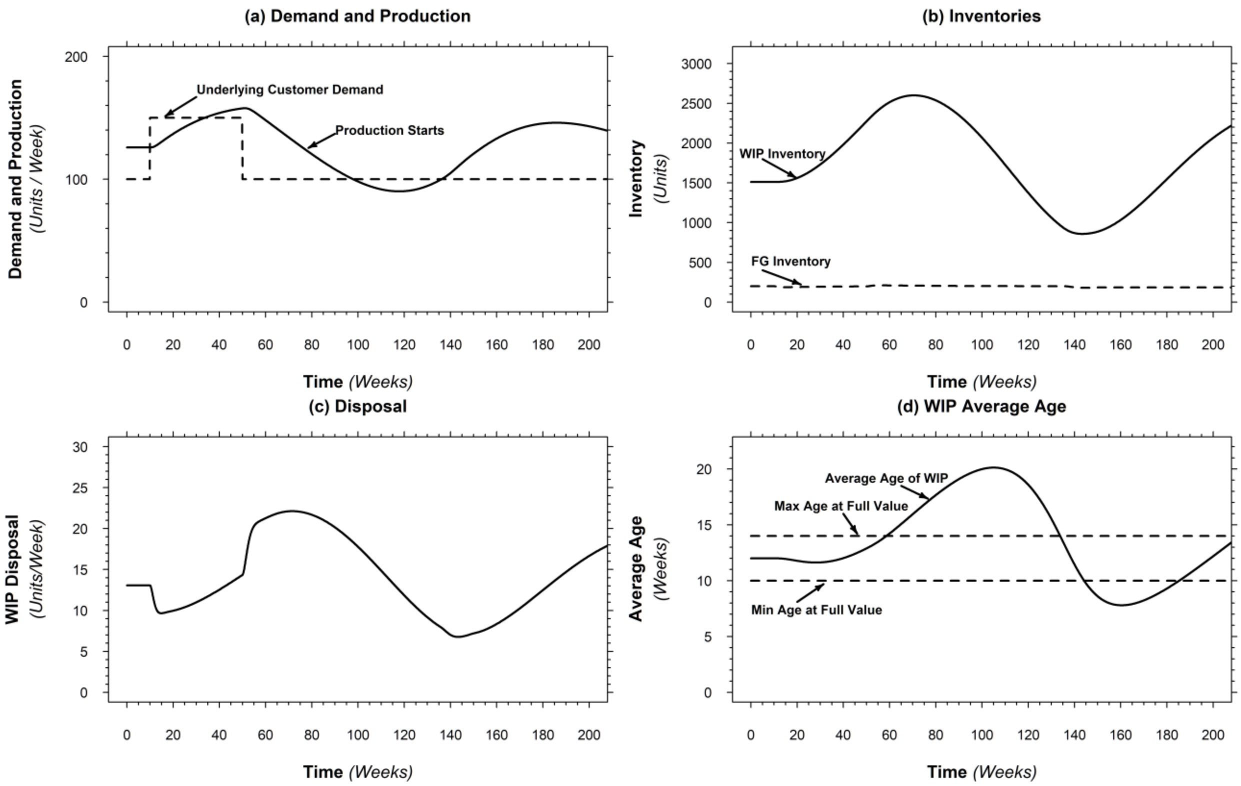


Figure 1-6. MNL Extended Model

Finally, Figure 1-7 incorporates a trapezoidal relationship between the average age or development time for the WIP goods and the ability for the producer to extract the full spot price in the marketplace. Now, holding goods in a WIP state beyond the maximum age at full value, or attempting to move goods into a finished goods state before the minimum age at full value reduces the price the producer can demand in the marketplace. As with the model used in Figure 1-6, with onset in the increase in consumer demand comes increasing willingness to pay by consumers, and the relative value of moving goods into a finished goods state also increases, even with a smaller ability to extract the full spot price from the market from goods that have been developed for less time. This results in an immediate drop in the disposal of units and a drop in the average age of goods under development. This moving of goods into a finished goods state earlier than under steady state conditions continues, resulting in a larger and quicker rundown of WIP inventories versus seen in Figure 1-6, and a correspondingly higher response in the form of later higher production starts. The net result is that the incorporation of this additional structure relating the development time for the WIP goods to the ability for the producer to extract the full spot price in the marketplace exacerbates the dynamics seen in the model with the MNL framework alone.

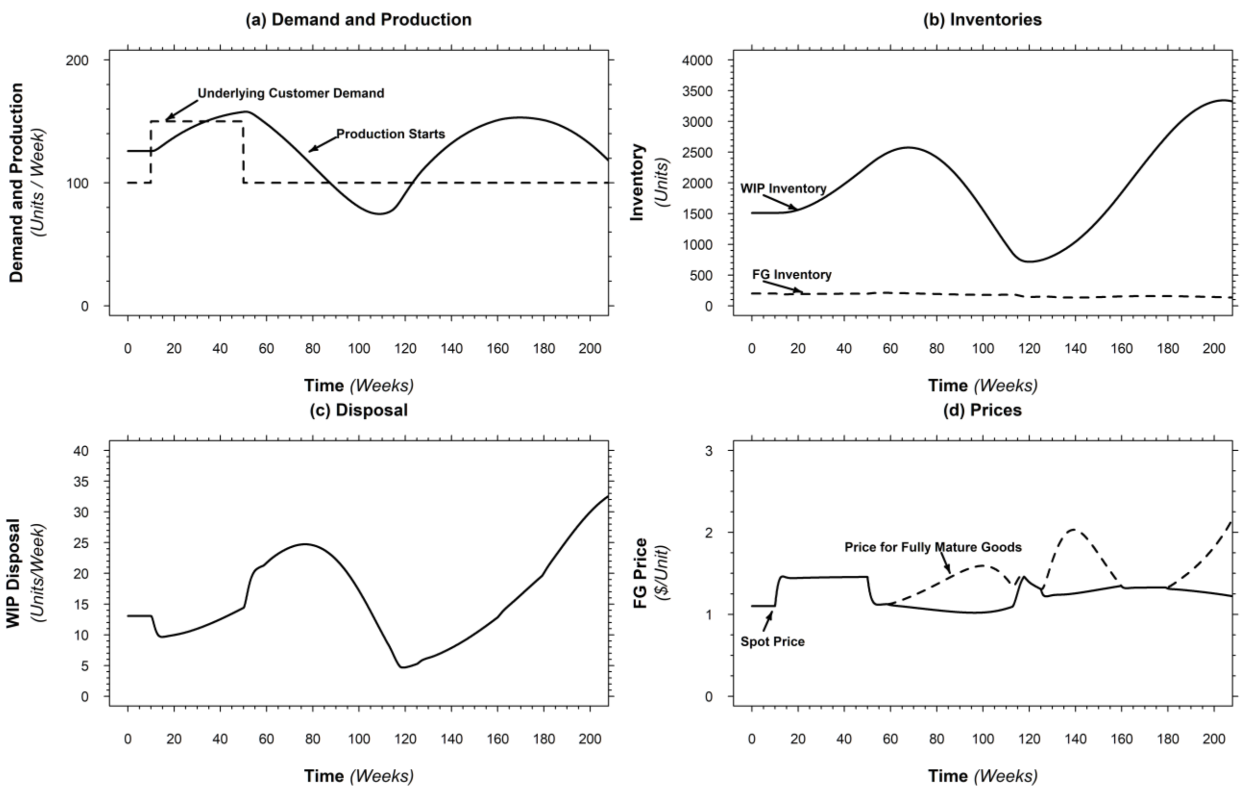


Figure 1-7. MNL with Age-Value Relationship Model

Perhaps most importantly beyond providing an example of the use of this modeling framework, the three scenarios illustrated above (baseline, with MNL choice modeling, and with MNL choice modeling and with age-value relationships) settle on three distinctly different qualitative modes of long run behavior. This is illustrated in Figure 1-8 for the inventory in the WIP stock, but similar patterns emerge for other key variables, including production starts and spot prices.

For the example here, with the parameters chosen and the trapezoidal relationship between average age of WIP goods and the effect on price, the baseline model settles quickly with negligible oscillatory behavior, while incorporating the MNL choice modeling frameworks generates oscillatory, but damped, patterns in production starts and inventories. Adding the relationship between development age and the price extracted to the marketplace can result in undamped oscillations over the long run.

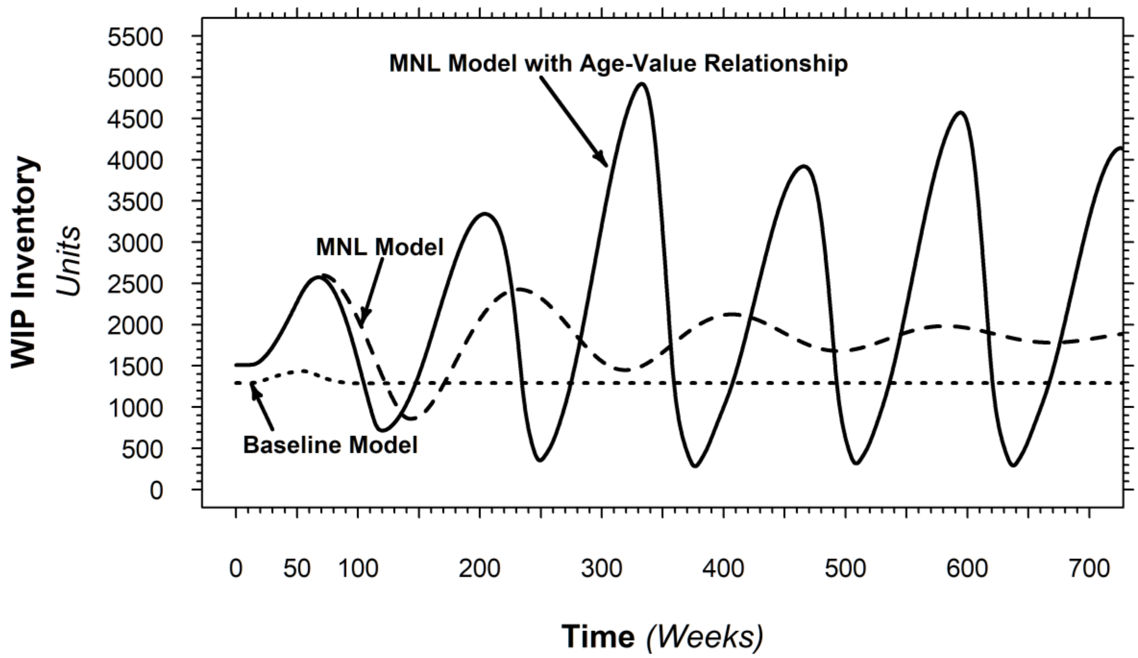


Figure 1-8. Long Horizon Comparison of Methodological Framework

The incorporation of the relationship between development age and the price extracted to the marketplace creates an additional feedback loop connecting the *Average Age WIP Inventory* to *Effect of WIP Age on Producer Price* in Figure 1-3, a connection which is otherwise absent. Figure 1-9 provides traces of behavior for parts of the model that interact in this new loop that generate the long-run oscillations seen in Figure 1-8 that are otherwise absent when not incorporating this relationship. Specifically, as parameterized in this example with the trapezoidal relationship between age and price, once the age of goods significantly exceeds the value at which the producer can expect full value, the expected prices begin to sharply drop. In response, after a delay, production starts are similarly curtailed based on future expectations of profitability. The underlying demand, after the initial rectangular pulse, remains and as a result goods are eventually sold down. As the average age of the WIP reduces back to a value at which they would receive full market value, this drives up the price expected by the producer which begins instigating newly increased production starts. This cycle continues.

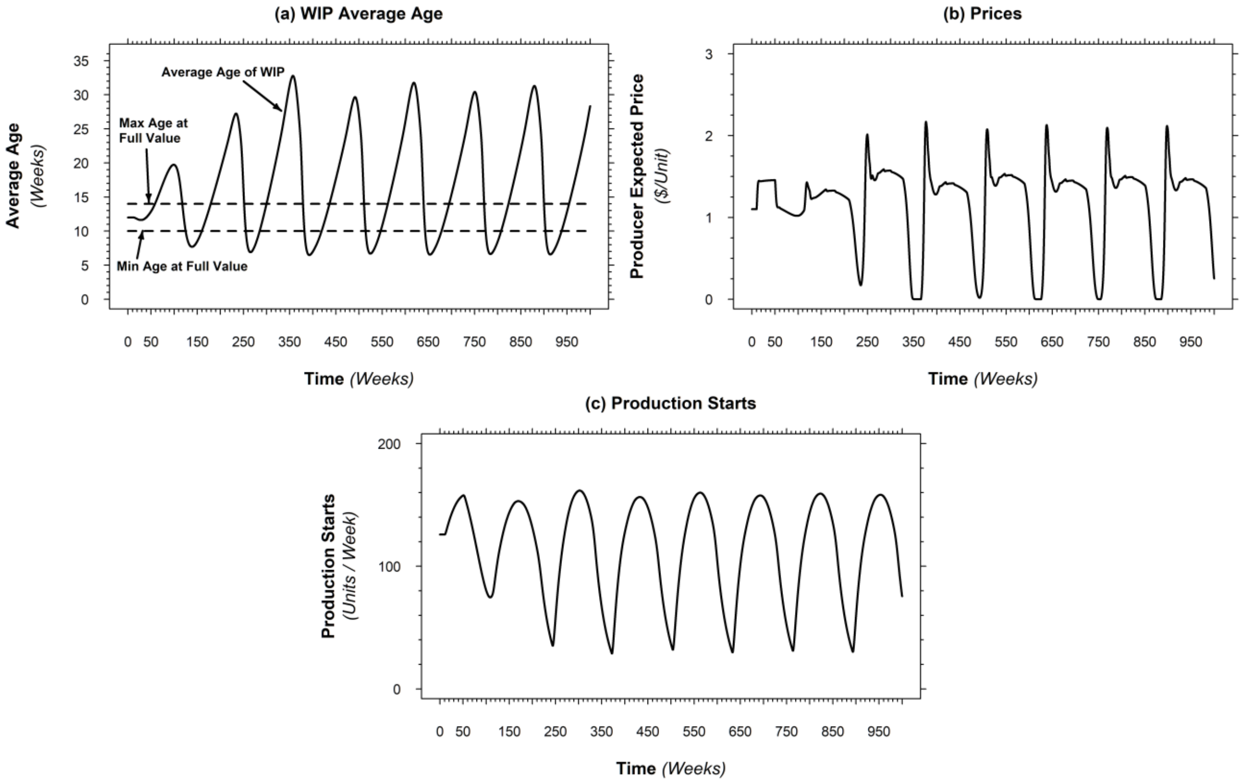


Figure 1-9. Detail on Feedback Generating Long-Run Oscillations in Model with Age-Value Relationship

These specific outcomes are a function of the specific parameter choices utilized in this model and were chosen to illustrate that the behavioral modes of the model can be materially different versus the traditional inventory models if one chooses *not* to incorporate these mechanisms. Of interest is that inclusion of these mechanisms also implies new policies.

In this parameterization, the long-run oscillatory behavior is sensitive to the time under which the producer updates production schedules, e.g., the delay that translates a desired production start rate into an actual production schedule, and the rate in which the producer incorporates the instantaneous spot pricing realized in the market into long-run expectations. The base model, and the model with the MNL framework but no age-value relationship, do not fundamentally change modes of behavior when adjusting these delays, while the full model does (with longer updates reducing or even eliminating the long-run oscillation). More specific commentary on these results is not the subject of this section of the paper, but it does reinforce that the presence of these mechanisms can materially change both the predicted modes of behavior in a model, and in turn illustrate policy levers that would have been otherwise hidden.

1.6 Vintaging Chains versus Aggregate Stocks

A valid critique of the structure presented in Figure 1-3 is that the work-in-progress (WIP) by the producer is treated as a single aggregate collection of units with an average effective age under production. This structure abstracts away from capturing the outflows from WIP to either finished goods or to purposeful destruction of goods at specific ages, or at specific points in the development process, and instead considers only a concept of average age of each of these dispositions. Additionally, the exact distribution of ages of material that is currently under WIP is abstracted away with only a mean value known.

While this aggregate framework may be a valid model in some contexts, in others it may be important, or at least of significant interest, to know an estimate of the distribution of ages of the WIP inventory and the relative volume each age cohort is contributing to the net dispositions. To capture these features, a vintaging chain, in which the aggregated stock of Work-in-Progress is disaggregated into a series of sequential sub-stocks, can be applied to extend the modeling framework developed above.

This specific trade-off between a fully aggregate model and a more subdivided series of connected models is by no means unique to the context of this chapter and has been explored extensively in prior literature in System Dynamics. The original population sector of *Limits to Growth* (Meadows et al., 1972) was built after an analysis of three different levels of aggregation and the relative tradeoffs and benefits of disaggregation into more precise age cohorts (Bongaarts, 1973). On the extreme end, direct comparison of fully agent-based models with fully aggregate models have helped illustrate relative utility or interchangeability of these two approaches, notably in the context of stochastic environments (Rahmandad & Sterman, 2008). Specific interchangeability and tradeoffs of differing levels of aggregation and disaggregation with real datasets have been explored (Fallah-Fini et al., 2013) along with investigations of the effects of cohort disaggregation to the point of mathematically continuous cohorting (Eberlein & Thompson, 2013).

While much of the work previously done does directly apply to the modeling framework developed in this chapter (Eberlein & Thompson, 2013), the analysis below makes the influence of the age-value relationship on the degree of disaggregation explicit. In doing so, this section reinforces that the choice of disaggregation ultimately is a free parameter left to the modeler, which should be based on what is behaviorally and physically realistic for the system being measured. As stated in the earlier section of this chapter that introduced the MNL framework, when the unit of analysis is individual goods under development, the framework used here generates probabilities of discrete and mutually exclusive disposition outcomes. When aggregating the unit of analysis up to a continuous flow of goods, this probability resolves into expected numbers of goods in each disposition class. This section of the chapter serves to provide an example of how to still subdivide the development process while still maintaining the underlying assumptions and utility of applying the framework developed above.

Consider the vintaging chain below in Figure 1-10. Here WIP is split into *N* evenly spaced vintages, followed by a single end-of-life cohort. This end-of-life cohort follows the same structure as the aggregate framework described above. The end-of-life cohort is unique in so much as it consists of those goods whose age no longer affects the price that can be extracted from the marketplace. For the trapezoidal example used above and shown in expression (3), this would be goods with ages older than d. All cohorts except the end-of-life have the same disposition options including moving along the chain furthering development.

The age for each numbered age cohort increases regularly along the chain, starting with some initial age of production starts (typically assumed to have a numeric value of 0 units of time). Using this structure, the age of each cohort is known, including the final aggregate end-of-life life cohort on average.

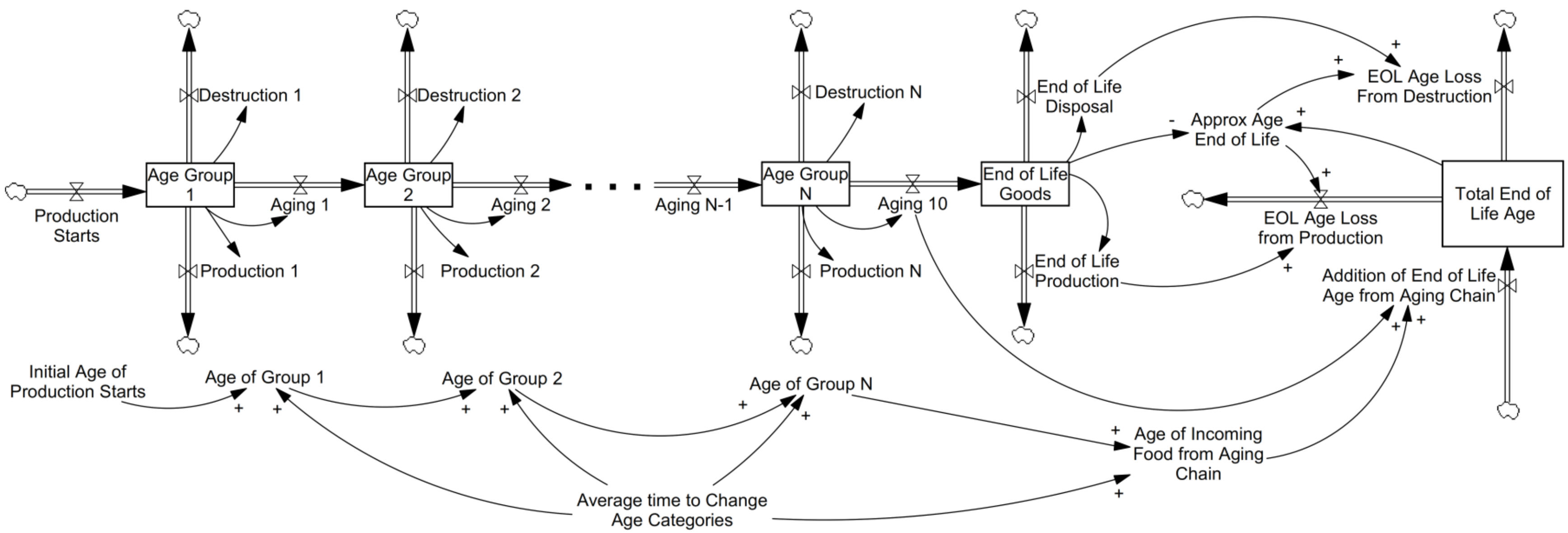


Figure 1-10. Core Vintaging Chain

Under the aggregate model framework, the value of the entire stock of WIP inventory is determined by the choice of age-value relationship. In the vintaging framework it can be applied to each cohort of ages. As an example, using the trapezoidal age-value relationship described in (3), in the vintaging extension we would expect that for low ages, almost all the inventories would be held to extract future value. At middle ages corresponding to the maximum full value, we would expect most of the inventory to be moved into a finished goods state. Finally, at high ages and especially in the terminal stock seen in Figure 1-10, we would expect most goods to be purposefully disposed.

The multinomial logistic (MNL) choice model, as illustrated in Figure 1-3, is applied to each part of this vintaging chain (though only one subpart is presented above). As presented here, the function that describes the Effect of WIP Age on Producer Price is the same across each vintage, though a trivial extension of this framework would allow for flexible value functions to be applied along the length of the chain.

Note that the structure developed here assumes that the average time to change cohorts categories is uniform. This formulation also assumes that economic valuation for the decisions around production and disposal are the same regardless of age. This is done here for clarity of exposition, and a more general model would apply the structure of Figure 1-3 separately, with possibly its own separate cost constructs, to each cohort sub-structure in Figure 1-10. However, the age-value relationship is being applied across the entire chain (e.g., with younger cohorts having a different economic value for being held onto and being allowed to mature versus later older cohorts).

As the MNL choice model ultimately depends on the relative size of valuations of each disposition, even having uniform costing for disposal and production but different realizations of the value of the expected value of production means that each cohort will experience different splits along each disposition route. Additionally, keeping this age-value relationship and associated costs consistent in its application along the chain allows for more direct comparison of the aggregate framework to the vintaging framework. As an example, consider the scenario utilizing the trapezoidal age-value function described in (3), and as parameterized in Table 1-1 below.

Table 1-1. Parameterization of Comparison of Vintaging and Aggregate Frameworks

Parameter Name	Description or Note	Value
Production Starts	Fixed and the same between the aggregate and vintaging framework	100 units/week
A	Minimum age of any value	2 weeks
B	Minimum age of full value	4 weeks
C	Maximum age of full value	6 weeks
D	Maximum age of any value	8 weeks
Number of Vintage Age Groups	Number of evenly aged vintaging groups, plus one end-of-life group	10 groups
Average Time to Change Age Categories	Chosen such that material older than time D moves into the end-of-life group	0.8 weeks
Processing Time	Average time for the producer, under either framework, takes to move material either into a finished goods state or to purposefully destroy it	3 weeks
Vintaging Framework Reference Price	Reference price by which the choices in the logistic model are evaluated at each vintaging age group	$1/unit
Spot Price	The exogenous (and here fixed) price that the market will pay for a full valued unit of production	$1/unit

Figure 1-11 shows how the disposition fractions determined by expression (2) evolve along this disaggregated development chain. Note that while purposeful disposal remains low, it is non-zero for each cohort, and during the ages that correspond to the maximum likely ability to extract the full spot price from the market (ages between 4 and 6 weeks as parameterized in Table 1-1), the likelihood of production (e.g. moving goods into a finished goods state) outpaces holding those goods in a WIP state.

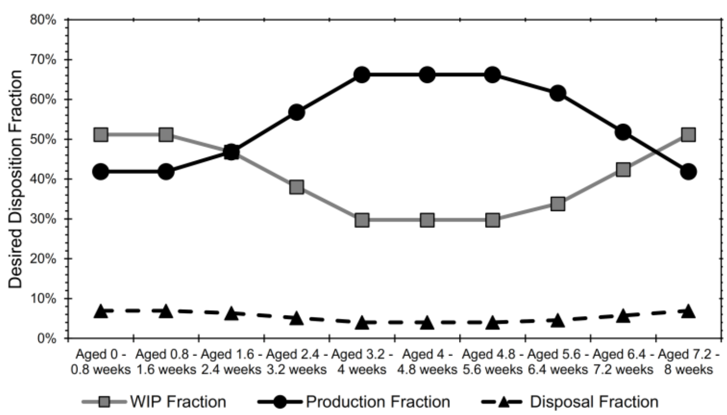


Figure 1-11. Example of the Evolution of Disposition Fractions in the Vintaging Framework

By comparing the structure for the vintaging framework in Figure 1-10 with that illustrated in Figure 1-3 for the aggregated framework, one key difference is that the concept of *Average Age of WIP Inventory* in the aggregated model is lost in the vintaging model. However, the two frameworks may be used largely interchangeably depending on the application and degree of utility of knowing the distribution of ages of goods in each disposition flow. This fits with the prior literature in this space discussed at the beginning of this section. Given the same number of production starts, it is possible to find an equivalent reference price that equates the net flow rates of each disposition route between these two models.

For example, using the parameters in Table 1-1 for a reference price of $1/unit for the vintaging framework, it is possible to achieve, in equilibrium, the same net flow of goods moved into finished goods inventory with the same average age and with the same net flow of purposeful destruction in the aggregate framework with a reference price of $ 1.30463/unit. It is important to note that this value depends on all the parameters enumerated above, including the number of age cohorts, their divisions, and the shape of the age-value relationship.

However, while this equivalence is maintained in steady state with fixed linear production starts, the two structures begin to diverge when the common production start rate is not fixed and constant for the specific reference price. Figure 1-12 illustrates this divergence, which focuses just on the net flows of purposeful destruction for illustrative purposes. The reference price that generates equivalent flows in in the case of fixed linear production starts generates qualitatively matching, but not exactly matching, flows in other cases of non-fixed production starts. It should be noted that for each of the inputs explored in Figure 1-12 there does exist a reference price that matches the two frameworks. For example, the reference price of $1.21432/unit matches the exponential growth case but causes divergences in behavior in an equilibrium state.

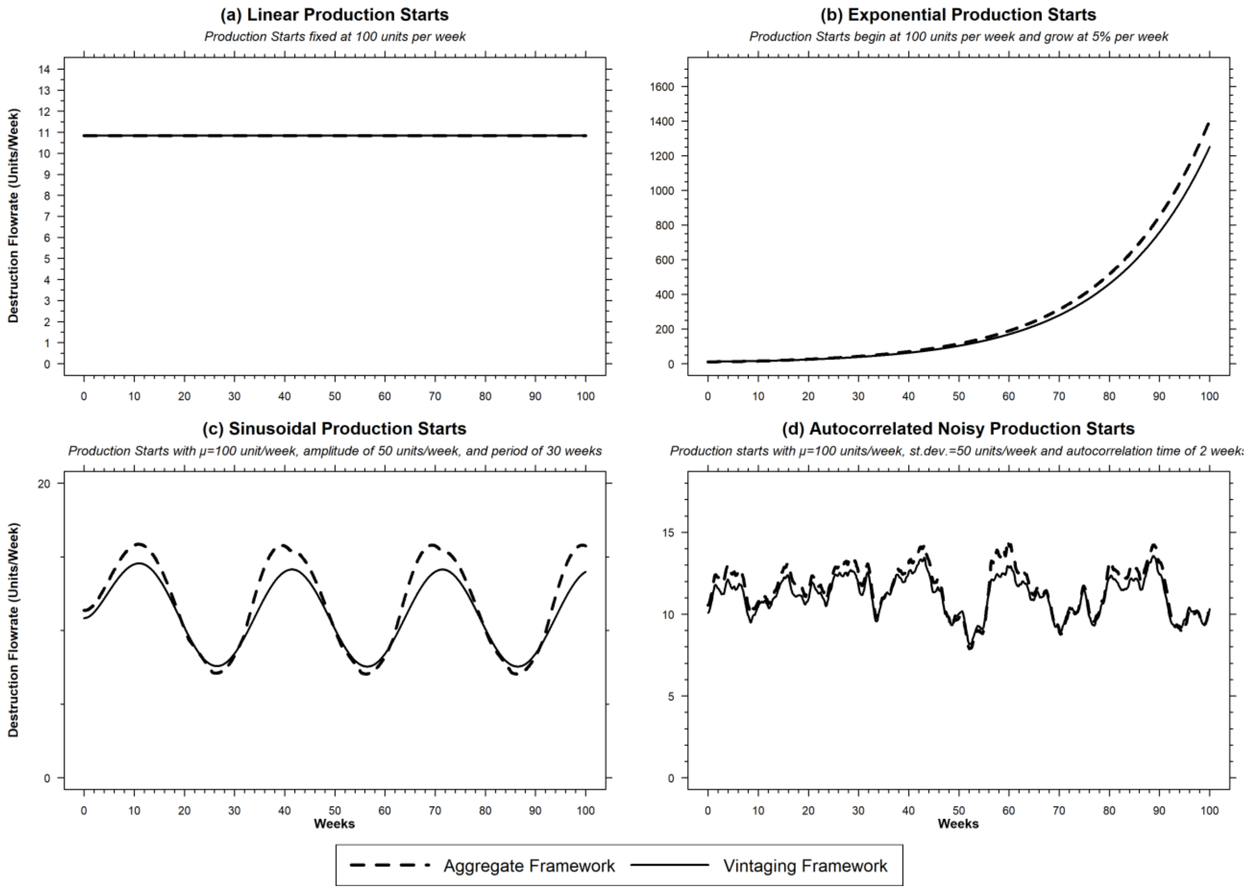


Figure 1-12. Comparison of Net Rates of Destruction between Aggregate and Vintaging Models

While this is only one example, it demonstrates that these two frameworks can be used interchangeably and equivalently when the limitations above are carefully considered. Specifically, it is most interchangeable when the general flow rates of interest are known, and the net flow into finished goods and destroyed goods is most important for downstream processes.

The purpose of this section has been to illustrate that the frameworks introduced in this chapter are not limited in application to a single aggregate stock of WIP but can also be applied to more disaggregated and detailed development cycles. Ultimately it is a choice of modeling freedom, and the choice should be based on both the reality of the system under investigation and the sources of value that come from the tradeoffs between a parsimonious model and more complete knowledge of the distribution of ages of production under a WIP state.

1.7 Discussion

This chapter presents a methodological contribution to the System Dynamics community that expands on how decision makers choose to move goods through development processes in a supply chain. By combining multinomial logistic (MNL) choice modeling with economic processes of price discovery, a more complete understanding of the behavioral features that determine the physical flow of goods through a supply chain can be explored, rather than simply assuming fixed or even multi-order delays in processing times.

As mentioned in the Literature Review, the underlying mathematical theory utilized by the MNL framework presented here already exists in the field of System Dynamics and has been used many times in consumer choice contexts such as modeling how the aggregate outcomes of individual choices resolve into population-level outcomes such as market share. However, the natural extension of these methods to supply chain models, which also contain decision makers weighing the relative value of different inventory disposition choices and which also has a rich history in the establishment of the field as far back as Forrester's original work for General Electric, has been largely ignored. Calls for more behaviorally grounded approaches within supply chain research have emerged recently and there have been recent calls to better integrate concepts, tools and frameworks between the OM and SD domains (Ghaffarzadegan & Larson, 2018; Größler et al., 2008; Morrison & Oliva, 2018; Sterman et al., 2015). This chapter in part answers those calls by leveraging the behaviorally grounded approach of System Dynamics with the tools and methods already available in the field to create richer models of supply chains that incorporate choice.

Using such a model supply chain, this chapter further illustrates how the modes of outcomes of otherwise identical systems can be materially different if one chooses *not* to incorporate the mechanisms developed herein. From a managerial perspective, the example also illustrated how specific policy interventions may be hidden or otherwise deemphasized when using a more traditional inventory management model. Models are ultimately simplifications of reality and tools that help understand how structural mechanisms and policy interventions combine to generate desired results or exacerbate already problematic behavior. In an environment in which producers have control over the inventory dispositions and development, and make those decisions based on the perceived greatest value of each option available to them, then abstracting away from this choice process both restricts the realism of the model and the richness of available policies.

The degree of aggregation or disaggregation of the stock of work-in-progress (WIP) was shown to be largely a free parameter left to the modeler and should be based on what is behaviorally and physically realistic for the system being measured. Both a fully aggregate and vintaging equivalent utilize choice modeling and economic concepts to determine the flow rates of goods through a supply chain and provide additional insight on the age (and thus corresponding value) of the goods in each disposition.

While the aggregate framework provides an average age across all disposition choices in a specific production activity, the vintaging framework allows for some additional insight on the distribution of ages, at the cost of significant additional modeling complexity and additional degrees of freedom in model design. For either framework, the distribution of ages (even if just the mean in the case of the aggregate framework) of material under production, or passed along any specific disposition path, emerges from the choices of the individuals, rather than being imposed exogenously.

While the context of economically motivated decision making of a producer in a supply chain was used here, specifically focusing on a stock of WIP goods, the methods and frameworks developed above could be applied to any member of a serial supply chain, each subject to their own process for determining the value of the choices available at each step in the supply chain. Additionally, while this chapter is steeped in the language of supply chains and production processes and was developed specifically to address a methodological shortcoming in traditional inventory and production management models, the frameworks used here are flexible enough to be applied to many other situations. So long as the problem under consideration allows for the decision maker to make some choice, and each possible choice can be enumerated, and the relative value of each disposition described, then the MNL choice framework can be applied to derive the probability of each of those dispositions.

As applied here, these probabilities of disposition outcomes were assumed to apply uniformly across many producers and thus are proportional to the expected total flows of each inventory disposition for continuous compartmental models like those used in the examples of this chapter. It is possible to apply this chapter to more disaggregated models of individuals operating in a supply chain and making mutually exclusive disposition decisions each period. Under such a model, the MNL choice framework would be utilized to form the probability of an individual choosing a specific disposition route, but only one such choice would be realized each period.

However, such an application at such a granular level of analysis should be taken with caution, as it begins to imply that the behavioral realism of this framework lies in directly describing the decision process of individuals, which it does not. Rather, the methods and mechanism introduced in this chapter have the most power and validity when describing aggregated expected outcomes of otherwise (to borrow the wording of McFadden 1974) qualitative, and 'lumpy' individual decision processes. To maintain the behavioral realism that System Dynamics as a field has emphasized since its beginning, these methods should be embedded within a larger model that incorporates additional delays and feedbacks found in real systems. Only then can this framework be used to capture the influences of the presence of *choice* in processes that were otherwise fixed and choiceless.

It is the hope of the author that this chapter contributes towards rigorous and behaviorally grounded modeling efforts in the future by providing a novel and useful framework for supply chain modelers, and others, to better incorporate choices that producers and managers are exposed to on a continual basis.

1.8 References to Chapter 1

Aloulou, F. (2018). The Application of Discrete Choice Models in Transport. In *Statistics—Growing Data Sets and Growing Demand for Statistics*. InTech. https://doi.org/10.5772/intechopen.74955

Anderson, E. G., & Fine, C. H. (1999). Business Cycles and Productivity in Capital Equipment Supply Chains. In S. Tayur, R. Ganeshan, & M. Magazine (Eds.), *Quantitative Models for Supply Chain Management* (pp. 381–415). Springer. https://doi.org/10.1007/978-1-4615-4949-9_13

Bauer, L. (2020). *The COVID-19 crisis has already left too many children hungry in America*. Brookings Institution Reports. https://www.brookings.edu/blog/up-front/2020/05/06/the-covid-19-crisis-has-already-left-too-many-children-hungry-in-america/

Bongaarts, J. P. (1973). A Review of the Population Sector in The Limits to Growth. *Studies in Family Planning*, *4*(12). https://doi.org/1965071

Chandukala, S. R., Kim, J., Otter, T., Rossi, P. E., & Allenby, G. M. (2007). Choice Models in Marketing: Economic Assumptions, Challenges and Trends. *Foundations and Trends® in Marketing*, *2*(2), 97–184. https://doi.org/10.1561/1700000008

Chen, H., Wu, O. Q., & Yao, D. D. (2009). On the Benefit of Inventory-Based Dynamic Pricing Strategies. *Production and Operations Management*, *19*(3), 249–260. https://doi.org/10.1111/j.1937-5956.2009.01099.x

Corkery, M., & Taffe-Bellany, D. (2020, April 18). The Food Chain's Weakest Link: Slaughterhouses. *New York Times*.

Corkery, M., Taffe-Bellany, D., & Kravitz, D. (2020, May 25). As Meatpacking Plants Reopen, Data About Worker Illness Remains Elusive. *New York Times*.

Croson, R., & Donohue, K. (2006). Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science*, *52*(3), 323–336. https://doi.org/10.1287/mnsc.1050.0436

Croson, R., Donohue, K., Katok, E., & Sterman, J. (2014). Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, *23*(2), 176–196. https://doi.org/10.1111/j.1937-5956.2012.01422.x

de Bok, M., de Jong, G., Tavasszy, L., van Meijeren, J., Davydenko, I., Benjamins, M., Groot, N., Miete, O., & van den Berg, M. (2018). A multimodal transport chain choice model for container transport. *Transportation Research Procedia*, *31*, 99–107. https://doi.org/10.1016/j.trpro.2018.09.049

Donohue, K., Katok, E., & Leider, S. (Eds.). (2018). *The Handbook of Behavioral Operations* (First). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119138341

Dorp, J. R. van, & Kotz, S. (2003). Generalized trapezoidal distributions. *Metrika*, *58*(1), 85–97. https://doi.org/10.1007/s001840200230

Eberlein, R. L., & Thompson, J. P. (2013). Precise modeling of aging populations. *System Dynamics Review*, *29*(2), 87–101. https://doi.org/10.1002/sdr.1497

Fallah-Fini, S., Rahmandad, H., Chen, H.-J., Xue, H., & Wang, Y. (2013). Connecting micro dynamics and population distributions in system dynamics models. *System Dynamics Review*, *29*(4), 197–215. https://doi.org/10.1002/sdr.1508

Forrester, J. W. (1961). *Industrial Dynamics*. Pegasus Communications.

Forrester, J. W. (1989). The History of System Dynamics. In *MIT System Dynamics Group Literature Collection*.

Ghaffarzadegan, N., & Larson, R. C. (2018). SD meets OR: a new synergy to address policy problems. *System Dynamics Review*, *34*(1–2), 327–353. https://doi.org/10.1002/sdr.1598

Gonçalves, P., Hines, J., & Sterman, J. (2005). The impact of endogenous demand on push-pull production systems. *System Dynamics Review*, *21*(3), 187–216. https://doi.org/10.1002/sdr.318

Graham, A. K. (1977). *Principles on the relationship between structure and behavior of dynamic systems*. Massachusetts Institute of Technology.

Greene, W. H. (2018). *Econometric analysis* (8th ed.). Pearson.

Größler, A., Thun, J. H., & Milling, P. M. (2008). System dynamics as a structural theory in operations management. *Production and Operations Management*, *17*(3), 373–384. https://doi.org/10.3401/poms.1080.0023

Hake, M., Dewey, A., Engelhard, E., Strayer, M., Dawes, S., Summerfelt, T., & Gundersen, C. (2021). *The Impact of the Coronavirus on Food Insecurity in 2020 & 2021* (Issue March).

Henserh, D. A., Rose, J. M., & Greene, W. H. (2005). *Applied choice analysis: A primer*. Cambridge University Press.

Hines, J. (2005). *Molecules of Structure: Building Blocks for System Dynamics Models Version 2.02*. Ventana Systems and LeapTec.

Hosseinichimeh, N., Rahmandad, H., Jalali, M. S., & Wittenborn, A. K. (2016). Estimating the parameters of system dynamics models using indirect inference. *System Dynamics Review*, *32*(2), 156–180. https://doi.org/10.1002/sdr.1558

Jalali, M. S., Rahmandad, H., & Ghoddusi, H. (2015). Using the Method of Simulated Moments for System Identification. In H. Rahmandad, R. Oliva, & N. D. Osgood (Eds.), *Analytical Methods for Dynamic Modelers*. The MIT Press. https://doi.org/10.7551/mitpress/9927.003.0007

Kampmann, C. E., & Oliva, R. (2009). Analytical Methods for Structural Dominance Analysis in System Dynamics. In R. A. Meyers (Ed.), *Encyclopedia of Complexity and Systems Science* (pp. 8948–8967). Springer New York. https://doi.org/10.1007/978-0-387-30440-3_535

Keith, D., Sterman, J., & Struben, J. (2017). Supply constraints and waitlists in new product diffusion. *System Dynamics Review*, *33*(3–4), 254–279. https://doi.org/10.1002/sdr.1588

Lee, H. L., Padmanabhan, V., & Whang, S. (1997). Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science*, *43*(4), 546–558. https://doi.org/10.1287/mnsc.43.4.546

Levi, R., Somya, S., & Zheng, Y. (2021). Artificial Shortage in Agricultural Supply Chains. *Manufacturing & Service Operations Management*, *24*(2), 746–765. https://doi.org/10.1287/msom.2021.1010

Marcus, R. (1991). Deterministic and stochastic logistic models for describing increase of plant diseases. *Crop Protection*, *10*(2), 155–159. https://doi.org/10.1016/0261-2194(91)90065-Y

Mass, N. J. (1975). *Economic cycles: An analysis of underlying causes*. Wright-Allen Press.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press.

Meadows, D. H., Meadows, D. L., Randers, J., & Behrens III, W. W. (1972). *The Limits to Growth; A Report for the Club of Rome's Project on the Predicament of Mankind*. Universe Books.

Morecroft, J. (1983a). A Systems Perspective On Material Requirements Planning. *Decision Sciences*, *14*(1), 1–18. https://doi.org/10.1111/j.1540-5915.1983.tb00165.x

Morecroft, J. (1983b). System dynamics: Portraying bounded rationality. *Omega*, *11*(2), 131–142. https://doi.org/10.1016/0305-0483(83)90002-6

Morecroft, J. (2015). *Strategic modelling and business dynamics: A feedback systems approach* (2nd ed.). John Wiley & Sons.

Morrison, J. B., & Oliva, R. (2018). Integration of Behavioral and Operational Elements Through System Dynamics. In *The Handbook of Behavioral Operations* (pp. 287–321). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119138341.ch8

Moxnes, E. (1990). Interfuel substitution in OECD-European electricity production. *System Dynamics Review*, *6*(1), 44–65. https://doi.org/10.1002/sdr.4260060104

Narayanan, A., & Moritz, B. B. (2015). Decision Making and Cognition in Multi-Echelon Supply Chains: An Experimental Study. *Production and Operations Management*, *24*(8), 1216–1234. https://doi.org/10.1111/poms.12343

Rahmandad, H., & Sibdari, S. (2012). Joint pricing and openness decisions in software markets with reinforcing loops. *System Dynamics Review*, *28*(3), 209–229. https://doi.org/10.1002/sdr.1473

Rahmandad, H., & Sterman, J. (2008). Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models. *Management Science*, *54*(5), 998–1014. https://doi.org/10.1287/mnsc.1070.0787

Revelt, D., & Train, K. (1998). Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics*, *80*(4), 647–657. https://doi.org/10.1162/003465398557735

Sterman, J. (1989a). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science*, *35*(3), 321–339. https://doi.org/10.1287/mnsc.35.3.321

Sterman, J. (1989b). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes*, *43*(3), 301–335. https://doi.org/10.1016/0749-5978(89)90041-1

Sterman, J. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World* (1st ed.). Irwin/McGraw-Hill.

Sterman, J., & Mosekilde, E. (1993). *Business cycles and long waves.* Alfred P. Sloan School of Management.

Sterman, J., Oliva, R., Linderman, K., & Bendoly, E. (2015). System dynamics perspectives and modeling opportunities for research in operations management. *Journal of Operations Management*, *39–40*(1), 1–5. https://doi.org/10.1016/j.jom.2015.07.001

Train, K. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.

