Why do nonlinearities matter? The repercussions of linear assumptions on the dynamic behaviour of assemble-to-order systems

The hybrid assembly-to-order (ATO) supply chain, combining make-to-stock and make-to-order (MTS-MTO) production separated by a customer order decoupling point (CODP), is well-recognized in many sectors. However, the study of ATO systems is very limited from a system dynamics perspective, in particular for those common nonlinearities present in the real-world supply chain systems. Nonlinear effects play an important, sometimes a dominant role in supply chain systems. Our aim thereby is to assess the impact of nonlinearities on the dynamic performance of the hybrid ATO system structure, including non-negative returns and capacity limits. We interrogate a generic nonlinear system dynamic model as an illustration of a typical hybrid ATO system, and benchmark against the well-known Inventory and Order Based Production Control System archetype. By adopting nonlinear control engineering and simulation approaches, we reveal that being aware of the ATO system's capacity and non-negative order constraints is very important, due to their significant impact on the system recovery speed and the CODP inventory level. Furthermore, a compromise between CODP inventory and capacity variations should also be evaluated, which is profoundly driven by the CODP inventory control policy. Future researches should consider the optimal trade-off design between CODP inventory and capacity as well as the exploration of delivery lead time dynamics.

Keywords: System dynamics, control engineering, nonlinearities, bullwhip, personal computer assemble-to-order systems, the IOBPCS family.

1. Introduction

Given the attractiveness of the assemble-to-order (ATO) strategy for companies, including increasing product variety, achieving quick response time and low cost, and benefiting the potential risk-pooling effect (Xiao, Chen, and Lee, 2012), academics and practitioners have become increasingly interested in analysing ATO systems. The ATO system is a hybrid production strategy that combines Make-to-stock (MTS) and Make-to-order (MTO) productions separated by a customer order decoupling point (CODP) (Naylor et al. 1999; Harrison et al. 2005) in the final assembly plant until the actual customized orders are received. This hybrid system is well-adopted by many manufacturing sectors, including personal computers (PC) (Katariya et al. 2014), semiconductors (Lin, Spiegler, and Naim, 2017), and printers (Tang and Tomlin 2008), to name but a few. From the stochastic modelling and analysis perspective, extensive academic studies can be found in the literature. The authors refer to Atan et al. (2017) for a comprehensive review.

However, the study of ATO systems is very limited from a system dynamics perspective. System dynamics plays a critical role in influencing supply chain performance under the volatile conditions of the current business environment (Spiegler and Naim 2017). Dynamic characteristics, particularly the bullwhip effect (Lee, Padmanabhan, and Whang,1997), are considered to be the main sources of disruption in the business world (Christopher and Peck 2004). The bullwhip effect refers to a phenomenon in which low variations in demand cause significant changes in upstream production for suppliers, with associated costs such as the ramp down and ramp up of machines, hiring and firing of staff, and excessive inventory levels (Wang and Disney, 2016). The PC industry and its associated semiconductor sectors have suffered severely from capacity unevenness, or the bullwhip effect (Karabuk and Wu 2003; Gonçalves, Hines and Sterman 2005), due to the characteristics of high levels of stochasticity and nonlinearity (Wang and Rivera 2008).

When confronted with system dynamics phenomenon such as bullwhip and inventory variance, the well-recognized inventory and order based production control system (IOBPCS) family, originally developed by Towill (1982), can be used, as the family models consist of general laws that represent many supply chain contexts (Lin et al. 2017), such as the well-known decision-making heuristic (Sterman 1989) that creates bullwhip (Lee, Padmanabhan, and Whang,1997), the order-up-to (OUT) policy (e.g. Wang et al. 2014) and remanufacturing systems (e.g. Zhou, Naim, and Disney, 2017). However, the IOBPCS family is traditionally used to represent MTS production systems, while limited effort has been made to model and analyse the dynamic behaviour of the ATO system. Furthermore, most IOBPCS based analytical studies assume that the system is completely linear, and thereby ignore those common nonlinearities present in the real-world supply chain systems, such as forbidden returns between suppliers and customers, capacity limits and shipment/inventory constraints, to name but a few. This has greatly limited the applicability of published results and has made it difficult to fully explain and describe oscillations caused by internal factors (Wang, Disney, and Wang, 2014). It has also been demonstrated that nonlinear effects play an important role in inventory systems, sometimes even a dominant role (Nagatani and Helbing, 2004). When linearity assumptions are removed complex dynamic behaviours are revealed. More importantly, oscillations generated internally by the system itself, rather than by the external environment, may arise.

As a result, this paper addresses the literature gaps and aims to develop a generic system dynamics model of an ATO system, to determine the impact of inherent nonlinearities on dynamic performance. We use the PC sector, a typical industry where the ATO strategy has been well-recognized and successfully implemented (Katariya et al. 2014), as an example to formulate the ATO model. Although the model is primarily based on the PC sector, it has general applicability to be easily adapted and extended to other sectors that employ an ATO strategy. Specifically, using combined control engineering and simulation approaches, we study the dynamic behaviour of the ATO system. Our contributions are twofold. First, we develop a dynamic model of ATO and investigate the impact of major control loops, including feedback inventory and feedforward forecasting policies, on the dynamic performance of ATO systems. Second, by adopting the nonlinear control engineering approach, namely the describing function method, we linearize the capacity and non-negative order constraints nonlinearities present in the ATO system, to determine the impact of nonlinearities on the dynamic performance. This offers analytical understanding about how the ATO system structure may characterise the dynamic oscillations and the possible strategy to avoid the poor dynamic behaviour, which would otherwise be missed if we relied only on linear assumptions.