Whelan, J., & Forrester, J. W. (1996). Economic Supply & Demand. *D-4388*, 7.

Yaffe-Bellany, D., & Corkery, M. (2020, April 11). Dumped Milk, Smashed Eggs, Plowed Vegetables: Food Waste of the Pandemic. *New York Times*.

Zhou, L. (2020, May 11). Coronavirus is exacerbating America's hunger crisis. *Vox*.

Chapter 2

Systemic Origins of Hunger Amidst Plenty During the Onset of the COVID-19 Pandemic in the United States

CHAPTER ABSTRACT

As the COVID-19 pandemic strained supply chains around the world, a striking, and seeming contradictory, outcome became apparent: Surges in rates of hunger occurred simultaneously with surges in food surplus and purposeful disposal. To better understand the systemic origins of this seeming imbalance, this chapter develops a dynamic model of a bifurcated food supply chain, from production through consumption, based on empirical observations drawn from the popular press and industry reports from the onset of the pandemic. The model considers both the physical flow of food from farmer to end consumer, along with the processes of price discovery and corresponding economic decisions made by the producer to ship, continue to grow, or purposefully dispose of food under cultivation. The resulting model combines three methods: compartmental differential equation modeling, inventory-based price discovery, and multinomial logistic choice modeling. The dynamics generated illustrate the origins of simultaneous paucity and plenty in a food supply chain and help suggest policy interventions. Policies that help reduce the upstream inventory stresses from a sudden decrease in downstream demand, or those that increase substitutability of end consumer goods across channels, are effective at mitigating the degree of purposeful food destruction and reducing economic stresses.

2.1 Introduction and Background

In the early months of the emergence the COVID-19 pandemic in the United States, there appeared to exist a fundamental paradox in both the production and demand for commodity foods. Farmers and wholesalers are dumping milk and tilling crops back into the ground (Bauer, 2020; Corkery et al., 2020; Corkery & Taffe-Bellany, 2020; Yaffe-Bellany & Corkery, 2020; Zhou, 2020), while food insecurity in the end consumer notably rose (Hake et al., 2021).

How can such states exist simultaneously? Some recent investigation of this has pointed towards fundamental structural features of modern supply chains (Durisin et al., 2020; Johnson, 2021; Lougee, 2020). This chapter contributes to this ongoing body of research by developing a model of a of a commodity supply chain which follows the production and disposition of food from an original producer, through a value-add supply chain, into two separate customer types. When exposed to shocks in demand patterns like those seen during the COVID-19 pandemic, this model helps illuminate the structural features that can lead to seemingly contradictory outcomes, such as a surge in hunger rates simultaneously with a surge in purposeful food destruction along the supply chain.

This food supply chain model is developed here purposely highly aggregated, focusing not on the interaction of various foods and even on the details of the coronavirus pandemic itself. Instead, it presents a generic supply chain for a generic commodity foodstuff, and it shows how the various changes in behavior seen during the pandemic (such as a drop in demand for bulk packaged foods, a reduction in the productivity and availability of staff at processing plants, and/or a general onset of a reduction in purchasing power in one or more classes of end consumers) can yield the imbalance described in the press of the time.

Differential equation compartmental models of supply chains have been a staple of operations management modeling since Jay Forrester first developed the tools and methods to describe industrial processes and business cycles in the 1950's and 1960's (Forrester, 1961). More recently, there has been a renewed interest in incorporating features of human decisions makers in supply chain models, and in doing so better define a concept of Behavioral Operations Management (Croson et al., 2013; Gino & Pisano, 2008; Hämäläinen et al., 2013).

2.2 Methods and Model Development

This chapter uses simulation modeling to develop a dynamic hypothesis of food imbalances in a supply chain subject to exogenous shocks like those brought on during the COVID-19 pandemic. The modeling methods utilized here incorporate the human decision-making features of this commodity food supply chain by allowing for the endogenous determination of dispositions of inventory and production in a supply chain via the application of multinomial logistic choice modeling to a compartmental differential equation model which tracks the relationship between the age, or development time, of goods under production and their corresponding market value.

Consider a group of farms producing a specific crop for later sale to a wholesaler, processor, or even directly to end consumers. Furthermore, assume these farms manage inventory both in the field and in storage based on the economic implications of the possible dispositions of their production, rather than acting as strict price-takers who perfectly fulfill all demand signals. Prices are assumed to be based on a commodity product, which implies the use of inventory-sensitive spot pricing (Chen et al., 2009; Sterman, 2000; Whelan & Forrester, 1996).

Generally described the model below explores a bifurcated food supply chain consisting of the following entities:

- A farmer, who is responsible for making decisions about how much to plant each time and how to manage his or her harvest.
- A wholesaler firm, which receives raw and unprocessed foodstuff from the farmer, and does some minimum value-added work to the food.
- Two different packaging processors
    - A CPG (consumer packaged goods) processor that received goods from the wholesaler and does extensive value-added rework to the food, packaging it in smaller consumer friendly forms for sale to the end consumer at some outlet like a grocery store.
    - A Bulk processor that receives goods from the wholesaler and does minor repacking of the food for sale directly to larger consumers like restaurants, governments, or schools.
- The end consumers, which include demand for both CPG and Bulk packaged food.

A general visual representation of these entities, and the physical flow of food, is shown in Figure 2-1.

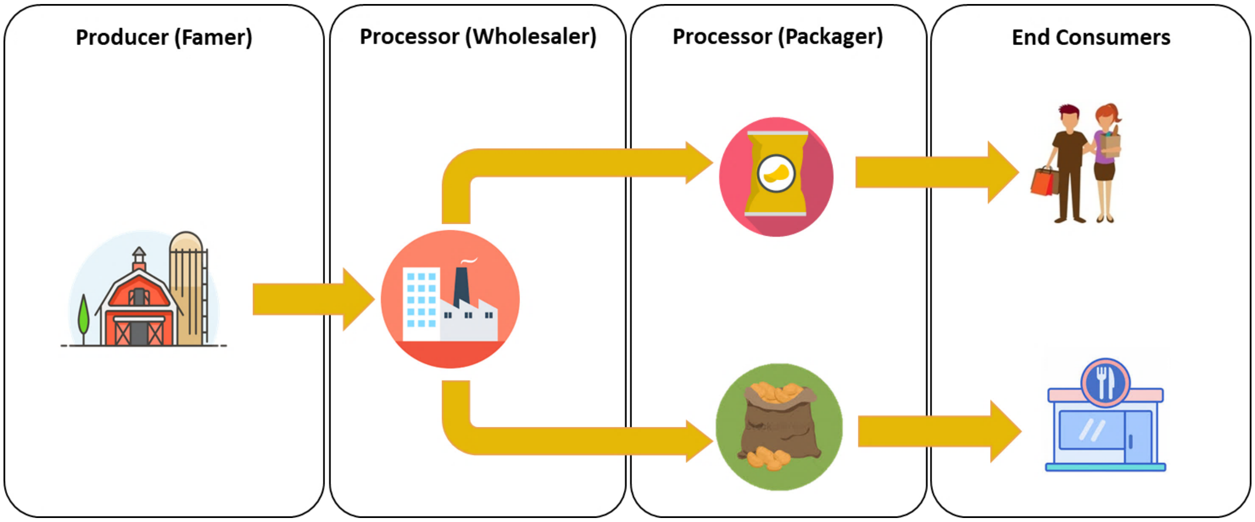


Figure 2-1. Visual Representation of the example food chain

The model developed here is purposefully bifurcated into two separate non-overlapping consumer types to both distinguish it from a standard linear and serial supply chain model, and to better capture the real world structural features that were at play during the onset of the COVID-19 pandemic. As discussed more in the sections below, the pandemic induced different shocks in demand to two different classes of consumers (CPG and Bulk), but which ultimately shared an upstream supplier.

The model was developed and run in the Vensim version 8.2 software package and is fully documented in Appendix B.

2.2.1 Physical Flows and Dispositions of Goods

The farmer[2] must decide about how much to plant and how to manage his or her storage capacity. The amount of food under cultivation at any given period is a function of the amount of production starts, less losses from harvesting, and from destruction (or other losses, either natural or purposeful). Similarly, the amount of food that has been harvested by the farmer and

[2] More generally, this could refer to any producer of a raw form of the food, which could also include someone raising livestock, or making milk as other examples. However, for this model the producer is conceptualized as someone growing a commodity vegetable like onions or potatoes.

is ready for shipment to the wholesaler is a function of the corresponding inflows and outflows into the farm storage.

As discussed in more detail in the prior Chapter, a unit of production started does not guarantee that it will be completed and represents a real opportunity cost during development. The dumped milk and destroyed crops referenced in the Introduction to this Chapter are examples of producers weighing the relative value expected from continuing development or holding onto finished goods against the value (less costs of disposal) that could be achieved by freeing up storage or production capacity. Under normal circumstances, one would expect this loss to be minimal across an entire supply chain. However, it cannot be ignored outright and ultimately motivates this modeling work.

These physical flows are visualized in the stock-and-flow diagram (Morecroft, 2015; Sterman, 2000) shown in Figure 2-2[3].

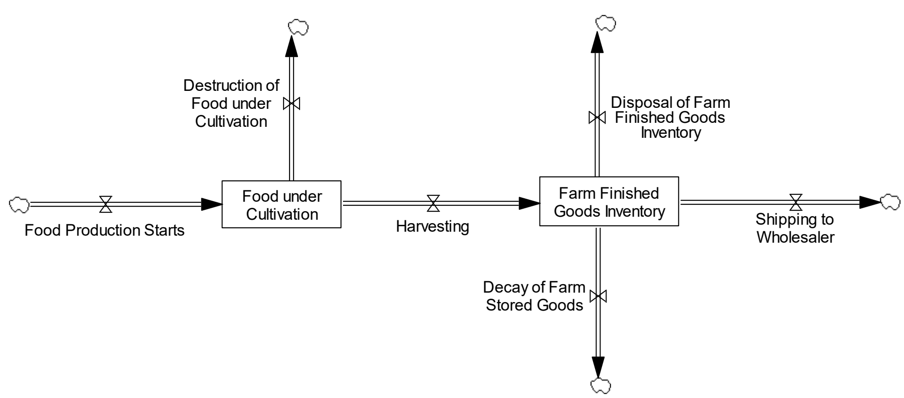


Figure 2-2. Farm – Physical Flows

Food from the farmer is shipped to a wholesaler, who performs some basic valued-added service to the goods before passing them along to entities that package the goods in either a CPG format (e.g., small packages like for grocery stores) or in a bulk format packaging

[3] In these diagrams, the boxes represent accumulations, which cannot be directly changed but rather are affected by the net of the inflows and outflows acting on that stock (Sterman, 2000). Here, these flows are represented by pipes and valves. For example, in Figure 2-2, the stock of 'Food under Cultivation' changes as a function of the sum of the inflows and outflows, or:

$$\frac{d}{dt}(Food\ Under\ Cultivation) = Production\ Starts - Harvesting - Losses$$

(e.g., with minimal packaging like for restaurants, schools, or governments). The wholesaler and these repackagers could, in some food supply chains, be the same. Here in this model they are separated out to specifically capture the dynamics described in the articles focusing on the processing restrictions (Corkery et al., 2020; Corkery & Taffe-Bellany, 2020).

Like the farm, the wholesaler is assumed to have two stocks of goods under management, those received from the farm, and those that have received some value-added work and are awaiting shipment to downstream customers. This is illustrated in Figure 2-3.

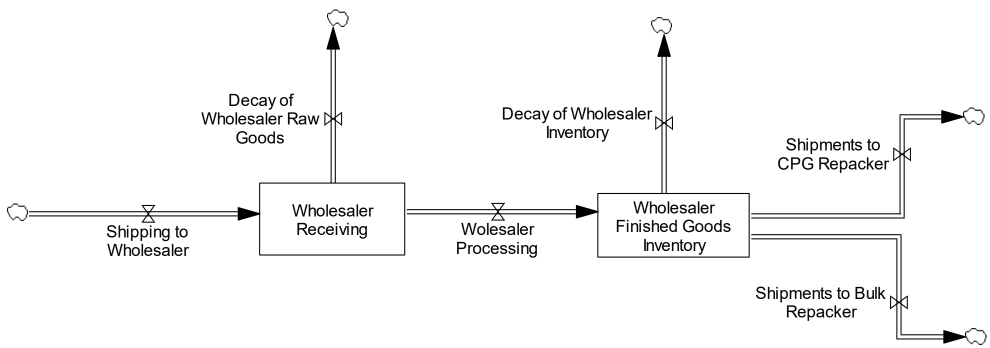


Figure 2-3. Wholesaler - Physical Flows

The decision the warehouse must make about filling orders to either the CPG focused processor or the Bulk packaging focused processor may be based on many considerations, from specific contracts with service rate guarantees and penalties to personal preferences for dealing with one client over another developed over years of business operations. For parsimony, this model abstracts away from those details and instead allows the wholesaler to simply ship what is demanded as it is demanded from either channel.

The only real heuristic that comes into play here is the choices that must be made when demand outstrips supply. To whom should the finished goods from the warehouse go? For simplicity, the model utilizes a directly proportional rule, splitting the total outflow available relative to the total demand in each channel. While more elaborate decision rules may be possible here it does not add meaningfully to the resulting dynamics.

As discussed above, the repackagers and the wholesalers could be considered one in the same under some circumstances and supply chain configurations. Therefore, many of the decision rules found in the earlier sectors, including managing physical space and adjusting labor practices, are abstracted away in this sector of the model for simplicity.

In essence, the repackagers convert the finished goods from the wholesalers into two distinct and non-substitutable forms. The CPG packaging process focuses on high value-added processing, separating out the food into smaller packages for sale to end consumers in a retail environment like a grocery store. Conversely, the Bulk packaging process does minimal processing to the goods received from the wholesaler, perhaps adding minimal packaging for sale to larger clients like restaurants, schools, or governments. The inventory in each channel can be represented as in Figure 2-4. The key difference between these channels is that they are exposed to two different demand streams.

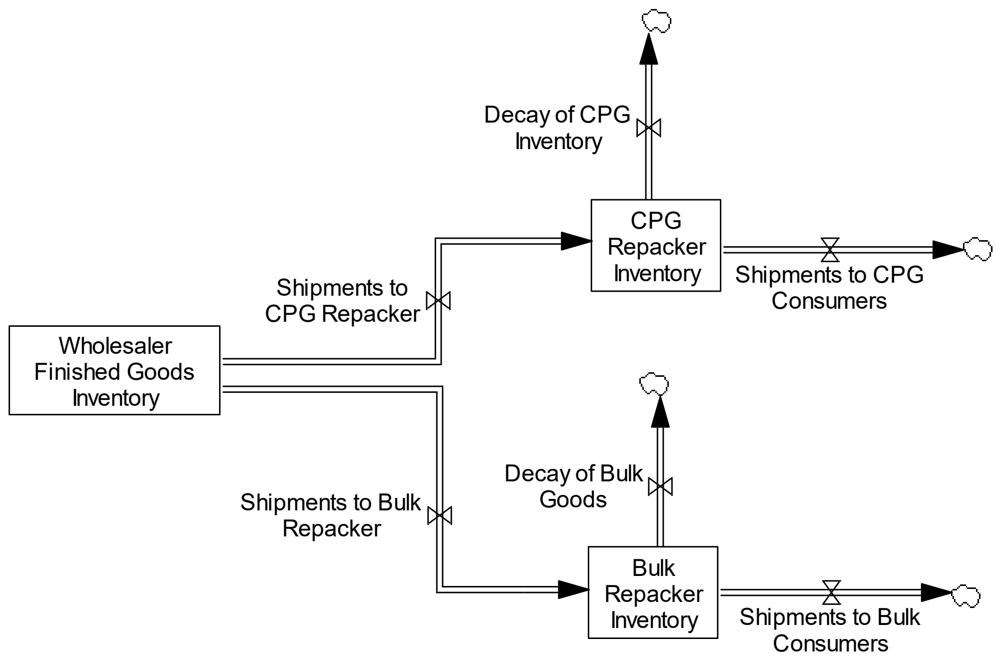


Figure 2-4. Repackager - Physical Flows

2.2.2 Applying a Dynamic Valuation of ‘Work in Progress’ to Commodity Goods

As stated in the prior chapter, the time that a unit of production is under development may have a meaningful economic impact on the final value of the product at hand, and thus the value of either holding or shipping inventory may change with a concept of time under development or age of the work-in-progress. Here, the influence of this decision is slightly more straightforward, as ‘work-in-progress’ maps directly towards maturation of the commodity good. Too short of a maturation time, and the farmer cannot expect to be able to sell the crop or good at full market price, or at all. Too long of a maturation time and the farmer has either locked up resources by holding inventory too long or may risk crops decaying away.

The co-flow structure used to keep track of *Average Age WIP Inventory* in the prior chapter can be directly mapped to this model. Figure 2-5 shows this mapping, forgoing the multinomial logistic choice framework pieces for concise presentation.

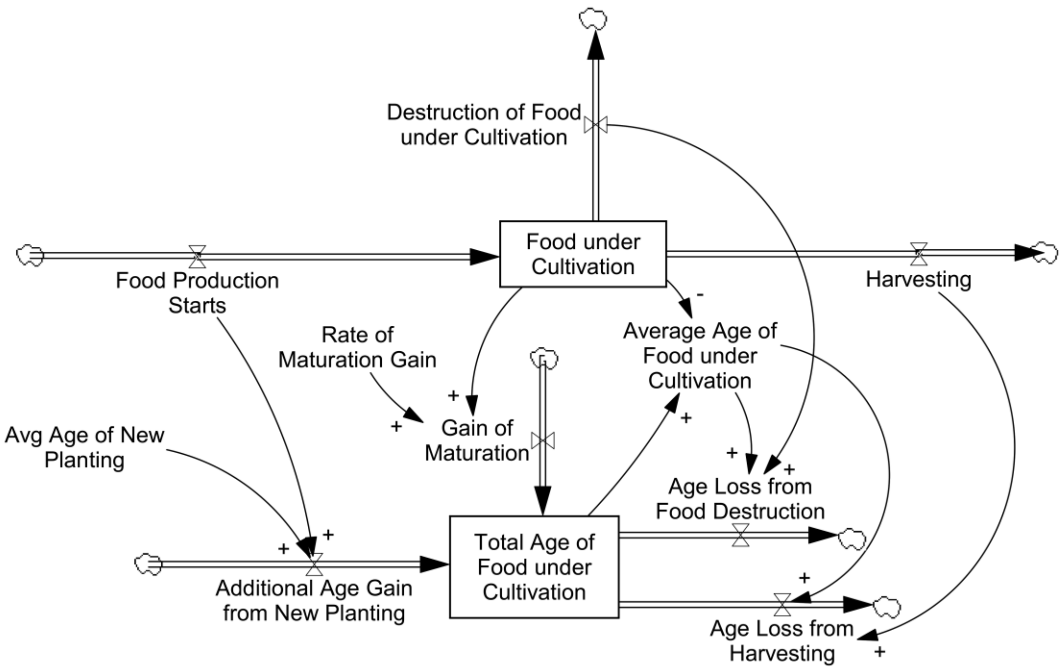


Figure 2-5. Keeping Track of the Average Age of Food Under Cultivation

The multinomial logistic (MNL) framework discussed in the prior chapter is an excellent modeling choice here. For this system, the farm has three choices to make with respect to crops that are maturing in the field: 1) Harvest and move into storage (for immediate or later selling to the wholesaler), 2) Keep in the ground to continue to mature (or decay), or 3) Dig up and destroy.

The MNL framework applies here because from the point of view of a single farmer that is considering a single unit of production, each of these dispositions are non-overlapping (a farm cannot simultaneously destroy, harvest, and continue to cultivate a single unit of food). However, for a larger model of a *system* of these decisions, the MNL framework can be used to represent the probability of a farmer choosing any of the above three options. In this example, and across many decisions, this probability becomes the proportion of the total work-in-progress inventory that is delegated to each of the possible disposition routes.

Following this framework, the stock of *Food Under Cultivation* (e.g., work-in-progress inventory) held by the farm is divided into three sub-groups, based on the relative economic value of each disposition. As a reminder from the prior chapter, the probability of any one unit of production being in a given sub-group, which in expectation is proportional to the total units that the producer will desire to be in any of these group, is given by the expression below:

$$P(X_i) = \frac{e^{\beta \pi_i}}{\sum_{I=1}^{N} e^{\beta \pi_I}} \tag{4}$$

As discussed in more detail in the prior chapter, the value on each disposition, π_i, can be taken to be the expected profitability of that disposition. Furthermore, the value that the producer places on profitability, β, can be taken as an inverse of a reference price for simplicity. Here, the reference price is chosen to be the long-run baseline stable price that the farmer typically receives for finished goods.

While the model development immediately above has focused on the valuation and inventory disposition decisions of the *Food Under Cultivation*, it can be readily applied as well to *Farm Finished Goods Inventory* in storage as well, creating a series of these MNL-based sub-models. Again, the farm has three choices: 1) Make inventory available for the wholesaler, 2) Keep finished goods in storage, or 3) Destroy goods. As with the *Food Under Cultivation* inventory stock, a multinomial logistic function is used, normalized with β values all chosen to be the inverse of a farm reference price.

For commodity food products, there is an ideal window of maturation time, or work-in-progress age, at which the food can receive its full economic value. Outside of this window, the producer or farmer can expect less than full value or even no value at all. To capture this price-value dynamic, a variety of analytic or empirical relationships could be explored (based on the reality of the exact commodity crop being developed) but for simplicity, consider the trapezoidal relationship between crop value and age (or maturation time) introduced in the prior chapter and restated in the expression below.

$$f\,(t) = \begin{cases} 0 & t \leq a \\ \left(\frac{1}{b-a}\right)t - \left(\frac{a}{b-a}\right) & a < t \leq b \\ 1 & b < t \leq c \\ \left(\frac{1}{c-d}\right)t - \left(\frac{d}{c-d}\right) & c < t \leq d \\ 0 & t > d \end{cases} \tag{5}$$

2.2.3 Production Starts and Capacity Management

The decision to plant by the farm considers the incremental profitability of an additional unit of production utilizing the variable costs of production and expects losses and gains from each of the possible inventory dispositions discussed above.

Production capacity utilization for the farm is a function of expected gross margin as a representation of expected profitability. Furthermore, utilization of existing capacity is unlikely to be at 100 percent when averaged across all pieces of owned capacity unless at very high levels of expected profitability. The exact shape of this relationship will vary by industry, and even by individual producer, or individual piece of owned unit of production capacity.

To qualitatively capture this behavior, consider the normal cumulative distribution function of utilizations versus expected profitability of a collection of different land (capacity) at different utilization depending on local factors. Figure 2-6 illustrates this curve under the assumption that under reference profitability, 50% of total production capacity is being utilized. The symmetric nature of this relationship is an assumption for this model and the real utilization relationship may be asymmetric.

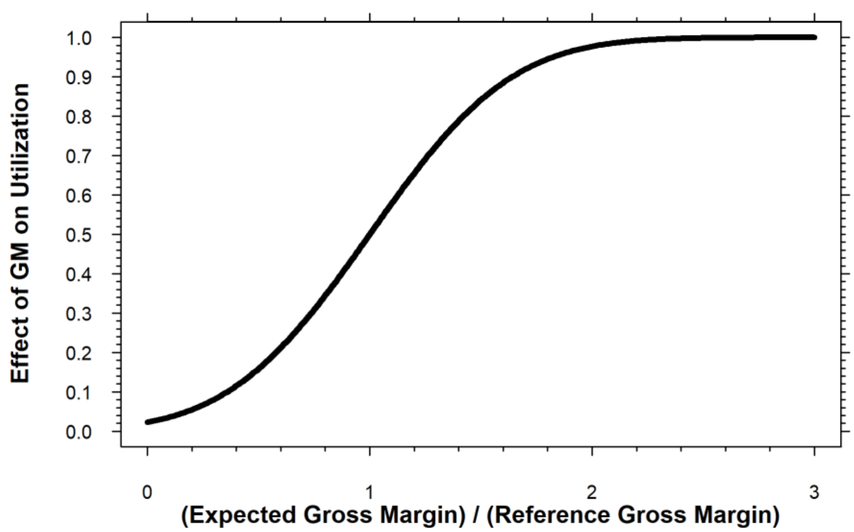


Figure 2-6. Farm Capacity Utilization versus Expected Gross Margin

2.2.4 Linking the Sectors and Defining the Market

As discussed in the prior chapter, there may be multiple ways to construct the interplay of supply and demand that determines the spot price at each interface point between the entities in this food supply chain. As this chapter explicitly considers commodity goods, and the use of inventory-sensitive spot pricing introduced in that prior chapter is still applicable here (again refer to Chen et al., 2009; Sterman, 2000; Whelan & Forrester, 1996 for examples of this price formation mechanism).

The multinomial logistic choice framework described in the prior chapter is utilized by the producer (here the farmer) in this supply chain model, with the actual spot price discounted based on the age-value relationship seen in expression (5), with the value of food under cultivation starting at zero percent of the full market spot price, then increasing linearly until it reaches an ideal maturation and can extract the full market spot price. After a time at full value, for higher of food under cultivation the value declines linearly back to zero percent of the market spot price.

One of the key features of the commodity pricing model utilized here is the effect of inventory coverage on pricing. In net, a model will capture the downward sloping relationship between additional inventory (beyond a set inventory coverage goal) and the price offered by the firm holding that inventory.

$$\begin{aligned} &Effect\ of\ [Entity]\ Inentory\ Coverage\ on\ [Entity]Price \\ &\qquad = [Entity]Inventory\ Ratio^{-Sensitivty} \end{aligned} \tag{6}$$

The *sensitivity* is a parameter that determines how much the price will raise or lower given a change in inventory coverage. As formulated here, *sensitivity* is assumed to be a positive value, with higher values corresponding to increasingly concave response functions.

Similarly, this model assumes that there is a long-run expected cost or margin that each customer expects, anchored to the long-run reference prices in the system. In this manner, prices in this model will return to this reference price over time. This is a simplifying assumption that provides an additional balancing mechanism that anchors the otherwise floating prices in the system.

$$\begin{aligned} &Effect\ of\ [Entity]\ Costs\ on\ [Entity]Price \\ &\qquad = \left(\frac{[Entity]\ Customer\ Expected\ Price}{[Entity]Reference\ Price}\right)^{-Sensitivity} \end{aligned} \tag{7}$$

The effect on demand due to expected gross margin does have some element of sensitivity to cost built in from the definition of gross margin. However, when considering an *expected* gross margin, the influence on demand is based on a smoothed view of previous prices (both costs for goods bought and the prices at which they were later sold). To affect demand based on the *instantaneous* spot price experienced by each entity, consider a linearly decreasing relationship that captures decreasing demand with increasing prices, with the slope of that relationship affected by some elasticity of demand. The functional form of this expression is seen below:

$$Effect\ of\ Price\ on\ Demand = MIN\left(Maximum\ Multiplier, MAX\left(0, 1 + Demand\ Curve\ Slope * \frac{Price - Reference\ Price}{Reference\ Demand}\right)\right) \tag{8}$$

Where:

$$Demand\ Curve\ Slope = \frac{-Reference\ Demand * Reference\ Elasticity}{Reference\ Price} \tag{9}$$

In the above, the spot price is used to determine the effect on instantaneous demand. This effect is purposefully designed to be immediate, in contrast to the effect from *expected gross margin* which is based on a smoothed concept of both prices and costs.

Combined, the relationships described above form an economic market for this commodity food. The decision that each entity in this food supply chain that affects the upstream entities is how much to order. As discussed above, this decision is based on the interplay of on-hand supply and expected future profitability.

Figure 1-4 in the prior Chapter captures much of the interactions described here for a generic N-entity supply chain. Figure 2-7 recasts this generic structure in the terminology used in this commodity supply chain and adds the effect of costs on price. It also distinguishes the farm (which controls its own supply) versus downstream entities. Note that Figure 2-7 is not a complete representation of this system, and instead focuses in on the key feedback loops that drive this ordering decision, abstracting away from other model structure such as inventory decay, purposeful destruction, and capacity utilization.

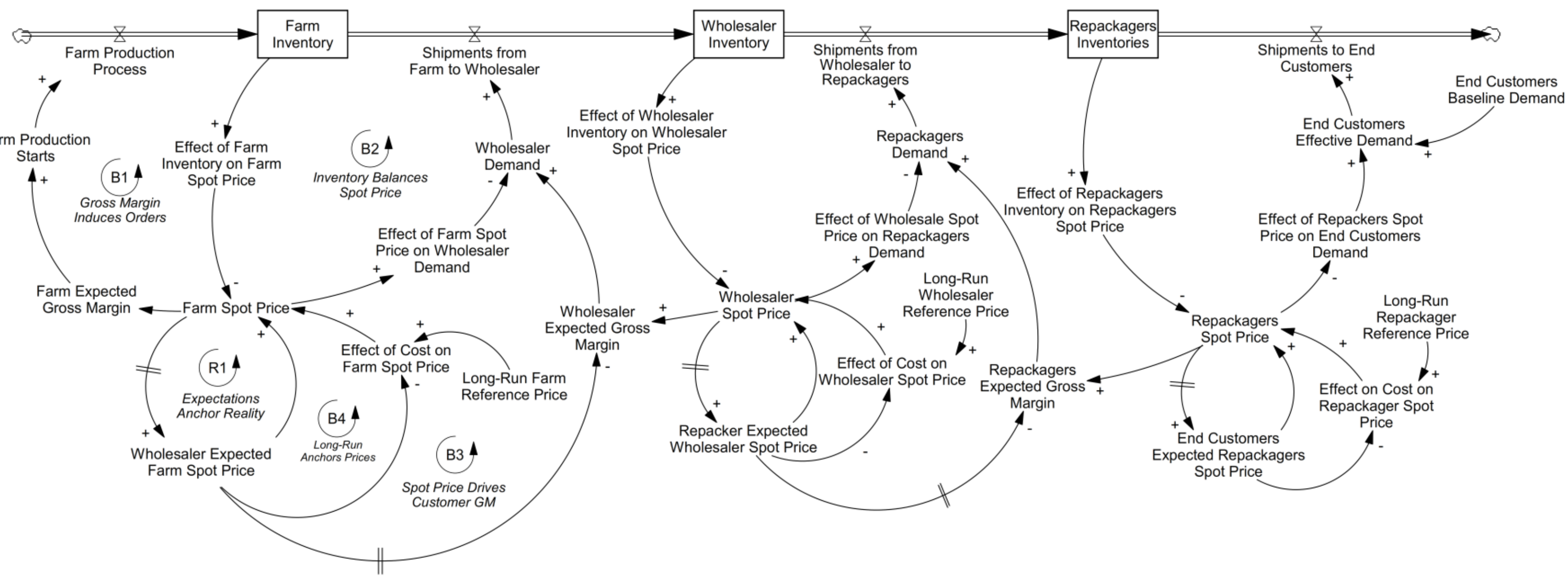


Figure 2-7. Simplified CLD Showing Key Drivers of Ordering Decisions

Each ordering decision is driven by the interplay of five feedback loops, four balancing and one reinforcing. The reinforcing mechanism is of special interest as it is a floating anchor, in which the spot prices experienced instantaneously in the system are anchored to the history of observed spot prices. In other words, price inflation or deflation is based on the existing expectations of the consumer in that link in the chain (Lee et al., 2022). The balancing mechanism described above that anchors spot pricing to long run expected prices directly offsets this reinforcing mechanism.

This model assumes that each entity, aside from the end consumer, is at least somewhat strategic, and able to plan ahead. Therefore, demand is based on a combination of the current spot price and the expected gross margin from the expected price. In other words, the customer in each link in this chain is basing his or her demand on expected future earnings but will modify demand given the reality of today. For both of these effects, higher prices drive down demand, which in turn drives up inventory, placing downwards pressure on prices and balancing the initial effect. The figures here only label these feedback loops for the first entity tin the supply chain (here the farm), but they are present at each link.

The supply chain considered here has a distinct bifurcation before the end consumer, segregating food stuffs into two distinct categories, namely Consumer Packaged Goods (CPG) and Bulk goods. The generic price formation and ordering decision structure for a linear, serial supply chain seen in Figure 2-7 is directly applied to the bifurcated supply chain used here in Figure 2-8.

While the difference between the linear serial supply chain and the bifurcated version may seem trivial, especially since the decisions being made by the Wholesaler are based on the aggregated demand induced by both channels in the bifurcated supply chain, it has the main structural effect of segregating the flows of goods into two non-substitutable paths. This occurs right before the end consumer, with one set of goods going towards the CPG channel and the other going towards the Bulk channel. This distinction in food categorization is important for the analyses below, which consider the effect of a major shock, disproportionally affecting one side of this bifurcated system.

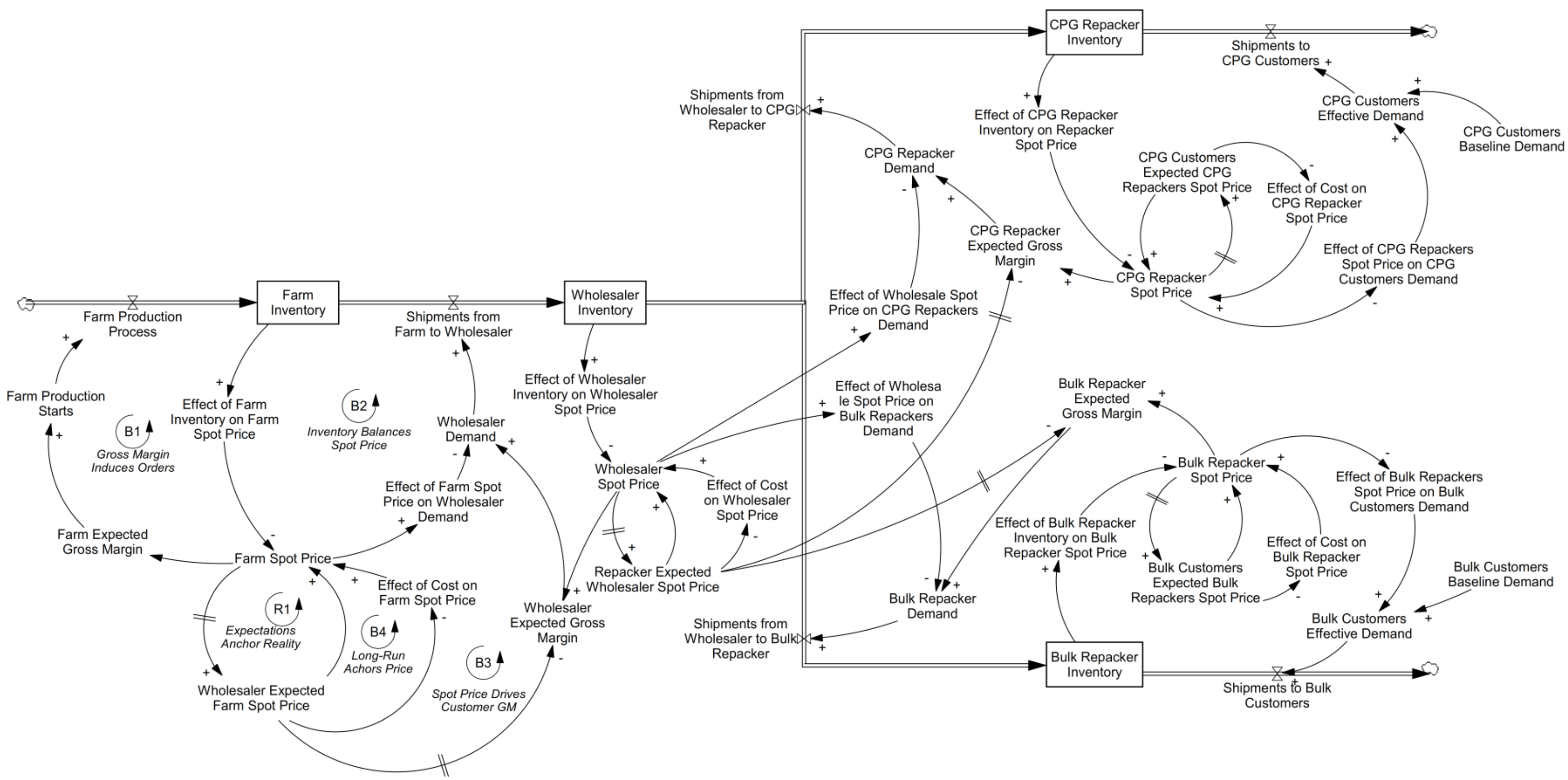


Figure 2-8. Simplified CLD of Drivers of Ordering Decisions in a Bifurcated Supply Chain

2.3 A Dynamic Hypothesis of Food Imbalances

During the onset of the novel Coronavirus in the US in early March of 2020, purchasers of bulk goods essentially stopped ordering from their wholesaler suppliers. This was driven in part by stay at home orders which reduced consumer demand originating at restaurants, and other bulk purchasers such as schools. Simultaneously, layoffs and financial hardship hit end consumers while SNAP benefits and other government programs were slow to respond and fill the reduction in consumer buying power (Bauer, 2020; Zhou, 2020).

Parts of this effect can be seen in a dataset tracking seated diners at restaurants utilizing the OpenTable reservation system as illustrated in Figure 2-9, which shows a near total evaporation in year-over-year seated diners in 2020, lasting for approximately 10 weeks, with a recovery to only 50% of the prior year's diners through the summer (OpenTable, 2021).

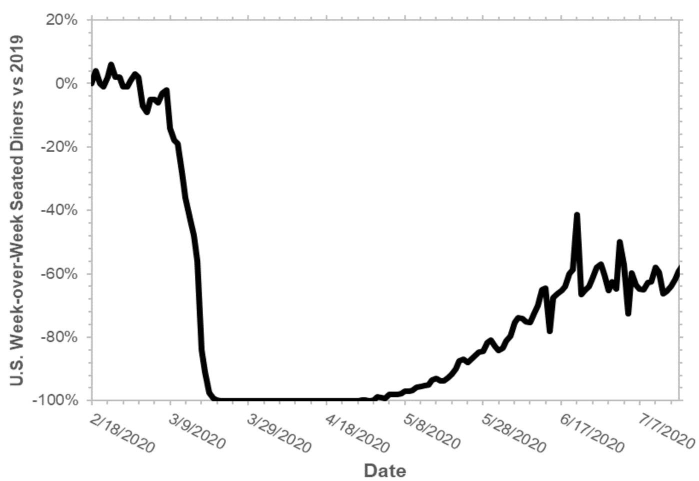


Figure 2-9. Drop in U.S. Year-Over-Year Demand for Seated Restaurant Diners (OpenTable, 2021)

While the above captures seated diners utilizing this one reservation platform, similar trends can be seen across industries in which basic foodstuffs are consumed. Consider the survey data shown in Figure 2-10, which is drawn from the April 2020 Mintel report on U.S. consumer grocery trends (Owen, 2020) and reflects wider consumption patterns beyond just those seen in Figure 2-9. In this survey data, a drop in business at restaurants occurs in an increase in consumption at end consumer outlets like grocery stores from consumers stocking up.

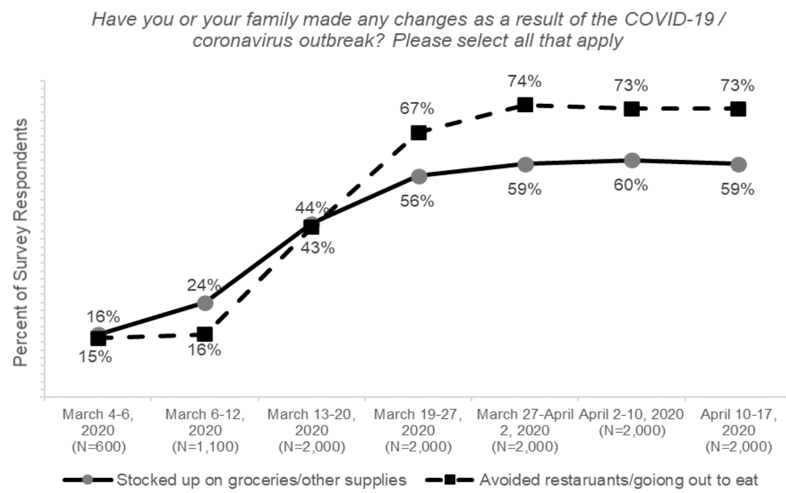


Figure 2-10. Changes in U.S. Food Purchase Behavior During the Onset of the Coronavirus Pandemic (Owen, 2020)

The model developed above was parameterized utilizing balanced demand for CPG and Bulk packaged goods, at a reference value of 100 food tons per week each, and conceptualized around a fast-growth commodity food crop such as leaf lettuce or green onions, with the parameter values of the trapezoidal function shown in expression (5) set to *a = 4 weeks, b = 10 weeks, c = 12 weeks,* and *d = 20 weeks* (*Growing Guides*, 2021). The cost structure of the simulation and reference profitability of each entity were chosen to be directionally consistent with that seen in other commodity food supply chains, with the tightest margins being experienced by the farmer or producer (Galen & Hoste, 2016).

As designed, this model allows for simultaneous shocks to different parts of the supply chains. However, the outcome of such simultaneous shocks, specifically a bottle neck in production processing combined with reduced CPG consumer spending power would trivially result in the outcomes described above, with reduced purchasing power and a bottle neck behind production.

Rather than show this expected outcome, this work considers a minimal scenario that can generate similar outcomes. Specifically, consider a scenario exploring *just* when demand for bulk packaged foods evaporates as restaurants and schools close. In this scenario, the demand for CPG and Bulk goods is initially balanced and equal. At week 10 (corresponding to approximately the middle of March in the simulation), the underlying purchasing power of this group of consumers drops to 50% of its former value and remains there for 20 weeks before returning to its prior value. In this specific scenario, consumer purchasing power for CPG packaged foods remains unchanged, as illustrated in Figure 2-11.

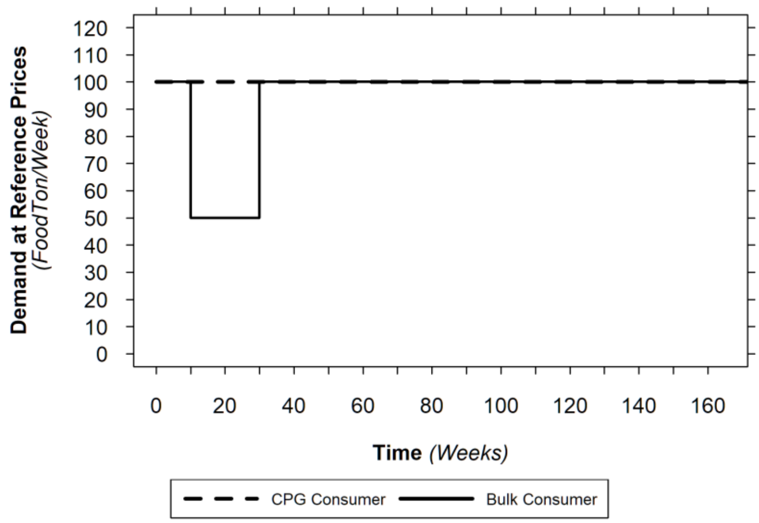


Figure 2-11. 50% Drop in Bulk Purchasing Power for 20 Weeks

This sudden reduction in total system demand by 25 percent causes a small but nevertheless present short run buildup of finished goods inventories throughout the supply chain even as initial production starts are rapidly cut as seen in Figure 2-12. However, the consequences of the bifurcated structure of this supply chain became clearer, resulting in a longer run buildup of goods just prior to the bifurcation at the wholesaler.