The rest of the paper is organised as follows. Section 2 gives a review of existing contributions and gaps, thus providing the motivation for the paper. Then using PC as an example, section 3 provides the model formulation for an ATO system by exploiting control block diagrams and associated difference equations. The IOBPCS family is used to benchmark the ATO model. The analysis of feedback and feedforward loops as well as the nonlinearities present in the ATO system can be found in Section 4. All findings and corresponding managerial implications are summarized in Section 5.

2. Literature review

2.1 Simulation studies of nonlinear supply chain dynamics

Besides the impact of feedback loops and delays as the main sources of demand amplification as claimed by Forrester (1958), he also calls attention to the importance of considering nonlinear models to represent industrial and social processes. 'Nonlinearity can introduce unexpected behaviour in a system' (Forrester 1961), causing instability and uncertainty. In supply chain system structures, nonlinearities can naturally occur through the existence of physical and economic constraints, for instance fixed and variable capacity constraints in the manufacturing and shipping processes, variable delays and variable control parameters (Spiegler et al. 2016a).

Capacity and non-negative order constraints are two most common nonlinearities present in real-world supply chain systems and a number of simulation studies have analysed the impact of them. Regarding capacity constraints, Cannella, Ciancimino, and Márquez (2008) explored the relationship between constrained capacity and supply chain performance. Hussaina, Khan, and Sabir (2015) analysed the influence of capacity constraint and safety stock on the bullwhip effect in a two-tier supply chain by using Taguchi experiment. Ponte et al. (2017) investigated the impact of capacity limit on bullwhip and fill rate in an OUT-replenishment policy environment. The general conclusion derived in above study is that the capacitated supply chains may benefit from an improved dynamic performance as compared to unconstrained ones, due to capacity limit acts as a production smoothing filter. However, Cannella et al. (2018) found that the capacity may negatively influence the supply chain performance under a load-dependent lead time environment, i.e. lead times is modelled as a nonlinear function depending on the current work in progress (WIP) at the manufacturer and its capacity saturation limit and responsiveness (as the ability of the system in delivering the same product within a shorter lead time).

Several studies focus on the impact of demand smoothing and information sharing under non-negative order constraint supply chain systems, see Cannella, Ciancimino, and Framinan (2011), Cannella et al. (2014) and Syntetos et al. (2011). They highlighted the benefit of demand smoothing and information sharing in reducing supply chain dynamics, but non-negative order constraints are not studied in detail. Furthermore, Chatfield, and Pritchard (2013) and Dominguez et al. (2015) conducted simulation study regarding the impact of forbidden return policy on dynamic performance. The authors indicated that permitting returns significantly increases the bullwhip effect, and some other factors such as configuration of the supply chain network (serial vs. divergent) may play an important role in influencing the impact of non-negative order policy on supply chain dynamics (Dominguez et al. 2015).

Despite many researchers offered deep understanding of the impact of nonlinearities on supply chain dynamics, only simulation methods have been recommended to analyses nonlinear supply chain models. However, simulating complex systems without having first done some preliminary mathematical analysis can be time intensive and lead to a trial-and-error approach that may hamper the system improvement process (Sarimveis et al. 2008, Lin et al. 2017).

2.2. Control engineering studies of nonlinear supply chain dynamics

Classic control theory techniques with feedback thinking and sufficient analytical tools are advantageous for analytically analysing supply chain dynamics (Sarimveis et al. 2008). The application of classic control theory in a production-distribution system can be traced back to Simon (1952). Through adopting classic control theory, Towill (1982) translated Coyle's (1977) causal loops and presented an IOBPCS in a block diagram form. The IOBPCS family has been extensively studied within the context of the linear-based MTS supply chain systems. Topics include stability (e.g. Warburton et al. 2004; Wang, Disney, and Wang 2012), forecasting (e.g. Li, Disney, and Gaalman, 2014) and supply chain resilience (e.g. Spiegler et al. 2012), to name but a few. However, linear assumptions are often criticized for failing to capture the nature of nonlinear attributes of the real supply chain systems with resource and physical constraints (Lin et al. 2017). Recent works that specifically address this concern by using nonlinear control engineering techniques are summarized in Table 1.