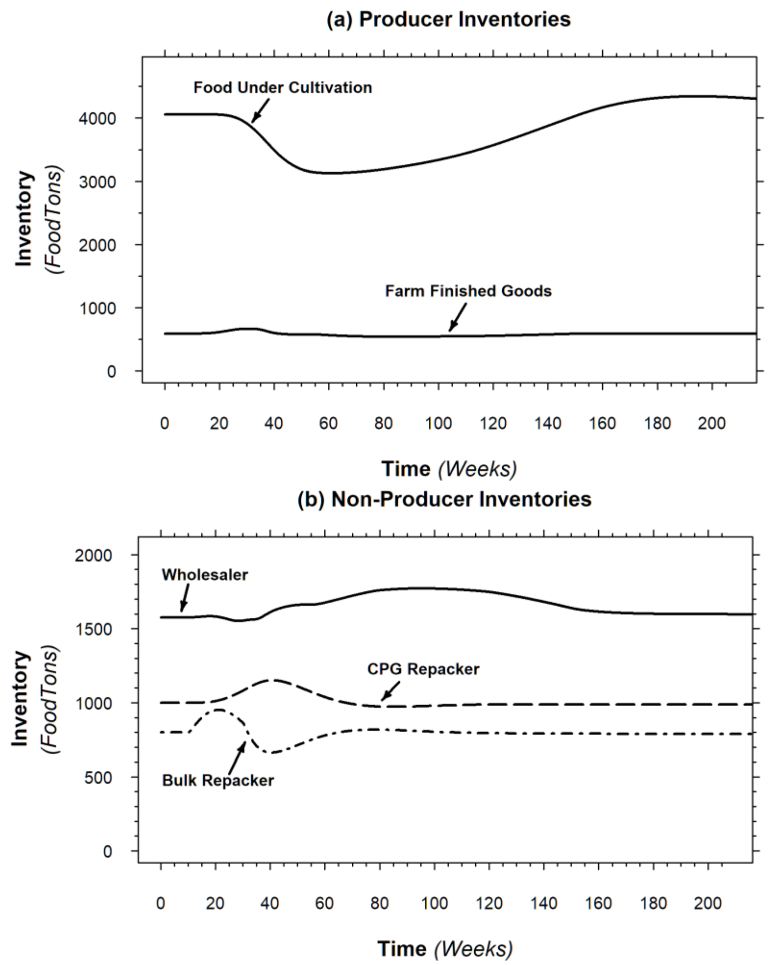


Figure 2-12. Supply and Inventories - 50% Drop in Bulk Purchasing Power for 20 Weeks

CPG items continue to be produced and consumed, but not at a rate to offset the loss from the diminished demand in bulk packaged foods. Spot prices for bulk repackaged goods drop with the collapse in demand. As demand from the point of view of the wholesaler is pooled across both channels, the general price the wholesaler can charge collapses due to the higher relative inventory position. This reduction in prices induces some short term increased demand in the CPG channel. This has the short-term effect of shifting some goods into the CPG channel that would have otherwise been routed to the bulk channel.

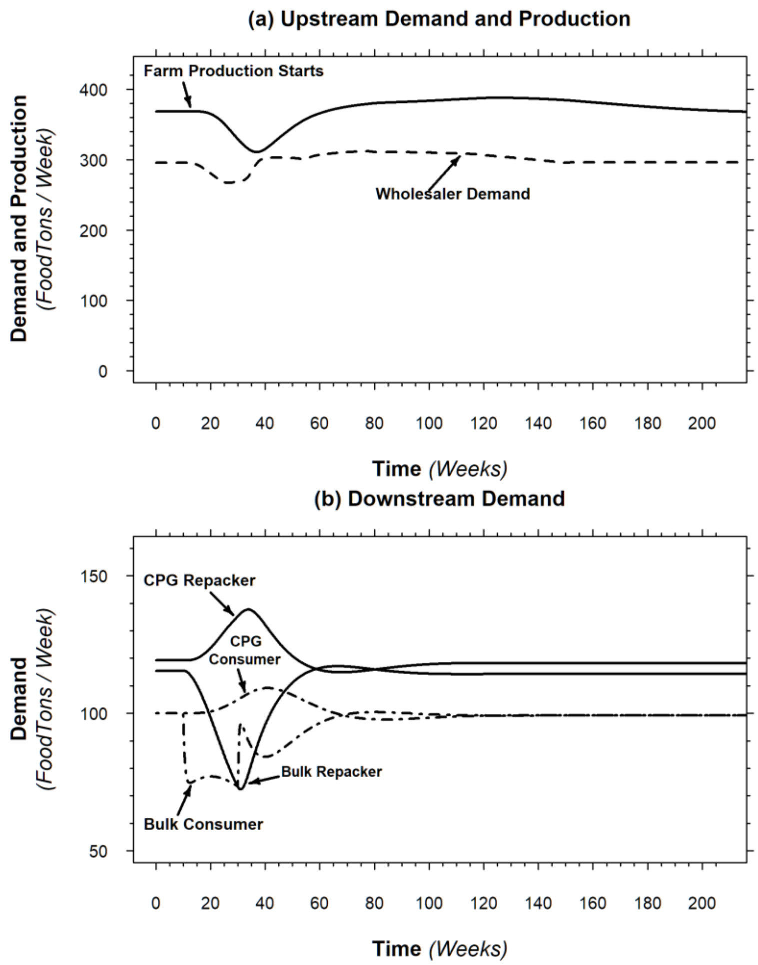


Figure 2-13. Demand and Production - 50% Drop in Bulk Purchasing Power for 20 Weeks

However, this price relief in the CPG channel is short-lived, as the drop in production starts from the farmer after the initial shock eventually results in constrained supply throughout the system. The resulting constraints cause prices for CPG consumers to temporarily rise to a marginally higher level than previously experienced before the reduction in bulk purchasing power as seen Figure 2-14. As parameterized here, prices throughout the system eventually settle back to the starting long-run prices.

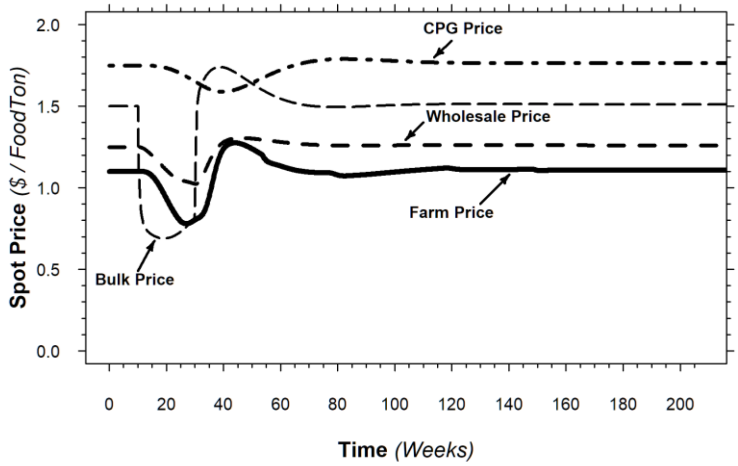


Figure 2-14. Spot Prices - 50% Drop in Bulk Purchasing Power for 20 Weeks

The spot price for the farm is affected not only by inventory values, but also by the quality of the food itself under cultivation. As seen in Figure 2-15 and as parameterized here, the age of the food under cultivation during the initial shock causes average food age to increase as the farm opts to hold crops in the field rather than move it through towards a finished goods state. Here, the shock subsides before the goods age to a point of reduced value. However, as demand recovers, but overall inventory availability is lower, the farm is economically motivated to harvest crops *earlier*, moving underdeveloped foodstuffs into the supply chain more quickly.

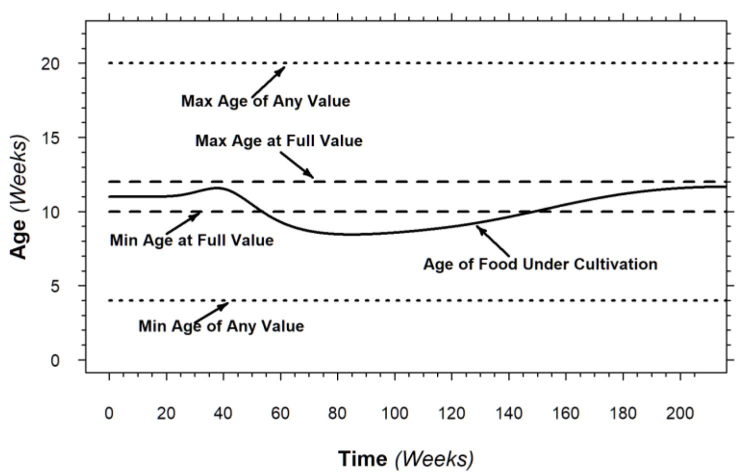


Figure 2-15. Food Maturation - 50% Drop in Bulk Purchasing Power for 20 Weeks

By inspection of Figure 2-15 alongside Figure 2-14 and Figure 2-13, the period of lower food quality coincides with the same period of increased CPG prices and reduced CPG demand. The net result is higher prices for less consumption, and the production of possibly inferior quality goods across the supply line during this transient period immediately following the shock.

Additionally, the choice model framework utilized allows for the investigation of the destruction of food across the supply chain, and the drop in prices alongside the drop in demand, coupled with a decrease in value of the food itself as it sits in the field, drives a surge in purposeful disposal of food alongside with the expected increased spoilage from unsold finished goods. However, as seen in Figure 2-16, the purposeful destruction and disposal of goods is only an extraordinary source of loss during the shock. In the longer run, spoilage remains the main source of food loss, especially for the perishable commodity foodstuffs being considered here.

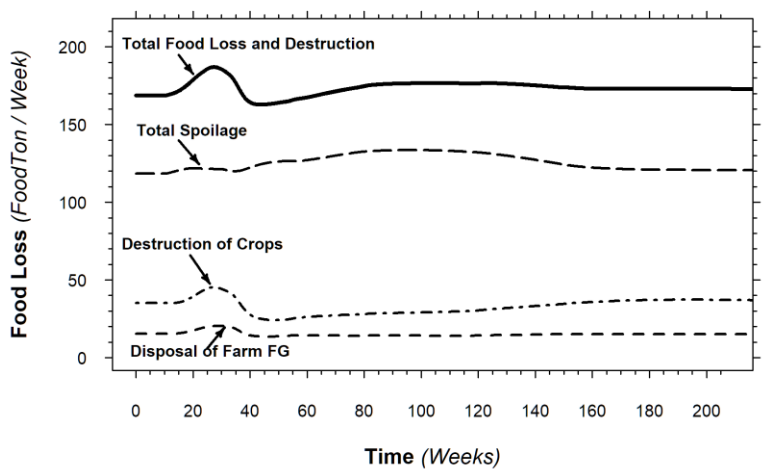


Figure 2-16. Disposal and Destruction of Food - 50% Drop in Bulk Purchasing Power for 20 Weeks

This spoilage results in a long but steady ‘second wave’ of loss seen Figure 2-16. The source of this spoilage is almost entirely driven by unsold goods that have built at the wholesaler, right before the bifurcation, but were never sold in a timely fashion. This build of unsold goods, and their ultimate loss, is a result of the spot pricing of the system which emerges from the farm producer selling goods at less than full value based on the age of the food under cultivation. This drives down prices from the farm and drives up corresponding demand from the wholesaler. However, the wholesaler is unable to sell these goods in a timely fashion and losses from spoilage build. When total demand across both arms of the bifurcated supply chain is reduced, the initial loss is slightly higher, there is no surge in demand across either arm from reduced prices, and this allows supply to normalize more quickly to a lower level with less spoilage, loss, and purposeful destruction.

The figures above illustrate how a drop in purchasing power on the bulk consumer side of this bifurcated supply chain *alone* can induce destruction of food with simultaneous higher prices for consumers so long as it is large enough and long enough. This results in precisely the scenario described above, with destruction of food upstream and reduced consumer purchasing power and consumption downstream. And while the example above is illustrative in nature, the reduction of bulk demand from governments, schools, restaurants was often larger than 50% and lasted longer than 20 weeks in some locations during the beginning of the pandemic in 2020, as seen in Figure 2-9.

2.3.1 MNL Formulation is Necessary, but Not Sufficient

The use of the MNL formulation is necessary to capture the purposeful disposal and destruction of foodstuffs by the farmer. However, it also materially changes the feedback structure of the system. As discussed in chapter 1, the introduction of the MNL choice framework in the production sector introduces a new balancing feedback loop between purposeful destruction and production starts. Using the terminology of this food supply chain model, this introduces feedback between *Destruction of Food under Cultivation* and *Food Production Starts* in the cultivation portion of this sector, and also between *Disposal of Farm Finished Goods Inventory* and *Harvesting*.

These mechanisms are essential to capture the key features of food destruction that underly the phenomena of interest for this chapter. Without these mechanisms, there is no destruction of crops, and the “Dumped milk, smashed eggs, [and] plowed vegetables” (Yaffe-Bellany & Corkery, 2020) does not occur. However, the addition of this framework alone is *not* the primary driver of much of the dynamics seen in the above model analyses, thought its inclusion does introduce two new balancing loops with delays and thus can contribute to instability and oscillation in the system.

As an example, consider the same system subjected to the same shock in demand in the bulk sector, but with the MNL formulation disabled. In this modified system, *food production starts* turn into harvesting and storage with a standard third-order delay formulation. Food is still allowed to decay in storage, but the farm does not purposefully destroy crops under cultivation, nor does it purposefully dispose of food under storage. Compare Figure 2-17 below, which shows demand and production with the MNL formulation turned off, to Figure 2-13 above which is the original analysis with this structure turned on.

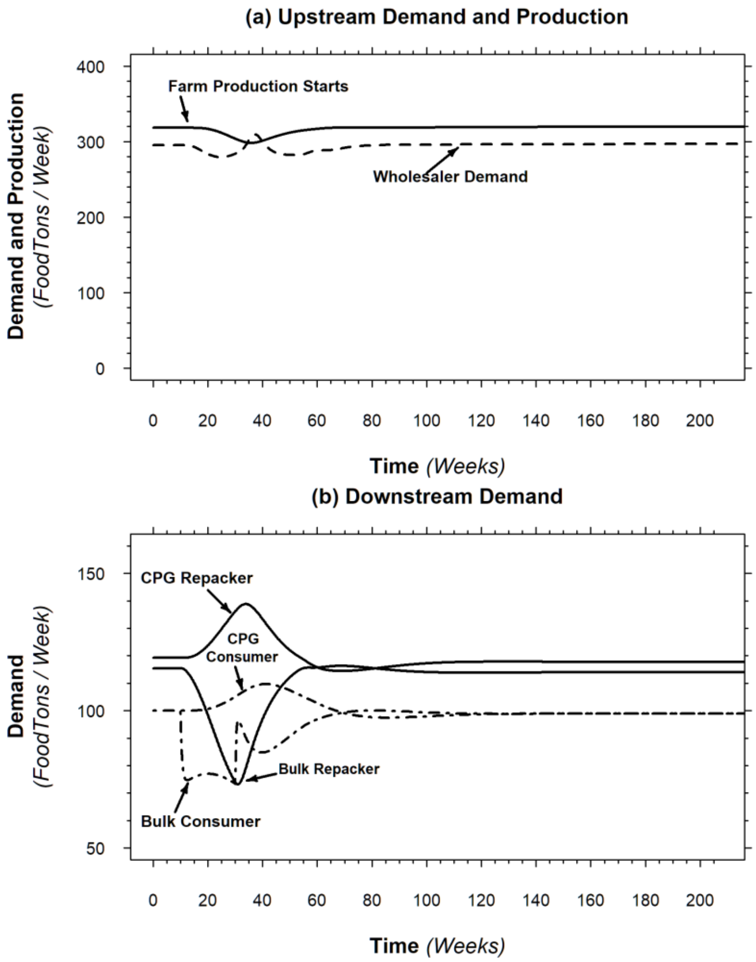


Figure 2-17. Demand and Production – MNL Formulation Turned Off

Perhaps unsurprisingly, the inclusion of the MNL formulation has the largest effect in the upstream sectors, nearer the farm where this formulation is being applied. However, past the Wholesaler, the influence of the MNL formulation is minimal on downstream demand patterns.

It is the interplay of supply and demand that drives material through the supply chain, and thus the influence of the MNL formulation in this model is limited to how it affects the supply available. However, removing this formulation also imposes the assumption that production starts are tied to finished goods availability, which is outright false here. Food was destroyed before entering a finished goods state and thus while this formulation is insufficient on its own without the price formation mechanisms, it is necessary to capture those observed outcomes.

To further emphasize this point, consider an alternative formulation that abandons price formation as well, and instead adopts the traditional demand-driven supply chain formulation (see chapter 18 of Sterman, 2000 for formulation specifics and examples). In this supply chain, demand from the end consumer and desired inventory coverage drives all demand signals in the supply chain. Figure 2-18 shows the demand and production for such a system.

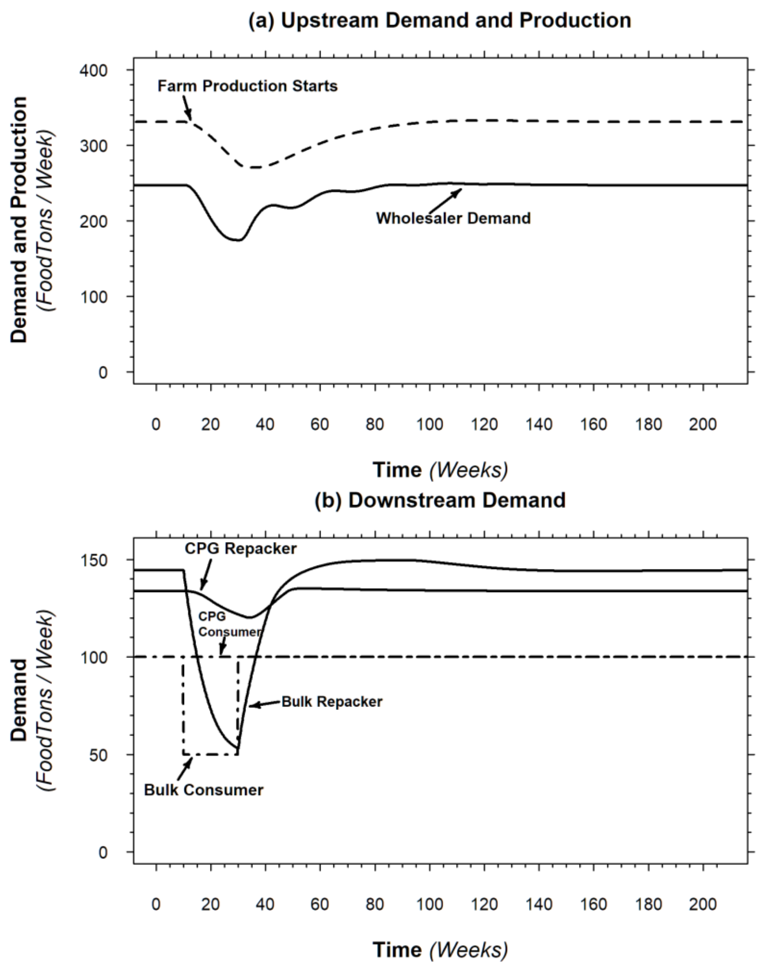


Figure 2-18. Demand and Production – Classically Driven Supply Chain with No Prices

Note that here, some assumptions must be made as some concepts of processing time are not present in the price formation and MNL model developed above. Those choices are unimportant relative to the observation that this traditional model breaks the connection between the two branches in this bifurcated supply chain.

In this simplified model, the feedback that connects the two branches via price formation at the wholesaler does not exist, and thus the CPG consumer is not affected *at all* by shocks on the bulk side of the supply chain. Again, this simplified modeling choice would be incorrect in the context of the problem being considered here, where a shock to the bulk demand side of the supply chain caused a ripple effect in inventory and prices that not only affected upstream producers, but also indirectly to CPG consumers.

2.4 Policy Interventions

The key behavior illustrated by the model is the paradox of hunger amidst plenty, which in the above dynamic hypothesis is illustrated when bulk demand is significantly reduced for an extended period of time. As shown in the scenario above, just having an exogenous reduction in bulk good demands from restaurants and school closures, *while still maintaining the same level of consumer demand*, is enough to cause a marked increase in food waste and spoilage. While the dynamics above are interesting, they beg the question of what policies are available to help mitigate the demanding consequences from surges in prices and food destruction.

2.4.1 Regulatory and Programmatic Interventions

In the United States, programs such as the Supplemental Nutrition Assistance Program (SNAP) are one governmental policy lever that can be used to restore consumer spending power for foodstuffs when faced economic hardship. As discussed in several of the referenced articles, changes to the SNAP program were under consideration during the early onset of the COVID-19 pandemic to do exactly that (Bauer, 2020; Zhou, 2020). But, as the scenario exploration above shows, even if the SNAP program were able to totally restore consumer spending in CPG space, food spoilage and destruction would still occur to a significant degree.

Given the analyses above, one possible set of policy interventions could be methods that mitigate the shock to the upstream producers and consumers from the sudden reduction in bulk consumption as best as possible. To investigate the value of such a strategy here, consider a modification to the wholesaler sector to allow for a regulatory agency or government to directly intervene and purchase raw food prior to processing. The regulator has some long run expected value of total demand, across CPG and Bulk channels, that slowly updates as actual demand is realized and compares that to the actual demand experienced.

Under this policy, the regulator attempts to fill the gap between the expected long run demand and the actual experienced demand, purchasing raw unprocessed food from the wholesaler. Additionally, suppose that the regulator is motivated to close the gap between perceived long run demand for food by the consumer and the actual demand. In this scenario, the regulator can directly provide the purchased food, with minimal or no additional processing, to the end CPG consumer to offset perceived food insecurity. The resulting purchases and disbursements by this policy are seen in Figure 2-19.

Figure 2-20 shows the effect of this intervention on the price formation in the system and food loss from disposal and spoilage. In comparison to the dynamics seen without the intervention, shown in Figure 2-14 and Figure 2-16 above, this policy helps stabilize upstream pricing at the farm and wholesaler. From the point of view of the wholesaler, the total demand remains close to constant, even though one side of the supply chain has severely lessened. The regulator here is viewed as imperfect and thus some shock is still felt, but the rapid collapse in prices at the wholesaler is not experienced to the same degree.

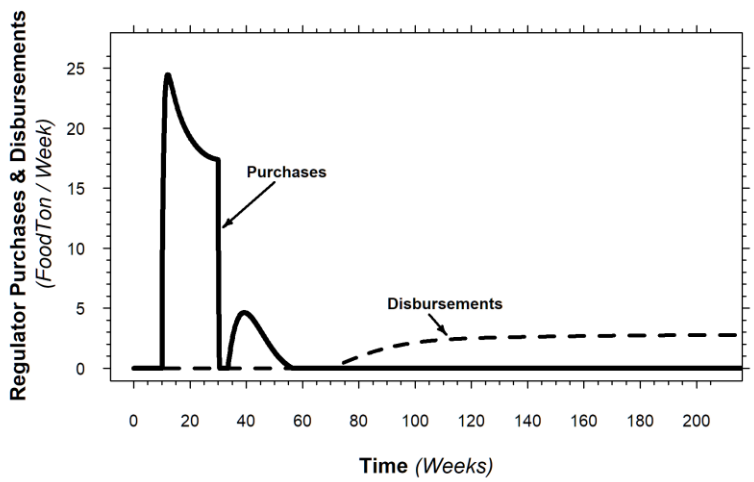


Figure 2-19. Policy Intervention – Regulator Purchase and Disbursement of Raw Foods

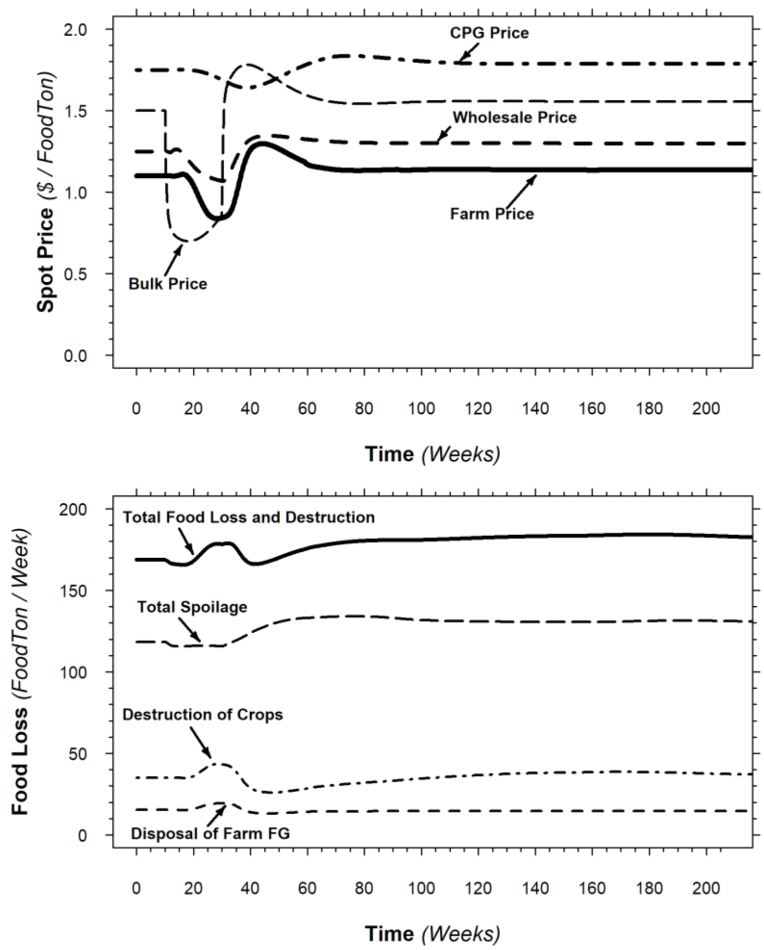


Figure 2-20. Regulator Intervention - Prices and Food Loss

However, total spoilage remains high, and the long ‘second wave’ of food loss remains under this intervention. Spoilage here comes from not only the original supply chain, but also from stocks purchased by the regulator.

2.4.2 Demand Pooling and Breaking Down the Silos in a Bifurcated Supply Chain

While the above policy does reduce the impact of the loss of demand from bulk consumers such as schools and restaurants, it relies on a third-party regulator to both act quickly and be willing to take on the costs of purchasing and distributing excess raw foods. As implied above, one of the key sources of waste in the above scenarios comes from food becoming ‘stuck’ in either side of the bifurcated supply chain, with CPG food and Bulk packaged food becoming non-substitutable.

Consider the same supply chain modeled above, but with either the CPG or Bulk consumer able to substitute goods from the other channel. Further assume that this substitution is preferentially limited, with either consumer preferring to first fill demand from his or her respective channel. Thus, substitution only occurs *after* demand at the existing spot prices is satisfied, and thereafter excess demand is supported by the prices in the other channel subject to the elasticity of demand formulation shown in expressions (8) and (9) above. As a simplifying assumption, assume the elasticity of demand for a good in either channel is the same for a given consumer. This can be easily relaxed and does not materially affect the dynamics explored below. Under normal stable conditions, allowing this substitution will have no effect, with CPG consumers fulfilling his or her demand from the CPG channel and similarly for the Bulk consumers and channel. However, when inventories are restricted or when pricing is so high as to not support the full underlying demand for goods, this allows consumers to substitute from the other channel if pricing and inventory availability there supports the excess demand. Stated differently, allow the bifurcated system shown in Figure 2-8 to be able to act as the linear system shown in Figure 2-7 with some assumptions.

For the scenario described above with a 50 percent reduction in Bulk purchasing power for 20 weeks, with the addition of substitution of goods between the two channels in this supply chain, the resulting prices and food loss is illustrated in Figure 2-21. Long run prices are stabilized and return to the same levels experienced before the shock in demand in the Bulk channel. Additionally, while there is still some food loss due to purposeful destruction of crops early on, it is markedly less than that experienced in either the scenario with no intervention or in the prior intervention via a third-party regulator.

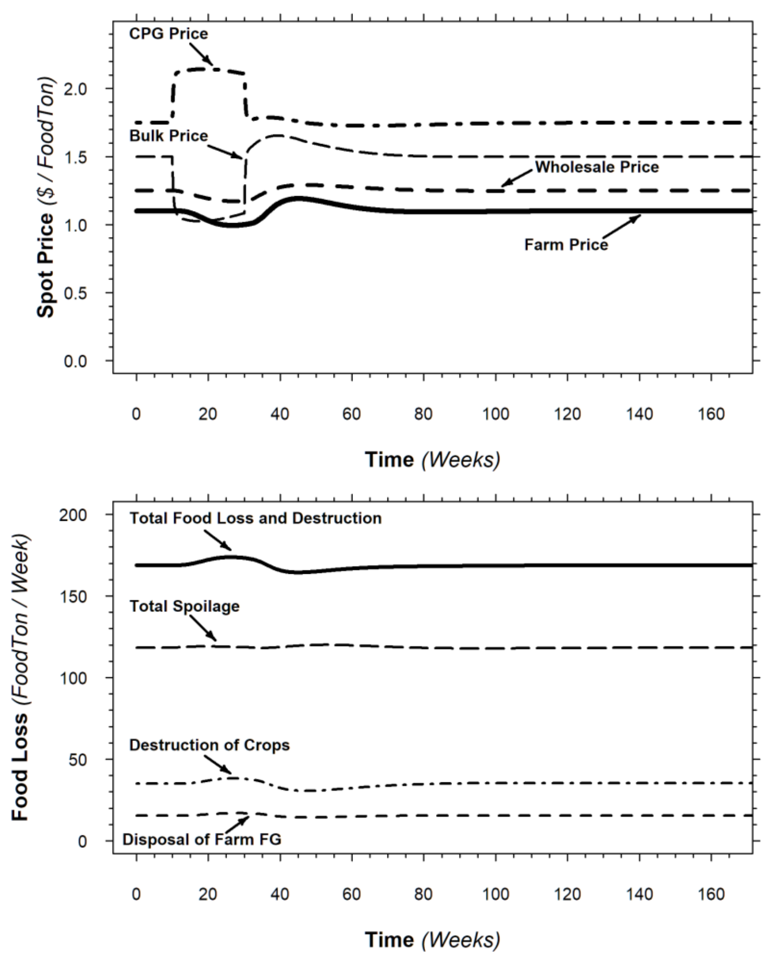


Figure 2-21. Pooled Demand - Prices and Food Loss

Comparison of the pricing seen in Figure 2-21 with that in the scenario without any intervention shown in Figure 2-14 and that in the other proposed intervention from a third-party regulator shown in Figure 2-20 illustrates a key difference in outcomes in the short run between these two models. When supply and demand are pooled, the long run outcomes are more stable with less price increases and food destruction, but in the short term a rush of Bulk consumers to the lower priced CPG channel causes a surge in prices for CPG consumers as seen in Figure 2-22.

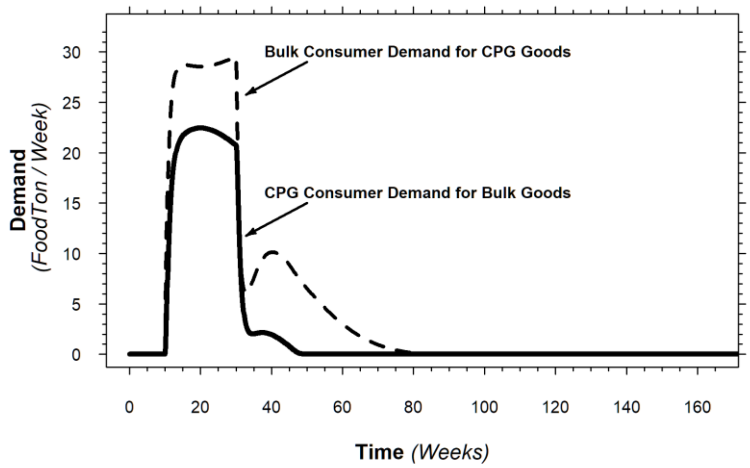


Figure 2-22. Pooled Demand – Demand Across Channels

The surge in CPG prices combined with reduced supply from increased demand from new customers causes CPG consumers to begin to purchase more from the Bulk channel. This both helps stabilize inventory that would have otherwise sat and decayed in this channel and helps stabilize prices in the short term. In net, increases mobility of consumers between channels, even if imperfect (as modeled here via the preferential demand fulfillment from each consumer's preferred channel) has a more robust stabilizing effect with less reliance on third-party intervention.

The above assumes that the underlying demand for bulk packaged goods is still the same as before the shock, with only the ability to purchase those goods by bulk consumers reduced by 50 percent. Under that circumstance, and with the preferential substitution, the total size of the reduction in demand experienced by upstream producers is reduced as some of the lost demand is recovered, albeit via another channel.

2.5 Discussion

In the analyses above, the concept of food insecurity for CPG consumers was only implicitly modeled via the increasing prices experienced in the CPG channel which emerged after a drop in demand in the Bulk channel. This is not only an implicit outcome of the structure of the food supply chain, but also exogenously exacerbated by increased unemployment or underemployment during the first weeks of the Coronavirus pandemic. Projections for 2020 show the number of food-insecure people in the United States rising to 45 million from 2019's

35.2 million, after several years of steady decline (Hake et al., 2021). As succinctly summarized by Feeding America in March of 2021:

> *"Before the start of the pandemic, the overall food insecurity rate had reached its lowest point since it began to be measured in the 1990s, but those improvements were being upended by the pandemic."* (Hake et al., 2021)

The analyses above do not impose an additional exogenous impact on the CPG consumers in the model, further reducing their ability to purchase goods. Rather this model highlights that the outcomes observed during this period, with food destruction occurring simultaneously with increasing food insecurity, can emerge from locally rational decisions made based on the economic realities faced by each entity in the supply chain. Moreover, this chapter emphasizes that it is the bifurcation of the supply chain itself that is a major contributor to the emergence of this apparent imbalance. As the end product becomes fixed in one form or another (either CPG or Bulk packaged in this model), and non-substitutable between channels, an exogenous shock to *only one side* of the supply chain can induce destruction of food with simultaneous higher prices for consumers.

Some prior work has shown that the variability and phase shift of order and inventory control signals through a supply chain subjected to an exogenous shock (such as in classic bullwhip scenarios) is reduced with increasing substitutability of products (Li et al., 2011). This chapter reinforces those observations, even though this model utilizes economic spot-price setting and profitability as the primary mechanism of moving goods through the supply chain, instead of simple order fulfillment. Even in this economic setting, a higher degree of substitutability between the channels in this bifurcated supply chain results in less inventory amplification, and in turn less waste.

However, perfect substitutability is a major assumption of the intervention explored here, as is the assumption that the demand lost from a reduction in purchasing power can be recovered from either channel. When silos between the types of end consumers exist, limiting substitutability, then interventions further upstream, closer to the original producer, become more effective. This was explored in the intervention in which a third-party regulator acts to relieve supply pressure between the producer (here the farmer) and the wholesaler that emerges from the sudden loss of demand from one side of the supply chain. The choice of policy is also a question of the goal. As seen in Figure 2-23, the regulator intervention does a superior job of reducing food loss during the early weeks of the shock. However, on a longer horizon, it can inject *more* waste by inducing artificial demand in the system for goods that ultimately spoil.

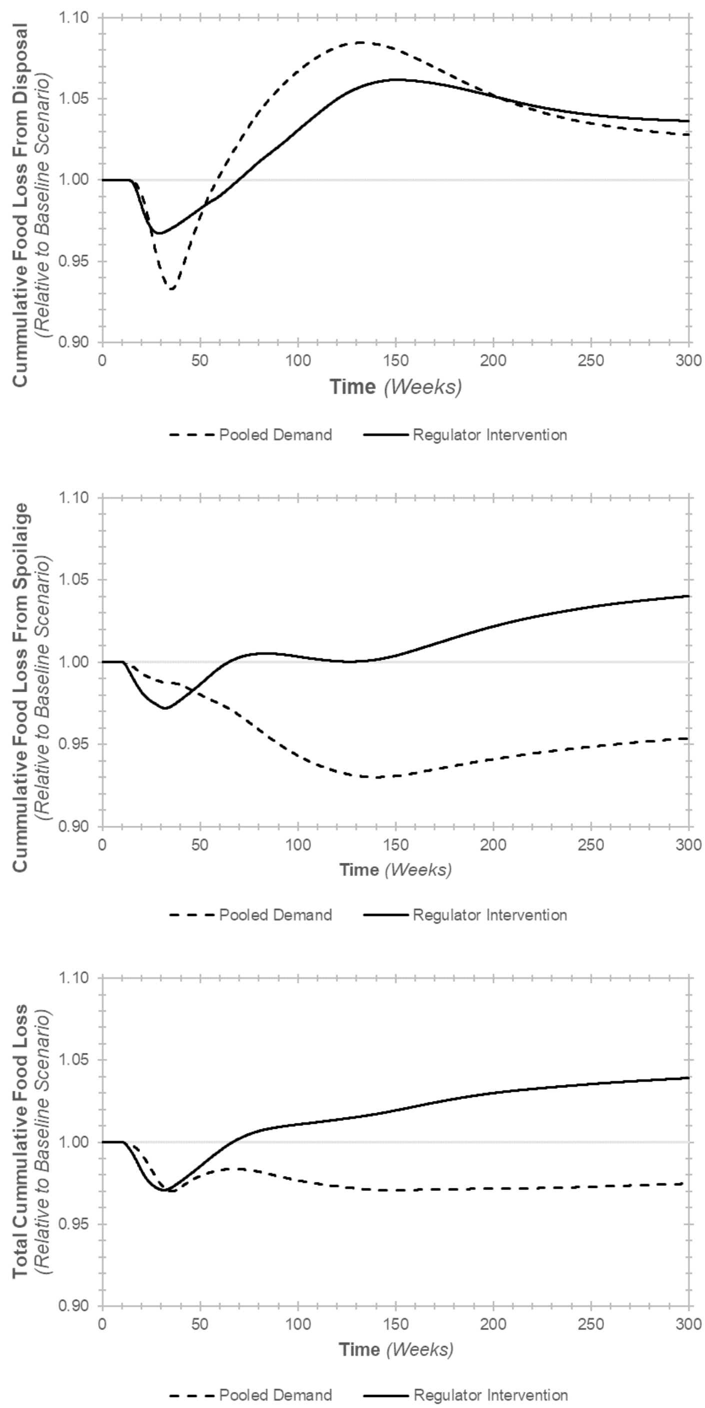


Figure 2-23. Policy Comparison – Long Run Food Losses

Also, there is a precedent in the United States of a third party acting to offset such losses directly via government or regulatory intervention: the Temporary Emergency Food Assistance Program (TEFAP). Under this program, United States President Jimmy Carter wanted to help raise the price of milk by 6 cents per gallon to offset inflation, and in order to do that authorized the federal government to *directly* purchase less perishable dairy products (cheese) from farmers and wholesalers (Blakemore, 2020; Malone, 2018). In this case, the perishable foodstuff (milk) underwent a process to reduce perishability (conversion to cheese). Such a change in perishability is not part of the model developed here but is a lever that a policy maker can consider. Another distinct advantage of this intervention, aside from smoothing the shock from a sudden evaporation in demand, is that this third-party regulator now can redistribute these goods and possibly help further alleviate food insecurity in the system. However, this presupposes that the regulator is both motivated and logistically capable of purchasing and redistributing the goods in question, while increasing substitutability requires no such outside party to act.

The model developed and explored in this chapter purposefully ignores or simplifies structures that may, in some contexts, be important to the dynamics that affect endogenous inventory disposition choices and price discovery. The producer or farmer is assumed to have adequate production capacity. In the specific scenarios explored here, the farm is met with an excess of inventory induced by a drop in demand, and thus is unlikely to consider expanding production capacity in the short and middle run dynamics explored here. A simple extension of this model would be to incorporate production capacity management based on long-run profitability expectations from the margin acquisition of additional land for cultivation.

Additionally, while the model does incorporate some aspects of storage constraints on the farm and wholesaler sectors, several of the articles in the press during the early months of the Coronavirus pandemic implied that the nominal storage is not necessarily a hard upper limit, and the farms may be able to try to store excessive amounts of goods in temporary locations (hallways, sheds, even in the farmhouses) in order to avoid excessive dumping (Yaffe-Bellany & Corkery, 2020). However, the specific dynamics of storage acquisition and adjustment (purchasing or leasing additional storage space as an example) were abstracted away in this model and subsequent analyses.

While this model arguably simplifies away many aspects of the supply chain, it also is highly parameterized, with multiple competing structures that may mask the true origins of some of the behaviors explored in the above analyses. For example, two competing price formation mechanisms, one based on inventory coverage and the other based on anchoring to long-run historic pricing, arguably make it more difficult to isolate operational and policy levers outside of those discussed above. Also, it is reasonable to assume that the inventory coverage spot price formation mechanism is not appropriate at all in environments with more contract-based pricing or otherwise more fixed pricing. This could even vary along the supply chain, for example with the producers being more subject to inventory-sensitive spot pricing and with consumers being subject to more fixed retail pricing. This can be approximated by varying the time constants in the price formation mechanisms in the model, or by varying the relative strength of the sensitivity parameters in the inventory coverage versus long-run anchoring mechanisms. While these are valid concerns, and could be used to propose alternative policies, this work focuses on behaviors and resultant policies that are hypothesized to be relatively independent of the effect of these mechanisms and instead emerge from the bifurcated nature of the supply chain.

This specific structure of the supply chain here, with a single producer, wholesaler, and then two consumer-facing and separated retail channels, is also a simplification. However, it matches the broad structure of the supply chains that motivated this work, and allows for exploration of a key sources of the dynamics of interest, namely the sequestration of finished goods in non-substitutable states. As explored in the analyses above, this non-substitutability is a key driver of instabilities generated when the supply chain is subject to an exogenous shock like that seen during the Coronavirus pandemic. For specific food products, or other supply chains, there may be more complex and subtle substitution effects between retail channels, or even multiple channels beyond the two explored here. Future application of the methods employed in this chapter to other supply chains should take those nuances into account.

Ultimately, this chapter helps illuminate how a seeming contradictory outcome of simultaneous food insecurity with food destruction during the first months of the COVID-19 pandemic can emerge from the interplay of price discovery and inventory disposition choices by entities in a simple supply chain. By combining multinomial logistic choice modeling with a model economic processes of price discovery, a more complete understanding of the behavioral features that determine the physical flow of goods through this food supply chain can be explored. Dynamics of the supply chain and features including the distribution of ages of food under production, or that passed along any specific disposition path, along with the proportions of food purposely disposed of emerge from the economically motivated choices of the individuals, rather than being imposed exogenously.

2.6 References to Chapter 2

Bauer, L. (2020). *The COVID-19 crisis has already left too many children hungry in America*. Brookings Institution Reports. https://www.brookings.edu/blog/up-front/2020/05/06/the-covid-19-crisis-has-already-left-too-many-children-hungry-in-america/

Blakemore, E. (2020). *How the U.S. Ended Up With Warehouses Full of "Government Cheese."* History.Com. https://www.history.com/news/government-cheese-dairy-farmers-reagan

Chen, H., Wu, O. Q., & Yao, D. D. (2009). On the Benefit of Inventory-Based Dynamic Pricing Strategies. *Production and Operations Management*, *19*(3), 249–260. https://doi.org/10.1111/j.1937-5956.2009.01099.x

Corkery, M., & Taffe-Bellany, D. (2020, April 18). The Food Chain's Weakest Link: Slaughterhouses. *New York Times*. https://www.nytimes.com/2020/04/18/business/coronavirus-meat-slaughterhouses.html%0A

Corkery, M., Taffe-Bellany, D., & Kravitz, D. (2020, May 25). As Meatpacking Plants Reopen, Data About Worker Illness Remains Elusive. *New York Times*. https://www.nytimes.com/2020/05/25/business/coronavirus-meatpacking-plants-cases.html

Croson, R., Schultz, K., Siemsen, E., & Yeo, M. L. (2013). Behavioral operations: The state of the field. *Journal of Operations Management*, *31*(1–2), 1–5. https://doi.org/10.1016/j.jom.2012.12.001

Durisin, M., Rembert, E., & Freitas, T. (2020). *A Tenth of the World Could Go Hungry While Crops Rot in Fields*. Bloomberg. https://www.bloomberg.com/news/features/2020-08-31/hunger-is-threatening-to-kill-more-people-than-covid-this-year

Forrester, J. W. (1961). *Industrial Dynamics*. Pegasus Communications.

Galen, M. V., & Hoste, R. (2016). Profit analysis in animal product supply chains: Exploratory research and proposal for a generic approach. *LEI Memorandum*, *2016*(052), 1–38.

Gino, F., & Pisano, G. (2008). Toward a theory of behavioral operations. *Manufacturing and Service Operations Management*, *10*(4), 676–691. https://doi.org/10.1287/msom.1070.0205

Growing Guides. (2021). The Old Farmer's Almanac. https://www.almanac.com/gardening/growing-guides

Hake, M., Dewey, A., Engelhard, E., Strayer, M., Dawes, S., Summerfelt, T., & Gundersen, C. (2021). *The Impact of the Coronavirus on Food Insecurity in 2020 & 2021* (Issue March, pp. 1–9). Feeding America. https://www.feedingamerica.org/research/coronavirus-hunger-research

Hämäläinen, R. P., Luoma, J., & Saarinen, E. (2013). On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems. *European Journal of Operational Research*, *228*(3), 623–634. https://doi.org/10.1016/j.ejor.2013.02.001

Johnson, B. (2021). In an Age of Abudance, Why do People Starve? *MIT Technology Review*, 74–79.

Lee, J., Powell, T., & Wessel, D. (2022, June 27). What are inflation expectations? Why do they matter? *Brookings Institute*. https://www.brookings.edu/blog/up-front/2020/11/30/what-are-inflation-expectations-why-do-they-matter/

Li, X., Song, L., & Zhao, Z. (2011). Quantifying the impact of demand substitution on the bullwhip effect in a supply chain. *Logistics Research*, *3*(4), 221–232. https://doi.org/10.1007/s12159-011-0060-y

Lougee, R. (2020). Has COVID-19 'broken' our food supply chain? *ORMS Today*. https://doi.org/10.1287/orms.2020.06.02

Malone, K. (2018). *Government Cheese: Well-Intentioned Program Goes Off The Rails*. National Public Radio. https://www.npr.org/2018/09/07/645459818/government-cheese-well-intentioned-program-goes-off-the-rails

Morecroft, J. (2015). *Strategic modelling and business dynamics: A feedback systems approach* (2nd ed.). John Wiley & Sons.

OpenTable. (2021). *The Restaurant Industry in Recovery*. https://www.opentable.com/state-of-industry

Owen, J. (2020). *Grocery Retailing: Including the Impact of COVID-19 April 2020*. Mintel. https://www.mintel.com/display/986958

Sterman, J. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World* (1st ed.). Irwin/McGraw-Hill.

Whelan, J., & Forrester, J. W. (1996). Economic Supply & Demand. *D-4388*, 7. https://ocw.mit.edu/courses/15-988-system-dynamics-self-study-fall-1998-spring-1999/resources/economics/

Yaffe-Bellany, D., & Corkery, M. (2020, April 11). Dumped Milk, Smashed Eggs, Plowed Vegetables: Food Waste of the Pandemic. *New York Times*.

Zhou, L. (2020, May 11). Coronavirus is exacerbating America's hunger crisis. *Vox*. https://www.vox.com/2020/5/11/21233063/food-banks-snap-coronavirus

Chapter 3

Simpler is (Sometimes) Better: A Comparison of Cost Reducing Agent Architectures in a Simulated Behaviorally Driven Multi-Echelon Supply Chain

CHAPTER ABSTRACT

Recent events have highlighted the real-world difficulty of managing bullwhip in multi-echelon supply chains. Existing streams of Operations Management and Supply Chain literature often focus on different parts of this problem and suggest distinct policy recommendations. This chapter directly compares the approaches implied by this prior work, including policies that range from myopic, limited information decision rules to more modern, but data-intensive machine learning methods. This comparison is focused on reducing costs in a simulated, behaviorally driven, multi-echelon supply chain by experimentally changing the features of polices, including complexity, adaptability, incentive structure, and information availability. Results show that relatively simple base-stock ordering policies can achieve high cost reductions and may be preferred when learning is not feasible. However, when dynamic learning is possible, directly incorporating behavioral assumptions in policies further reduces costs. For some conditions, locally focused and limited information decision rules can be cost reducing globally. Under plausible conditions, decision makers in supply chains with other behavioral actors need not be perfectly rational and can be locally focused while achieving global benefits.

3.1 Introduction

One of the more studied consequences of the interaction human decision making and supply chain structure is the emergence of instability as embodied by the 'bullwhip effect' in multi-echelon supply chains. Bullwhip refers to the increasing amplitudes in both orders and on-hand inventory positions of members of a multi-echelon supply chain the further one moves away from a source of order variability (see X. Wang & Disney, 2016 for an excellent review of research on the origins of this effect from both a structural and behavioral perspective from the early 1960s through the mid-2010s).

Given the over sixty-year history of studying this phenomenon explicitly, it could be easy to dismiss bullwhip as a 'classical' problem in Operations Management. However, the bullwhip effect remains a current source of excessive strain on real world inventory management systems via stock outs, and unnecessary capital reservation through safety stock building (Ellram, 2010). This phenomena is also not restricted to any one industry, but present in varying forms whenever ordering decisions being made in moderately complex and interlinking environments such as multi-echelon supply chains (Lee et al., 2004; Sterman, 1989, 2000). Even more recently, worldwide experiences with supply chain shortages and excesses induced by the Coronavirus pandemic for consumer goods, foodstuffs and medical supplies, have catapulted the term 'bullwhip' into the popular consciousness (Bamakan et al., 2021; Evenett, 2021; Hockett, 2020; Johnson, 2021; Shih, 2020; Stank et al., 2021).

A supply chain manager faced with bullwhip could, correctly, expect that the classical optimal control approaches in multi-echelon supply chains are well understood and well-studied. However, this prior work often ignores or over-simplifies the non-rational components of supply chain management, features a manager would likely be well aware of from personal experience. The balance between managing a supply chain under rational assumptions, versus managing a supply chain under behavioral expectations may cause the manager to consider more complex methods that could bridge this divide, or at least allow the manager to learn more accurately his or her environment and respond accordingly, including dynamic learning policies or more exotic general machine learning methods

Thus, the problem facing managers in modern supply chains facing real and unexpected bullwhip is how to balance these different policy approaches. Indeed, many of the pertinent streams of research were developed in parallel or in isolation from each other (control theory, contracting, psychology, behavioral decision sciences, general purpose machine learning, etc.).

The influence of policy features that minimize costs in a multi-echelon supply chain have been pursued by various subfields within Operations Management, but they have focused on different parts of the overall puzzle, often resulting in distinct recommendations. The manager could adopt a static base-stock policy as the suggested 'optimal' policy from Operations Research; or a heuristic implied by Behavioral Operations Management that removes noise in the ordering decision while fully accounting for the supply line. Additionally, this manager could actively incorporate more information about the system in which they are embedded and move away from a static to a dynamic policy. The manager could also consider changing his or her incentive structure to better encompass others. But with limited time and resources, it is unclear which of the above options, either alone or in combination, should be the priority of this decision maker.

This chapter directly compares these suggested policy features side-by-side. In doing so, this chapter provides meaningful managerial insights on where to focus limited time and attention. It also has implications for the wider Operations Management research community by revealing the research streams that generate the most incremental value in this supply chain context.

To do this policy feature comparison, this chapter introduces a simulation of a multi-echelon supply chain, based on a real-world game used to teach supply chain concepts in educational settings and calibrated based on real human players from historic runs of that game. This chapter then compares the effectiveness of reducing inventory management costs by introducing a policy (also referred to as an 'agent' in certain settings in this text as well) with different structural features at different positions in this simulated system

In developing agents that allow for comparing these different policy architectures this chapter presents two minor methodological contributions to cost-reducing supply chain agent modeling: an extension of existing deep-Q network (DQN) architectures that builds from recent prior literature in this space, and also a model-predictive learning architecture that incorporates features of control theory with features of behavioral modeling.

However, these are secondary to the main contribution of this chapter, namely the observation that the most complex agent architectures are not necessarily significantly better than simpler and more directly interpretable alternatives. For environments where learning about the surrounding supply chain is not feasible, the simple base-stock polices described by the early literature (Clark & Scarf, 1960) are cost reducing, even if the rest of the supply chain breaks many of the underlying assumptions of that work. For environments where learning is possible, incorporating behavioral assumptions of the other actors in the supply chain improves cost reductions, but does not necessarily require complete, or in some cases any, knowledge of the states of those other actors. Simply having a model of the system that can be dynamically updated is sufficient to develop a cost reducing policy.

The rest of this chapter is organized as follows: The Literature Review section provides a brief expansion on the relevant streams of research that imply different cost reducing policies in a serial multi-echelon supply chain. The Design and Methods Section gives details on simulation setting and the specific policy features that are tested. The Results section provides the outputs of statistical analyses that were performed to compare the relative cost reducing benefit introduced by the different policy architecture choices. Finally, the Discussion section provides an overview of how these results inform the decision making process of managers in supply chains and provides some additional notes on the limitations of this simulation setting.

3.2 Literature Review

Much of the emphasis of Supply Chain Management and Operations Research has been on intra-supply chain relationships, as multi-echelon supply chains emerged from partnerships of smaller firms towards the end of delivering a product or service to an end consumer from subcomponents sources from all over the world.

As mentioned above, classical optimal control approaches in multi-echelon supply chains are well understood and well-studied. Work by Clark and Scarf demonstrated that an optimal control policy can be applied via a base-stock ordering system when the final customer demand distribution is known (Clark & Scarf, 1960). Stated simply, this system places orders necessary to achieve a given and fixed ‘base-stock’ quantity of units, inclusive of inventory on-hand, on backorder, and ordered but not yet received. Their algorithm was later generalized and operationalized to both multi-echelon supply chains with imperfect local information and stationary demand patterns (Chen, 1999; Chen & Samroengraja, 2009; Lee et al., 2004).

The field of Operations Management has increasingly followed its peers in economics, marketing, and finance by expanding these rational approaches by recognizing the influence of human heuristic-based decision rules and incorporating these behavioral observations into the models of supply chains and inventory management (Gino & Pisano, 2008).

Recent work has endeavored to more clearly and explicitly define ‘Behavioral Operations Management’ as a subfield of Operations Management and Operations Research (Croson et al., 2013; Cui & Wu, 2018; Gino & Pisano, 2008; Größler et al., 2008). Research here can focus on more detailed representation of human decision heuristics, but often emphasizes defining the gap between optimality and observed reality of decision makers over specific policy interventions (see Bendoly, 2013; Huang et al., 2013; Morrison & Oliva, 2018; J. Sterman, 1989 among others).

A key behavioral bias that leads to bullwhip is commonly identified as 'supply-line underweighting' (Croson & Donohue, 2006; Narayanan & Moritz, 2015; Sterman, 1989) and emerges as part of a larger anchoring and adjustment heuristic employed by decision makers in a multi-echelon supply chain (Sterman, 1989; Tversky & Kahneman, 1974). Here, decision makers systematically under-weigh the value of inventory ordered but not yet received (i.e., in their supply-line) relative to on-hand inventory and backorders. Mitigation of bullwhip has focused on adjusting the structure of the supply chain itself, the information availability along the supply chain (Croson et al., 2014; Croson & Donohue, 2006; Wu & Katok, 2006), and on the instruction and training strategies of supply chain managers (Croson et al., 2014; Martin et al., 2004; Wu & Katok, 2006). While mitigation is possible, the underlying ordering heuristics that drive the emergence of bullwhip remain in many of these studies.

Much of this prior literature, both 'rational' and 'behavioral' decision making takes place in a dynamic context, but often identifies or prescribes an ordering or inventory management rule that is static in so much as it does not vary in time. As discussed in the introduction to this chapter, a manager in a dynamic supply chain may very well consider a policy that does vary with time, or at least updates along with the manager's understanding of his or her environment.

Model predictive control (MPC) schemes (see Åström et al., 2001; Seborg et al., 2016 for overviews) would allow for such online learning and adjustment, and have been used in operational contexts (see Ciocan & Farias, 2012; Pannek & Frazzon, 2014; Secomandi, 2008; Vossen et al., 2022 for recent supply chain applications of MPC). MPC methods generally are used as a process control method, relying on a dynamic model of a given process and a set of constraints to generate a control scheme to achieve a specific control target. However, the dynamic model used in these systems often emphasize physical and information systems, and abstract away or outright exclude the influences of human decision making.

More recently, there has been an expansion of the use of neural network architectures that emphasize policy outcomes at the expense of simplicity and interpretability. These methods are likely tempting to any manager as they have been both emphasized in the popular business and management press (Kelber, 2020; PYMNTS, 2022; Tarafdar et al., 2019; Urban et al., 2020), and also explored in academic literature (see Chaharsooghi et al., 2008; Oroojlooyjadid et al., 2021 for two recent supply chain examples).

However other fields that Behavioral Operations Management borrow heavily from such as psychology and other behavioral sciences have already noted that more complex models do not necessarily generate more robust or accurate outcomes. Simple linear models even with imperfectly (but consistently applied) weights can outperform human decision makers (Dawes & Corrigan, 1974; Kahneman et al., 2021; Yu & Kuncel, 2020), and simple models have also been found to outperform more complex models in timeseries forecasting tasks that match this supply chain decision setting (Makridakis et al., 2020; Makridakis & Hibon, 2000).

3.3 Simulation and Modeling Framework

The Beer Game (Sterman, 1989) is a classical inventory management dynamic system learning tool and provides the context for the simulation and policies developed here. This is a multi-agent decentralized supply chain and is modeled much like real decentralized serial inventory management system.

First developed by Jay Forrester at MIT, the game has been used since the 1950s to illustrate system thinking concepts and the prevalence of the bullwhip effect. The original purpose of the simulation was to illustrate the difficulty of rational thinking during time-delayed and non-linear information feedback loops, the value of information sharing, and most classically the bullwhip effect in inventory management. Since its first use in the classroom, this framework that has also been used extensively in related supply chain and behavioral research (notably Chaharsooghi et al., 2008; Croson & Donohue, 2006; Narayanan & Moritz, 2015; Oliva et al., 2022; Oroojlooyjadid et al., 2017; Sterman, 1989; Sterman & Dogan, 2015 among others).

For this chapter, a discrete time model of the Beer Game was developed in both the R and Python scripting languages. R was used as the primary analysis and simulation environment, while Python was used exclusively for training more generalized machine learning methods explained in more detail in the sections below. More specifics about the mechanics of the Beer Game are provided in Appendix C. The model was made as both a self-contained simulation of the system over a given time horizon, and as a callable function that takes a given state-action pair and returns an updated state, given an ordering rule for the entities in the system. The order rule utilized by each entity in the supply chain itself is modular, and is able to generate environments with agents based on multiple models from prior work, notably (Croson

& Donohue, 2006; Oliva et al., 2022; Sterman, 1989; Sterman & Dogan, 2015) and also classical and fully rational base-stock replenishment strategies (Clark & Scarf, 1960).

The stated goal of the game is to reduce the amount of *total cost of the entire team* over some time horizon *T,* subject to some known inventory holding and backorder/stockout costs. Backorders do not expire under the traditional interpretation of this game and must be filled from existing stock prior to meeting any new demand.

$$Cost_{Team} = \sum_{t=1}^{T} \sum_{entity=1}^{N} \left(C_{bo} * Backorders_{t,n} + C_{inv} * \ Inventory_{t,n} \right) \quad (10)$$

Typically, real human players are placed into this system to make inventory management decisions. Within a few rounds of ordering, the bullwhip in inventory and backorders appears, amplifying over time along the simulated supply chain as each player acts to reserve inventory to satisfy his or her own myopic forecasts and needs. As discussed above, exact solutions for optimal ordering quantities in similar serial supply chains have been developed, such as the base-sock method (Chen & Samroengraja, 2009; Clark & Scarf, 1960), but require all agents to be acting rationally and consistently, and for specific costing structures to be present (notably increasing costs along the supply chain). Additionally, while these optimal ordering methods presume stationary customer order patterns, which this simulation satisfies, the human participants themselves have no knowledge *a priori* of the distribution of the customer order pattern.

In the Beer Game, the players (acting as inventory managers) have only one operational decision to make each period: how many units to order. How the inventory manager processes the information available to him or her in the environment, including either actual or perceived states, has been the subject of much of the prior work discussed in the Literature Review section above. The policies that this chapter explores ultimately need to arrive at the same singular ordering decision. As in the prior literature, these policies vary in complexity and the size of the observable space, including both rationally optimizing and heuristically satisfying fundamental architectures. The system these ordering decisions are embedded in is dynamic, especially when considering the reactions of non-perfectly rational players in the supply chain to the ordering signals sent by the agents.

However, a fully dynamic response may not necessarily be needed in all circumstances. Indeed, the original base-stock policies discussed in the sections above suggest static responses, at least to stationary input signals. For that original base-stock agent, the manager need only consider the gap between total inventory on hand and on order (e.g., previously ordered but not received) less orders received but not filled, and some desired inventory level. The prior work in this field focused in part on setting that level to minimize costs. But, when considering behaviorally driven systems, one could consider how the policy, and in turn how the other players in the supply chain, may either discount or inflate any of these inputs into the ordering decision.

A strategic manager may order not based on the order received in each time period, but based on the long-range orders they expect to receive. A manager with low confidence in his or her upstream supplier's ability to fill orders accurately and completely may discount the orders placed and not received. A manager may also weigh different aspects of these inputs to his or her ordering decision differently, placing unequal weight on orders outstanding versus orders received or any other aspect of the observable space.

On the other end of spectrum, a manager may make no decision whatsoever, and simply pass through orders received each period, making no attempt to manage on hand inventory or backlogs.

Of course, many other methods by which a manager arrives at an order decision could be considered and have in the prior 60 years of research in this space. But, the ordering heuristic introduced in (Sterman, 1989) and further described (Martin et al., 2004), compactly covers the scenarios discussed above, and is presented in (11) and (12) below:

$$O_t = MAX\left(0, \widehat{L_t} + \alpha_S(S' - S_t - \beta\, SL_t) + \varepsilon_t\right) \tag{11}$$

$$where\ \widehat{L_t} = \ \theta L_t + (1-\theta)\hat{L}_{t-1} \tag{12}$$

In the above, *O* is the order placed at time *t* given the information observed in the right-hand side of the above expression. In that expression $\hat{L}$ is a smoothed interpolation of the expected outflow of inventory, subject to a smoothing parameter *θ*. *SL* refers to the total inbound supply line of inventory heading towards the player. *S* is the current on-hand inventory (or stock), and *S'* is a parameter that can be considered analogous to the desired or goal on-hand inventory of the player. Thus, we have an expression with four parameters: θ, α, β, and S'. As conceptualized in (Sterman, 1989), the above parameters are bounded as $0 \leq \theta, \alpha, \beta \leq 1$ and $0 \leq S'$, and that paper also provides fitted values for expressions (11) and (12) for a set of real human teams, along with a set of parameter values that best fit the overall behavior of all teams.

Note that when $\alpha = 1$, $\beta = 1$, $\theta = 1$ then expressions (11) and (12) collapse to the classical Clark and Scarf base-stock ordering rule, with S' being the desired inventory level. Conversely when $\alpha = 0$, $\beta = 0$, $\theta = 0$, these expressions collapse to an ordering scheme in which no inventory management occurs whatsoever, the entity just passes through orders.

Of course the points in this multidimensional ordering rule that map to human ordering decisions fall in between these bounds, and prior work has specifically emphasized the influences of supply chain underweighting (generally when $\beta < 1$, though also when $\alpha < 1$ as well) on the bullwhip effects central to this paper (Sterman, 1989). For the results presented here, the data drawn from order traces collected in the fall of 2021 and spring of 2022 of real players of the beer game was fit using expressions (11) and (12) and combined with the data in Sterman 1989.

The error from fitting the real order history to the simplified model in (11) and (12) can be expressed at both the parameter level (e.g., bootstrapped confidence intervals on fitted parameter values) or the overall model fit level (e.g., via root mean squared error between predicted and observed order traces). For the order traces collected in the fall of 2021 and spring of 2022, both error measures are obtainable from the fitting process. However, for the models developed in the Sterman 1989 paper, only the overall model root mean square error (*RMSE*) is reported, not parameter-level measures of error, and the original order traces and fitting process are not readily available. Thus, for this paper, the *RMSE* for each fit of order behavior was used to provide a measure of noise to the order decision. An alternative analysis could be done, omitting the original data from Sterman 1989 and focusing just on models derived from the more recent runs of the Beer Game, and using bootstrapped errors on parameter estimates. From limited testing, the results below are qualitatively similar using either RMSE or parameter level errors, so the RMSE method was used here to maximize the number of available models to test the policies in this paper.

3.4 Policy Feature Construction

As discussed in the Introduction and Literature Review sections, the cost-reducing policy the agent follows could either be based on early 'rational' base-stock modeling literature, or more recent heuristic 'behavioral modeling literature. Furthermore, given either approach, the agent could take on a static 'single shot' policy, or a dynamic 'learning' policy.

'Single-Shot' cost reducing agents operate by fixing their ordering decision rule *a priori*, either by fixing some generally cost-reducing value of the four parameters of expressions (11) and (12) (e.g. ‘behavioral’ agents), or by following a base-stock replenishment policy with a fixed order-up to level (e.g. ‘base-stock’ agents).

Alternatively, model-predictive ‘learning’ agents are assumed to follow a scheme similar to that found in model predictive control literature (Åström et al., 2001; Sutton & Barto, 2014), starting from the assumed cost reducing rule utilized by the ‘single-shot’ agents defined above, but then adjusting the parameters of that decision rule over time based on observations of the evolution of the environment.

3.4.1 ‘Single-Shot’ Agents

As stated in several of the articles in the literature review above the manager may first try to compensate for his or her own supply chain underweighting. Using the language of the behavioral ordering heuristics used here this would be equivalent to setting $\beta = 1$ in expressions (11) and (12), while leaving other aspects of his or her ordering rule unchanged.

However, given an assumption about the ordering heuristics being used by the rest of the supply chain, the manager could go further and actively optimize all four parameters in expressions (11) and (12) in his or her ordering rule. For these ‘single-shot’ agents, costs for the entire team of four entities are reduced by fixing the ordering parameters of all entities save one (the agent) in a nested set of these expressions and finding a set of parameters for the remaining entity that are cost reducing. This routine is illustrated in Figure 3-1.

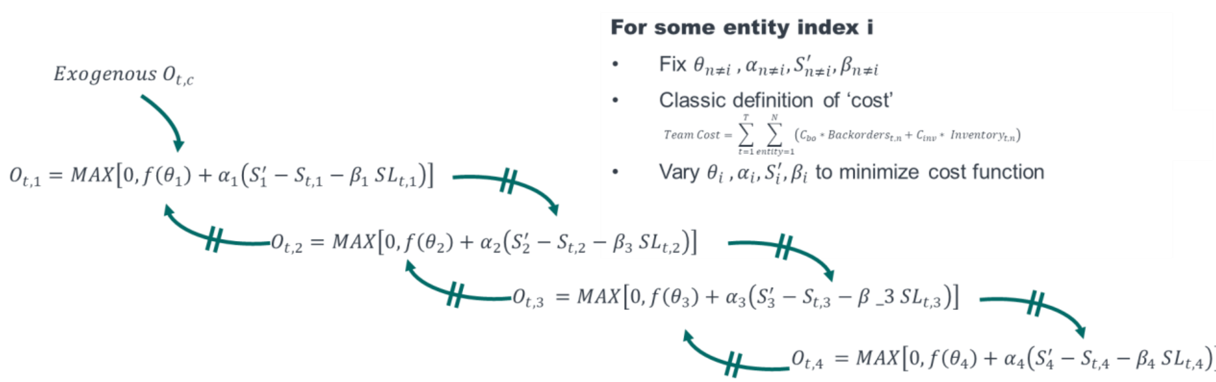


Figure 3-1. Cost Minimization Routine for the Model-Based Approach

The general learned parameters for a horizon of *t = 52* periods utilizing a bounded BFGS method to reduce costs (Byrd et al., 2005) against the ‘average’ team reported in Sterman, 1989. For the examples used in this chapter, the pertinent values were found for gaussian normally drawn input of customer orders (mean orders 10 with standard deviation of 4 units) as used in (Chen & Samroengraja, 2009), with the combination of values that were on average most reducing over 50 draws of the order distribution. Table 3-1 shows the parameters that, on average, were resulted in the lowest cost when applied at each position in the supply chain in isolation in the average team reported in Sterman 1989.

Table 3-1. Learned Parameters for the Behavioral Single Shot Cost Reducing Agent for the Average Sterman 1989 Team Exposed to Normally Drawn Customer Orders

Entity Optimized	Fitted Parameters			
	θ	α	β	S'
N/A (baseline)	0.360	0.2600	0.340	17.00
0 (Retailer)	0.021 *(0.014, 0.029)*	0.086 *(0.077, 0.095)*	1 *(1, 1)*	12.22 *(11.16, 13.29)*
1 (Warehouse)	0.178 *(0.111, 0.244)*	0.009 *(0.001, 0.018)*	1 *(1, 1)*	23.09 *(21.10, 24.99)*
2 (Distributor)	0.132 *(0.126, 0.139)*	0.0004 *(0.0001, 0.0006)*	0.969 *(0.966, 0.972)*	55.91 *(55.64, 56.17)*
3 (Factory)	1 *(1, 1)*	1 *(1, 1)*	0.999 *(0.997, 1)*	16.92 *(16.58, 17.24)*

Note: fitted parameters are reported with 95% confidence intervals

Note that the above is only cost reducing if the manager happens to be in a supply chain that matches the average team reported in this prior literature. If the manager had *a priori* knowledge of the ordering heuristics being used by the other members of the supply chain, he or she could repeat the above optimization on a per-team (or per-supply chain) basis. As discussed in more detail below, this would provide a lower floor on cost (or upper ceiling on performance), as this would both correctly assume the structural features of the supply chain decision making processes and allow for a situationally-specific policy. However, this is not necessarily feasible and thus while it provides a good point of comparison is viewed as trivial in the context of this paper.

Alternatively, the agent could assume that all other entities in the supply chain are fully rational, following a simple base-stock ordering policy. In which case, the ideal base-stock value for the agent itself to follow is a function of the information structure and cost structure of the system and the distribution of the input orders. As discussed in recent work, the information and costing structure used here does not perfectly match the criteria needed for direct application of the Clark and Scarf algorithm for optimal base-stock values, but that same work provides a blueprint for obtaining near-optimal values via a grid-search method (Oroojlooyjadid et al., 2021). Using that same method, we can obtain the average most cost reducing base-stock values for this specific model supply chain.

For the same normally drawn customer order string input order string, the average most minimizing cost over 50 draws was determined to be (0,24,8,5). In other words, on average and for the four entity supply chain modeled here, when all four positions were following a base-stock policy simultaneously, the costs are minimized when Position 1 (Retailer) has a base-stock level of 0 units of inventory, Position 2 (Wholesaler) has a level of 24 units, Position 3 (Distributor) has a level of 8 units, and Position 4 (Factory) as a level of 5 units.

3.4.2 Model Predictive 'Learning' Agents

These agents build on the structure developed in the 'single-shot' agents above by incorporating a concept of learning. This agent still has a static order structure at its root (either based on a heuristic rule like that of expressions (11) and (12), or based on a base-stock replenishment policy) but iteratively estimates the parameters of an assumed model of its environment (the other players in the supply chain), and then optimizes its own ordering rule parameters over a given horizon. This routine is summarized in the pseudocode in Figure 3-2.

```
t = 0
      Assume Structural and Dynamic Model of System
      Define Agent position in System model
      Define observable space for Agent
      Populate initial parameter assumptions
      Define calibration memory and optimization horizon

for t in 1: horizon
Calibrate System Model given history
      ArgMin {System Parameter Estimate} |
            Error (Expected space of simulation of System Model,
            Actual obs space)
      Return estimated parameters of System Model

Optimize forward given System Model estimate
      ArgMax {Agent Decision Rule} |
            Over t:(t+opt horizon): Reward from t:horizon given
            System Model estimate
```

Figure 3-2. MPC Pseudocode

In the routine in Figure 3-2, step of 'populate initial parameter assumptions' is a matter of modeling choice, but one that may be important, as early and unverified assumptions about the agent's environment will affect the early ordering decision being made, and in turn affect the early evolution of the system. While many plausible choices exist, for this chapter we assume that the manager starts with a best-case assumption about the surrounding environment. Thus, for base-stock agents, at $t = 0$ the agent would assume that the surrounding system is following the average best base-stock policy (determined here to be base stock values of Retailer = 0, Wholesaler = 24, Distributor = 8, and Factory = 5), and then update that assumption in subsequent steps.

For the manager or agent following a heuristic policy, the initial assumption is that the system consists of other entities that are following the average from prior literature (specifically Sterman, 1989 here). Thus, at $t = 0$ the agent follows a policy with parameters from Table 3-1 and on subsequent time steps, updates that assumed model of the system and repeats the forward optimization.

3.5 Design of Experiment

The model predictive learning architecture introduces several new hyperparameters (calibration memory and optimization horizon), and structural choices around the 'observable space' for the agent. In other words, this raises questions about how information conditions affect the performance of the agent.

The calibration step is highly dependent on the observation space of the agent (am I calibrating based on a full set of knowledge about the inventories of all other entities, or do I only see my own inventory?), and the optimization step is highly dependent on the objective of the agent (do I care about *total* supply chain costs, or only *my own* costs?).

This choice of calibration information availability and optimization goal is not only of mechanical interest, but a real point of concern in prior literature in this space. Higher degrees of information availability directly map to concepts of supply chain integration, and entire niche industries have emerged around integrated information exchange systems in supply chains. Prior literature on model predictive control has the embedded assumption that additional information about the state space can only improve outcomes via a more accurate representation of the 'reality' of the system by implying that the main tradeoff in these schemes is between computational time and system performance (Mayne, 2014; Seborg et al., 2016). Additionally, the choice of motivation of the agent, either selfish and myopic on its own outcomes, or non-myopic and team focused, is not only a subject of intense interest in supply chain and economic literature but also a key point of discussion in the classroom when the Beer Game is used in an educational setting.

Thus, the Learning Agent provides an opportunity to test the value of the combination of these factors, adding incentive structure and information availability as additional experimental controls in addition to the fundamental architectures of base-stock versus behavioral. Table 3-2 displays the full factorial conditions for features for this learning policy. When appropriate, subsequent figures in this chapter may use the notation in Table 3-2 to refer to specific experimental features. For example, a policy that is based on a base stock ordering rule, attempting to reduce costs for the entire supply chain, but doing so with low information, would be referred to as *R-T-L*.

Table 3-2. Conditions with the Full Factorial Design of Experiment on Learning Agents

(1) Rule Complexity	(2) Incentive Structure	(3) Information Availability
'Rational' Base Stock Ordering (R)	Self (S)	Minimal (M)
Behavioral Ordering (B)	Team (T)	Low (L)
		Classical (C)
		High (H)

For this chapter, the information conditions are summarized below:

- The 'Minimal' condition, the information available as part of the calibration of the agent is the most restrictive, with the agent only having access to directly verifiable and visible information
    - On-hand inventory position of the agent itself and no others
    - The most recent order placed by the agent itself and no others
- For the 'Low Information' condition, the calibration of the agent is given a slightly larger set of information, including information that it itself is not directly in control of, namely inbound shipments
    - On-hand inventory position of the agent itself and no others
    - Inbound shipments to of the agent itself and no others
    - The orders placed by the agent itself and no others

- For the ‘Classical Information’ condition, the calibration of the agent is based on information that a human player in the traditional playing of Beer Game could view at any given moment, and matches the information available to the players in used in the datasets used to fit the models:
    - All on-hand inventory positions of all four entities
    - All inbound shipments to all entities
    - The orders placed by the agent itself and no others
- For the ‘High Information’ condition, the calibration of the agent is omniscient and has access to effectively every state variable in the system, including those only imputed and not directly observable (e.g., backorder):
    - All on-hand inventory positions of all four entities
    - The backorder positions of all entities
    - All inbound shipments to all entities
    - The entire order and information flows of all four entities

In the routine described above, the agent must first assume a model of the world. Next the agent uses its available information to fit a best estimate of that model. It then projects forward and optimizes its own ordering behavior given that estimated model. The ‘Rule Complexity’ feature in Table 3-2 is really a simplification of what could be, a true full factorial design, two features: the structure of the rules that the agent assumes applies to the rest of the supply chain, and separately the structure of the rules that the agent itself uses for its own ordering. The analysis here considers that if the agent is assuming that the other entities in the supply chain are following a base-stock ordering policy, it will structure its own ordering rule similarly. Likewise, if the assumption is that a more complex heuristic ordering rule is being followed by the other entities in the supply chain, the agent will choose a structurally similar ordering rule.

This simplification serves a practical purpose insomuch as it simplifies this analysis. But this also follows from the underlying assumptions of the base-stock ordering policy, which only has guarantees of optimality if the entire supply chain is following that rule. Thus, if the agent assumes that the other entities are following a base-stock rule, the best policy should be presumed to be a base-stock policy. Additionally, the base-stock policy can be viewed as a special case of most heuristic policies, and specifically expressions (11) and (12) used here when $\alpha = \beta = 1$, and $\theta = 0 \; or \; 1$.

For completeness of experimental description, the ‘base-stock’ and ‘behavioral’ architectures are described in more detail below:

- Base-Stock:
  - The agent assumes that the other entities in the supply chain are using base-stock ordering rules,
  - The agent matches the same base-stock structure of this rule for its own response.
  - Each time step, the agent fits a window of prior observed information about the system, and its own ordering history, to a model that assumes all other agents are following a base-stock policy, updating an estimate of the base-stock value being sought by those other entities. The agent then optimizes its own base-stock value over a forward horizon that would minimize team costs under its fitted model.
- Behavioral Ordering:
  - The agent assumes that the other entities in the supply chain are using an ordering rule that can be described by a behavioral heuristic, in this case the rule from Sterman 1989.
  - The agent matches the same structure of this rule for its own response.
  - Each time step, the agent fits a window of prior observed information about the system, and its own ordering history, to a model that assumes all other agents are following this heuristic ordering rule, updating an estimate of the parameters in that rule for all other entities that would generate the observed behavior of the system. The agent then optimizes the parameters of its own ordering rule over a forward horizon that would minimize team costs under its fitted model.

3.6 Results

The results presented below are based on the cost reduction achieved by the policy agent placed into each of the four positions in the 49 simulated supply chains as described above. 12 of these models of supply chains come directly from literature (specifically Sterman, 1989). Order traces for real runs of the game were used to create the remaining simulated supply chains via online runs of the Beer Game at MIT as part of various executive and graduate-level classes. These runs occurred twice in August of 2021, with 12 teams in one run and 22 teams in another run, and in June of 2022 for an additional three teams.

Of these 49 teams, three were dropped (approximately 6%) from the analyses below as outliers due to poor fits of the behavioral model to the observed ordering. The performance of the agent at each position in isolation were also explored, but the overarching differences between the agent being in one position versus the others was minimal aside from at the start of the supply chain, which is discussed more below.

As used to calibrate the ‘single-shot’ agents, the simulated supply chain was subjected to a normally drawn input order signal of mean 10 and standard deviation 4 in a manner similar to (Chen & Samroengraja, 2009), and the order decision reached by each entity subject to a normally drawn noise term with mean zero and standard deviation drawn from the RSME of the fitting process described in the sections above. The normally drawn input signal from Chen & Samroengraja 2009 was chosen here to specifically emphasize that these results are emergent from the policy architectures, not any specific feature of the customer demand signal.

For the remaining results presented below, the performance of each policy architecture is presented as the percentage cost reduction induced along the entire supply chain because of introducing that policy, versus the costs that were seen in a baseline simulation of that same supply chain with no such policy. The choice of this baseline is discussed in more detail in Appendix C, but as this chapter is intended to highlight policy features that are applicable in behaviorally driven supply chains, the use of a behaviorally driven baseline is appropriate here.

3.6.1 Behavioral versus Base-Stock

Figure 3-3 shows a box-and-whisker plot of the cost reduction observed after the introduction of the agent following specific policy feature choices, aggregated across all four positions in this simulated supply chain. Of the 16 possible combinations of conditions for the learning agent shown in Table 3-2, only two are illustrated in Figure 3-3, alongside the three static ‘single-shot’ agent policies (one with base-stock and two with behavioral policies). Only two of the full 16 conditions for the learning agent were chosen for illustration here for compact presentation, and because these two specific learning conditions mirror the conditions used in the single-shot policies most closely (specifically the team-level optimization objective, though concepts of information availability do not apply to the static agents).

For the *Behavioral Static Agent (Setting β = 1)* policy, this is the simplest policy suggested from existing behavioral operations management literature, where in the policy is to follow whatever the existing behavioral heuristic would have been for that entity in that position, but overwriting the value β in expression (11) to be 1. At the other extreme, for the *Behavioral Static (Separate Opt Per Team)* policy, a separate optimization was conducted for each possible team and position the agent could find itself in. This provides a performance ceiling (or cost floor) for these analyses.

For the *Behavioral Static (Opt for Average Team)* agent, this follows the policy outlined in Table 3-1, which is the best static policy for the 'average' team suggested by prior literature. Here, it is assumed the agent following this policy has no knowledge a priori about the team they find themselves in and thus cannot pick the 'best' policy suggested by the *Behavioral Static (Opt for Average Team)* policy and thus simply assumes the policy . For the *Base-Stock Static* agent, the policy is to follow the base stock rule determined to be cost reducing on average given the characteristics of the customer orders. As discussed above, this is a base stock level by position of (0, 24, 8, 5) for this normally drawn customer order pattern.

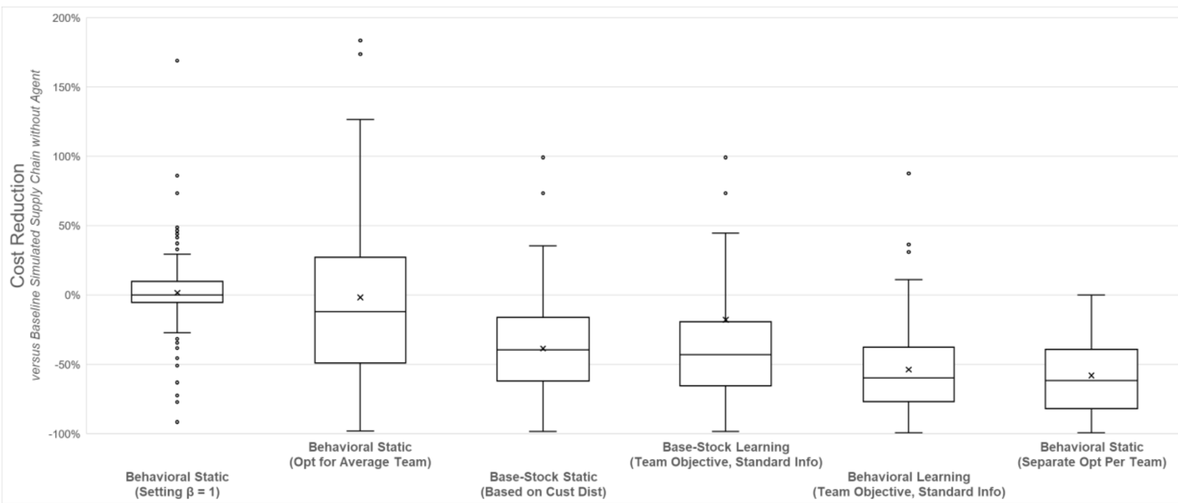


Figure 3-3. Static vs Learning Agent Induced Cost Reduction Across all Positions

Perhaps the most surprising outcome of these initial results is the effectiveness of the static base-stock agent. This is the original policy given nearly 60 years ago, and even in this behavioral context in which many of the assumptions that would imply true optimality of this rule are broken, this policy performs *better* than the static behavioral policy that is assuming that it is operating in an 'average' environment. In other words, when learning is absent, and *a priori* knowledge of the responses of the other entities in the supply chain is unavailable, a static base-stock policy generally outperforms a static average behavioral policy. This simulated environment imposes the condition that the other entities *do not* follow a base stock policy, and therefore this outcome implies that most of the benefit that emerges comes from this simple policy approach, not from some other feature of the supply chain itself.

This is perhaps not surprising as the base-stock policy can be viewed as a special hyper-rational case of the behavioral policy, but is noted nevertheless. What is noteworthy here is that the static base stock policy is *simpler* than the static behavioral rule in so much that it can be collapsed down to only one parameter, versus the four in the behavioral case. Thus, for the case when dynamic learning does not occur, the simpler policy outperforms the more complex.

Furthermore, simply following the first-order rule suggested by behavioral operations literature to just set β = 1 is not sufficient by itself to meaningfully reduce supply chain cost. Similarly, statically optimizing for an expectation of encountering the 'average' team from literature does reduce costs on average, but also greatly increases the spread of outcomes, increasing the probability of introducing a destabilizing policy.

Figure 3-3 has another interesting implication, namely that for non-learning static policies, the simpler base-stock approach has greater cost reductions and conversely with learning policies a behavioral approach may have greater cost reductions. This is borne out by the ANOVA analyses shown in Table 3-3 and Table 3-4.

Table 3-3. ANOVA for Static Single-Shot Agents based on Rule Complexity

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Rule Complexity (Behavioral vs Base-Stock Static Agents)	1	18.17	18.174	81.95	2e-16***
Residuals	550	121.97	0.222		

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3-4. ANOVA for Learning Agents based on Rule Complexity

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Rule Complexity (Behavioral vs Base-Stock Learning Agents)	1	18	18.322	8.77	0.0.00308**
Residuals	2942	6142	2.088		

Note: *p<0.1; **p<0.05; ***p<0.01

The opposing effect of the base-stock versus behavioral architecture choice in the learning agents versus the static agents invites further analysis. Figure 3-4 shows the box-and-whisker plots of cost reduction for *all 16* experimental combinations described in Table 3-2. The prior analysis has shown differences between base-stock and behavioral rule complexity for these learning agents, but Figure 3-4 implies mixed effects for degree of information availability and possibly different effects by optimization objective as well.

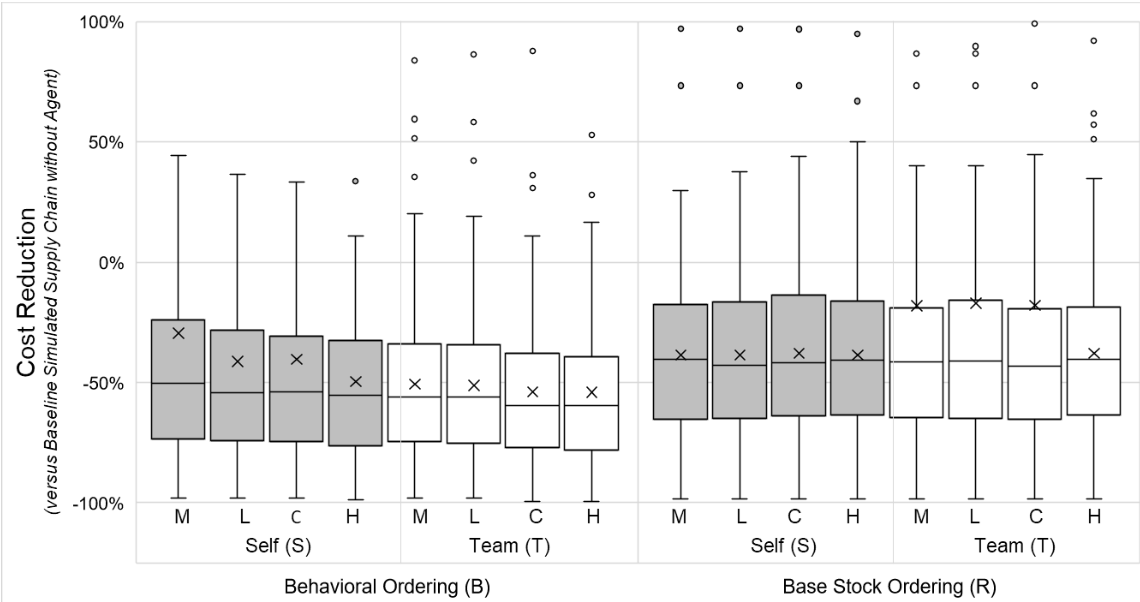


Figure 3-4. Learning Agent Induced Cost Reduction Across all Positions

To further explore these differences by policy feature in learning agents Table 3-5 shows a series of regression models against all the feature combinations described in Table 3-2. As all the independent variables here are factors, the summation of the regression coefficients in each of these regressions represents the predicted percent cost reduction of an agent with all the features used in the model, while the constant indicates the predicted percent cost reduction of the architecture being compared to the policy feature combination chosen as the baseline for comparison. While multiple choices are possible for this baseline policy, the analysis presented here chooses the simplest rule complexity, using the same combination of information and optimization features that would be present under the 'standard' conditions present in the original Beer Game. Specifically, the baseline is chosen to be the base-stock agent, utilizing a team-level optimization objective, with the standard level of information available in its calibration.

Model (1) in Table 3-5 focuses on the main effects of the features in Table 3-2, controlling for the specific team in the simulation.

$$Model\ (1):\ C_i = \beta_0 + \beta_1 f_{Behavioral} + \beta_2 f_{Self\ Focused} + \beta_3 f_{MinimalInfo} + \beta_4 f_{LowInfo} + \beta_5 f_{HighInfo} + \sum_N \beta_{Team_N} f_{team_N} + \epsilon_i \quad (13)$$

The constant is statistically significant and predicts an average reduction in cost of nearly 50 percent by simply using a general base-stock learning agent, with the standard amount of information about its environment and optimization based on team-level outcomes. The shift to an agent that correctly assumes its environment has behaviorally ordering others (vs the base-stock), further reduces the cost on average by an additional 16 percent, inverting the result seen in the static policies.

What is most striking about the regression results in Model (1) of Table 3-5 is the *lack* of significance of other feature choices. The main-effects of moving towards a self-focused reward (e.g., having an agent that is attempting to reduce the cost for only itself and not the entire supply chain) are insignificant. Similarly, the difference in moving from the standard information set towards a more minimal and hyper limited set of information for the calibration, or conversely towards a higher near omniscient set of information is also insignificant.

In Model (1), the influence of the agent in all four positions in the supply chain is aggregated. However, the significantly different base-stock values found in the above analyses for an idealized supply chain by position (Retailer = 0, Wholesaler = 24, Distributor = 8, Factory = 5) imply possibly different effects along the supply chain.