Authors	The type of system	The assessment criteria	Nonlinear control engineering method	Key insights
Jeong, Oh, and Kim (2010)	MTS (Forrester model)	Stability, bullwhip and inventory variance	Matsubara time delay theorem; Small perturbation theory	Explore the effect of different capacity levels on the factory's production rate, unfilled orders.
Wang and Disney (2012) and Wang, Disney, and Wang (2014)	MTS (the order-up-to policy)	Bullwhip and Inventory variance	Eigenvalue methods	Explore the stability boundaries of a piecewise linear inventory control system (non-negative order constraint) and identify a set of behaviours in the unstable region.
Wang et al. (2015)	MTS (the order-up-to policy)	Bullwhip	Describing function	Identify the effect of non-negative order nonlinearity on the bullwhip effect in responding sinusoid demand, and propose strategies (forecasting, low ordering frequency) to mitigate bullwhip effect.
Spiegler et al. (2016a)	MTS (Empirical UK grocery model)	Bullwhip and Inventory variance	Describing function	Identify the influence of demand characteristics (frequency and amplitude) caused by shipment and truckload constraints on system dynamic behaviour, such as backlog, inventory and system's resilience.
Spiegler et al. (2016b)	Forrester model (Forrester 1961)	Bullwhip, inventory and shipment variance	Taylor series expansion with small perturbation theory; Matsubara low order modelling (Matsubara 1965)	Propose a simplification technique to provide a better visualization and understand of the variable interactions in the Forrester's model. also, the linearization approaches offer further insights due to the possible derivation of system's transfer function and local stability boundaries.
Wang and Gunasekaran (2017)	MTS and remanufacturing	Bullwhip and environmental dynamics	Taylor series expansion with small perturbation theory	Investigates the impact of production, environment, and demand variations on the dynamics and economic performance of sustainable supply chain systems. Their findings suggest that supply chain sustainability is essential to the continuous improvements of supply chain performance
Spiegler and Naim (2017)	MTS (APIOBPCS)	Bullwhip, inventory variance and stability (Limit Cycle)	Describing function	Investigate the effect of non-negative order and shipment constraints on the dynamic performance of the APIOBPCS model. The phenomenon called 'limit cycle', triggered by a non-negative nonlinearity is also explored.
This study	**ATO**	**Bullwhip, inventory variance**	**Describing function**	**Analytically explore bullwhip and inventory variance in the nonlinear ATO system with capacity and non-negative order constraints**

Table 1. Summary of applying nonlinear control engineering approaches in studying supply chain dynamics

Although recent studies contribute to the understanding of the effect of nonlinearities on the dynamic behaviour of the production-inventory system, there are several common limitations. Jeong, Oh, and Kim (2010) only use simulation to analyse the effect of different capacity constraints on the dynamics behaviour, despite efforts to linearize a part of the model. Also, most studies consider the impact of different nonlinearities on dynamic performance of the system individually. For instance,

Wang and Disney (2012), Wang, Disney, and Wang (2014) and Wang et al. (2015)'s studies are limited to the analysis of the non-negative order constraint on the replenishment order, and Spiegler et al.'s (2016a, 2016b) analysis is limited to the capacity constraint. In particular, no previous analytical study has considered the impact of capacity and non-negative order constraints simultaneously when orders are placed to the supplier. Furthermore, all studies solely explore the dynamic performance of MTS-based production control system by utilizing bullwhip and inventory variance as the main performance indicator, while, to our best knowledge, no previous work has analytically explored nonlinear ATO systems with capacity and non-negative constraints.

2.3. ATO system dynamics

From a system dynamics perspective, existing literature puts major emphasis on the dynamic modelling and analysis of CODP by developing and simulating the hybrid MTS-MTO model (Hedenstierna and Ng 2011; Choi, Narasimhan, and Kim, 2012; Wikner et al. 2017). Specifically, by decoupling generic FD (forecasting-driven) and CD (customer-driven) models, Hedenstierna and Ng (2011) evaluated the dynamic consequences of shifting the position of the CODP and found that the ideal position depends on the frequency of demand. However, their model is linear and lacks more realistic representations, such as capacity constraints and availability of material. Choi, Narasimhan, and Kim (2012) developed a system dynamics simulation model from Lee and Tang's (1997) model and their experiences gained through a case study in a Korean automobile manufacturer. In contrast to Hedenstierna and Ng (2011), their model represents complex variable relationships, but their simulation results are limited to Korean global automobile companies. Wikner et al. (2017) conceptually developed a hybrid MTS-MTO model that can represent a typical ATO system by decoupling the customer orders at the final assembly plant. By using system dynamic simulation, they highlight the significant impact of capacity constraint at the downstream of CODP on backlog and CODP inventory dynamics, although the conceptual model does not explicitly consider the upstream capacity limit.

Within the context of the PC sector, the focus of our study, limited effort has been made to study the dynamics of ATO systems. Berry and Towill (1992) developed causal loop diagrams to explain the 'gaming' that yields bullwhip in electronics supply chains, including semiconductor production, while Berry, Towill, and Wadsley (1994) undertook simulation modelling of a generic electronics industry supply chain to highlight the opportunities afforded by different supply chain reengineering strategies to mitigate bullwhip. However, their model did not explicitly represent the CODP and nonlinearities in the hybrid ATO system. Gonçalves, Hines, and Sterman (2015) developed a system dynamics simulation model to explore how market sales and production decisions interact to create unwanted production and inventory variances in the Intel hybrid ATO supply chain. Although these system dynamics simulations contribute to the representation of a real system by incorporating nonlinear components and complex structures, their trial-and-error approach limits the system improvement process. Lin, Spiegler, and Naim (2017) overcome such limitations by analytically exploring the Intel's hybrid ATO model using the linear control engineering approach. The analytical insights including the stability region as well as the root causes of the bullwhip effect are derived and verified by simulation tests. However, their study is limited to a linear analysis, which cannot represent the realistic supply chain system.

Overall, two main limitations are identified regarding the studies of the ATO system from system dynamics perspective. First, most studies do not consider the impact of nonlinearities, especially the capacity and non-negative order constraints, on the dynamics of the ATO system. Second, simulation is the primary choice for most studies and thereby gives little analytical insight or guidance in understanding the system control policies and structures to reduce supply chain dynamics. We aim to address these gaps by dynamic modelling and analysis of the nonlinear ATO system using PC supply chain as an example.

3. Modelling the ATO system.

3.1 The PC supply chain description

As visualized in Figure 1, there are usually four major echelons for a PC company supply chain: component production (e.g. semiconductor fabrication), sub-assembly, final assembly and distribution/dealer (Berry, Towill, and Wadsley 1994; Naylor, Naim, and Berry 1999; Huang and Li 2010; Katariya et al. 2014). From the material flow perspective, the component and sub-assembly echelons offer 'commodities' required by final PC final assembly, and the corresponding lead times is measured in terms of weeks. As the material flows downstream, production moves from automated production to highly manual operations. Final assembly of a PC is a largely manual process to allow quick changeovers and high levels of flexibility, with the corresponding lead times is measured in terms of days. The final products are either shipped to a number of the company owned distribution centres or directly to authorized dealers/final customers.
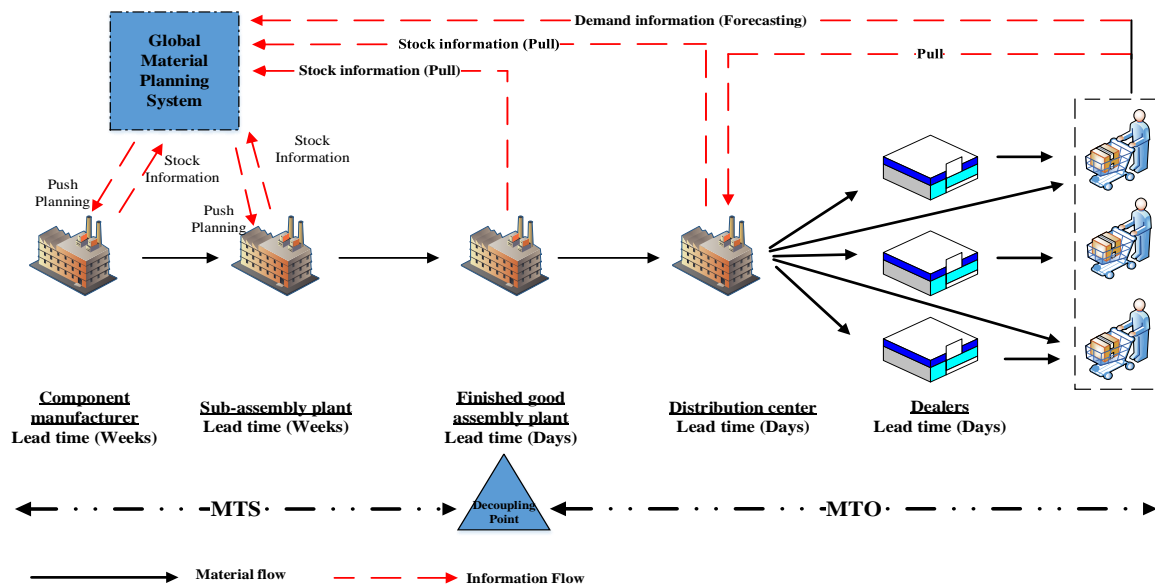

Figure 1. A general material and information flow of PC supply chains (Naylor, Naim, and Berry 1999).

Regarding information flows, companies adopted the hybrid production planning and control strategy, i.e. the CODP is implemented in the final assembly echelon, which separates the downstream MTO and upstream MTS production and the whole supply chain is an ATO structure. The replenishment of downstream distribution and retailers is driven by actual customer orders, while the upstream production of CODP is based on a forecast generated via a global material logistics system that links all manufacturing units in the supply chain, which use the stock and forecast information from downstream echelons of CODP to generate a bill of materials. As a result, the design of the hybrid system balances customer responsiveness and cost efficiency. Note that such an ATO system structure illustrated by Figure 1 can be modified as a general one to be applied in other sectors by considering the location of CODP and the characteristic of pull and push planning. For instance, the semiconductor internal fabrication and final assembly has a similar ATO structure in which upstream wafer production uses long-term forecasting from end customer and downstream final assembly pull rate as the desired wafer start rate, while the downstream final assembly and distribution is directly pulled by end customer orders (Lin, Spiegler, and Naim 2017).

3.2 Modelling the PC supply chain.

The materials/information flows of a PC company's supply chain is modelled at an aggregate level. The model is restricted to one supply chain player per echelon and this corresponds to the minimum number of echelons for players required to analyse its dynamic behaviour. The entire supply chain is modelled as a two-echelon system, i.e. a sub-assembly and final assembly system connected by CODP inventory to represent the typical hybrid ATO structure. The downstream distribution/sales and marketing echelons are not considered in this study, since the orders can be directly transferred to final assembly plant via on-line shopping. Also, the upstream component echelon is not considered due to the same ordering policy adopted in the sub-assembly echelon, i.e. MRP replenishment rule. All notations used in this paper are presented as follows:

$AINV_{AS}$: PC parts inventory at the final assembly plant
$AINV_{AS}^{*}$: Targeted PC parts inventory at the final assembly plant
$AINV_{SA}$: PC parts inventory at the sub-assembly plant
$AINV_{ASadj}$: PC parts inventory adjustment at the final assembly plant
AVCON: Averaged consumption rate
BL: Backlog level
BL*: Target backlog level
BL_{ADJ}**:** Backlog adjustment
CL: Capacity limit
CONS: Customer demand rate
$COMRATE_{AS}$**:** Completion rate for final assembly
$COMRATE_{SA}$**:** Completion rate for sub-assembly
$ORATE_{AS}$**:** Order rate for final assembly plant
$ORATE_{SA}$**:** Order rate for sub-assembly plant
S: Actual shipment
S*: Desired shipment
S_{MAX}**:** Maximum shipment
τ_{A}: Time to smooth customer demand
τ_{AS}: Final assembly delay
τ_{BL}**:** Time to adjust backlog discrepancies
τ_{I}**:** Time to adjust raw inventory error at the final assembly plant
τ_{DD}**:** Desired order fulfilment delay (e.g. final assembly, order processing and delivery delay)
τ_{SA}**：** Sub-assembly delay
τ_{SA}'**:** Estimated sub-assembly delay, which is assumed as $\tau_{SA}' = \tau_{SA}$, in line with John, Naim, and Towill (1994)
WIP: Work-in-process inventory level for the sub-assembly system
WIP*: Desired WIP
WIP_{ADJ}**:** WIP error adjustment
s: Laplace transform operator
$\triangle T$**:** Time interval between samples
IOBPCS: Inventory and Order Based Production Control System
VIOBPCS: Various Inventory and Order Based Production Control System
APVIOBPCS: Automatic Pipeline and Various Inventory and Order Based Production Control System

Modelling the final assembly echelon

The final assembly plant can be described in terms of two inter-linked control structures (Berry, Towill and Wadsley 1994; Mason-Jones, Naylor and Towill 2000). The first structure focuses on the control of physical final assembly transformation under the pure order-driven strategy. The second is responsible for the replenishment of raw materials (PC parts) as the inputs ($AINV_{AS}$) into the transformation process. To model the first control structure, the relationship between incoming orders and the replenishment of $AINV_{AS}$ should be captured. In most cases, the exogenous demand into the supply chain system begins when end customers decide the PC configurations with dealers. Dealers will electronically transmit orders to order desks in the company's sales and marketing organizations, the order desks check the availability of PC parts stock in the decoupling point, i.e. the final assembly echelon. If all required $AINV_{AS}$ are available, confirmation of orders including expected delivery information is confirmed and the plant starts to the assembly and ships to customer by quoted lead times (τ_{DD}). From the aggregate perspective, this is reflected by S* in each period. However, if $AINV_{AS}$ is insufficient, the final assembly plant can only assembly all PC parts they currently have on hand and this is reflected by S_{MAX}.

The first order lag approach (Sarimveis et al. 2008) can be used to model the MTO based final assembly process. Specifically, depending on the availability of $AINV_{AS}$, the output of first order delay, i.e. S, is determined by

$$S(t) = Min\left(S^*(t), S_{MAX}(t)\right) \qquad (1)$$

If required $AINV_{AS}$ are available for immediate final assembly, $S=S^*$, the difference between inflow CONS and outflow S^* is calculated as a measure of BL. i.e. a kind of work-in-progress orders, WIP (Wikner 2003):

$$BL(t) = BL(t-1) + CONS(t) - S(t) \qquad (2)$$

The output S* is the result of the fraction of WIP ($1/\tau_{DD}$). In other words, τ_{DD} is the average delay of the production unit. As suggested by Atan et al. (2017), a fixed τ_{DD} is a realistic assumption due to high flexibility and reliable delivery time for the final assembly process.

$$S(t) = S^*(t) = \frac{BL(t)}{\tau_{DD}} \qquad (3)$$

Under such conditions, all incoming customized orders can be fulfilled by quoted τ_{DD}, that is, customers need to wait for physical lead times only. However, if insufficient $AINV_{AS}$ constrains S*, the final assembly can only ship S_{MAX} estimated by current $AINV_{AS}$ and τ_{DD}.

$$S(t) = S_{MAX} = \frac{AINV_{AS}(t)}{\tau_{DD}} \qquad (4)$$

The second inter-linked control structure considers the replenishment of $AINV_{AS}$ resulted by $COMRATE_{AS}$. $COMRATE_{AS}$ is the result of delayed $ORATE_{AS}$ (transport delay between sub-assembly and final assembly plant). A first order lag is used to model it (Sipahi and Delice 2010):

$$COMRATE_{AS}(t) = COMRATE_{AS}(t-1) + a\left(ORATE_{AS}(t) - COMRATE_{AS}(t-1)\right) \quad (5)$$

$$\text{Where } a = \frac{1}{\left(1 + \frac{\tau_{AS}}{\Delta T}\right)} \quad \text{(Towill 1977)}$$

Where $ORATE_{AS}$ is determined by the minimum between desired Pull $ORATE_{AS}$ from the final assembly echelon and the feasible Push $ORATE_{SA}$ from the sub-assembly echelon:

$$ORATE_{AS}(t) = Min\left(Pull\ ORATE_{AS}(t),\ Push\ ORATE_{SA}(t)\right) \qquad (6)$$

In other words, only smaller signals can pass the Min function as the feasible $ORATE_{AS}$. If there are enough finished PC parts in the sub-assembly echelon, the customer's orders pull the replenishment of $AINV_{AS}$, otherwise the sub-assembly plant pushes all feasible $AINV_{SA}$ to meet the final assembly requirement as soon as possible. Pull $ORATE_{AS}$ aims to eliminate gaps for $AINV_{AS}$ and BL. More reliable S as a proxy is also used for deciding Pull $ORATE_{AS}$ and a non-negativity constraint is given to avoid negative order rate for the final assembly:

$$Pull\ ORATE_{AS}(t) = Max\big(0,\ AINV_{ASadj}(t) + S(t) + BL_{ADJ}(t)\big) \qquad (7)$$