Model (2) in Table 3-5 incorporates these effects by position, again with fixed effects by specific simulated team. Here, the policy feature baseline is expanded to specifically refer to a base-stock agent, optimizing on team-level outcomes, calibrating using the standard information set, but at position 1 (the retailer) in the supply chain.

$$Model\ 2:\ C_i = \beta_0 + \beta_1 f_{Behavioral} + \beta_2 f_{Self\ Focused} + \beta_3 f_{MinimalInfo} + \beta_4 f_{LowInfo} + \beta_5 f_{HighInfo} + \beta_6 f_{Position2} + \beta_7 f_{Position3} + \beta_8 f_{Position4} + \sum_N \beta_{Team_N} f_{Team_N} + \epsilon_i \quad (14)$$

The main effects observed in Model (1) are largely maintained here, with a notable lack of influence of both optimization objective and information availability. However, as expected there are notable differences by position in the supply chain. An agent placed in Position 2 or 4 is significantly less able to reduce costs versus one placed in Position 1 or Position 3. Generally, as the agent is moved further upstream from the customer order signal, the cost reducing influence of the agent is diminished. Position 1 is unique in that it affects the flow of information into the supply chain from the external customer, and thus the cost reducing effect of a policy at this position is expected. Additionally, while Position 3 is in the middle of the supply chain, that position is typically the source of most of the team costs in real-world runs of the Beer Game, on which this simulation is based.

Finally, Model (3) introduces the interaction effect of a policy with complexity and the optimization objective, while still including the effects of specific position in the supply chain and maintaining team fixed effects.

$$\begin{aligned} \mathit{Model}\ 3\!:\ C_i = \beta_0 + \beta_1 f_{Behavioral} + \beta_2 f_{Self\ Focused} + \beta_3 f_{MinimalInfo} + \beta_4 f_{LowInfo} \\ + \beta_5 f_{HighInfo} + \beta_6 f_{Position2} + \beta_7 f_{Position3} + \beta_8 f_{Position4} \\ + \beta_9 (f_{Behavioral} * f_{NonMyopic}) + \sum_N \beta_{Team_N} f_{Team_N} + \epsilon_i \end{aligned} \qquad (15) \qquad (16)$$

Model (3) shows that the incorporation of a team-level optimization objective *can* be valuable if and only if the agent is following a behavioral policy (but only marginally so when considering the sum of the effects of β_1 and β_9). Additional interactions were explored but not included here as they were insignificant. The terms associated with the information states are maintained in these results as they are main effects from Table 3-2 and their insignificance is a surprising result of these analyses.

Table 3-5. Feature Influence for Learning Agents

	Dependent variable:		
	System Cost Change after Introducing Agent		
	(1)	(2)	(3)
β_0: Constant	-0.499***	-0.621***	-0.551***
	(0.185)	(0.190)	(0.192)
β_1: Behavioral Agent	-0.158***	-0.158***	-0.298***
	(0.052)	(0.052)	(0.073)
β_2: Self Focused	-0.017	-0.017	-0.157**
	(0.052)	(0.052)	(0.073)
β_3: Minimal Information	0.033	0.033	0.033
	(0.073)	(0.073)	(0.073)
β_4: Low Information	0.005	0.005	0.005
	(0.073)	(0.073)	(0.073)
β_5: High Information	-0.075	-0.075	-0.075
	(0.073)	(0.073)	(0.073)
β_9: Behavioral x Self Focused			0.280***
			(0.104)
β_6: Position 2		0.175**	0.175**
		(0.073)	(0.073)
β_7: Position 3		0.112	0.112
		(0.073)	(0.073)
β_8: Position 4		0.199***	0.199***
		(0.073)	(0.073)
Team Fixed effects	Yes	Yes	Yes
Model Baseline	Base-Stock Agent, Team Focused, Standard Information	Base-Stock Agent, Team Focused, Standard Information, at Position 1	Base-Stock Agent, Team Focused, Standard Information, at Position 1
Observations	2,944	2,944	2,944
R^2	0.069	0.072	0.074
Adjusted R^2	0.053	0.055	0.057
Residual Std. Error	1.408 (df = 2893)	1.406 (df = 2890)	1.405 (df = 2889)
F Statistic	4.307*** (df = 50; 2893)	4.238*** (df = 53; 2890)	4.304*** (df = 54; 2889)
Note:			*p<0.1; **p<0.05; ***p<0.01

This model-predictive learning agent can perform this cost reduction not necessarily by perfectly learning the ordering rules that govern its environment, but rather by simply having a model to learn in the first place. Figure 3-5 shows a measure of the error between the estimation of the environment learned by the agent and the actual environment ordering rules. As seen in the top of this figure, even when exposed to near perfect information about its environment, the agent may eventually learn a near accurate representation of its environment. However, when compared to the minimal information case as seen on the bottom of that same figure, the minimal case is both more likely to learn an incorrect representation of the environment but also is significantly less stable in its underlying model. However, even with this unstable and less accurate model of its surroundings, this minimal information case with a relative erroneous model of the environment *still* can perform statistically similarly to the near total information case.

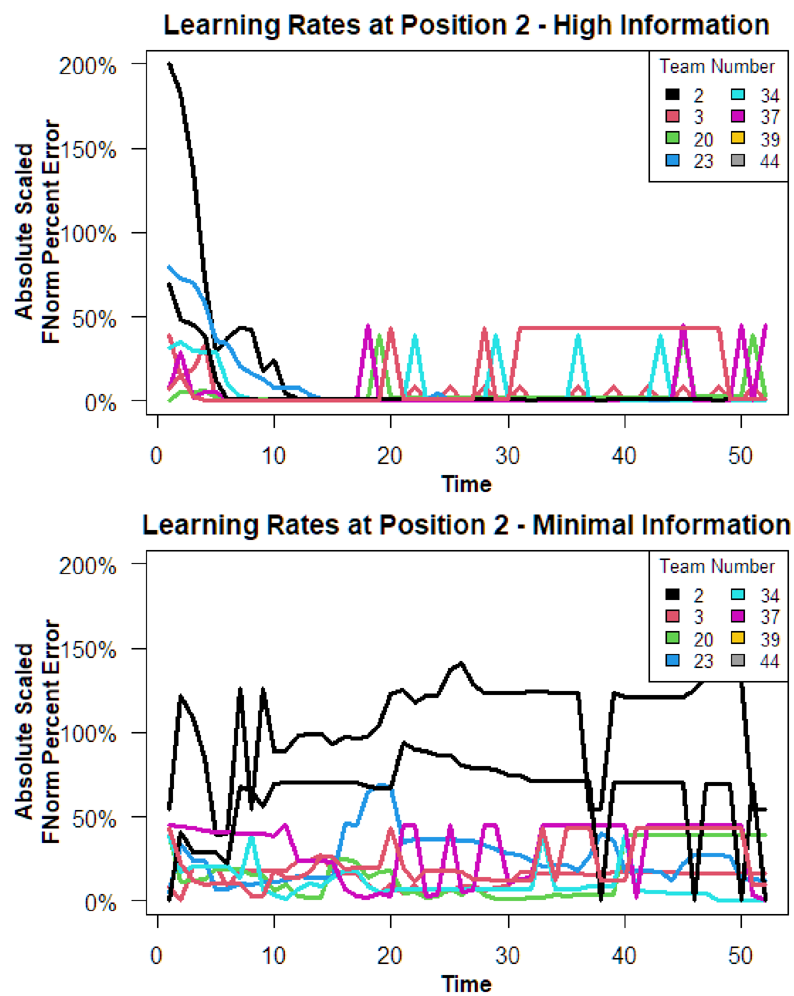


Figure 3-5. Sample of Learning Rates for Model-Predictive Learning Agents

3.6.2 Comparison to More Complex Machine Learning Methods

The results presented above compare features between static pre-trained policies with a relatively small number of free parameters, to those that follow the same architecture but are allowed to dynamically learn over time. More complex policies are of course plausible, including those that discover the relationship between the observable state space, the action taken by the agent, and the reward directly via repeated interactions.

The Beer Game itself has been used as a training environment in such reinforcement learning applications, most notably utilizing various modifications of Deep Q-Network architectures, including modifications to allow for independent training across entities utilizing pooled reward schemes (Chaharsooghi et al., 2008; Opex Analytics, 2018; Oroojlooyjadid et al., 2021). The Beer Game as a model of a multi-echelon supply chain presents a challenge to direct application of DQN architecture, challenges that are often also found in real supply chains. Specifically challenges emerge from 1) the true ‘full state’ of information is unknown to any one entity, 2) rewards are communal and not realized until the end of the time horizon, 3) DQN architectures can be ‘over optimistic’ in even mildly noisy environments (Thrun & Schwartz, 1993), and perhaps most importantly 4) the current overarching quality of the system, e.g. whether bullwhip is in progress or if the supply chain is stable, matters almost as much if not more than any specific action.

In order to address the above issues, most notably the final point in the above list, this chapter presents a DQN architecture for use in multi-echelon supply chains like the Beer Game that has the following general architecture: 1) An ‘order-plus’ action space (Oroojlooyjadid et al., 2021) which both allows for unbounded ordering in absolute terms and follows from observations in the model-based approach above, 2) a dual DQN network (Z. Wang et al., 2016) that separately maintains a value function estimation for both the current overarching combined state of the system and separately for each state-dependent action, 3) an observation space defined over a window of prior state observations corresponding to the signal delay in the system, 4) a combination of epsilon-greedy and Boltzmann exploration policies (Wiering, 1999), and finally 5) three sequential dense layers with ReLu activations of 256, 128, and 64 free parameters respectively for a total of 448 free parameters for the ‘single-shot’ version of this policy architecture.

The environment itself is built on the same framework utilized in the other architectures described above, with the functionalized form of the Beer Game translated into the commonly used opensource Gym research framework developed by OpenAI (Brockman et al., 2016). This environment allows for training against all positions in the simulated supply chain against randomly bootstrapped assemblages of human-like players, whose ordering rules are drawn from classical supply chain literature (Sterman, 1989). Additionally, this agent can be trained against random (but bounded) simulation horizons to avoid over-learning endgame-dependent policies, and even noisy realizations of ordering decisions. An illustration of this framework is shown in Figure 3-6.

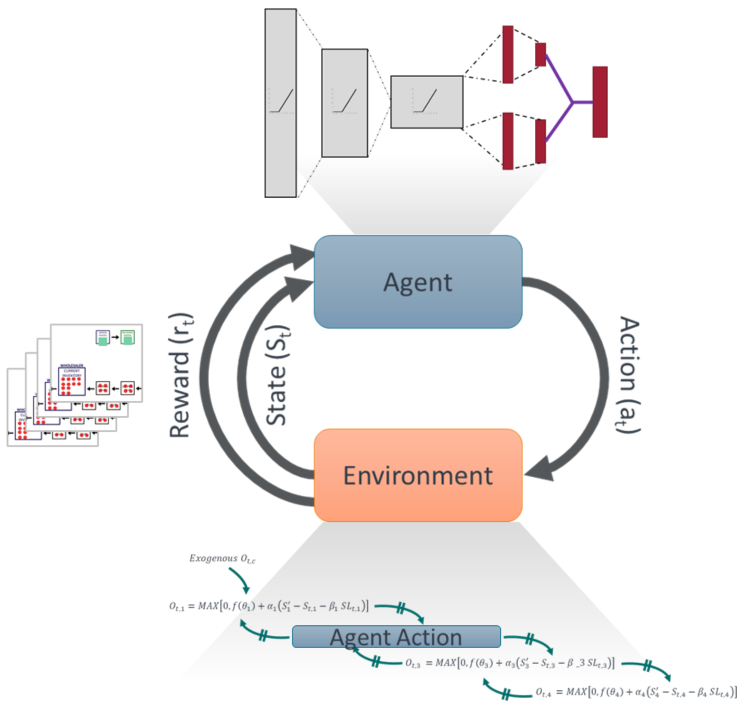


Figure 3-6. Cost Minimization Framework for the Model-Free DQN Approach

The above structure is closer to the 'single-shot' agents described above in so much as the agent is unaware of its surroundings, but instead tunes a set of free parameters ahead of time that allow it to map a current observed state to a desired action. This DQN can be further improved by developing a model aware DQN structure, which is identical to the DQN described above, but in addition to the base set of observations about its environment also incorporates estimates of the ordering parameters of the other agents in the supply chain. In this manner, this environment-aware DQN can then be used directly in the model-predictive learning structure described above, with the calibration based on history sill occurring, but then this estimate of the system being used as a state input into the DQN.

However, while development of this more complex method does provide some methodological contribution, it does not necessarily provide more cost reducing benefit, as seen in Figure 3-7. The DQN is able to learn a model of its environment with sufficient success to act as cost-reducing policy but does no better than the Behavioral Learning agents developed above.

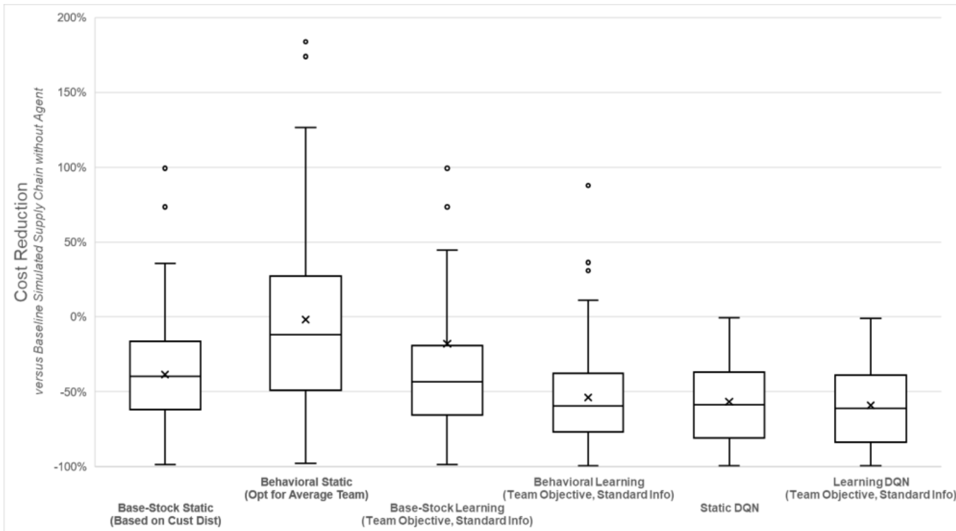


Figure 3-7. Static vs Learning vs DQN Agent Induced Cost Reduction Across all Positions

Indeed, the relative cost reduction, if any, gained by these agents comes at the cost of orders of magnitude of more complexity versus the simpler methods. One of the key differences between each of the above-described architectures is the degree of complexity of the agent. For the ‘single-shot’ agents, their parameters are determined *a priori* and held fixed over the course of the simulation. For the base-stock ‘single-shot’ agent this means that agent has only one parameter, the fixed desired safety-stock, while for the behavioral version of this same agent, the number of parameters is equal to the number of free parameters in the behavioral heuristic rule, which for this chapter is the four from expressions (11) and (12). For the model-predictive learning agents, the number of free parameters is slightly more difficult to enumerate as the agent is updating the value of the underlying parameters determining its order response each period. For a simple approximation of complexity, we can treat each of these decision points as another set of parameters. Finally, for the DQN the number and value of free parameters is fixed *a priori* like in the ‘single-shot’ case, but much larger. For the specific DQN architecture used here, 448 parameters in total across three layers.

Figure 3-8 illustrates this tradeoff between complexity and performance by recasting Figure 3-7, focusing on the median cost reduction of each architecture type as a function of complexity. Note the logarithmic scale for the x-axis.

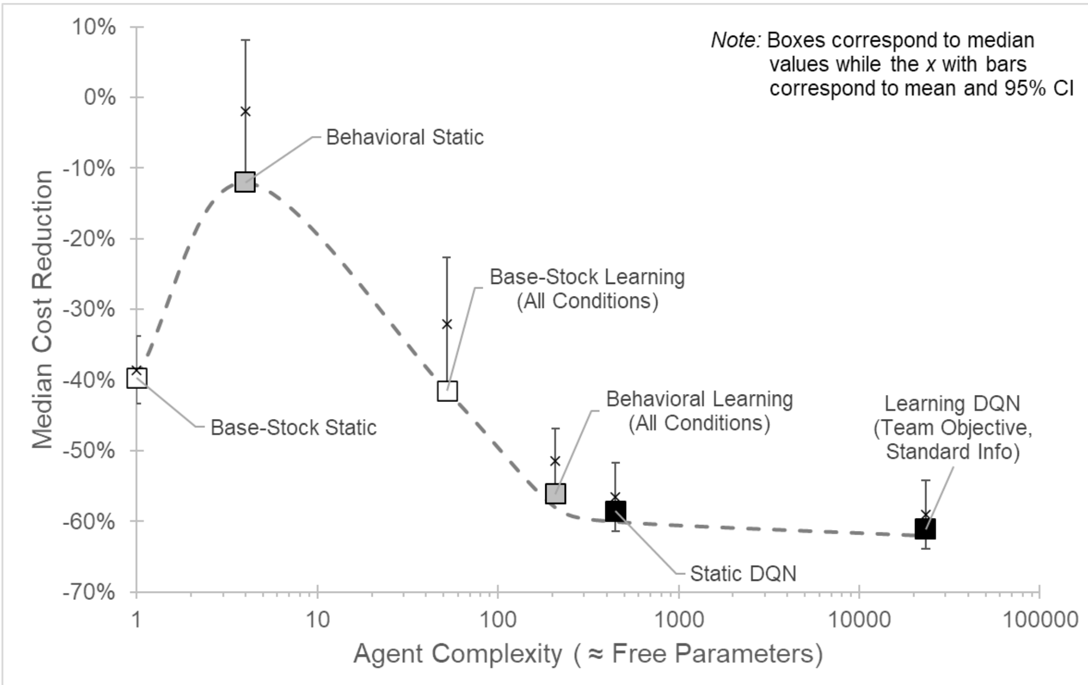


Figure 3-8. Median Cost Reduction of Policies as a Function of Approximate Complexity

3.7 Discussion

The most striking outcome of the results above is that the relatively naïve ordering rules that simply follow a fixed order response policy achieve most of the cost reduction that can be expected, especially when dynamic learning is not possible. This follows from the related work in psychology and behavioral economics referenced in the introduction to this chapter that have noticed that consistent and simple policies can often perform as well as more complex ones. As stated by others, Behavioral Operations Management “requires an operations context” while also acknowledging the presence of “potentially non-hyper-rational actors in [that] operational context” (Croson et al., 2013). Thus, it should not be surprising that these results, placed in a realistic operational context, follow similar observations from other fields in which non-hyper-rational decision makers exist.

Furthermore, the fact that the policies that assumed behavioral responses performed better in general, when dynamic learning is possible, than those that assume base-stock responses from their fellow entities in the supply chain is not surprising. At minimum, this is a result of simply better matching the underlying system in which the agent is placed, and as stated above base-stock policies are a hyperrational special case of the behavioral policies explored here.

While not included in this chapter for compact presentation, it should be noted that even the behavioral-based agent, when also combined with model-predictive learning structures, is able to perform well when placed in an environment with perfectly rational base-stock other agents. This is because such an agent, while still assuming behavioral responses from the environment, can adapt quickly. Stated differently and as seen in the results above, even learning a ‘wrong’ model of the environment can result in improved performance. However, this is not necessarily a full exploration of the explosion of the ‘Rule Complexity’ factor seen in Table 3-2, which would consider the rule complexity followed by the agent separately from the assumed rule complexity followed by the rest of the supply chain separately. As discussed in the main text above, the simplification here follows from the assumptions that a manager following either policy would likely make. The non-diagonal combinations of these two sub-features (when the structure of the policy is purposefully mismatched with the assumed policy of others) could also interact with the underlying actual decision rule being used by the supply chain. That interaction effect is abstracted away from in this work in part by simplifying these two dimensions into the single ‘Rule Complexity’ feature. However, this does open up interesting future avenues of exploration.

The learning method itself applied in this chapter is novel insomuch as it provides one of the few examples of model-predictive learning methods applied in a dynamic supply chain setting that this author has been able to find. That this control mechanism is applied in a behavioral context, in which the model itself is uncertain and based on a combination of the physics of the system and the assumed psychology of the other players, is also a general extension of model-predictive control methods. The results here show that such model-predictive learning architectures do result in better performance than the simplest single-shot methods, though this improvement is secondary to the initial improvement received simply by having a stable policy to begin with.

These observations also allow for some discussion on the tradeoff spent on academic research into the areas associated with each of these policy architectures. Much work has been devoted recently in expanding the use of general-purpose machine learning to complex nonlinear systems such as the supply chain modeled here. While that work is no doubt valuable, this research implies that the largest improvements come from simpler, and more interpretable, policies.

That is not to say that there is no value in developing these more complex tools, but rather in recognizing the tradeoffs inherent in the resources expended in their design, context-specific customization, and on-the-ground application. In developing the observations central to this chapter, the author has also built recent work illustrating the application of DQN structures in a multi-echelon supply chain setting. This was done leveraging insights from prior literature developed from directly applying similar structure to the Beer Game (notably Chaharsooghi et al., 2008 and Oroojlooyjadid et al., 2021), but while also leveraging a dueling architecture to overcome some of the computational difficulties previously faced. However, the resulting tool is specifically trained to this environment, dependent on specific state-space observations and a proper meshing of its outputs into the surrounding environment.

While this is of methodological interest in so much as it contributes to a growing body of literature on the mechanics of integrating DQN architecture into supply chain settings, the greater contribution of this chapter is the observation from above. Namely, that significantly simpler and more directly interpretable polices can achieve similar levels of performance to these more complex methods.

'Stable' here does not mean 'static.' The 'one-shot' architecture is both stable and static, while the model-predictive architecture is stable but dynamic. Stable in this context refers to consistent application of a rule to transform a set of observed inputs into a given decision or output. The 'one-shot' agent defined in this chapter has a pre-determined rule, but one that is follows consistently as the system progresses. The model-predictive learning agent similarly has an assumption that a stable response rule will reduce costs, but not that it necessarily knows that rule a priori.

Of additional interest in this model-predictive context is that the traditional tradeoff of computational time from better model fidelity and performance outcomes may not be as stark as previously implied in other MPC applications. Here, even the most minimal information about the state of the system was adequate to achieve a flexible enough policy that could adapt and shift over time in response to a changing environment.

For this simple linear supply chain, this is because the system itself is nearly fully defined by the small amount of information to which the agent has access. There are no losses of physical material, and the agent has no direct control over how the other entities in the system will respond to orders and shipments, only respond to how their own on-hand inventory is changing over time as a function of their order signals. More information about the state of the system is advantageous insomuch as it allows the manager to more rapidly settle on a stable model of his or her environment.

Additionally, the tradeoff that does exist between model accuracy and overall performance, conditional on information availability, becomes clearer when considering the number of instances in which the agent is destabilizing. In the analyses presented here, such destabilizations were still rare, though did occur more frequently with both less information about the overall state of the system and greater divergences between the assumed model and the underlying reality, and between the goal of the agent and the goal of the actual cost structure. Thus, the tradeoff here is not represented in terms degrees of performance and computational cost, but rather between external validity of the model-predictive process and absolute risk.

Specifically, to what degree is the manager limited to some physical or behavioral reality in a real system? Does the manager have access to the entire state of the environment, or just a more siloed view? For a fully integrated supply chain with entities held to a performance standard defined by overall team outcomes, then perhaps the more near-omniscient policy for this manager with global goals is realistic. Additionally, this chapter assumes a simple linear and serial supply chain, in which each step of the process has one customer and one supplier. In a more realistic branching supply chain, additional information about *which* supplier is best able to fill orders may be valuable but is outside the scope of this chapter.

Indeed, for a manager in a supply chain that is not integrated and rather simply part of a value-add chain of independent organizations, then the locally focused ‘greedy’ goal with limited information may be the only realistic option. What is interesting here is that it is possible, in fact likely, to still be able to implement a model-predictive learning policy in this most restrictive context and still achieve cost reductions versus no policy.

This chapter shows that the manager can, generally, be locally-focused and achieve global benefits, but not always. Having more information to discern an accurate model of the supply chain is not necessarily needed in the long-run, nor is having a more global the goal, but both correspond to reduced risk of a destabilizing outcome.

In net, a supply chain management policy that first is consistent in its application of a rule, even a simple and static base-stock ordering mechanism first proposed over sixty years ago, can achieve meaningful benefits, which can also be viewed as a hyperrational special case of many behavioral rules. For managers that find themselves embedded in the middle of a multi-echelon supply chain, explicitly incorporating models of human behavior can further improve results especially if learning about the surrounding supply chain is not possible. Learning about the surrounding environment and adjusting allows for flaws in initial assumptions to be revealed and corrected in time, but when such learning is not feasible then base-stock policies remain the best choice. While more complex and exotic general machine learning methods like the DQN presented here are potentially valuable, they come at a large start-up cost and additional complexity. Ultimately, it is the simpler policies that are (sometimes) the better choices for achieving cost reduction in even complex and behaviorally driven Operations Management environments.

3.8 References to Chapter 3

Åström, K., Albertos, P., Blanke, M., Isidori, A., Schaufelberger, W., & Sanz, R. (Eds.). (2001). *Control of Complex Systems*. Springer London. https://doi.org/10.1007/978-1-4471-0349-3

Bamakan, S. M. H., Malekinejad, P., Ziaeian, M., & Motavali, A. (2021). Bullwhip effect reduction map for COVID-19 vaccine supply chain. *Sustainable Operations and Computers*, *2*, 139. https://doi.org/10.1016/J.SUSOC.2021.07.001

Bendoly, E. (2013). Real-time feedback and booking behavior in the hospitality industry: Moderating the balance between imperfect judgment and imperfect prescription. *Journal of Operations Management*, *31*(1–2), 62–71. https://doi.org/10.1016/j.jom.2012.06.003

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). *OpenAI Gym*. http://arxiv.org/abs/1606.01540

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (2005). A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM Journal on Scientific Computing*, *16*(5), 1190–1208. https://doi.org/10.1137/0916069

Chaharsooghi, S. K., Heydari, J., & Zegordi, S. H. (2008). A reinforcement learning model for supply chain ordering management: An application to the beer game. *Decision Support Systems*, *45*(4), 949–959. https://doi.org/10.1016/j.dss.2008.03.007

Chen, F. (1999). Decentralized supply chains subject to information delays. *Management Science*, *45*(8), 1076–1090. https://doi.org/10.1287/mnsc.45.8.1076

Chen, F., & Samroengraja, R. (2009). The Stationary Beer Game. *Production and Operations Management*, *9*(1), 19–30. https://doi.org/10.1111/j.1937-5956.2000.tb00320.x

Ciocan, D. F., & Farias, V. (2012). Model Predictive Control for Dynamic Resource Allocation. *Mathematics of Operations Research*, *37*(3), 501–525. https://doi.org/10.1287/moor.1120.0548

Clark, A. J., & Scarf, H. (1960). Optimal Policies for a Multi-Echelon Inventory Problem. *Management Science*, *6*(4), 475–490. https://doi.org/10.1287/mnsc.6.4.475

Croson, R., & Donohue, K. (2006). Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science*, *52*(3), 323–336. https://doi.org/10.1287/mnsc.1050.0436

Croson, R., Donohue, K., Katok, E., & Sterman, J. (2014). Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, *23*(2), 176–196. https://doi.org/10.1111/j.1937-5956.2012.01422.x

Croson, R., Schultz, K., Siemsen, E., & Yeo, M. L. (2013). Behavioral operations: The state of the field. *Journal of Operations Management*, *31*(1–2), 1–5. https://doi.org/10.1016/j.jom.2012.12.001

Cui, T. H., & Wu, Y. (2018). Incorporating Behavioral Factors into Operations Theory. In *The Handbook of Behavioral Operations* (pp. 89–119). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119138341.ch3

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*(2), 95–106. https://doi.org/10.1037/h0037613

Ellram, L. M. (2010). Introduction to the forum on the bullwhip effect in the current economic climate. *Journal of Supply Chain Management*, *46*(1), 3–4. https://doi.org/10.1111/j.1745-493X.2009.03178.x

Evenett, S. J. (2021). *How big of a vaccine surplus will the US have?* Brookings Institution Reports2. https://www.brookings.edu/blog/future-development/2021/05/04/how-big-of-a-vaccine-surplus-will-the-us-have/

Gino, F., & Pisano, G. (2008). Toward a theory of behavioral operations. *Manufacturing and Service Operations Management*, *10*(4), 676–691. https://doi.org/10.1287/msom.1070.0205

Größler, A., Thun, J. H., & Milling, P. M. (2008). System dynamics as a structural theory in operations management. *Production and Operations Management*, *17*(3), 373–384. https://doi.org/10.3401/poms.1080.0023

Hockett, M. (2020). The Pandemic's Bullwhip Effect on Food Manufacturers' Inventory. *Food Manufacturing*. https://www.foodmanufacturing.com/supply-chain/article/21140181/the-pandemics-bullwhip-effect-on-food-manufacturers-inventory

Huang, T., Allon, G., & Bassamboo, A. (2013). *Bounded Rationality in Service Systems*. *15*(2), 263–279. https://doi.org/10.1287/msom.1120.0417

Johnson, B. (2021). In an Age of Abudance, Why do People Starve? *MIT Technology Review*, 74–79.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. William Collins.

Kelber, J. (2020, September 1). Are Neural Networks and Deep Learning the Next Big Supply Chain Trends? *Flexis*. https://blog.flexis.com/neural-networks-supply-chain-trends

Lee, H. L., Padmanabhan, V., & Whang, S. (2004). Information distortion in a supply chain: The bullwhip effect. *Management Science*, *50*(12 SUPPL.), 1875–1886. https://doi.org/10.1287/mnsc.1040.0266

Makridakis, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions and implications. *International Journal of Forecasting*, *16*(4), 451–476. https://doi.org/10.1016/S0169-2070(00)00057-1

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, *36*(1), 54–74. https://doi.org/10.1016/j.ijforecast.2019.04.014

Martin, M. K., Gonzalez, C., & Lebiere, C. (2004). Learning to make decisions in dynamic environments: ACT-R plays the beer game. In M. Lovett, C. Schunn, C. Lebiere, & P. Munro (Eds.), *Proceedings of the Sixth International Conference on Cognitive Modeling: ICCCM 2004: Integrating Models* (Vol. 420, pp. 178–183). Lawrence Erlbaum Associates Publishers.

Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica*, *50*(12), 2967–2986. https://doi.org/10.1016/j.automatica.2014.10.128

Morrison, J. B., & Oliva, R. (2018). Integration of Behavioral and Operational Elements Through System Dynamics. In *The Handbook of Behavioral Operations* (pp. 287–321). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119138341.ch8

Narayanan, A., & Moritz, B. B. (2015). Decision Making and Cognition in Multi-Echelon Supply Chains: An Experimental Study. *Production and Operations Management*, *24*(8), 1216–1234. https://doi.org/10.1111/poms.12343

Oliva, R., Abdulla, H., & Gonçalves, P. (2022). Do Managers Overreact When in Backlog? Evidence of Scope Neglect from a Supply Chain Experiment. *Manufacturing & Service Operations Management*. https://doi.org/10.1287/msom.2021.1072

Opex Analytics. (2018). *The Beer Game*.

Oroojlooyjadid, A., Nazari, M., Snyder, L., & Takáč, M. (2017). A Deep Q-Network for the Beer Game: A Deep Reinforcement Learning algorithm to Solve Inventory Optimization Problems. *Arxiv*, 1–38.

Oroojlooyjadid, A., Nazari, M., Snyder, L. V., & Takáč, M. (2021). A Deep Q-Network for the Beer Game: Deep Reinforcement Learning for Inventory Optimization. *Manufacturing & Service Operations Management*, msom.2020.0939. https://doi.org/10.1287/msom.2020.0939

Pannek, J., & Frazzon, E. (2014). Supply Chain Optimization via Distributed Model Predictive Control: Supply Chain Optimization via Distributed Model Predictive Control. *PAMM*, *14*(1), 905–906. https://doi.org/10.1002/pamm.201410433

PYMNTS. (2022, April 11). *Healthcare AI Better at Back Office Than Back Surgery* [PYMNTS]. https://www.pymnts.com/healthcare/2022/healthcare-artificial-intelligence-back-office/

Seborg, D. E., Edgar, T. F., Millichamp, D. A., & Doyle III, F. J. (2016). *Process Dynamics and Control* (4th ed.). Wiley.

Secomandi, N. (2008). An Analysis of the Control-Algorithm Re-solving Issue in Inventory and Revenue Management. *Manufacturing & Service Operations Management*, *10*(3), 468–483. https://doi.org/10.1287/msom.1070.0184

Shih, W. (2020). *COVID-19 And Global Supply Chains: Watch Out For Bullwhip Effects*. Forbes. https://www.forbes.com/sites/willyshih/2020/02/21/covid-19-and-global-supply-chains-watch-out-for-bullwhip-effects/?sh=2fa2b2467195

Stank, T., Goldsby, T., & Saunders, L. (2021). Commentary: Caution—Bullwhip Effect Ahead. *The Wall Street Journal*. https://www.wsj.com/articles/commentary-cautionbullwhip-effect-ahead-11623664801

Sterman, J. (1989). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science*, *35*(3), 321–339. https://doi.org/10.1287/mnsc.35.3.321

Sterman, J. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World* (1st ed.). Irwin/McGraw-Hill.

Sterman, J., & Dogan, G. (2015). "I'm not hoarding, I'm just stocking up before the hoarders get here." Behavioral causes of phantom ordering in supply chains. *Journal of Operations Management*, *39–40*(1), 6–22. https://doi.org/10.1016/j.jom.2015.07.002

Sutton, R., & Barto, A. (2014). *Reinforcement Learning: An Introduction* (2nd ed.). The MIT Press.

Tarafdar, M., Beath, C. M., & Ross, J. W. (2019). Using AI to Enhance Business Operations. *MIT Sloan Management Review*, *60*(4), 37–44.

Thrun, S., & Schwartz, A. (1993). Issues in Using Function Approximation for Reinforcement Learning. *Proceedings of the 1993 Connectionist Models Summer School*, 255–263.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. https://doi.org/10.4324/9781912282562

Urban, G., Timoshenko, A., Dhillon, P., & Hauser, J. R. (2020). Is Deep Learning a Game Changer for Marketing Analytics? *MIT Sloan Management Review*, *61*(2), 71–76.

Vossen, T. W. M., You, F., & Zhang, D. (2022). Finite-horizon approximate linear programs for capacity allocation over a rolling horizon. *Production and Operations Management*, *31*(5), 2127–2142. https://doi.org/10.1111/poms.13669

Wang, X., & Disney, S. M. (2016). The bullwhip effect: Progress, trends and directions. *European Journal of Operational Research*, *250*(3), 691–701. https://doi.org/10.1016/j.ejor.2015.07.022

Wang, Z., Schaul, T., Hessel, M., Van Hasselt, H., Lanctot, M., & De Frcitas, N. (2016). Dueling Network Architectures for Deep Reinforcement Learning. *33rd International Conference on Machine Learning, ICML 2016*, *4*(9), 2939–2947.

Wiering, M. (1999). *Explorations in efficient reinforcement learning*. University of Amsterdam.

Wu, D. Y., & Katok, E. (2006). Learning, communication, and the bullwhip effect. *Journal of Operations Management*, *24*(6), 839–850. https://doi.org/10.1016/j.jom.2005.08.006

Yu, M. C., & Kuncel, N. R. (2020). Pushing the Limits for Judgmental Consistency: Comparing Random Weighting Schemes with Expert Judgments. *Personnel Assessment and Decisions*, *6*(2), 1–10. https://doi.org/10.25035/pad.2020.02.002

Appendices and Supporting Materials

Appendix A Chapter 1: Dynamic Supply Chains with Endogenous Dispositions

A.1 Model Availability

Accompanying the main chapter and this Appendix are the full models available as .mdl files, along with supporting data files to illustrate the how specific analyses were run and figures generated. The .mdl files can be open and run using Vensim software, developed by Ventana Systems, Inc. A free version of the Vensim software for personal use, along with a standalone model viewer, is available from Ventana Systems, Inc. These model files can be obtained directly at:

https://github.com/jpain3/MIT-Disseration/tree/main/chapter-1

The models were originally developed in Vensim version 8.2 and revised in version 9.1.1. As of Vensim version 9.0, the visual style of this software package has changed significantly. The .mdl files are still fully viewable these later versions of the software, but layout and text changes may make viewing the model slightly more difficult. The author suggests, if using Vensim version 9.0 or later, to view the associated .mdl files in the 'Traditional Sketch' or 'Old Sketch' layout. In Vensim version 9.0 this can be toggled via the Tools menu as seen below. All screenshots of software menus or images depicting the model layout or menu screens in the Appendix were done in this 'Traditional' view scheme.

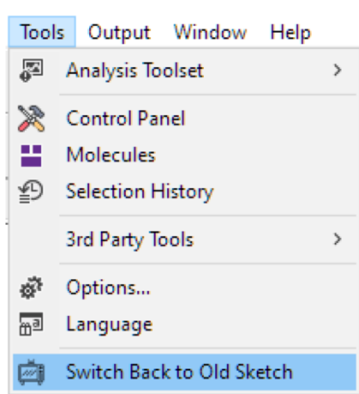


Figure A-1. Switching to 'Old Sketch' in Vensim 9.0 and later

Ventana Systems, Inc provides detailed documentation on the Vensim software, including how to manipulate and examine specific formulations. However, the reader may quickly explore the influence of parameter choices on the model via the SyntheSim mode on the main Dashboard view of the model. This can be accessed by pressing the corresponding button in the top toolbar of the software as seen below: 

For the supporting model comparing the core methodological framework, the .mdl file is divided into views: an overview Dashboard, a view of the full model itself, and several views that detail specific reporting or supporting structures. Different views can be accessed via the buttons in each view, or via the view menu. Examples of these two views (but not the entirety of these views) are provided below.

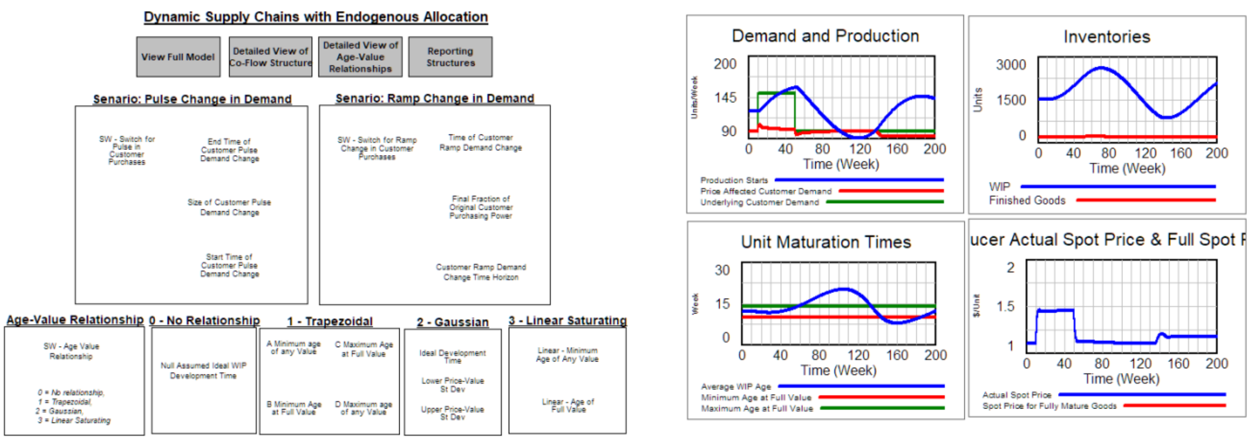


Figure A-2. Example of the Dashboard View of the Methodology Comparison Model

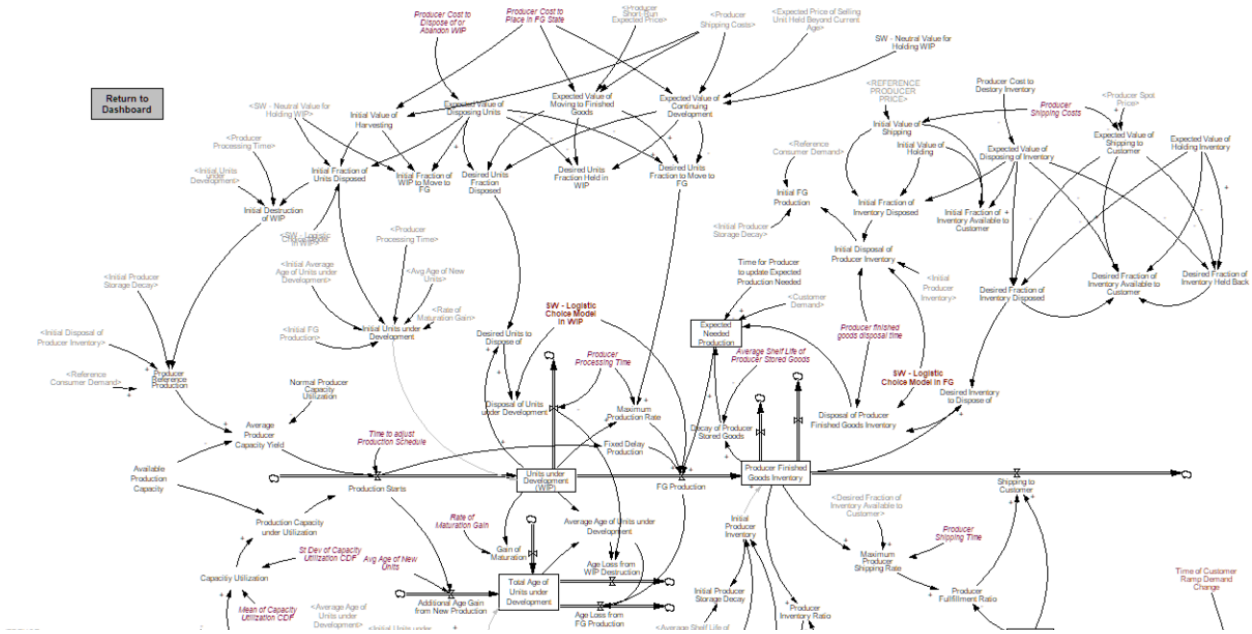


Figure A-3. Detail of Aggregate Framework Embedded in Full Model View of the Methodology Comparison Model

As a note, this model view is presented entirely and largely has no hidden structure or hidden causal connections. The model is still provided here for the interested reader and allows for the reader to investigate in detail how the frameworks developed in the main chapter are practically applied as subcomponents in a larger model, along with alternative value-age relationships.

For the model illustrating the details of the aggregate and vintaging framework, the presentation is designed to allow for comparison of the outputs of both frameworks when subjected to the same inputs. The .mdl file is divided into several views, most notably an overview Dashboard, and detailed views of each framework. Different views can be accessed via the buttons in each view, or via the view menu. Examples of these views (but not the entirety of these views) are provided below.