Where $AINV_{ASadj}$ is the $AINV_{AS}$ feedback loop based on the discrepancies between $AINV_{AS}*$ and $AINV_{AS}$ adjusted by τ_I:

$$AINV_{ASadj}(t) = \frac{1}{\tau_I} \cdot \big(AINV_{AS}^*(t) - AINV_{AS}(t)\big) \quad (8)$$

$$AINV_{AS}^*(t) = S(t) \cdot \tau_{AS} \quad (9)$$

and BL_{ADJ} is the backlog control loop adjusted by τ_{BL}:

$$BL_{ADJ}(t) = \frac{1}{\tau_{BL}} \cdot \big(BL(t) - BL^*(t)\big), \quad BL^*(t) = CONS(t) \cdot \tau_{DD} \qquad (10)$$

Usually the sub-assembler can supply PC parts according to the planned requirements to satisfy Pull $ORATE_{AS}$. However, a loop exists to indicate an inability to meet demand. If part of the required parts stock is insufficient for immediate transport to the final assembly plant, a re-defined delivery date is given to end customers. This means the plant will delay the final assembly due to late arrive of required parts from the upstream sub-assembly plant. In this situation, the supplier submits their best can do commitment and push out all feasible $AINV_{SA}$ to meet the downstream requirement:

$$Push\ ORATE_{AS}(t) = AINV_{SA}(t) \qquad (11)$$

Modelling the sub-assembly echelon

The upstream sub-assembly echelon operates as the MTS state driven by a forecast of customer demand for various PC configurations. The output from this process is a 12 months' parts requirement generation instruction fed into upstream echelons. Also, each plant in the downstream echelons of CODP is responsible for feed backing their stock and order-in-process information, known as "coverage", to sub-assembly and component echelons. Once the instruction has been generated, the system automatically calculates how many sub-assembly and component part numbers are required and when they are required.

As a result, the upstream echelons are a typical MRP-based ordering system and the well-established APVIOBPCS (Wang et al. 2014) can be used to model the upstream sub-assembly system. Specifically, for each replenishment cycle, $ORATE_{SA}$ is determined by:

$$ORATE_{SA}(t) = Max\left(0, Min\left(\begin{matrix} AVCON(t) + AINV_{SAadj}(t) \\ + FWIP_{ADJ}(t), \\ CL \end{matrix}\right)\right) \quad (12)$$

AVCON(t) is the feedforward forecasting policy and well-recognized exponential smoothing can be adopted, although other forecasting methods such as moving average (Dejonckheere et al. 2002) and damped trend forecasting (Li et al. 2014) can be considered:

$$AVCON(t)=AVCON(t\text{-}1)+c\big(CONS(t)\text{-}AVCON(t\text{-}1)\big), c=\cfrac{1}{1+\cfrac{\tau_A}{\triangle T}} \quad (13)$$

$AINV_{SAadj}$ is the finished good inventory adjustment loop within the sub-assembly plant based on the discrepancies between $AINV_{SA}*$ and $AINV_{SA}$:

$$AINV_{SAadj}(t)=\frac{1}{\tau_{AINV}} \cdot \big(AINV_{SA}{}^{*}(t)\text{-}AINV_{SA}(t)\big) \qquad (14)$$

Where $AINV_{SA}*$ is based on sub-assembly estimated lead time ($\boldsymbol{\tau_{SA}}' = \boldsymbol{\tau_{SA}}$) and pull $ORATE_{AS}$, i.e. safety stock is calculated by the amount of PC parts as raw materials required by downstream final assembly and covered by averaged sub-assembly lead time:

$$AINV_{SA}{}^{*}(t)=\tau_{SA}\cdot Pull\ ORATE_{AS}(t) \qquad (15)$$

and $AINV_{SA}$ depends on the accumulation between the replenishment from $COMRATE_{SA}$ and the actual depletion of $ORATE_{AS}$ (minimum between Pull and Push $ORATE_{AS}$);

$$AINV_{SA}(t)=AINV_{SA}(t\text{-}1)+COMRATE_{SA}(t)\text{-}\ ORATE_{AS}(t) \qquad (16)$$

Also, the dynamic role of WIP inventory in the sub-assembly system is considered in an MRP ordering system, which can be interpreted as products queue at the detailed level. In line with John, Naim and Towill (1994)'s standard modelling approach, a fraction of WIP error (WIP_{ADJ}) is corrected based on the difference between WIP* and WIP:

$$WIP_{ADJ}(t)=\frac{1}{\tau_{WIP}} \cdot \big(WIP^{*}(t)\text{-}WIP(t)\big) \qquad (17)$$

Where WIP* depends on AVCON and $\boldsymbol{\tau_{SA}}$, and WIP is accumulative level between $COMRATE_{MTS}$ and $ORATE_{MTS}$. Furthermore, a first order lag is used to model the physical sub-assembly lead time, which can be interpreted as a production smoothing element representing how slowly the production units adapts to changes in $ORATE_{AS}$ (Wikner 2003):

$$WIP^{*}(t)= \tau'_{SA}\cdot AVCON(t);\ WIP(t)=WIP(t\text{-}1)+ORATE_{SA}(t)\text{ - }COMRATE_{SA}(t)$$

$$COMRATE_{SA}(t)= COMRATE_{SA}(t\text{-}1)+c\big(ORATE_{SA}(t)\text{-}COMRATE_{SA}(t\text{-}1)\big),\ c=\cfrac{1}{(1+\cfrac{\tau_{SA}}{\triangle T})} \quad (18)$$