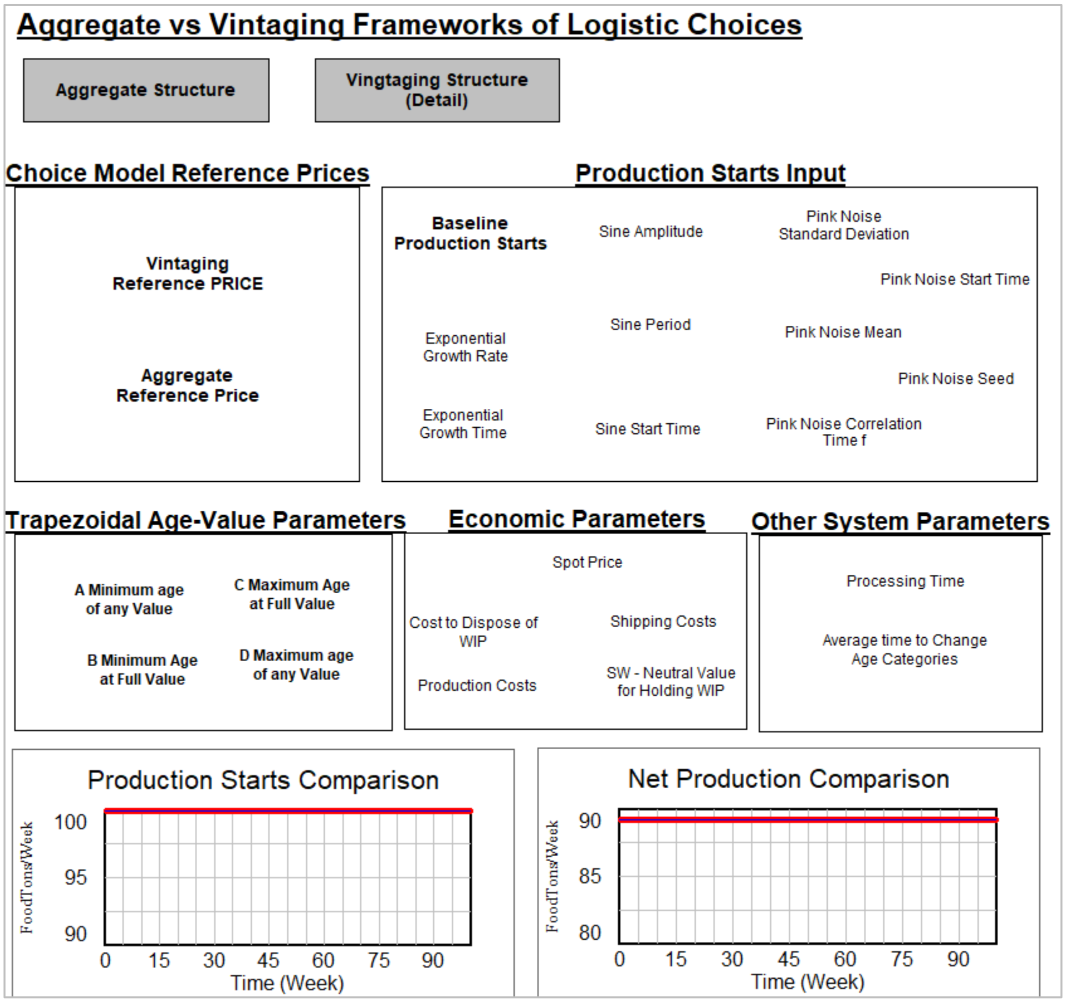


Figure A-4. Partial Example of the Dashboard View of the Framework Comparison Model in Vensim

Note that the vintaging framework is presented for 10 age cohorts, and this is fixed by design. This is to make the presentation of the framework direct and easy to interpret without the need for subscripting or array formulations. This presentation can be greatly simplified via array approaches but doing so hides the underlying interplay of choices in the vintaging structure. However, the cost of this choice is a highly cluttered display of the fully connected model, along with difficulty in adjusting the number of cohorts. This fully connected view is present in the .mdl Vensim file, but the author encourages readers to focus on the detailed view of the beginning and end of the illustrative vintaging chain for ease of understanding.

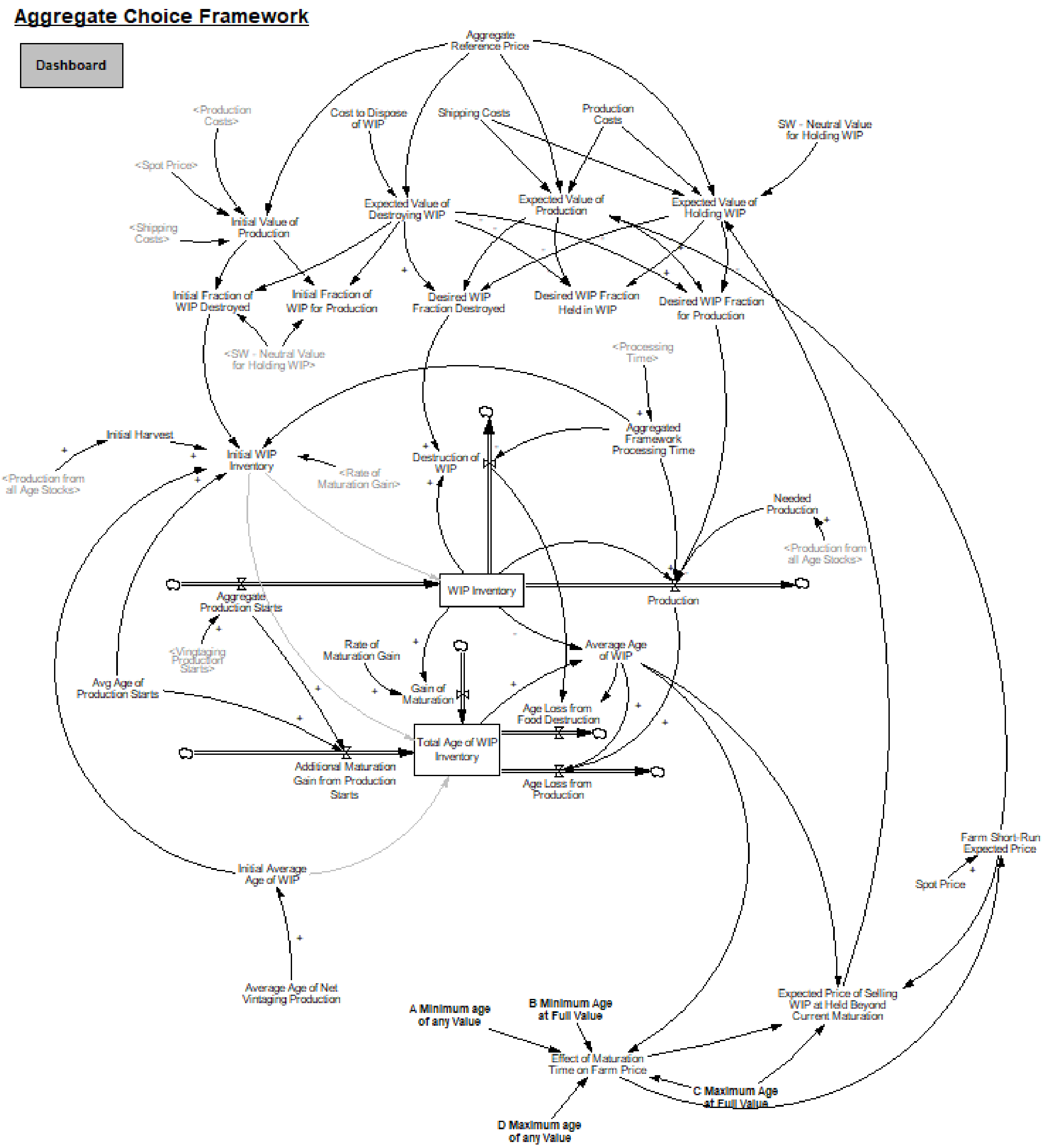


Figure A-5. Detail of Aggregate Framework View in the Framework Comparison Model

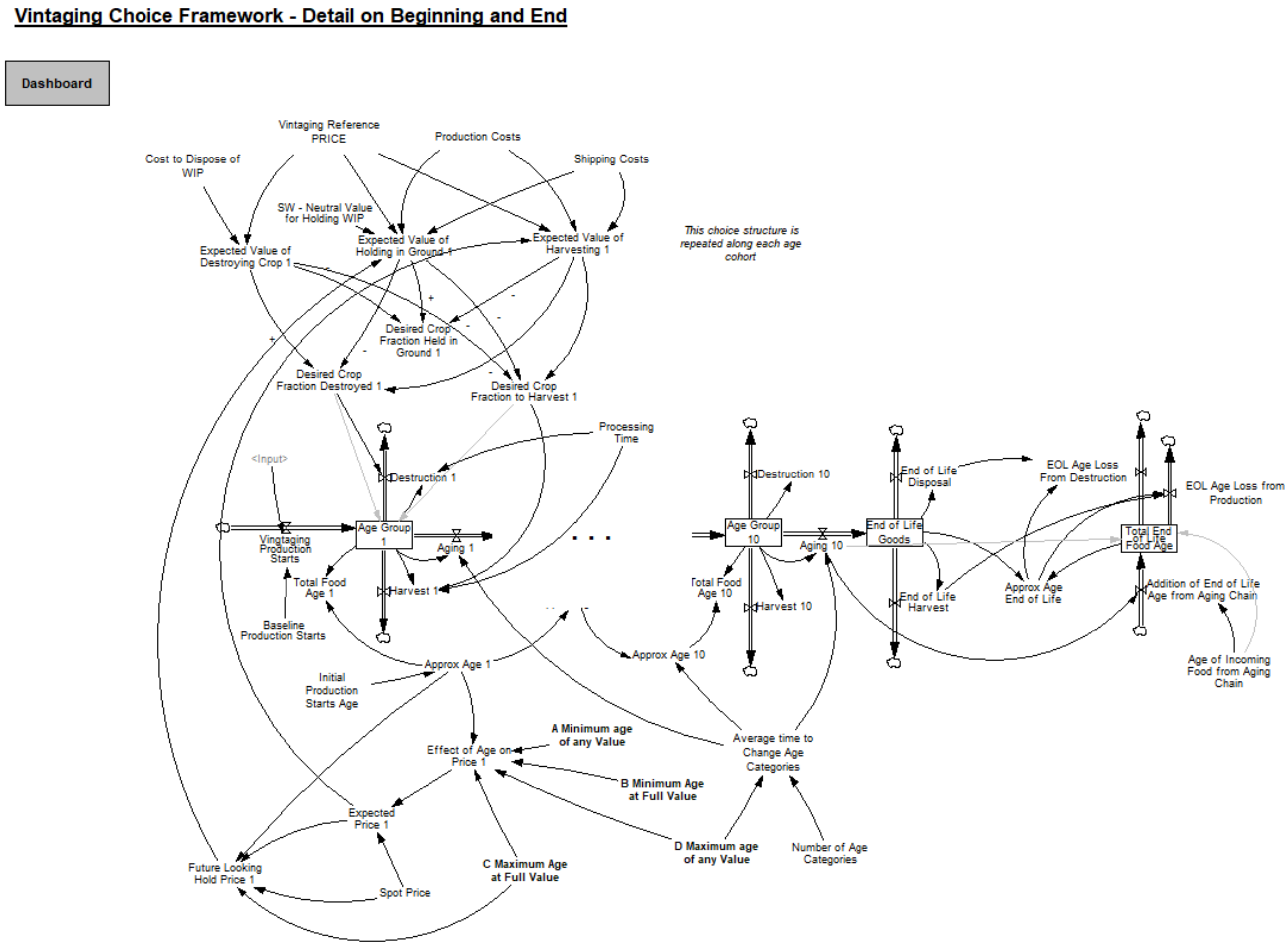
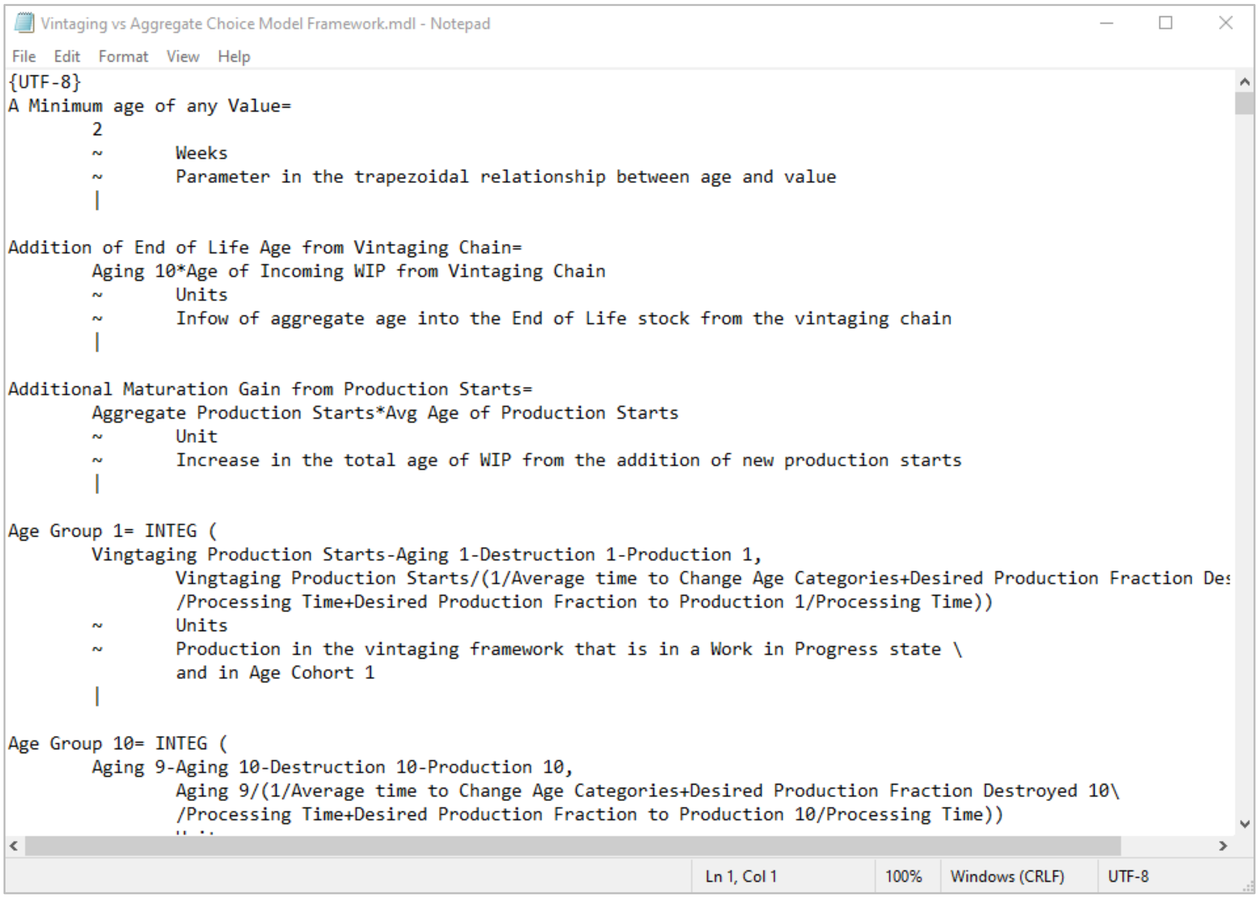


Figure A-6. Detail of Vintaging Framework View in the Framework Comparison Model

Furthermore, the .mdl files provided may be opened in any program that is able to read UTF-8 encoding and the formulations directly viewed in plaintext. Examples of programs that can open the .mdl file for direct viewing in plaintext include Notepad in the Windows operating system and Textpad in the Macintosh operating system. An example of this view of the model file is seen in Figure A-7.



Vintaging vs Aggregate Choice Model Framework.mdl - Notepad

File Edit Format View Help

```
{UTF-8}
A Minimum age of any Value=
        2
        ~       Weeks
        ~       Parameter in the trapezoidal relationship between age and value
        |

Addition of End of Life Age from Vintaging Chain=
        Aging 10*Age of Incoming WIP from Vintaging Chain
        ~       Units
        ~       Infow of aggregate age into the End of Life stock from the vintaging chain
        |

Additional Maturation Gain from Production Starts=
        Aggregate Production Starts*Avg Age of Production Starts
        ~       Unit
        ~       Increase in the total age of WIP from the addition of new production starts
        |

Age Group 1= INTEG (
        Vingtaging Production Starts-Aging 1-Destruction 1-Production 1,
                Vingtaging Production Starts/(1/Average time to Change Age Categories+Desired Production Fraction Des
                /Processing Time+Desired Production Fraction to Production 1/Processing Time))
        ~       Units
        ~       Production in the vintaging framework that is in a Work in Progress state \
                and in Age Cohort 1
        |

Age Group 10= INTEG (
        Aging 9-Aging 10-Destruction 10-Production 10,
                Aging 9/(1/Average time to Change Age Categories+Desired Production Fraction Destroyed 10\
                /Processing Time+Desired Production Fraction to Production 10/Processing Time))
```

Ln 1, Col 1 | 100% | Windows (CRLF) | UTF-8

Figure A-7. Example of Viewing the Supporting .mdl File in Notepad on Windows

The models developed for this chapter are also fully documented utilizing the SDM-Doc tool described in (Martinez-Moyano, 2012). The output from this documentation tool is available alongside the .mdl files.

A.2 Formulation Details for the Methodological Comparison Model

In the main chapter, a simplified supply chain model that allows for the switching on and off of methodologies is used as an illustrative example of the use of the frameworks developed in a larger model. The sections below provide additional detail on the development of that comparison model, focusing on details that are not necessary to illustrate the frameworks developed in the main chapter, but are still of interest in the dynamics in this overarching system. As a note, some portions of the explanatory text from the main chapter are repeated below where needed to create a self-contained description of this model.

The model consists of a producer who manages goods flowing through two stocks: Work in Progress and Finished Goods. The figure below provides a high-level visual overview of the model, with each subsequent section providing more operational detail.

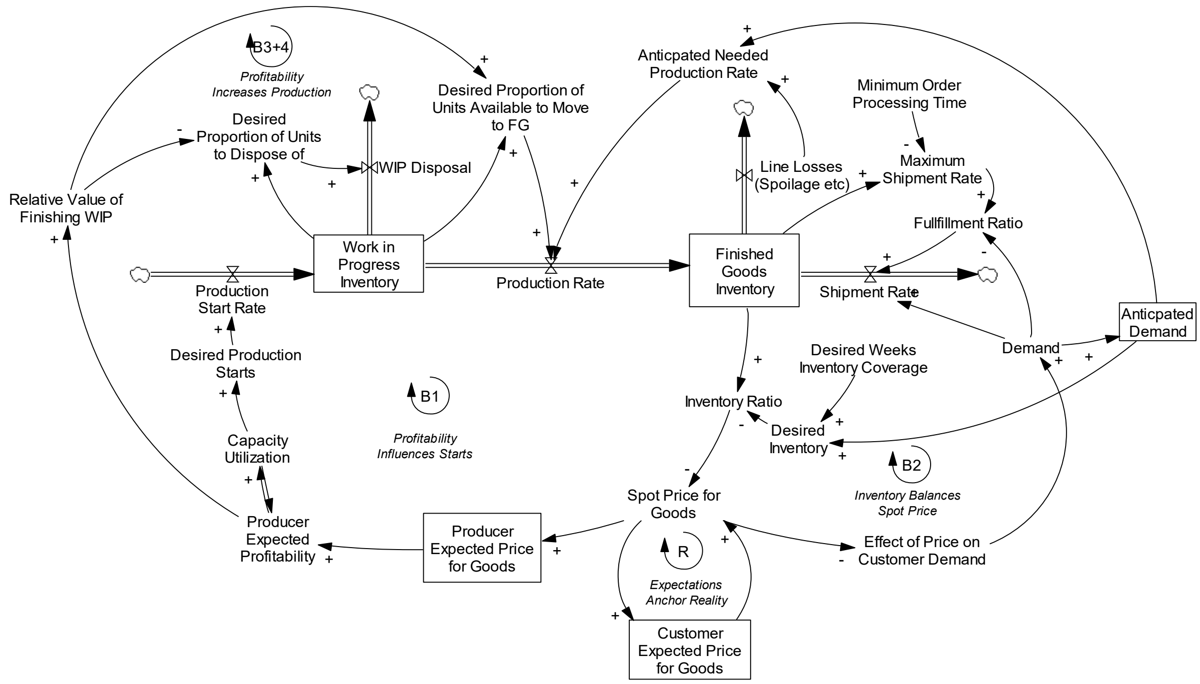


Figure A-8. Overview of Methodological Comparison Model

A.2.1 Defining the Market

While there may be multiple ways to construct the interplay of supply and demand that ultimately forms the spot price at each interface point in the market illustrated in Figure A-8, the loops defined as B2 and R utilize inventory-sensitive spot pricing (Chen et al., 2009; Sterman, 2000; Whelan & Forrester, 1996). These effects also feedback into B1 and even affect the loops that determine the relative value of each inventory disposition formed in B3+4.

The core of this economic model is two balancing loops across each entity in the supply chain, with spot pricing driving either demand or supply.

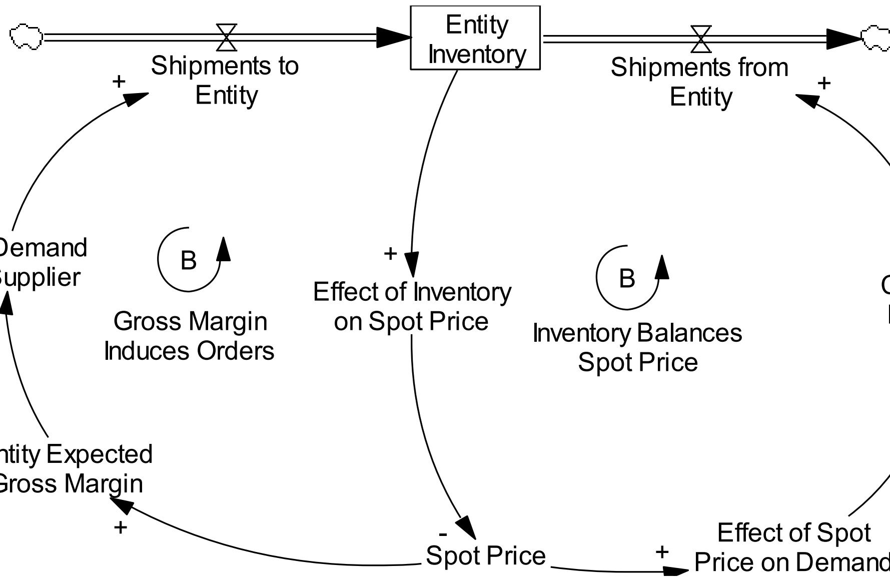


Figure A-9. Core Two Balancing Loops Inventory-Based Spot Prices

However, the above entity may exist in a chain of upstream and downstream entities, each ordering from their suppliers and selling to their own customers. This effects the 'Expected Gross Margin' and introduces another balancing loop. Additionally, the spot price is fundamentally anchored to what the market expects it to be, and this introduces a reinforcing loop around the spot price and the expected prices. These two new loops, in the context of the larger supply chain, are seen below:

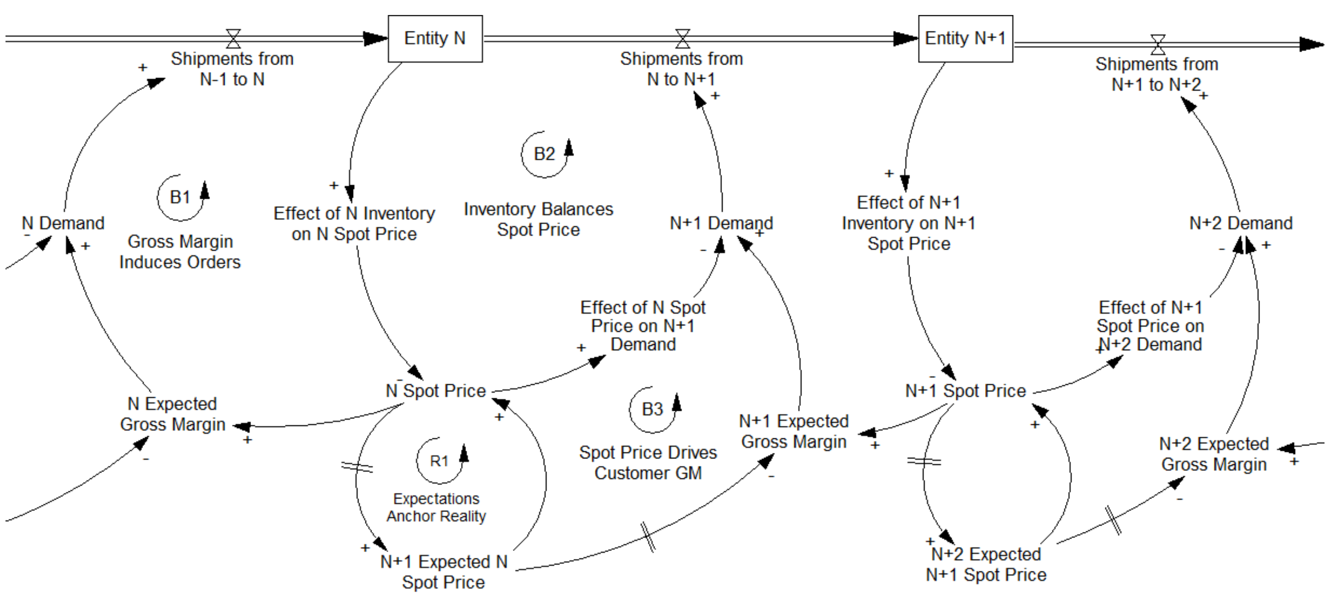


Figure A-10. Ordering and Price Setting is Nested in Larger Interconnected Supply Chain

A.2.1.1 Effect of Inventory Coverage on Price

One of the key features of the pricing model visually summarized above is the effect of inventory coverage on pricing. In net, a model will capture the downward sloping relationship between additional inventory (beyond a set inventory coverage goal) and the price offered by the firm holding that inventory.

$$\begin{aligned} &Effect\ of\ [Entity]\ Inentory\ Coverage\ on\ [Entity]Price \\ &\quad = [Entity]Inventory\ Ratio^{-Sensitivty} \end{aligned} \tag{17}$$

The *sensitivity* is a parameter that determines how much the price will raise or lower given a change in inventory coverage. As formulated here, *sensitivity* is assumed to be a positive value, with higher values corresponding to increasingly concave response functions

Another feature explored in the above formulation is a ‘cap’ on the maximum multiplier that inventory coverage could have on price. I.e., if inventory coverage approaches 0 (there is no inventory to sell), then the effect on the price will approach infinity. This does not happen as the increase in spot prices drives down demand from downstream customers, preventing the final marginal units of inventory from ever being sold in practice.

A.2.1.2 Effect of Expected Gross Margin on Demand

The concept of expected gross margin can be used to influence production in the case of the producer, and demand in the case of all other entities in the supply chain, with increases in expected Gross Margin assumed to induce greater production or demand.

There may be multiple methods of incorporating this relationship here, including truncated sigmoidal functions and directly applying table functions. For this example, consider a simple truncated linear representation that meets the following criteria:

1. Passes through the point of (1,1) on a normalized scale
2. Is truncated at an upper maximum multiple on demand
    a. This assumes that it is infeasible for an entity will ever request more than some multiple of its reference demand at any expected future profit level
    b. This could be due to several possible factors not explicitly modeled such as storage space constraints, transportation limitations, or risk of spoilage).
3. Is truncated at a lower level of demand

a. In other words, it is bounded at a minimum acceptable gross margin, which could be greater than 0%

b. Paratactically this means the line passes through the point of *(Minimum Normalized GM, 0)*

Given points 1 and 3 above, the slope of the line is defined fully by the specification of the minimum acceptable gross margin at which any demand or production will exist, and the definitions of the reference gross margin and corresponding reference demands. Examples of what this curve looks like can be seen in Figure A-11 below.

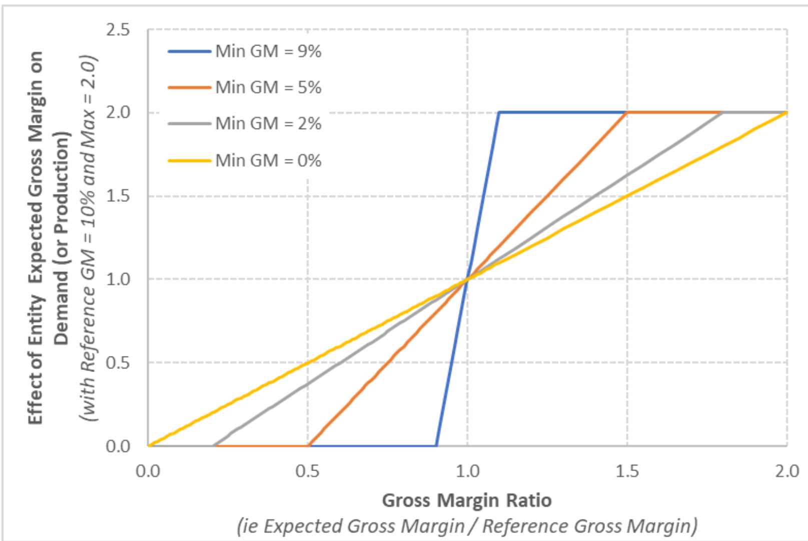


Figure A-11. Examples of the Formulation of Demand versus Expected Gross Margin

It should be emphasized that this curve is based on the *expected* gross margin to influence demand, which is in turn can be based on smoothed perceptions of previous prices that the entity has experienced.

A.2.1.3 Effect of Spot Prices on Demand

The above effect on demand due to expected gross margin does have some element of sensitivity to cost built in from the definition of gross margin. However, it is purposely done in relationship to an expected gross margin based on a smoothed view of previous prices (both costs for goods bought and the prices at which they were later sold).

To affect demand based on the instantaneous spot price experienced by each entity, consider a linearly decreasing relationship that captures decreasing demand with increasing prices, with the slope of that relationship affected by some elasticity of demand. The functional form of this expression is seen below:

$$\begin{aligned} &Effect\ of\ Price\ on\ Demand \\ &\quad = MIN\Big(Maximum\ Multiplier, MAX\Big(0{,}1 + Demand\ Curve\ Slope \\ &\quad * \frac{Price - Reference\ Price}{Reference\ Demand}\Big) \end{aligned} \tag{18}$$

Where:

$$Demand\ Curve\ Slope = \frac{-Reference\ Demand * Reference\ Elasticity}{Reference\ Price} \tag{19}$$

An example of the shape of this function for various values of elasticity are seen below, where *e* refers to the Reference Elasticity in expression (9) above.

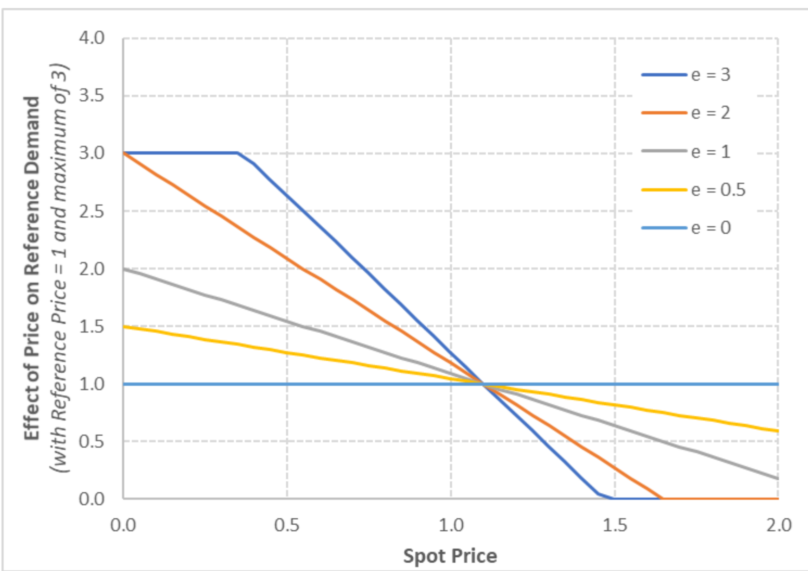


Figure A-12. Examples of the Formulation of Demand versus Spot Price

Here the spot price is purposefully used to determine the effect of this instantaneous demand. This is designed to be immediate, in contrast to the effect from expected gross margin which is based on a smoothed concept of both prices and costs.

Combined, the relationships described in the economic market for this commodity good in which the new modeling framework presented here can be applied.

A.2.2 Production Starts and Capacity Management

The producer considers two different conceptualizations of profitability: the incremental profitability of an additional unit of production (utilizing just the variable costs of production), and the expected profitability from expanding production capacity (utilizing a fully allocated cost of production).

As a note, in this example this relationship utilizes the 'Producer Expected Price' which is the spot price smoothed over a short time range. The producer considers the price relative to expected costs to form a gross margin estimation when making capacity change decisions. Here, this expected gross margin utilizes a *fully allocated* unit cost. This expected gross margin determines the effect on desired capacity.

As discussed in other System Dynamics models of commodity markets (notably chapter 20 of (Sterman, 2000)) utilization is a function of expected gross margin. Furthermore, utilization of existing capacity is unlikely to be at 100% when averaged across all pieces of owned capacity unless at very high levels of expected profitability. The exact shape of this relationship will vary by industry and even by individual producer or individual piece of owned unit of production capacity. To qualitatively capture this behavior, consider a function which approximates a curve approaching the CDF of a collection of different land (capacity) at different utilization depending on local factors. One such curve, and the one utilized in this example is shown in Figure 2-6.

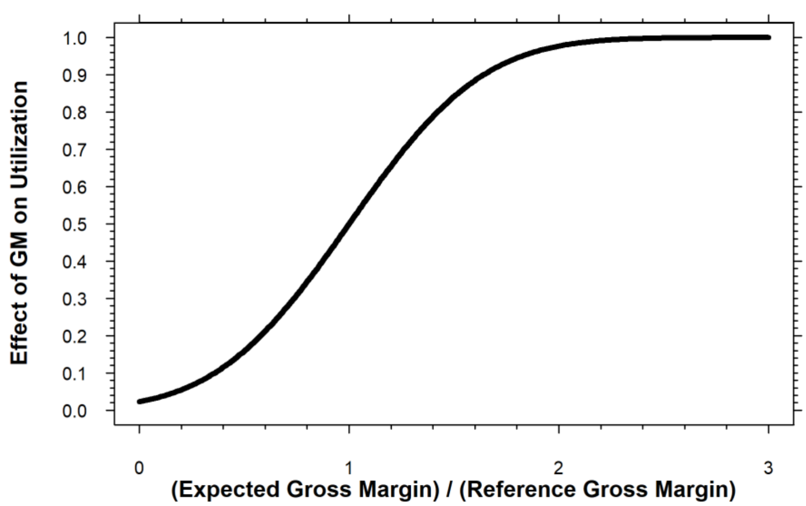


Figure A-13. Producer Capacity Utilization versus Expected Gross Margin

A.2.3 Discounting the Spot Price based on Development Age

A central piece of the framework presented in the central paper is the relationship between the age of the goods being produced and the value they derive in the marketplace. For the example in the main chapter, consider a relationship of the same trapezoidal functional form as that described in expression (22) and illustrated in Figure A-14. Furthermore, the example below utilizes the single aggregate stock of work-in-progress inventory instead of a more granular vintaging framework as described in the main chapter.

The quantification of the opportunity cost capturing the tradeoff between time that a unit of potential inventory spends under production (or development) versus the amount of economic value the producer can expect to get from its eventual sale is explored in more detail in the sections below.

$$Effect\ of\ Age\ on\ Price = f(Average\ Age\ of\ WIP) \tag{20}$$

$$\begin{aligned} &Producer\ Spot\ Price \\ &\quad = Effect\ of\ Age\ on\ Price * Spot\ Price\ for\ Full\ Mature\ Goods \end{aligned} \tag{21}$$

In the above, the 'Spot Price for Full Mature Goods' is defined via the method described in expression (6), and is a function of the inventory coverage of the producer.

A.2.4 Quantifying the Age-Value Relationship

As discussed in more detail in the main chapter, this relationship that defines 'Effect of Age on Price' is context-specific can vary depending on the product under development and how the market values that product as function of the development or maturation time.

For the example used in the model of a supply chain of a commodity product, this relationship can be summarized as first having a low value that rises until it reaches a peak of full value at an ideal maturation time, and then declines as it sits in the field either further maturing past its prime or even decaying.

To capture the above dynamics, a table function could be employed but for simplicity consider a trapezoidal relationship between crop value and age (or maturation time). This relationship utilizes four parameters to capture when a crop first has any economic value, the range over which it has full economic value, and the age above which it again has no economic value.

$$f(t) = \begin{cases} 0 & t \le a \\ \left(\frac{1}{b-a}\right)t - \left(\frac{a}{b-a}\right) & a < t \le b \\ 1 & b < t \le c \\ \left(\frac{1}{c-d}\right)t - \left(\frac{d}{c-d}\right) & c < t \le d \\ 0 & t > d \end{cases} \tag{22}$$

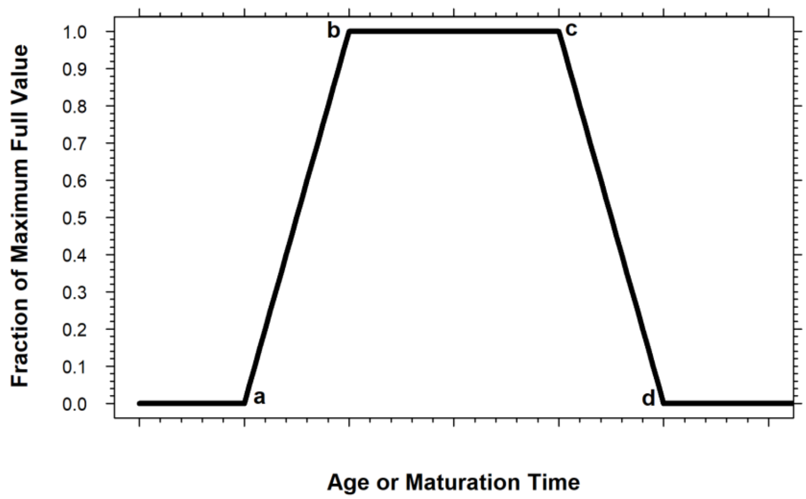


Figure A-14. Example of Trapezoidal Function Discounting the Value of Crops based on Maturation

Note that the expression above assumes a linear change from minimum to maximum value, and constant maximum value between points *c* and *d*. More general trapezoidal shapes are possible that do not necessarily have these features (for example see (Dorp & Kotz, 2003)) and may be more appropriate in specific contexts, but this formulation is sufficient here.

Under the aggregate model framework, which is used in this example supply chain model, the value of the entire stock of work-in-progress inventory is derived by the formulation above. If the vintaging model were used, it would be applied to each cohort of ages.

A.2.5 A Multinomial Logistic model of Inventory Dispositions

The section immediately below largely restates material in the main chapter. It is repeated here to allow for a self-contained narrative of the development of the methodological comparison model.

In the methodological comparison model, we consider that the producer has three choices to make with respect to units that are actively under development (WIP) 1) Terminate development and move into a finished goods state (for immediate or later sell to the customer), 2) Keep under development to continue to mature (or decay), or 3) Terminate development and destroy or dispose.

From the point of view of a single producer, each of these dispositions are binary (for example a farm cannot simultaneously destroy, harvest, and continue to cultivate a single unit of food). Under a model of a single producer, this economic decision becomes a straight-forward assessment of the expected value of each disposition route (for example weighing the of the costs of shipping and storing goods versus the costs of destroying it, offset by the value that would come from selling if it were sold). However, for a larger model of a system of producers, it is more appropriate to utilize a *multinomial logistic model*, to represent the probability of a producer choosing any of the above three options.

For some relative economic value π_i for choice X_i, the probability of choosing X_I is given by the expression below:

$$P(X_i) = \frac{e^{\beta \pi_i}}{\sum_{I=1}^{N} e^{\beta \pi_I}} \tag{23}$$

In the above, β is the weight the producer places on the concept of economic value. Under a full logistic model that we could fit to observed data, this becomes a free parameter. Here, we have no observed data, but rather a conceptual model. Thus, to simplify the model overall, we can fix values of β to be the inverse of some reference price for the producer (e.g., the price at which a producer sells its goods under normal steady state conditions). This has the advantage of allowing the relative values of each choice, π_i, to be expressed in terms of prices and monetary values, while allowing the expression above to properly reduce to a dimensionless probability.

$$\beta_i = \frac{1}{reference\ price}\ \forall i \tag{24}$$

A.2.6 Valuing WIP Dispositions

As discussed above, how the producer derives the relative values of each of the choices is a matter of modeling freedom and should ideally be based on observations of how real produces value these choices. The advantage of the logistic model is that changing these assumptions only changes the relative value of each choice, and thus the relative proportion of the crop delegated to each option, but not the underlying model.

This is the most straight forward valuation in the model and is simply the cost of destroying the units under development. The act of ceasing production and destroying goods is not considered ‘free’ and has a cost assigned to it in the model as an exogenous parameter. This could be expanded by applying a ‘mental resistance’ or ‘sunk cost fallacy price’ to further discourage the disposal of WIP, if evidence supports it. As a note, as modeled here, the value of disposing of goods is *always* negative. While the other options can be more negative, even if they are strictly positive, some portion of the crop is nevertheless destroyed each period under the multinomial logistic model.

$$\pi_{destroying\ WIP} = -Producer\ Cost\ to\ Dispose\ of\ WIP \tag{25}$$

If the producer is to finished development and store the units in the same area (or hold up the same production capacity) while not actually adding value to the good, they would do so under the expectation that they would receive their current expected price for the goods, less the costs for moving into an FG state, less the eventual costs for shipping to the customer.

$$\begin{aligned}\pi_{Producing} = {} & Producer\ Short\ Run\ Expected\ Price - Cost\ to\ move\ into\ FG\ state \\ & - Shipping\ Costs\end{aligned} \tag{26}$$

Combined with the above logistic model, this gives a fraction of the units under development that could be made available, at most, for shipping.

The ‘Production Rate’ flow in the model developed in the main chapter is based on both the *expected* future need of goods to fulfill demand from the wholesaler and anticipated spoilage or loss in storage.

$$\begin{aligned}& Expected\ Needed\ Production \\ & \quad = SMOOTH(Anticpated\ Customer\ Demand \\ & \quad + Expected\ Loss\ of\ FG, Time\ to\ Update\ Expected\ Production\ Need)\end{aligned} \tag{27}$$

If the producer were fully willing to meet customer demand and replace any goods previously destroyed or spoiled or other loss in storage, the above alone would move the goods from production starts through to finished goods. However, the value of the units under production and available to be moved into a finished goods state is limited by the logistic model described above. Thus, expression for Production Rate from the main chapter can be recast into this example context as seen in expression (28) below.

$$\begin{aligned} &Production\ Rate \\ &= \min\left(1, \frac{Maximum\ Production\ Rate}{Expected\ Needed\ Production}\right) \\ &* Expected\ Needed\ Production \end{aligned} \tag{28}$$

The actual number of units left in the WIP state is defined by how many units are destroyed and how many units are moved into a finished goods state each period. However, the probability that a producer will choose to destroy, or complete development is also dependent on how the producer values keeping units under development. There are two possible ways to capture the value of leaving work-in-progress alone to continue to age, both of which are explained below:

The first option is both the easiest conceptually, and perhaps the most robust because it introduces the least number of additional assumptions: that keeping the units under development in a WIP state has a null value. In many logistic models, there is a ‘null choice’ or simply a choice of zero value, often used to represent not making a choice at all (e.g., between a red car and a blue car I choose to not buy a car today). For this model, the relative value of holding crops is 0.

$$\text{Option 1: } \pi_{holding\ units\ in\ WIP} = 0 \tag{29}$$

The other option is more behaviorally complex, but more realistic. Here, the producer is assumed to be forward looking, anticipating getting the maximum value from his or her production that could be expected.

$$\begin{aligned} \text{Option 2:} \quad \pi_{holding\ units\ in\ WIP} = &\ Producer\ Future\ Looking\ Price - \\ &Prodcution\ Costs - Shipping\ Costs \end{aligned} \tag{30}$$

Under this model, the producer is assumed to know the shape of the relationship between age and value discussed above and can expect the maximum fraction of the value of his or her production if the maturation time is lower than the ideal maturation time, but nothing better than the current value for maturation times higher than the ideal value.

$$Producer\ Future\ Looking\ Expected\ Price = \begin{cases} Short\ Run\ Expected\ Price & if\ \mathrm{T_{maturation}} < \mathrm{T_{ideal}} \\ Short\ Run\ Expected\ Price * Effect\ of\ Age\ on\ Price & o.w. \end{cases} \tag{31}$$

Ultimately, the choice of option 2 causes the producer to reserve more units in the WIP state each period, as the value of the goods is viewed higher than null.

A.2.7 Valuing FG Inventory Dispositions

While the model development immediately above has focused on the valuation and inventory disposition decisions of work-in-progress production, it can be readily applied as well to finished goods inventory in storage as well. Again, the producer has three choices: 1) Make inventory available for the customer, 2) keep finished goods in storage, or 3) dispose of finished goods. As with the work-in-progress inventory, a multinomial logistic function can be used, normalized with β values all chosen to be the inverse of a producer reference price. As a note, the inclusion of this feature has negligible impact on the example system parameterized in the main chapter but is included for completeness for the reader to experiment with.

As with the previous sector, the value of destroying finished goods can be assumed to be some simple value. It is possible to expand this valuation by considering how destroying inventory frees storage space, but rather than complicate the valuation here, those considerations are rolled into the valuation of holding goods.

$$\pi_{destroying\ finished\ goods} = -Producer\ Cost\ to\ Destroy\ Inventory \tag{32}$$

The value of making inventory available to ship is simply the current spot price, less the cost of shipping those goods. Note that the current spot price is affected by the maturation of the units as described above.

$$\pi_{Shipping} = Spot\ Price - Shipping\ Costs \tag{33}$$

As with the choice to hold WIP to further mature, there are two ways to look at the valuation of holding inventory rather than destroying or shipping it, either with a null value or with a more forward-looking model of valuation.

$$\text{Option 1: } \pi_{Hold\ Inventory} = 0 \tag{34}$$

For the forward-looking estimation, consider the that the opportunity cost of storing an additional unit of goods for an additional unit of time increases with finite storage space, and the only feasible method of storing additional units of goods when storage is full is to acquire additional space at some costs. This is captured in the relationship below:

$$\begin{aligned} &Cost\ to\ Hold\ Based\ on\ Free\ Space \\ &\quad = Farm\ Holding\ Costs + Fraction\ of\ Storage\ Full \\ &\quad * (Farm\ Holding\ Costs + Costs\ to\ Acquire\ Storage) \end{aligned} \tag{35}$$

Furthermore, by holding the finished goods, the producer must be expecting not the current spot price, but some future estimate of the price for their goods. Combined, this gives the following alternative option for valuing holding inventory in this model:

$$\text{Option 2: } \lambda_{Hold\ Inventory} = Short\ Run\ Expected\ Price - Cost\ to\ Hold\ Based\ on\ Free\ Space \tag{36}$$

Either of the two options of valuation above presuppose a decision to acquire storage space if full. Thus, we can consider that the producer has a desired total storage space that is approximately equal to the actual finished goods inventory, with perhaps an additional allowance for free space for comfort or other purposes.

$$Desired\ Storage\ Space = \frac{FG\ Inventory}{1 - Producer\ Desired\ Fraction\ Free\ in\ Storage} \tag{37}$$

The producer will then actively work to adjust the actual storage space to the desired storage space, though perhaps in an asymmetric manner. Specifically, I hypothesize that the producer will be quick to add space but slow to divest it.

$$\begin{aligned} &Storage\ Space \\ &= SMOOTH\Bigg(Desired\ Storage\ Space, \\ &\begin{cases} Time\ to\ Add\ to\ Storage\ Space & if\ Desired\ Storage\ Space > Farm\ Space \\ Time\ to\ Reduce\ Storage\ Space & o.w. \end{cases}\Bigg) \end{aligned} \tag{38}$$

A.3 Limits of the MNL Formulation under Extreme Conditions

Consider expression (2) above, restated below for the reader's convivence:

$$P(X_i) = \frac{e^{\beta\pi_i}}{\sum_{I=1}^{N} e^{\beta\pi_I}} \tag{39}$$

The extreme case where $\beta = 0$ of expression (39) is trivial to determine, and results in the simple equal allocation to all dispositions, independent of the economic value of those dispositions.

$$P(X_i)_{\beta=0} = \frac{e^{0\pi_i}}{\sum_{I=1}^{N} e^{0\pi_I}} = \frac{1}{\sum_{I=1}^{N} 1} = \frac{1}{N} \tag{40}$$

While the other extreme case where $\beta \to \infty$ can be determined intuitively as discussed in the main chapter (here, all goods are in the disposition with the highest economic value with probability 1 and all others with probability 0), the derivation requires a few more steps.

First, consider the extreme value of the highest valued disposition, e.g., $\pi_M > \pi_{i \neq M}$.

$$P(X_M) = \frac{e^{\beta\pi_M}}{e^{\beta\pi_M} + \sum_{\substack{I=1 \\ I \neq M}}^{N} e^{\beta\pi_I}} \tag{41}$$

Next divide (41) by the leading term of the denominator, which by definition is $e^{\beta\pi_M}$

$$P(X_M) = \frac{e^{\beta\pi_M} \,/\, e^{\beta\pi_M}}{\left(e^{\beta\pi_M} + \sum_{\substack{I=1 \\ I \neq M}}^{N} e^{\beta\pi_I}\right) / \, e^{\beta\pi_M}} = \frac{1}{1 + \sum_{\substack{I=1 \\ I \neq M}}^{N} e^{\beta(\pi_I - \pi_M)}} \tag{42}$$

As $\pi_M > \pi_{i \neq M}$, therefore $\pi_I - \pi_M < 0 \; \forall I$. The exponential terms in the denominator of (42) will all tend to 0 as $\beta \to \infty$ given that $\pi_I - \pi_M$ is strictly negative. Finally, therefore:

$$P(X_M)_{\beta\to\infty} = \frac{1}{1 + \sum_{\substack{I=1 \\ I \neq M}}^{N} e^{-\infty}} = \frac{1}{1} = 1 \tag{43}$$

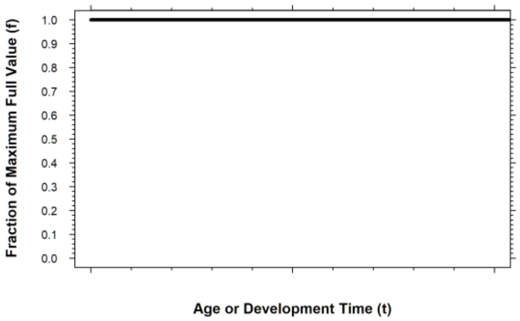
Similarly, when considering any other disposition choice with a value that is not the maximum, the probability *must* be 0 given the outcome of (43), or could also be shown directly to be 0 by the same procedure as above, still dividing by the leading term in the denominator of (41), but noting that in (42) the numerator will now be 0 for all values of dispositions not the maximum.

A.4 Some Alternative Functional Forms for the Age-Value Relationship

The example in the main chapter and the development of the formulations above assume a trapezoidal relationship between the age of work-in-progress and the value that the producer can extract. As discussed in the main chapter, multipole alternative shapes could be feasible in different contexts, and even the core trapezoidal shape explored in this chapter can take on more complex configurations (Dorp & Kotz, 2003).

The methodological comparison model .mdl file that accompanies this chapter allows for uses to experiment with several of the relationships seen in Figure A-15. Note that the *Asymmetric Gaussian* can be made to be arbitrarily close to an s-curve by setting σ_u to arbitrarily high values.

Null Relationship (Fixed Value)

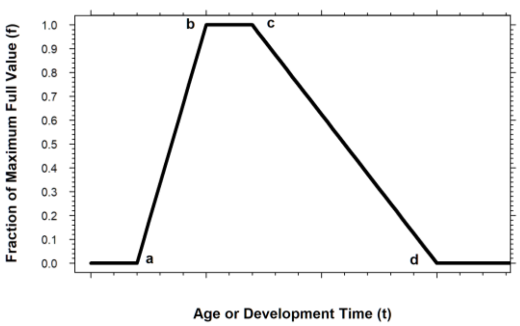


$$\boldsymbol{f(t) = 1}$$

Linear and Saturating Relationship

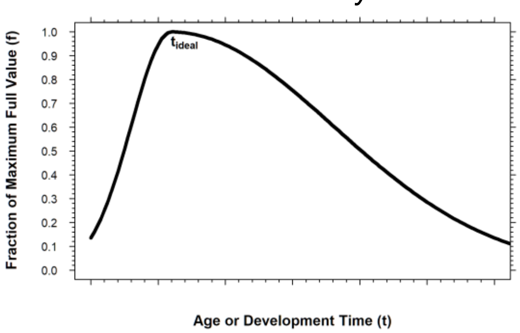
$$f(t) = \begin{cases} 0 & t \le a \\ \left(\dfrac{1}{b-a}\right)t - \left(\dfrac{a}{b-a}\right) & a < t \le b \\ 1 & t > b \end{cases}$$

Trapezoidal Relationship



$$f(t) = \begin{cases} 0 & t \le a \\ \left(\dfrac{1}{b-a}\right)t - \left(\dfrac{a}{b-a}\right) & a < t \le b \\ 1 & b < t \le c \\ \left(\dfrac{1}{c-d}\right)t - \left(\dfrac{d}{c-d}\right) & c < t \le d \\ 0 & t > d \end{cases}$$

Asymmetric Gaussian Relationship



$$f(t) = e^{\left(-\frac{1}{2}\right) * \left(\frac{t - T_{ideal}}{\sigma}\right)^2}$$
$$\begin{cases} \sigma = \sigma_l & if\ t < T_{ideal} \\ \sigma = \sigma_u & o.w. \end{cases}$$

Figure A-15. Examples of Price-Value Relationships

A.5 Parameterization of the Methodological Comparison Model

The detailed parameterization of the example supply chain model used in the main chapter is largely omitted for the sake of space. This was done also because the focus of that section of the chapter was not about the influences of specific parameter choices, but rather to illustrate how the methodology introduced can generate fundamentally different modes of behavior for otherwise identically parameterized models. The .mdl file included with the chapter comes parameterized in the same manner as used for the main chapter, but those parameter values are explicitly listed below as well.

Below are the values used for the age-value relationship, which for the chapter utilized the trapezoidal relationship, which is described in expression (5) above.

Table A-1. Parameterization for the Age-Value Relationship in the Methodological Comparison Model

Parameter Name	Description or Note	Value
A	Minimum age of any value in the Trapezoidal relationship	4 weeks
B	Minimum age of full value in the Trapezoidal relationship	10 weeks
C	Maximum age of full value in the Trapezoidal relationship	14 weeks
D	Maximum age of any value in the Trapezoidal relationship	30 weeks
Initial average WIP Age	To ensure a steady state at the initialization of the model, this value should be within (inclusive) of B and C above. Chosen to be the average of those two numbers for simplicity of exposition	12 weeks

The structure used to track the average age of WIP inventory utilized in this example model is most similar to the ‘Coflow with Experience’ structure discussed in detail in *Molecules of Structure* (Hines, 2005), though the unit balancing takes on a different form to be dimensionally consistent with the rest of the system. Note that any structure that captures the development time of the cohort of interest could be used to relate this time to the economic value that could be extracted from the goods.

As stated in the main chapter in the formulation for the co-flow that monitors the average age of work-in-progress inventory, the ‘Average Age of New Production Starts’ may have a value of 0 units of time, or some other non-negative value. For example, when applying to employee experience in a firm, an employee may arrive with some pre-existing experience. However, in the context of a food producer planting crops it may be safe to assume that newly planted crops arrive with no pre-existing maturation. The model is flexible to allow for this assumption to be relaxed based on specific circumstance (for example, buying partially matured nut trees or fully matured sows rather than starting from seeds or piglets).

Furthermore, while it may be generally safe to assume that the 'Rate of Age Gain' is constant and unitary (i.e., 1 week/week or 1 years/year or similar). The formulation itself does allow for some flexibility if, as an example, a fertilizer was applied to speed maturation, or a drought hit and slows maturation down.

In general, the structure here is most similar to the 'Coflow with Experience' structure discussed in detail in *Molecules of Structure* (Hines, 2005), though the unit balancing takes on a different form to be dimensionally consistent with the rest of the system. Note that any structure that captures the development time of the cohort of interest could be used to relate this time to the economic value that could be extracted from the goods.

Table A-2. Parameterization for the Co-Flow Structure Monitoring Average WIP Age

Parameter Name	Description or Note	Value
Rate of Maturation Gain	Rate at which unit under development gains age.	1 Weeks/Week
Average Age of New Units	The average maturation of brand-new production starts.	0 Weeks

Throughout the model there are time constants that affect the rate at which entities in this model supply chain either incorporate information and update estimations or limit the rate at which they can perform actions. These values were chosen to be behaviorally realistic (for example the producer incorporating price changes into their forward projection affecting production starts much more quickly than the customer adjusting their demand in response to those same price fluctuations), but again the primary purpose of this model is not to explore the sensitivity to these parameters but rather illustrate that different modes of behavior emerge when utilizing the methodological contributions illustrated in the chapter.

Table A-3. Parameterization for Time Constants

Parameter Name	Description or Note	Value
Time to adjust Production Schedule	Average time to adjust the actual unit production starts (or planting schedule) to the desired value.	26 Weeks
Time to adjust expected mixed variable costs	Time, on average, for the producer to update its expectation of the mixed variable costs it will typically incur per unit production started.	2 Weeks
Producer Processing Time	The desired typical time it takes for the producer to process WIP goods, either by disposing of them or moving them along into finished goods inventory.	4 Weeks
Producer finished goods disposal time	Time, on average, for the producer to dispose of stock of finished goods awaiting shipment to the customer	4 Weeks
Time for customer to adjust demand	The time, on average, for the demand from the Customer to change based on the indicated demand	4 Weeks
Time for producer to adjust short-run expected price	Time, on average, for the producer to incorporate the spot price into its expected price	1 Week
Time for customer to adjust short-run expected price	Time, on average, for the customer to incorporate the producer spot price into its expected price	24 Weeks
Average Shelf Life of Producer Stored FG	The average time the unit that has been moved to a finished goods state, but is still being stored at the producer, can sit in storage before spoiling and being disposed of. For non-foodstuffs this could be an average obsolescence time	24 Weeks

The parameters below were used to from the price formation mechanism at use in this model and as described in the Effect of Inventory Coverage on Price and Effect of Spot Prices on Demand sections of this Appendix.

Table A-4. Parameterization for the Effects of Inventory Coverage and Elasticities

Parameter Name	Description or Note	Value
Producer Desired Inventory Coverage	Weeks supply of inventory the producer wants to have on hand	2 Weeks
Sensitivity of Producer Price to Producer Inventory Coverage	Factor affecting how 'steep' the inverse relationship between inventory coverage and price is. Note that based on this formulation, this is assumed a *positive* value here for the expected inverse relationship. Higher *positive* values of this factor imply more sensitivity.	2 *(dimensionless)*
Elasticity of Customer Demand	Under an assumption of a linear demand curve near the reference prices and demand, this is the negative value of the slope of that curve. Note that this parameter is assumed to take on a *positive* value under the default assumptions of decreasing demand with increasing spot prices. High *positive* values of the factor create a steeper, but still negatively sloped, demand curve.	1 *(dimensionless)*

As described in the main chapter and partially restated in this Appendix, the multinomial logistic choice model depends on the relative difference in perceived value of each disposition option. Thus, each option must have some manner by which that value can be calculated. For this specific model, this value is simply determined by comparing the expected profit from each disposition route along the supply chain. The costs and baseline revenue values used are given below.

Table A-5. Parameterization for Producer Costing

Parameter Name	Description or Note	Value
Reference Producer Price	Price at which the Producer experiences its reference Gross Margin and Reference Planting. Sets the default profitability expectations in steady state for the producer	\$1.1/unit
Producer Raw Material Costs	Raw cost per unit (i.e., the variable cost) the producer endures	\$0.05/unit
WIP Development Costs	Cost of developing a single unit of productions starts for a single unit of time that the producer endures	\$0.01/unit/week
Producer Cost to Dispose of or Abandon WIP	The cost incurred by the producer to dispose of a unit being actively developed in a WIP state. Note that this could not only be the actual material cost (labor and equipment) but also could be extended to include physiological costs from sunk cost fallacy or similar resistance to disposing units that have already had resources invested in their development.	\$2/unit
Producer Cost to Place in FG State	Costs, per unit, that the producer incurs to move a unit from the WIP to the FG state.	\$0.1/unit
Producer Shipping Cost	Costs, per unit, that the producer incurs to process and ship goods for the Customer.	\$0.1/unit
Producer Cost to Dispose of FG Inventory	The cost incurred by the producer to dispose of a unit that is being stored after production and before shipping to the customer. Note that this could not only be the actual material cost (labor and equipment) but also could be extended to include physiological costs from sunk cost fallacy or similar resistance to disposing units that have already had resources invested in their development and storage.	\$1/unit

A.6 References to Appendix A

Chen, H., Wu, O. Q., & Yao, D. D. (2009). On the Benefit of Inventory-Based Dynamic Pricing Strategies. *Production and Operations Management*, *19*(3), 249–260. https://doi.org/10.1111/j.1937-5956.2009.01099.x

Dorp, J. R. van, & Kotz, S. (2003). Generalized trapezoidal distributions. *Metrika*, *58*(1), 85–97. https://doi.org/10.1007/s001840200230

Hines, J. (2005). *Molecules of Structure: Building Blocks for System Dynamics Models Version 2.02*. Ventana Systems and LeapTec. http://www.mindseyecomputing.com/molecule.pdf

Martinez-Moyano, I. J. (2012). Documentation for model transparency. *System Dynamics Review*, *28*(2), 199–208. https://doi.org/10.1002/sdr.1471

Sterman, J. D. (2000). *Business Dynamics—Systems Thinking and Modeling for a Complex World*. McGraw- Hill Higher Education. https://www.worldcat.org/oclc/42771322

Whelan, J., & Forrester, J. W. (1996). Economic Supply & Demand. *D-4388*, 7. http://ocw.mit.edu/courses/sloan-school-of-management/15-988-system-dynamics-self-study-fall-1998-spring-1999/readings/economics.pdf

Appendix B Chapter 2: Systemic Origins of Hunger Amidst Plenty During the Onset of the COVID-19 Pandemic in the United States

B.1 Model Availability

Accompanying the main chapter and this Appendix are the full models available as .mdl files. The .mdl files can be open and run using Vensim software, developed by Ventana Systems, Inc. A free version of the Vensim software for personal use, along with a standalone model viewer, is available from Ventana Systems, Inc. These model files can be obtained directly at:

https://github.com/jpain3/MIT-Disseration/tree/main/chapter-2

Ventana Systems, Inc provides detailed documentation on the Vensim software, including how to manipulate and examine specific formulations. However, the reader may quickly explore the influence of parameter choices on the model via the SyntheSim mode on the main Dashboard view of the model. This can be accessed by pressing the corresponding button in the top toolbar of the software as seen below:

For the supporting Food Supply Chain model, the .mdl file is divided into views, consisting of an overview Dashboard, a view of the full model itself, and several views highlighting the physical flows of goods through each entity in this supply chain. Different views can be accessed via the buttons in each view, or via the view menu. Examples of these two views (but not the entirety of these views) are provided below.

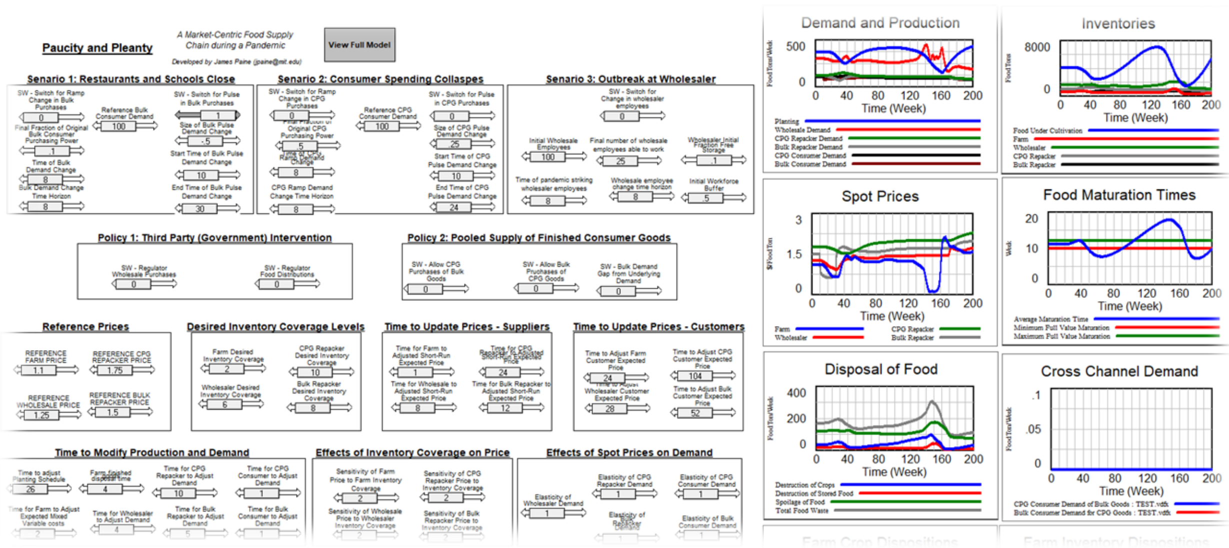


Figure B-1. Example of the Dashboard View of the Food Supply Chain Model

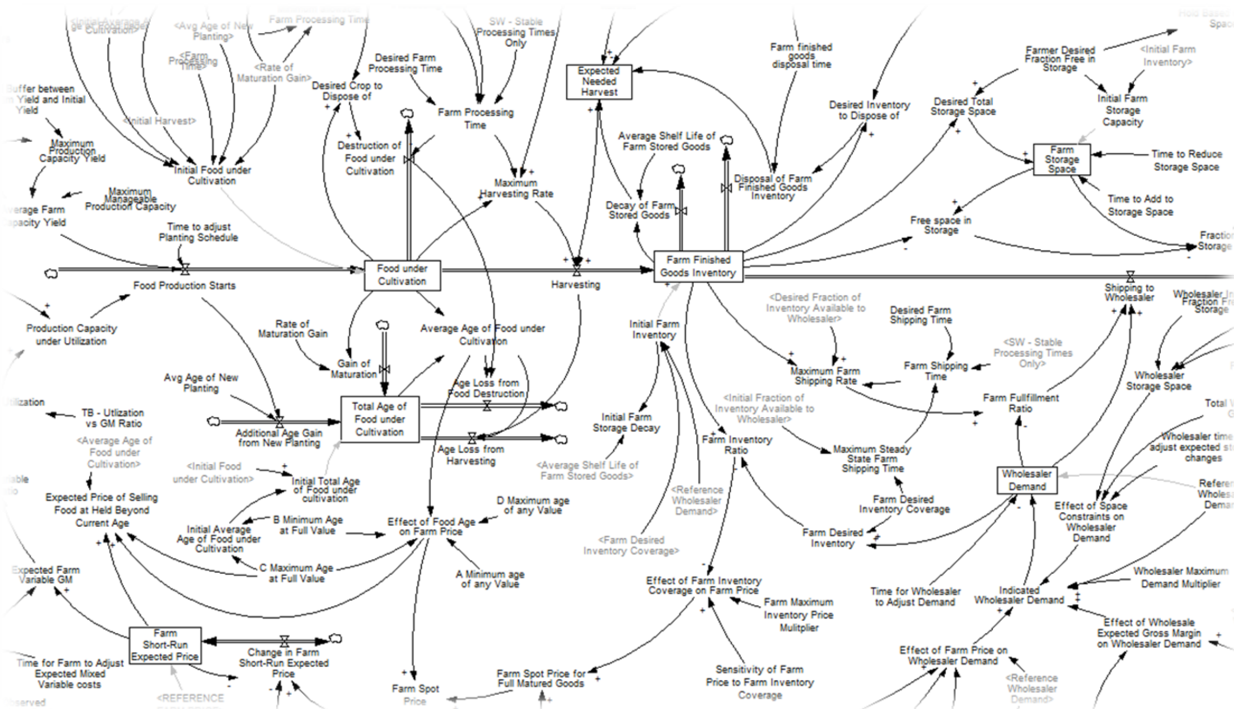


Figure B-2. Detail of Aggregate Framework Embedded in Full Model View of the Food Supply Chain

As a note, the full model view is presented entirely and largely has no hidden structure or hidden causal connections. Furthermore, the .mdl file provided may be opened in any program that is able to read UTF-8 encoding and the formulations directly viewed in plaintext. Examples of programs that can open the .mdl file for direct viewing in plaintext include Notepad in the Windows operating system and Textpad in the Macintosh operating system. An example of this view of the model file is seen in Figure B-3.

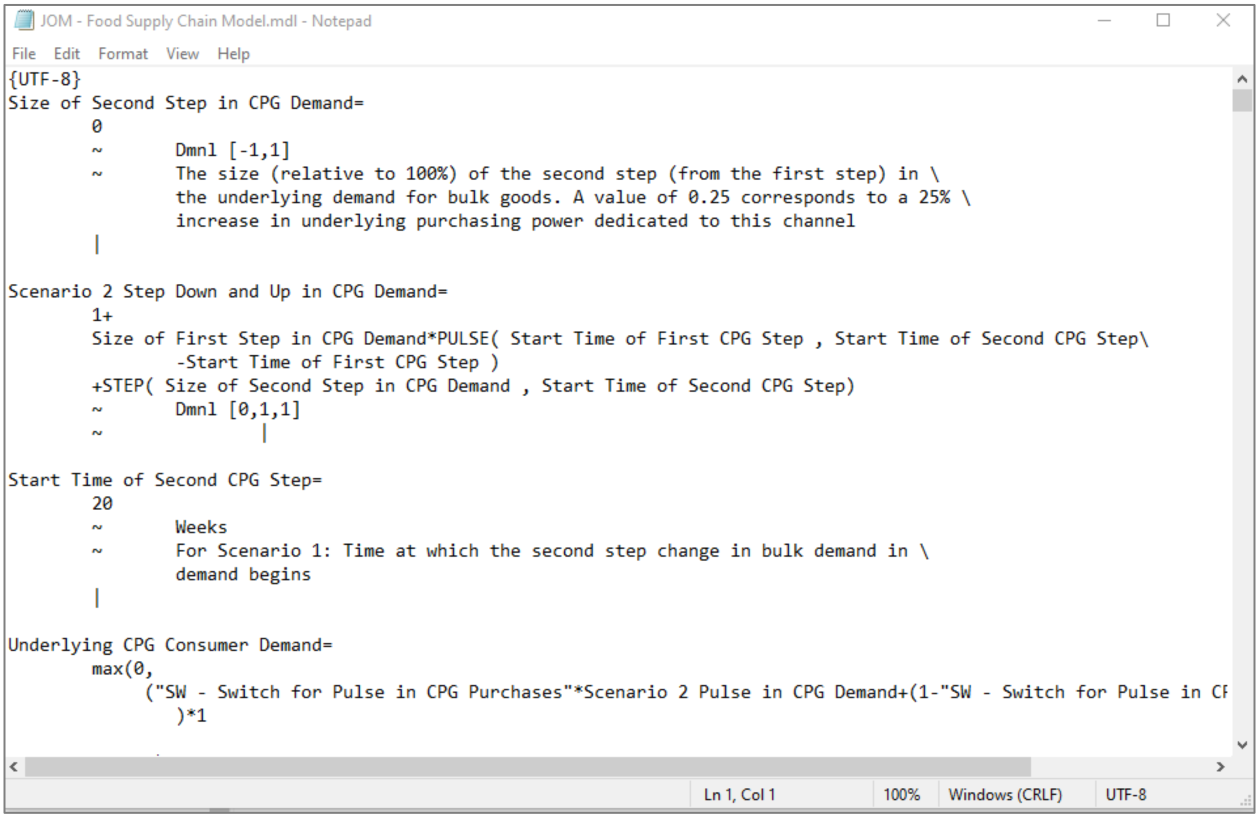


Figure B-3. Example of Viewing the Supporting .mdl File in Notepad on Windows

The models developed for this chapter are also fully documented utilizing the SDM-Doc tool described in (Martinez-Moyano, 2012). The output from this documentation tool is available alongside the .mdl files.

B.2 Formulation Details for the Supporting Food Supply Chain Model

The sections below provide additional detail on the development of that food supply chain model, focusing on details on the formulations developed for the food supply chain model seen in the main chapter. Please note that some portions of the explanatory text from the main chapter are repeated below where needed to create a self-contained description of this larger model.

The example model described below explores the application of this modeling framework to a hypothetical bifurcated food supply chain consisting of the following entities:

- A farmer, who is responsible for making decisions about how much to plant each period and how to manage his or her harvest
- A wholesaler firm, which receives raw and unprocessed foodstuff from the farmer, and does some minimum value-added work to the food
- Two different packaging processors
    - A CPG (consumer packaged goods) processor that received good from the wholesaler and does extensive value-added rework to the food, packaging it in smaller consumer friendly forms for sale to the end consumer at some outlet like a grocery store
    - A Bulk processor that receives goods from the wholesaler and does minor repacking of the food for sale directly to larger consumers like restaurants, governments, or schools.
- The end consumers, which include demand for both CPG and Bulk packaged food

B.2.1 Core Market Formation Mechanisms

The definition of this marketing being based on a commodity product implies the use of inventory-sensitive spot pricing (Chen et al., 2009; Sterman, 2000; Whelan & Forrester, 1996), and thus similar market formation mechanisms as that used in chapter 1 can be used here as well. The core of this economic model remains the same two balancing loops across each entity in the supply chain, with spot pricing driving either demand or supply as illustrated Figure A-9 above, chained in sequence as illustrated in Figure A-10.

Similarly, the core mechanisms explained in more detail in Appendix A above are used directly in this more expansive use case in this chapter. Please refer to those sections for more details on the mechanics of the market and price formation. They are linked below for convenience:

- A.2.1 Defining the Market
- A.2.2 Production Starts and Capacity Management
    - Note the additional structural feature of *Yield is Decreasing with Additional Arable Land* expanded on below

- A.2.3 Discounting the Spot Price based on Development Age
- A.2.4 Quantifying the Age-Value Relationship
- A.2.5 A Multinomial Logistic model of Inventory Dispositions
- A.2.6 Valuing WIP Dispositions
- A.2.7 Valuing FG Inventory Dispositions

B.2.2 Yield is Decreasing with Additional Arable Land

To a first approximation, one could consider the net incremental productivity added by acquiring new land constant and fixed. However, that creates a scenario in which the farm is able to infinitely expand so long as the gross margin is justified (i.e., the additional operating expenses of the land are covered by the new production). A more realistic model though would capture that the net productivity would decrease with additional land under management. In other words, there are decreasing *returns to scale* with respect to land and productivity.

While there are multiple ways to model this relationship, here a linear decay model is used where the farm has a known maximum amount of land that they could get any yield out of. Figure B-4 shows the shape of this new function, which I ultimately used in the model for both the reasons named above and simplicity.

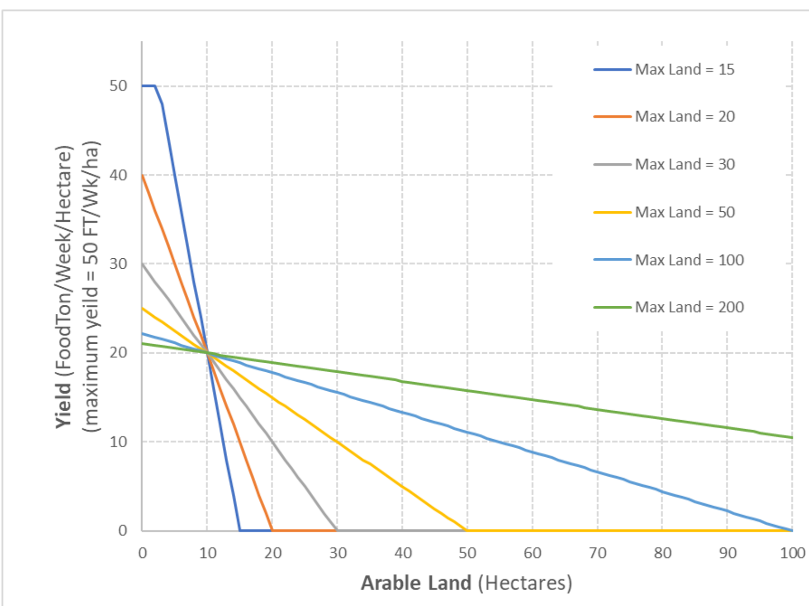


Figure B-4. Linearly Decaying Relationship between Yield and Arable Land

B.3 Alternate Combined Color Figures From the Main chapter

For the sake of presentation in print, the figures in the main chapter are all presented without color. Also, to avoid excessively cluttered diagrams, figures of production, demand, and prices were often split into two sections (such as ‘upstream’ and ‘downstream’, or ‘producer’ and ‘non-producer’).

Below are selected key figures exported directly from the Vensim .mdl model viewer and based on the same datasets as those used for the main chapter and are based on the scenario of a 50% drop in bulk consumer demand for a period of 20 weeks.

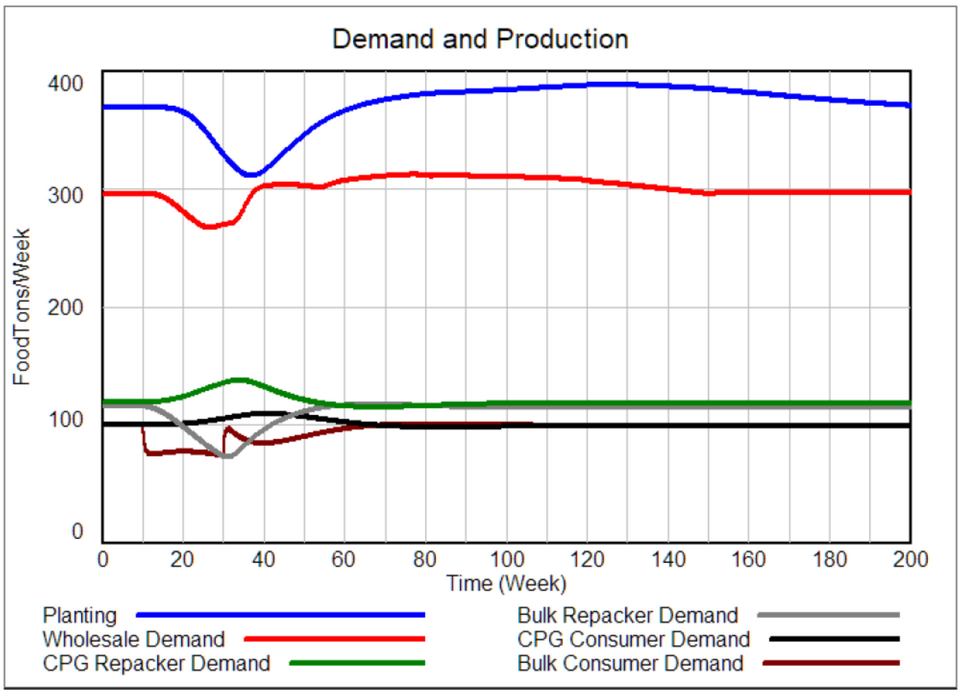


Figure B-5. 50% Drop in Bulk Purchasing for 20 Weeks – Demand and Production

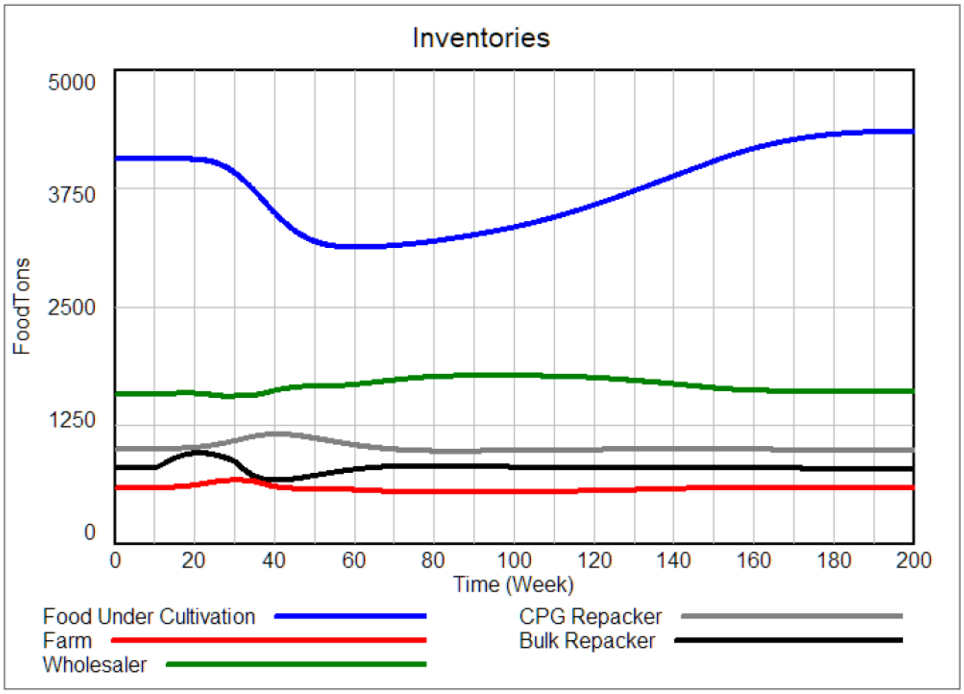


Figure B-6. 50% Drop in Bulk Purchasing for 20 Weeks – Inventories

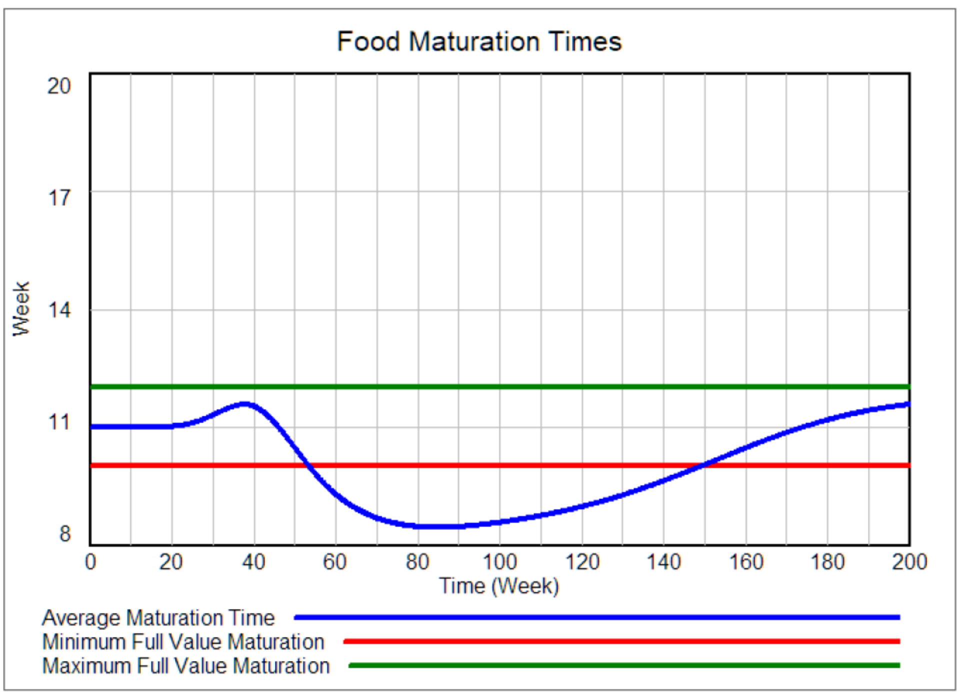


Figure B-7. 50% Drop in Bulk Purchasing for 20 Weeks – Food Maturation

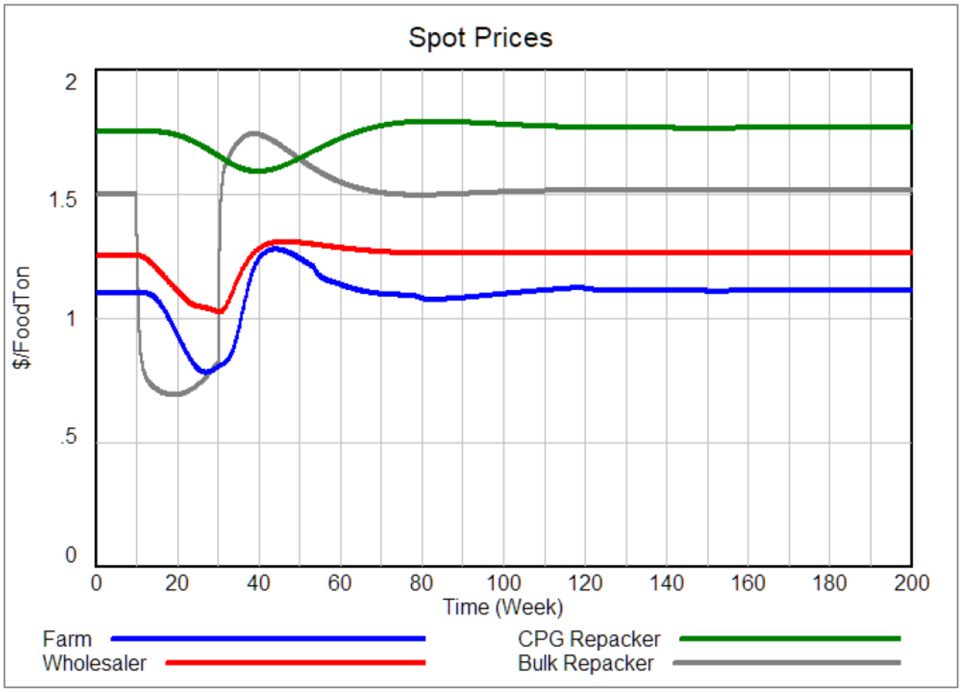


Figure B-8. 50% Drop in Bulk Purchasing for 20 Weeks – Spot Prices

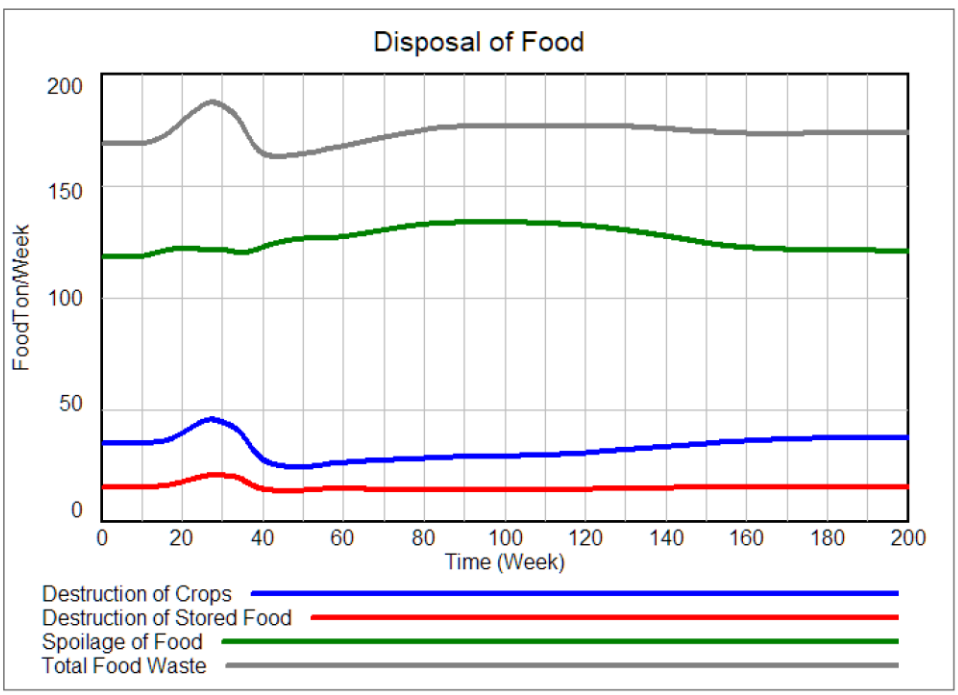


Figure B-9. 50% Drop in Bulk Purchasing for 20 Weeks – Food Loss and Disposal

B.4 Reference to Appendix B

Chen, H., Wu, O. Q., & Yao, D. D. (2009). On the Benefit of Inventory-Based Dynamic Pricing Strategies. *Production and Operations Management*, *19*(3), 249–260. https://doi.org/10.1111/j.1937-5956.2009.01099.x

Dorp, J. R. van, & Kotz, S. (2003). Generalized trapezoidal distributions. *Metrika*, *58*(1), 85–97. https://doi.org/10.1007/s001840200230

Martinez-Moyano, I. J. (2012). Documentation for model transparency. *System Dynamics Review*, *28*(2), 199–208. https://doi.org/10.1002/sdr.1471

Sterman, J. D. (2000). *Business Dynamics—Systems Thinking and Modeling for a Complex World* (Vol. 53, Issue 4). McGraw- Hill Higher Education. https://doi.org/10.1057/palgrave.jors.2601336

Whelan, J., & Forrester, J. W. (1996). Economic Supply & Demand. *D-4388*, 7. http://ocw.mit.edu/courses/sloan-school-of-management/15-988-system-dynamics-self-study-fall-1998-spring-1999/readings/economics.pdf

Appendix C Chapter 3: Simpler is (Sometimes) Better: A Comparison of Cost Reducing Agent Architectures in a Simulated Behaviorally Driven Multi-Echelon Supply Chain

C.1 Model and Code Availability

All code used to produce the results in the main chapter are available at:

https://github.com/jpain3/MIT-Disseration/tree/main/chapter-3

This includes the code, in Python, used to train the DQN structures described in the main chapter, along with code, in R, used to apply those trained TensorFlow objects. Note that the DQN was originally trained using TensorFlow (Abadi et al., 2016) version 2.3.0. The DQN model objects *must* be used in an environment that similarly uses TensorFlow version 2.3.x or is backwards compatible with objects trained in that environment. The Python scripts also includes the custom OpenAI gym environment (Brockman et al., 2016) used to train the DQN agent. To train the agent, the package Keras-RL2 (McNally, 2019) was used as a front-end API manager for TensorFlow. Keras-RL2 is an extension of the Keras package (Chollet, 2015) and was used specifically because it better implements the dueling reward function structure central here (Wang et al., 2016).

To support use of these objects trained objects and to replicate the training process, the code repository includes yaml objects as well, which contain snapshots of the supporting packages and similar supporting infrastructure used in Python while training these DQN agents. Note that while these yaml objects do include all necessary packages for recreating the training environment, they may contain superfluous packages as well. The author has attempted to trim down these unnecessary packages but makes not guarantees.