Based on Equations (1)- (18), the entire ATO model is presented in Figure 2 in block diagram form, using the Laplace s domain. The model consists of basic elements such as flow, stock, decision policies, feedback and delays. Also, there existing two switches (final assembly and sub-assembly echelons) defined by Min functions that separate the system into different operational statuses depending on the feasible inventory, i.e. the $AINV_{SA}$ in sub-assembly echelon and the $AINV_{AS}$ in the final assembly system. Only one signal can be passed based on to the comparison between desired Pull and feasible Push, since the desired hybrid ATO system is only operated if there is enough inventory at each echelon. Having developed the model, it is important to verify the logic and correctness of the model (Sargent 2013). This verification process is undertaken via simulation on Matlab[TM]. Although we do not show the full verification results, part of the simulation analysis is reported in Table 2. The verification result shows the hybrid ATO model is logical and correct.
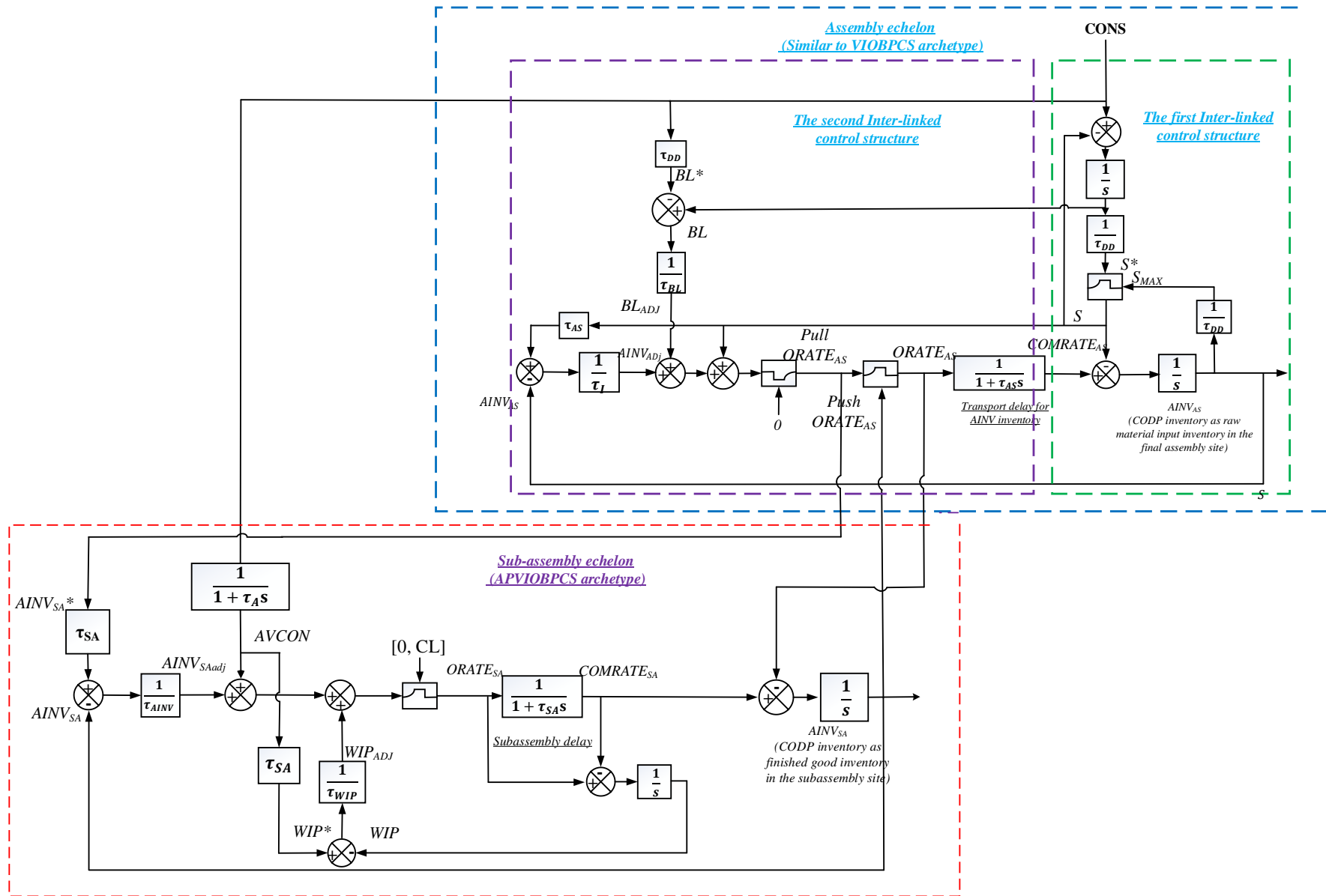
Figure 2. System dynamics model for the ATO supply chain.