The R scripts include self-contained functions to simulate the Beer Game over a given time horizon with variable values of information delays, shipping delays, costing, and order input types. All R scripts include code at the beginning to install any needed packages that are beyond the vanilla installation of R. Note that these scripts were primarily developed and run using RStudio as the IDE (RStudio Team, 2020), and may contain references to graphical objects (such as progress bars or window status values) that may not be pertinent in all environments, especially headless clusters or similar decentralized computing systems.

The code is intended to act as a flexible framework for future research and allows for multiple models of human ordering based on prior literature to be substituted into each position in the supply chain. As of this publication, the framework supports the following ordering schemes:

- Base-Stock replenishment – This does not calculate the optimal base stock policy like that seen in (Clark & Scarf, 1960) but rather follows a fixed replenishment policy based on a given bast-stock value for the position in the supply chain
- (Sterman, 1989) – This is the mechanism used in the main chapter and follows a four-parameter ordering scheme with anchoring and adjustment of expectations of future ordering. This ordering scheme is also derived most directly from the context in which the real-world runs of the Beer Game on which this paper is built were derived.
- (Sterman & Dogan, 2015) – This paper was based on stationary and known orders, and introduces a more complex rule built on other similar research (Croson et al., 2014) and allows for the desired supply line to shift over time.
- (Oliva et al., 2022) – All four variants of the model utilized in this paper are available as options in the framework here, but models 3 and 4 notably vary from the (Sterman, 1989) model described above in that they allow the response to differ when agents are in a backlog state.

For all the models described above, the framework allows for entity-level parameters to be supplied (like those fitted to real world ordering behavior for the Sterman ’89 rule used in the main chapter). If no parameters are supplied, the code utilizes the ‘average’, or ‘baseline’, or ‘best-fit’ values reported in the corresponding original paper.

C.2 Order Data Availability

For the 49 simulated teams used throughout the main chapter as a testing bed for the agents, 11 come directly from the original presentation of the four parameter ordering rule used throughout the analyses here, and 1 additional team modeled on ‘average’ performance of those other 11 (Sterman, 1989), for 12 historic teams. These teams do not have specific order traces, and instead were presented as estimated models using the four-parameter ordering rule in Sterman ’89. Order traces for real runs of the game were obtained from online runs of the Beer Game at MIT as part of various executive and graduate-level classes. These runs occurred twice in August of 2021, with 12 teams in one run and 22 teams in another run, and in June of 2022 for an additional 3 teams.

For these more recent runs, order traces are available with the team names and exact dates of the games obfuscated for individual privacy. Additionally, these teams have been fitted to the Sterman ’89 ordering rule to provide the total 49 teams used in the main paper. The code used to perform this fit is also available at the repository linked above.

C.3 Detail on the Beer Game

The simulation environment used in this work is based on the board-game configuration of the Beer Game as illustrated in Figure C-1 below. This figure also shows the typical starting layout for the game as used in numerous previous studies using this modeling framework (Croson & Donohue, 2006; Narayanan & Moritz, 2015; Sterman, 1989), which is started with 12 units of inventory on hand for each player, and 4 units of inventory in transit at each stage in the shipping system, and 4 orders moving through the order chain.

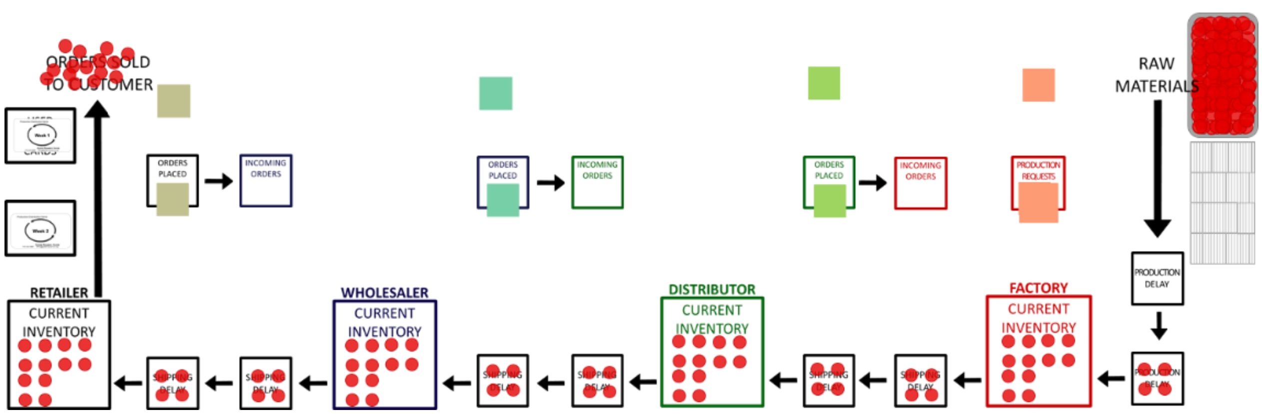


Figure C-1. Example of Beer Game Board Layout

Each round of the game proceeds as follows:

1. Receiving inventory and advance shipping delays – Each entity receives the units in the shipping delay immediately to their right. The contents of the furthest shipping delay to the right are moved up.

2. Fill orders – Entity 1 (retailer) views the customer order, all others examine the 'incoming orders' and orders, inclusive of any outstanding backorders, are filled to the extended inventory allows
3. Record inventory or backlog
4. Advance order slips – the order slips further to the left are moved up
5. Place orders – Each entity decides what to order and places it the 'orders placed' box to their right

The stated goal of the game is to reduce the amount of *total cost of the entire team* over some time horizon *T,* subject to some known inventory holding and backorder/stockout costs. Backorders do not expire under the traditional interpretation of this game and must be filled from existing stock prior to meeting any new demand. In the prior studies referenced above, and in this work, the cost of holding inventory, C_{inv}, is \$0.50 per unit per period, and the cost of backorders, C_{bo}, is \$1.00 per unit per period.

$$Cost_{Team} = \sum_{t=1}^{T} \sum_{entity=1}^{N} \left(C_{bo} * Backorders_{t,n} + C_{inv} * \ Inventory_{t,n} \right)$$

C.4 Applicability to Alternative Ordering Rules

The framework allows for the cost-reducing agent to be placed in the supply chain in any of the positions, and for its ordering rules to be defined separately from those used by the other entities. Thus, a DQN agent can be placed in a supply chain run by Sterman '89 behaving agents, or base-stock agents, or any other of the available ordering schemes. For the model-predictive learning agent, this is taken a step further, and the assumption of the agent about its surroundings can be further defined. For simplicity of exposition, the in main chapter the core architecture of the agent and its assumptions about its environment were kept the same, with a base-stock responding agent assuming base-stock responses from the other agents. This follows from the idea that for an agent to assume that a base-stock response would be near optimal it must also assume the other agents around it is behaving similarly. Conversely, if the agent itself is using a behavioral response model this presupposes that the agent is assuming behavioral responses from its peers in the supply chain.

However, while this matching of architecture and assumptions makes intuitive sense, it can be relaxed in the framework developed here, and the agent can assume any one of the above listed ordering rules are being used by other entities in the supply chain, and in turn use any of those rules (plus the DQN structures) to respond. Furthermore, the underlying reality that the agent is placed in, e.g., the *actual* rules being followed by the other entities can take on any of the above forms and does not have to match the assumptions the agent is making.

While not the focus on the main chapter, this framework also allows for some additional observations, namely that the model-predictive learning method can perform well even when making fundamentally flawed assumptions about its environment. Table C-1 shows the results from placing the model-predictive learning agent (with a calibration memory of 10 time steps and a forward horizon of 30 timesteps as in the main chapter) that assumes the other entities int the supply chain is following the Sterman ’89 ordering heuristic. In reality, they are following Model 3 from Oliva et. .al ’22, a related but still fundamentally different ordering rule.

Table C-1. Model-Predictive Learning Agent Assuming Sterman ’89 but in Oliva et al ’22 Average Model 3 Environment subject to Step Input

		Model-Predictive Learning Agent
	Baseline Costs	26257
Agent Position	**1 (Retailer)**	869 (-96.7%)
	2 (Wholesaler)	1040 (-96.0%)
	3 (Distributor)	2859 (-89.1%)
	4 (Factory)	12274 (-53.3%)

C.5 Model-Predictive Learning Agent Hyperparameters

The model-predictive learning agent introduced in the main chapter has several hyperparameter assumptions that are implied by the pseudocode used in the second that introduces that agent. In addition to the assumptions used to model the environment in which it resides, which is discussed in this Appendix in the sections above, this agent also has a calibration memory over which to fit an estimate of that model, and an optimization horizon over which to plan based on that calibrated model.

C.5.1 Calibration Memory and Optimization Horizon

In the main paper, all results for the model-predictive learning agent are based on a calibration memory of 10 units of time and a forward optimization horizon of 30 units of time. These numbers are somewhat heuristically chosen based on multiple trials of the agent during development, and also somewhat from support in related literature (notably the memory of 10 units used in the semi-optimal baseline developed in (Moritz et al., 2021)). However, some of the tradeoffs from choosing differing values of these hyperparameters can be directly explored. As discussed above, the choice of position 2 (wholesaler) in the supply chain was chosen for this analysis to isolate the main feature choices of the agents versus other idiosyncrasies. The hyperparameter choices are muted at this position in the supply chain as well.

However, when a model-predictive learning agent is placed at the beginning of the supply chain, in position 1 (retailer) and exposed to non-stationary order inputs then the influence of these hyperparameters, especially the forward optimization horizon, can be more significant. Figure C-2 shows the total team costs incurred in the simulated four entity supply chain, all using the average ordering rule reported in Sterman '89, with a model-predictive learning agent placed at position 1, the retailer, and exposed to either deterministically linearly increasing orders, or to noisily increasing orders. Figure C-3 zooms in on the specific case of the optimization horizon equaling 35 in the noisy non-stationary case to illustrate how the orders being placed by the agent in position 1 influence the overall inventory positions taken during the simulation.

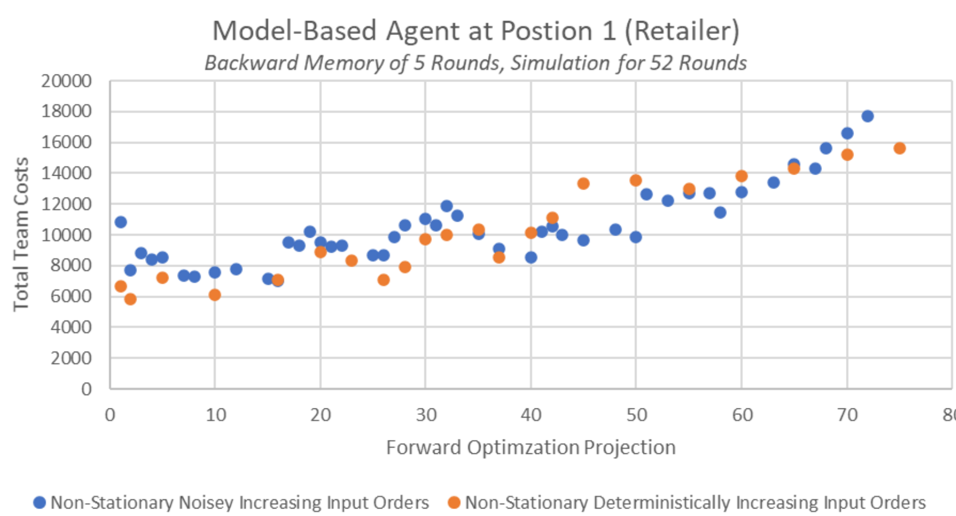


Figure C-2. Total Team Costs With MP Agent at Position 1 with Non-Stationary Increasing Orders

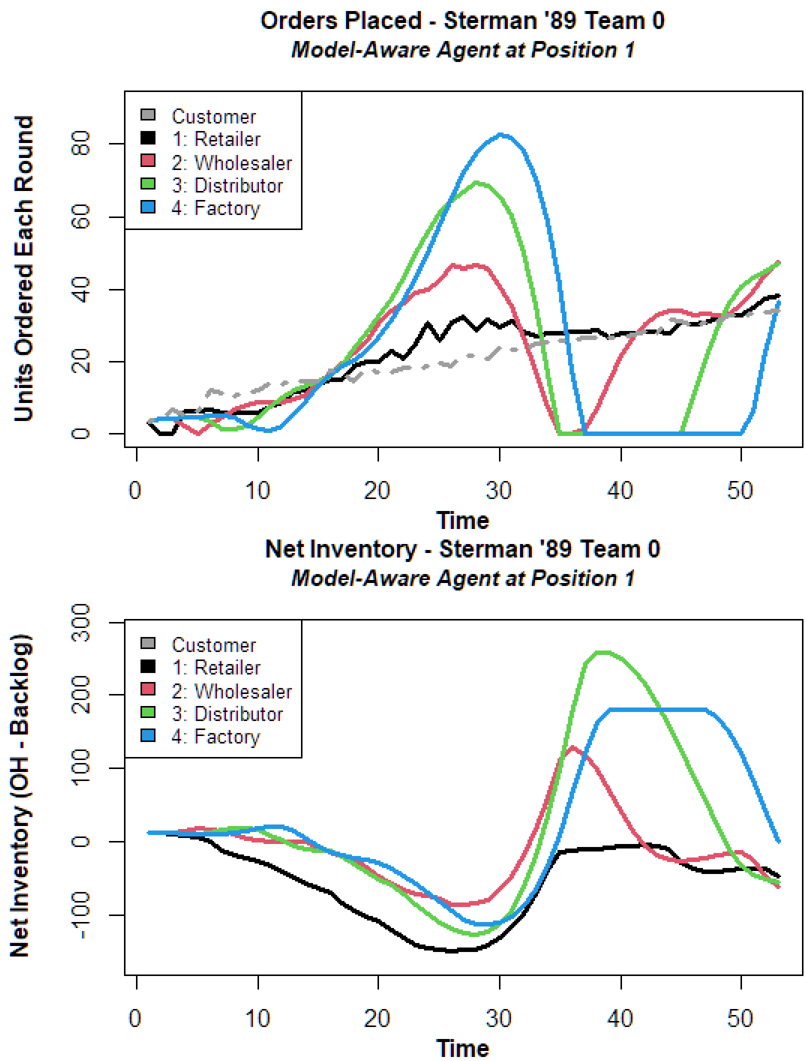


Figure C-3. Orders and Inventory with Agent at Position 1 and Horizon of 35

Figure C-2 would seem to imply that, generally, a smaller optimization horizon results in lower costs for the team. In other words, a greedier agent in position 1, one that only considers the immediate future, reduces overall team costs. However, consider Figure C-4 which zooms in on the case with optimization horizon equal to 10. While the average team costs over the 52 week simulation are objectively lower than in the case of the longer optimization horizon above, the agents behavior is arguably much worse. By attempting to minimize the costs that are incurred by having a destabilized supply chain, the agent effective ignores the increasing orders from the end customer, and eventually gets into a position later in the simulation where matching customer orders and incurring temporary disruptions in the supply chain are too costly over the near horizon.

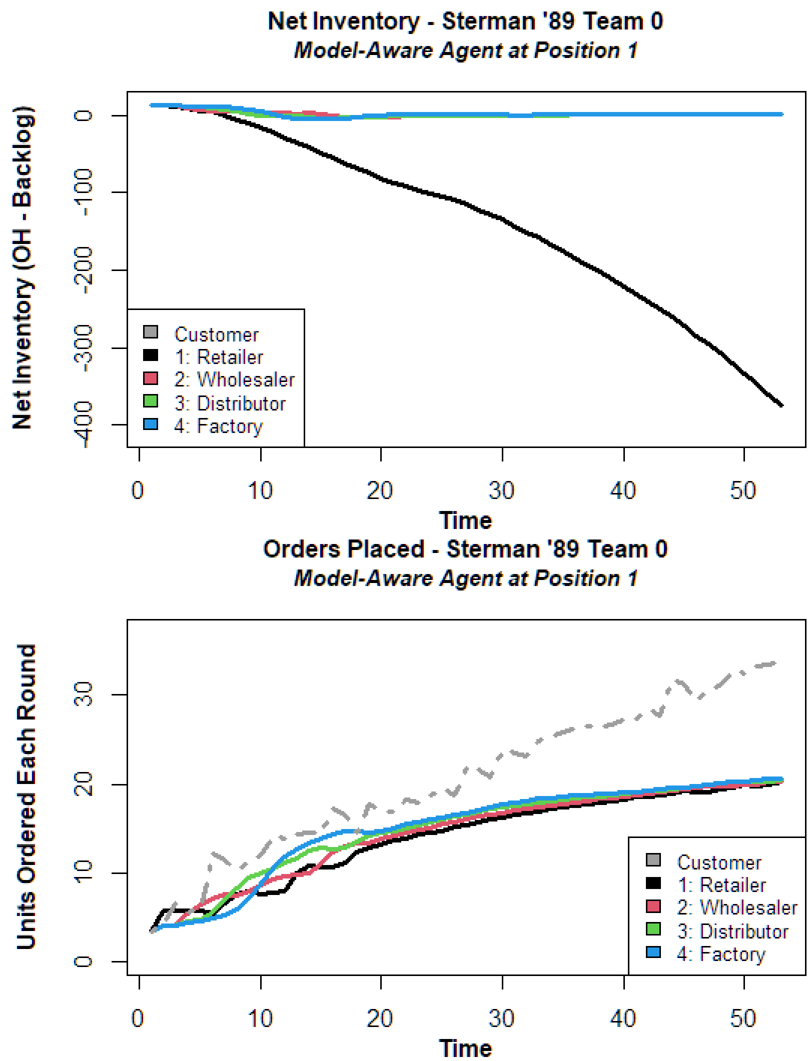


Figure C-4. Orders and Inventory with Agent at Position 1 and Horizon of 10

Such pernicious outcomes as function of hyperparameter choices are interesting, and of concern for specific scenarios but secondary to the focal points of the main chapter and thus left for this Appendix.

C.5.2 Matched vs Mismatched Environmental Assumption and Agent Response

As discussed elsewhere in the main chapter and in this Appendix, the most straightforward structural assumption of the model-predictive learning agent is to match the assumption of the environmental ordering rules with the agent's ordering rules. If the agent assumes the other entities in its environment are rational base-stock actors, then the best course of action for that agent is to also respond in a base-stock manner. Similarly, if the agent assumes other entities are not necessarily rational and following some other ordering heuristic, then using a behavioral response itself at minimum grants the agent more degrees of freedom to form its own ordering policy.

Figure C-5 shows box-and-whisker plots for the cost reduction for an agent placed at various spots the supply chain versus the baseline of no agent present for the same 49 teams used elsewhere in these analyses (the original 11 Sterman ’89 teams plus the average Sterman ’89 team plus 37 additional teams fitted from real order data from three separate runs of the Beer Game in 2021 and 2022), subject to the step input signal from the original Sterman 89 paper. Note that the truth’ of the environment in which the agent is acting is behaviorally-driven (all other agents are using the Sterman ’89 ordering rule). Note that this implies that the primary benefit comes from matching the assumption of the environment to the truth, independent of the agent architecture. However, this is less clear in the middle positions of the supply chain, and even reversed in some middle positions. The relationship when matching the ordering rule of the agent to the assumed environment is however consistent among all four positions.

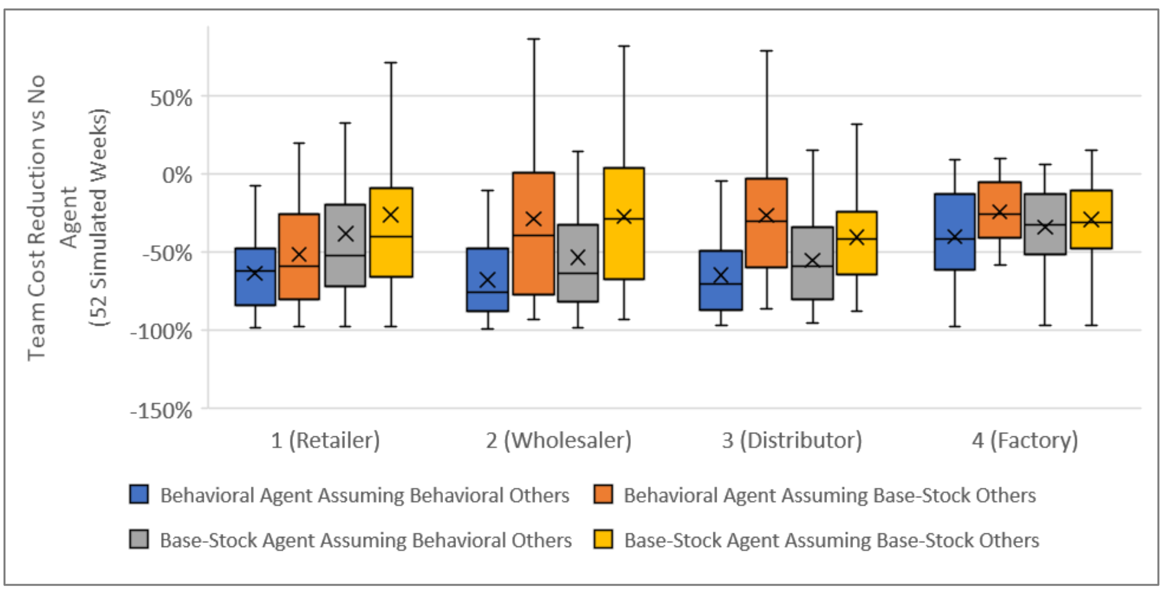


Figure C-5. Orders and Inventory with Agent at Position 1 and Horizon of 10

C.6 Choice of Baseline for Analyses

The base-stock values presented in the main chapter were derived for a team that all follow static base-stock ordering rules, and that perform best *on average* for 50 draws of the order distribution. However, it is possible to obtain even better performance for specific draws of the order distribution as an even more restrictive estimated upper bound on performance (or here lower bound on cost incurred). Starting from the values determined above, a further grid search of +/- 10 units was done for each of the draws of the order string actually faced by the example teams used in the analyses below to determine an order-string specific estimated lower bound on cost that could have been achieved by a team subject to the same order pattern and following a base-stock replenishment policy.

Perhaps unsurprisingly the costs incurred for the baseline simulated behaviorally ordering teams were significantly higher than that which would have been incurred by a similar supply chain in which all four entities were following the average cost-reducing base-stock replenishment policy, and much higher than those following an estimated idealized base-stock policy for the specific order pattern each team faced. As seen in Figure C-6, the median cost incurred by the 49 simulated behavioral teams was 13,898 while for the static base-stock team exposed to the same 49 input order patters experienced a median cost of 6,419, and for the case of order-string specific base stock rules a median cost of 3,211.

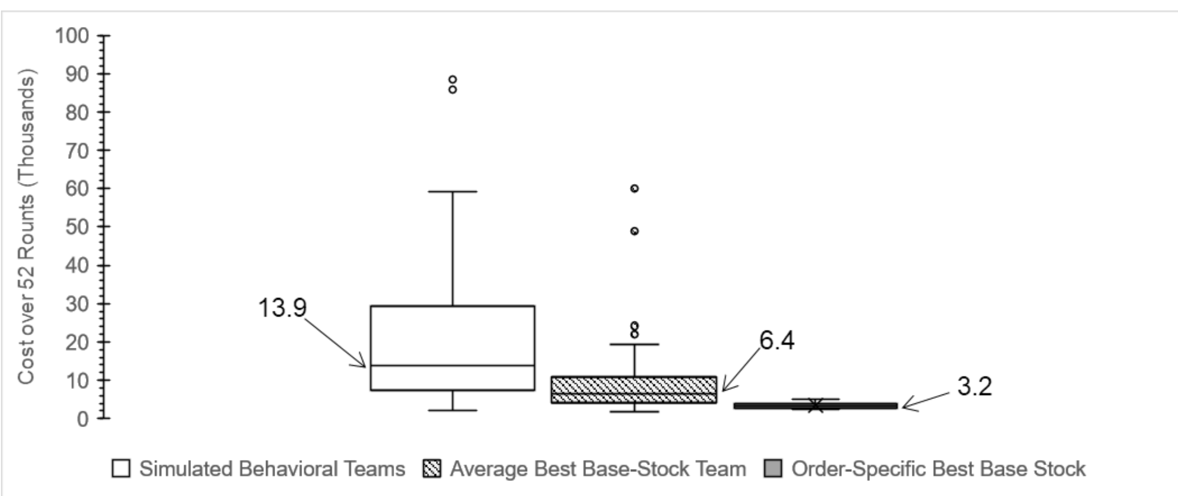


Figure C-6. Baseline Simulated Costs of Behavioral Teams vs Base-Stock Teams

However, the purpose of this research is not to simply restate that the behavioral responses result in poorer performance than an optimal decision rule. Rather, this work *assumes* that the other entities in supply chains are acting in a behavioral manner and asks what features of an agent placed into such an environment can help reduce costs overall. Thus, the results in Figure C-6 provide two different baselines of comparison for any such agent. While the full team of base-stock agents provides a realistic floor of costs, the simulated behavioral teams do not necessarily represent a true upper bound on costs. Indeed, one or two teams within that sample perform reasonably similarly to the base-stock teams even with simulated behavioral response rules. Therefore, it is reasonable to expect that an agent placed into these already well performing teams could even be *destabilizing* and introduce additional costs.

C.7 Robustness Check for Different Heuristic Parameter Combinations

As a review, the main paper focuses on agent response in the context of other entities utilizing the ordering heuristic introduced in (Sterman, 1989) and (Martin et al., 2004), and summarized in expressions (44) and (45) below:

$$O_t = MAX\left(0, \widehat{L_t} + \alpha_S(S' - S_t - \beta\, SL_t) + \varepsilon_t\right) \tag{44}$$

$$where\ \widehat{L_t} = \theta L_t + (1 - \theta)\hat{L}_{t-1} \tag{45}$$

This was used to fit 49 'human-like' agent models using real world order data. However, while the above parameters are bounded as $0 \leq \theta, \alpha, \beta \ \leq 1$ and $0 \leq S'$, the actual distribution of parameters in the fitted data do not uniformly fall in these ranges as seen in Figure C-7.

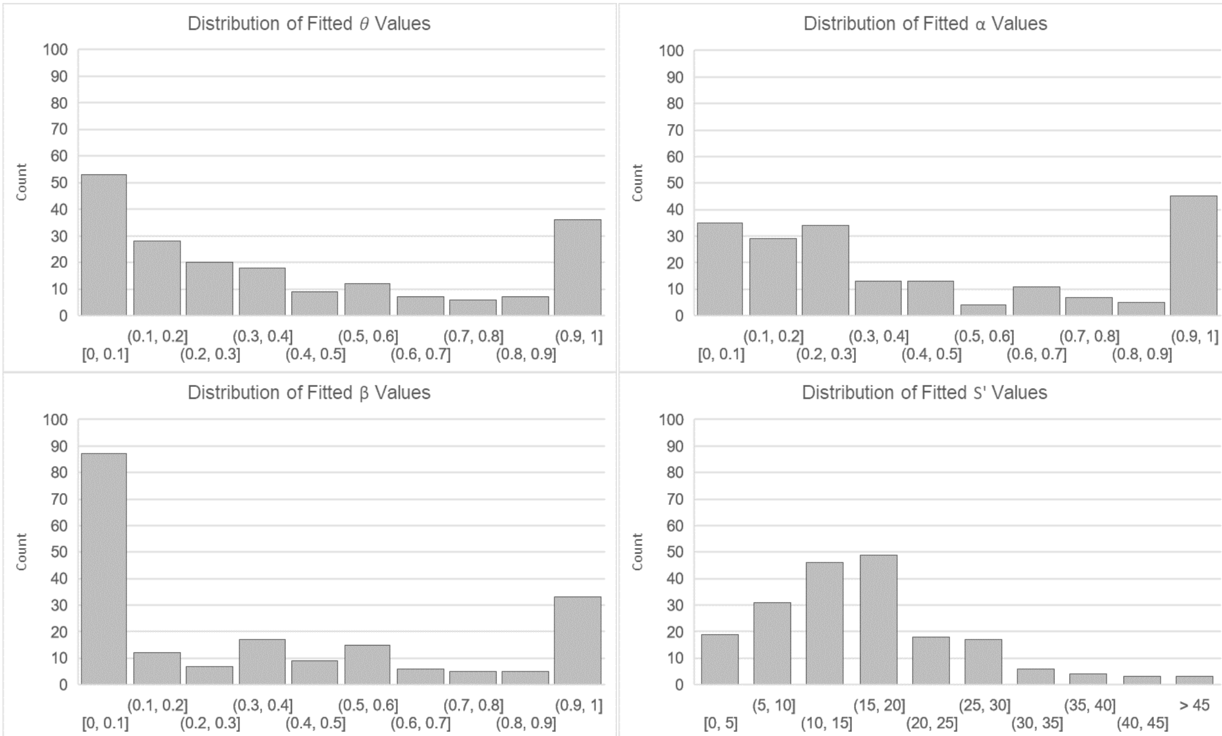


Figure C-7. Distribution of Fitted Sterman '89 Order Parameters

Given the semi-bimodal distributions of θ, α, and β, and the clustering of S' (but which can vary orders of magnitude greater), an argument could be made that the policies explored in this work apply only to situations in which the distribution of the parameters follows a similar pattern to the above.

To explore this possibility, and thus also test the robustness of the observations in the main text, consider the extreme points of the parameter space and how they map to other order policies:

- *Fully Rational:* α = 1, β = 1, θ = 1, S' = Clark and Scarf base-stock value. This is the classical full rational base-stock system described in the original Clark and Scarf paper.
- *Pass Through:* α = 0, β = 0, θ = 0. This is a rule mapping to no inventory management whatsoever, the entity just passes through orders.
- *Full Supply Chain Underweighting*: α = 1, β = 0, θ = 1, S' = Clark and Scarf base-stock number. This maps to an extreme case of the Sterman '89 rule wherein the agent is totally forgetful of the supply line.

- *Partial Supply Chain Underweighting*: α = 1, β = 0.5, θ = 1, S' = Clark and Scarf base-stock number. Here, the agent is partially underweighting the supply line by 50%, and was chosen specifically because it is in the middle of the bimodal distribution shown for this parameter in Figure C-7.

In addition to the above extreme points, mean-zero gaussian noise is added to the final order signal to match the process used in the main paper. Here, there is no *RMSE* measure of error to use for this noise, so the standard deviation of the noise was varied along $\sigma = [0, 2.5, 5.0, 7.5, 10]$. Combined with the four overarching conditions above, this creates a new set of 20 simulated teams in which to test these policy architectures (four main conditions with five sets of noise each).

The same style analysis as shown in the main chapter was performed here, only also incorporating the main effects of the values of α, β, and θ. As α only took on values of 0 or 1, its influence on this regression is colinear with that of β by examination of (44), and thus is not shown in the outputs below separately.

Table C-2. Feature Influence: Behavioral and Learning Agent at Position 2

	Dependent variable:		
	Performance Improvement over Baseline		
	(S1)	(S2)	(S3)
Behavioral Agent	-0.125^{***}		-0.127^{***}
	(0.043)		(0.043)
Learning Agent		-0.082	-0.090
		(0.066)	(0.065)
θ	0.336^{**}	0.297^{**}	0.344^{**}
	(0.134)	(0.134)	(0.134)
β	-0.224	-0.221	-0.224
	(0.178)	(0.180)	(0.178)
σ	-0.015	-0.019	-0.014
	(0.015)	(0.015)	(0.015)
Constant	-0.521^{***}	-0.505^{***}	-0.442^{***}
	(0.029)	(0.062)	(0.065)
Observations	360	360	360
R^2	0.036	0.017	0.041
Adjusted R^2	0.025	0.006	0.028
Residual Std. Error	0.387 (df = 355)	0.391 (df = 355)	0.386 (df = 354)
F Statistic	3.314^{**} (df = 4; 355)	1.575 (df = 4; 355)	3.034^{**} (df = 5; 354)
Note:			*p<0.1; **p<0.05; ***p<0.01

The magnitude and significance of the base-line model (a simple base-stock static order replenishment model), remains largely unchanged in these extreme supply chain models. Similarly, the value of incorporating explicit behavioral models is still valuable, though arguably simply because of the additional degrees of freedom those policies enjoy from additional parameters. Increasing noise does directionally reduce the performance of these agents, but not significantly so. Similarly, when faced with less rational and more behaviorally driven environments, the results from the main chapter stand.

Of interest, the influence of ‘learning’ in these extreme environments effectively drops out, though it was of notable significance in the results in the main chapter. Thus, while these robustness results reinforce the fundamental applicability of the overall observations of the main chapter (namely that most of the benefit is derived from having a base stock policy, which is improved by adding behavioral features), it also emphasizes that under certain conditions more complex policies are less applicable.

C.8 Additional Differences Across Supply Chain Positions

As stated in the main text of this chapter, it is reasonable to expect, or at least test for, different policies that emerge for an agent (or manager) in the middle versus the terminal positions in the supply chain. In addition to the tests done in the main text, there is another minor, but notable difference in the influence of an agent placed in the Retailer position in the supply chain when considering the extreme case where the threshold in the expression below is taken to the extreme minimum of $threshold = 0$.

$$Outlier \equiv \frac{Cost_{Agent_{Team_i}} - Cost_{Baseline_{Team_i}}}{Cost_{Baseline_{Team_i}}} > Threshold$$

In other words, looking *only* at the first position in the supply chain (the ‘Retailer’ here), and *only* at those policies that have no destabilizing influence whatsoever.

For this position and extreme threshold value, the introduction of a behavioral agent or even a dynamic learning agent has no significant improvement over the simplest static base-stock agent. In other words, for a manager at the *beginning* of the supply chain, dynamically learning about his or her downstream partners, or more accurately incorporating behavioral features into her model no more valuable than the simple static base-stock policy first described in the 1960’s, even in an environment that breaks many of the underlying assumptions of that early literature.

Table C-3. Behavioral and Learning Agent at Position 1 with Threshold = 0

	Dependent variable:			
	Performance Improvement over Baseline at Position 1			
	(1)	(2)	(3)	(4)
Behavioral Agent	-0.028		-0.028	0.027
	(0.018)		(0.018)	(0.053)
Learning Agent		-0.028	-0.028	0.003
		(0.028)	(0.028)	(0.040)
Behavioral x Learning				-0.062
				(0.056)
Constant	-0.577***	-0.566***	-0.552***	-0.580***
	(0.013)	(0.026)	(0.028)	(0.038)
Observations	792	792	792	792
Log Likelihood	-14.44	-15.2	-13.94	-13.33
McFadden ρ^2	0.081	0.033	0.113	0.151
R^2	0.003	0.001	0.004	0.006
Adjusted R^2	0.002	0.00002	0.002	0.002
Residual Std. Error	0.247 (df = 790)	0.247 (df = 790)	0.247 (df = 789)	0.247 (df = 788)
F Statistic	2.533 (df = 1; 790)	1.019 (df = 1; 790)	1.768 (df = 2; 789)	1.581 (df = 3; 788)
Note:				*p<0.1; **p<0.05; ***p<0.01

C.9 Stability of Different Architecture Choices

The main analyses about the influence of different agent architectures and information choices were all based on those runs that mixed both stabilizing and destabilizing agents. When incorporating *all* teams including those originally dropped in the main analysis as outliers, of the 3,528 total runs (49 runs each for each across 18 different feature and information configurations, for 882 runs per each of four positions in the supply chain), 407 total resulted in *higher* costs being realized with the presence of the agent versus the baseline performance of the simulated team. Similarly, among the subset of 3,118 runs that had model-predictive learning features, 284 were destabilizing.

Table C-4 and Table C-5 show the fraction of the 3,118 runs in each condition category of model-predictive learning agents that were destabilizing, resulting in more incurred costs versus the baseline simulated teams. While the absolute numbers of destabilizing runs are low (no more than 6 of 49 runs in each condition in each position in the supply chain), it is nevertheless interesting to observe that generally higher information availability results in lower occurrence of destabilizing runs. This is especially true for non-myopic agents that endeavor to reduce costs across the entire supply chain. For myopic agents focused only on their own cost reduction, there is an interesting trade-off implied among the lowest information availability states. The lowest information state, which as defined above is only the on-hand inventory of the agent and its own placed orders, is less likely to be destabilizing than providing a small additional amount of information in the calibration process (specifically the inbound shipments from the agent's supplier). This countertrend is only represented by a single difference in occurrence and thus cannot be said to be significant but should be noted for future inquiry.

Table C-4. Percent Destabilizing for Myopic Agents

		Information State			
		Minimal	*Low*	*Standard*	*High*
Agent Type	*Base-Stock (Myopic Agent)*	12.2%	12.2%	13.3%	10.7%
	Behavioral (Myopic Agent)	9.7%	11.7%	7.7%	5.6%

Note: N = 196 for all conditions

Table C-5. Percent Destabilizing for Non-Myopic Agents

		Information State			
		Minimal	*Low*	*Standard*	*High*
Agent Type	*Base-Stock (Non-Myopic Agent)*	10.7%	11.7%	10.2%	13.3%
	Behavioral (Non-Myopic Agent)	5.1%	5.1%	2.6%	3.1%

Note: N = 196 for all conditions

C.10 DQN Agent Architecture and Hyperparameters

The DQN agent introduced in the main chapter serves two purposes: 1) to provide a minor methodological contribution by providing another viable DQN approach towards managing ordering decisions in multi-echelon supply chain, specifically in the beer game, and specifically utilizing the dueling reward function architecture (Wang et al., 2016), and intends to extend directly from other recent work notably on architectures that use transfer learning (Oroojlooyjadid et al., 2021). 2) Provide a 'high complexity' point for comparison of other, often significantly less complex, agent policy architectures.

As the DQN itself is secondary to the central argument of this chapter, and recent prior literature by Oroojlooyjadid et al 2021 has provided a recent detailed assessment of the DQN architecture in this environment in general, only a curiously overview of the agent is provided in the main chapter. In the supporting material that accompanies this Appendix is all the code used to train both the model-free and model-aware versions of the DQN. To restate from the main chapter both DQN agents have the following general architecture: 1) An 'order-plus' action space (Oroojlooyjadid et al., 2021) which both allows for unbounded ordering in absolute terms and follows from observations in the model-based approach above, 2) a dual DQN network (Wang et al., 2016) that separately maintains a value function estimation for both the current overarching combined state of the system and separately for each action, 3) an observation space defined over a window of prior state observations corresponding to the signal delay in the system, 4) a combination of epsilon-greedy and Boltzmann exploration policies (Wiering, 1999), and finally 5) three sequential dense layers with ReLu activations of 256, 128, and 64 free parameters respectively for a total of 448 free parameters.

With respect to the hyperparameters of the system, the one of most interest is the amount of training steps employed and how this affects the spread in performance between the model-free version and the model-aware version of the agent. Figure C-8 shows the average cost improvement as a function of training steps when the agent is placed in position 1 (the retailer) for both versions the DQN. Note that this is inclusive of *all* 49 different teams, including destabilizing outcomes, and subject to the classical step input from Sterman '89. While there is *some* improvement from the Model-Aware agent versus the Model-Free agent (which is also seen in the main chapter), this improvement is relatively minimal once sufficient training steps occur. Indeed, the primary value of the model aware agent is under smaller training iterations. This supports similar observations made in the main chapter above the model-predictive learning agent as well, namely that additional information about the environment is less useful than expected. Here we hypothesize, given enough training data, the model-free agent can still determine a sufficient estimate of the state of the entire system from its own state variables without needing to be told an explicit representation of that system.

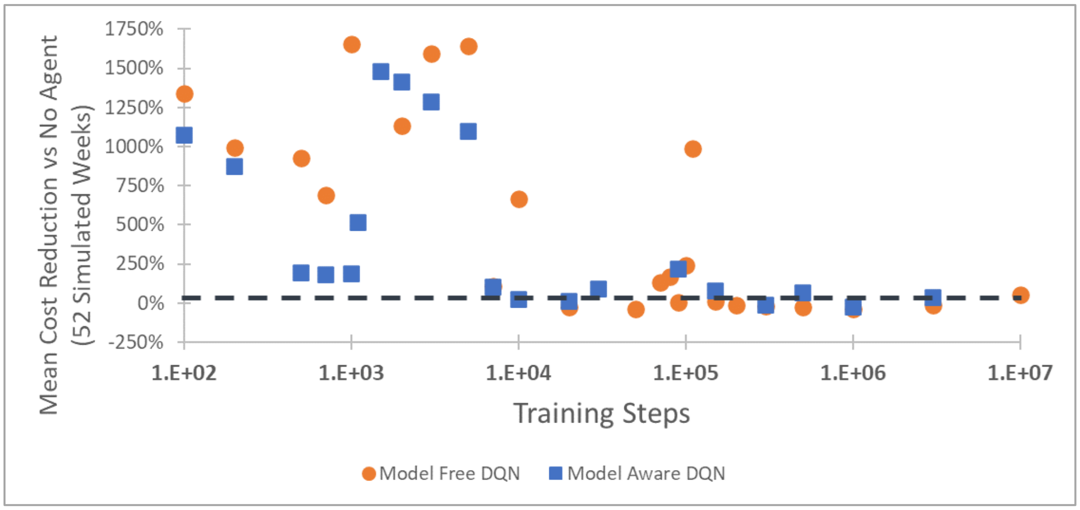


Figure C-8. Overall DQN Performance at Position 1 (Retailer) versus Training Steps

C.11 References to Appendix C

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., … Zheng, X. (2016). TensorFlow: A System for Large-Scale Machine Learning. *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). *OpenAI Gym*. http://arxiv.org/abs/1606.01540

Chen, F., & Samroengraja, R. (2009). The Stationary Beer Game. *Production and Operations Management*, *9*(1), 19–30. https://doi.org/10.1111/j.1937-5956.2000.tb00320.x

Chollet, F. (2015). *Keras*. GitHub. https://github.com/fchollet/keras

Clark, A. J., & Scarf, H. (1960). Optimal Policies for a Multi-Echelon Inventory Problem. *Management Science*, *6*(4), 475–490. https://doi.org/10.1287/mnsc.6.4.475

Croson, R., & Donohue, K. (2006). Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science*, *52*(3), 323–336. https://doi.org/10.1287/mnsc.1050.0436

Croson, R., Donohue, K., Katok, E., & Sterman, J. (2014). Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, *23*(2), 176–196. https://doi.org/10.1111/j.1937-5956.2012.01422.x

Martin, M. K., Gonzalez, C., & Lebiere, C. (2004). Learning to make decisions in dynamic environments: ACT-R plays the beer game. In M. Lovett, C. Schunn, C. Lebiere, & P. Munro (Eds.), *Proceedings of the Sixth International Conference on Cognitive Modeling: ICCCM 2004: Integrating Models* (Vol. 420, pp. 178–183). Lawrence Erlbaum Associates Publishers.

McNally, T. (2019). *Keras-rl2*. GitHub. https://github.com/tayloarmcnally/keras-rl2

Moritz, B. B., Narayanan, A., & Parker, C. (2021). Unraveling Behavioral Ordering: Relative Costs and the Bullwhip Effect. *Manufacturing & Service Operations Management*, *November*. https://doi.org/10.1287/msom.2021.1030

Narayanan, A., & Moritz, B. B. (2015). Decision Making and Cognition in Multi-Echelon Supply Chains: An Experimental Study. *Production and Operations Management*, *24*(8), 1216–1234. https://doi.org/10.1111/poms.12343

Oliva, R., Abdulla, H., & Gonçalves, P. (2022). Do Managers Overreact When in Backlog? Evidence of Scope Neglect from a Supply Chain Experiment. *Manufacturing & Service Operations Management*. https://doi.org/10.1287/msom.2021.1072

Oroojlooyjadid, A., Nazari, M., Snyder, L. V., & Takáč, M. (2021). A Deep Q-Network for the Beer Game: Deep Reinforcement Learning for Inventory Optimization. *Manufacturing & Service Operations Management*, msom.2020.0939. https://doi.org/10.1287/msom.2020.0939

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. http://www.rstudio.com/

Sterman, J. (1989). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science*, *35*(3), 321–339. https://doi.org/10.1287/mnsc.35.3.321

Sterman, J., & Dogan, G. (2015). “I’m not hoarding, I’m just stocking up before the hoarders get here.” Behavioral causes of phantom ordering in supply chains. *Journal of Operations Management*, *39–40*(1), 6–22. https://doi.org/10.1016/j.jom.2015.07.002

Wang, Z., Schaul, T., Hessel, M., Van Hasselt, H., Lanctot, M., & De Frcitas, N. (2016). Dueling Network Architectures for Deep Reinforcement Learning. *33rd International Conference on Machine Learning, ICML 2016*, *4*(9), 2939–2947.

Wiering, M. (1999). *Explorations in efficient reinforcement learning*. University of Amsterdam.