Verification test	Details	Verification process	Verification results
Family member and parameters	Behaviour reproduction for cognate system and be consistent with system data and description	1.Regarding the final assembly system, we use the similar Intel supply chain model (Lin et al. 2017) to reproduce its dynamic behaviour in responding a step demand increase by utilizing the same system parameter settings, i.e. $\tau_{AS} = \tau_I = \tau_{BL} = 2\tau_{DD} = 4$. 2. For the sub-assembly system, order-up-to policy (i.e. $\tau_{SA} = \tau_A/2 = 8$, $\tau_{AINV} = \tau_{WIP} = 1$) is used to check whether the dynamic behaviour is consistent with Dejonckheere et al. (2003).	1. Dynamic behaviour of the final assembly is consistent with the Intel hybrid supply chain model e.g. maximum overshoot/undershoot, rising time and setting time. 2. The dynamic performance of the order-up-to policy can be reproduced.
Boundaries and Structure	Include all important factors and be consistent with system description	Related empirical works including Kapuscinski et al. (2004), Katariya et al. (2014) and Huang and Li (2010) are used to check the consistency regarding the system framework and important factors of the ATO supply chain.	1. The ATO system dynamic model is consistent with empirical descriptions characterized by combined order- and forecasting-driven production, and material and information CODP. 2. All important factors are included for the system dynamic model. Also, the model is cross-checked by corresponding Intel supply chain (Lin et al. 2017), APVIOBPCS and VIOBPCS archetypes (Edghill and Towill 1990; John et al. 1994; Dejonckheere et al. 2003).
Extremities	Model is logical for extreme values	1. We check whether the dynamic performance of the final assembly system is consistent with the VIOBPCS archetype (Edghill and Towill 1990) if $\tau_{BL} = \tau_{DD} = \infty$ 2. For the supplier manufacturing part, we increased the value of τ_{WIP}, τ_{AINV} and τ_A to extreme conditions to see whether the system can generate the expected dynamic outcome.	1. The dynamic behaviour of the final assembly system is consistent with corresponding performance in the original VIOBPCS if the backlog and shipment loops are removed. 2. The extreme values of τ_A, τ_{AINV}, and τ_{WIP} will lead to the expected dynamic performance in responding to a step demand increase. For example, the infinite τ_{AINV} will remove the inventory feedback loop, which result in the permanent inventory drift in responding a step increase as expected.

Table 2. The verification of the PC system dynamics model.

4. Dynamic analysis of the ATO system

In this section the dynamics of the ATO system are analysed by using combined nonlinear/linear control engineering and system dynamics simulation. First, we analyse the completely linear system by assuming all nonlinearities are inactive in order to understand the impact of feedback and feedforward loops on the dynamic behaviour of the ATO system. This gives the fundamental dynamic insight of the ATO system regarding the adjustment of corresponding control on the dynamic behaviour of two inventory ($AINV_{AS}$ and $AINV_{SA}$) and production capacity fluctuations ($ORATE_{SA}$).

We then analyse the impact of nonlinearities present in the ATO system by classifying the nonlinearities and identifying appropriate simplification and linearization methods. The 'filter' demand input signals (Towill, Zhou, and Disney 2007), or sinusoidal input, evident in the PC industry (Lin, Spiegler, and Naim, 2017) is used to assess the dynamic performance of the ATO system. Analysing system dynamics models via the 'filter lens' or sinusoidal input allows important dynamic properties of the system to be investigated, including the natural frequency (ω_n) and damping ratio (ζ). The former determines how fast the system's output oscillates during the transient response, while the latter describes how oscillations in the system decay with time. Table 3 categorizes nonlinearities present in the ATO system and illustrates corresponding simplification/linearization methods adopted in this study. As a result of such simplification and linearization, the original ATO model can be simplified as ***a truly ATO state structure*** illustrated in Figure 2, including only capacity and non-negative order nonlinearities. Its corresponding block diagram, based on simplification of multi-valued discontinuous nonlinearity (Table 3), is presented in Figure 3.

Type of nonlinearity in this study	Main characteristics	Simplification/linearization methods
Single-valued discontinuous nonlinearity: 1) Non-negative order constraint i.e. Equation (7) and (12) 2) Capacity constraint in the supplier manufacturing plant, i.e. Equation (12).	Sharp changes in output values or gradients in relation to input (e.g. piecewise linear function). Single-valued nonlinearities are also called memory-less, which means that the output value does not depend on the history of the input (Spiegler et al. 2016a).	1. The describing function method is used to linearize such nonlinearities in responding sinusoidal demand input. 2. Characteristics equations analysis derived from linearized transfer function is conducted. System dynamic simulation is used for verification.
Multi-valued discontinuous nonlinearity: 1) Shipment constraint, i.e. Equation (1) 2) CODP inventory constraint, i.e. Equation (6).	In contrast to the single-value nonlinearity, the output value of multi-valued discontinuous nonlinearity does depend on the history of the input. e.g. changes in manufacturing strategies depending on foreign exchange rate directions.	Two multi-valued nonlinearities (i.e. switches) govern different operational states of nonlinear ATO supply chains depending on the feasible $AINV_{AS}$ and $AINV_{SA}$. However, we only analyse the ***truly ATO state*** by assuming: $$S = S^*$$ $$ORATE_{AS} = Pull\ ORATE_{AS}$$ This scenario also fits the reality that PC supply chains maintain the ATO state by ensuring enough CODP inventory (Lin, Spiegler, and Naim, 2017).

Table 3. Types of nonlinearities present in the ATO system and corresponding linearization and simplification method.
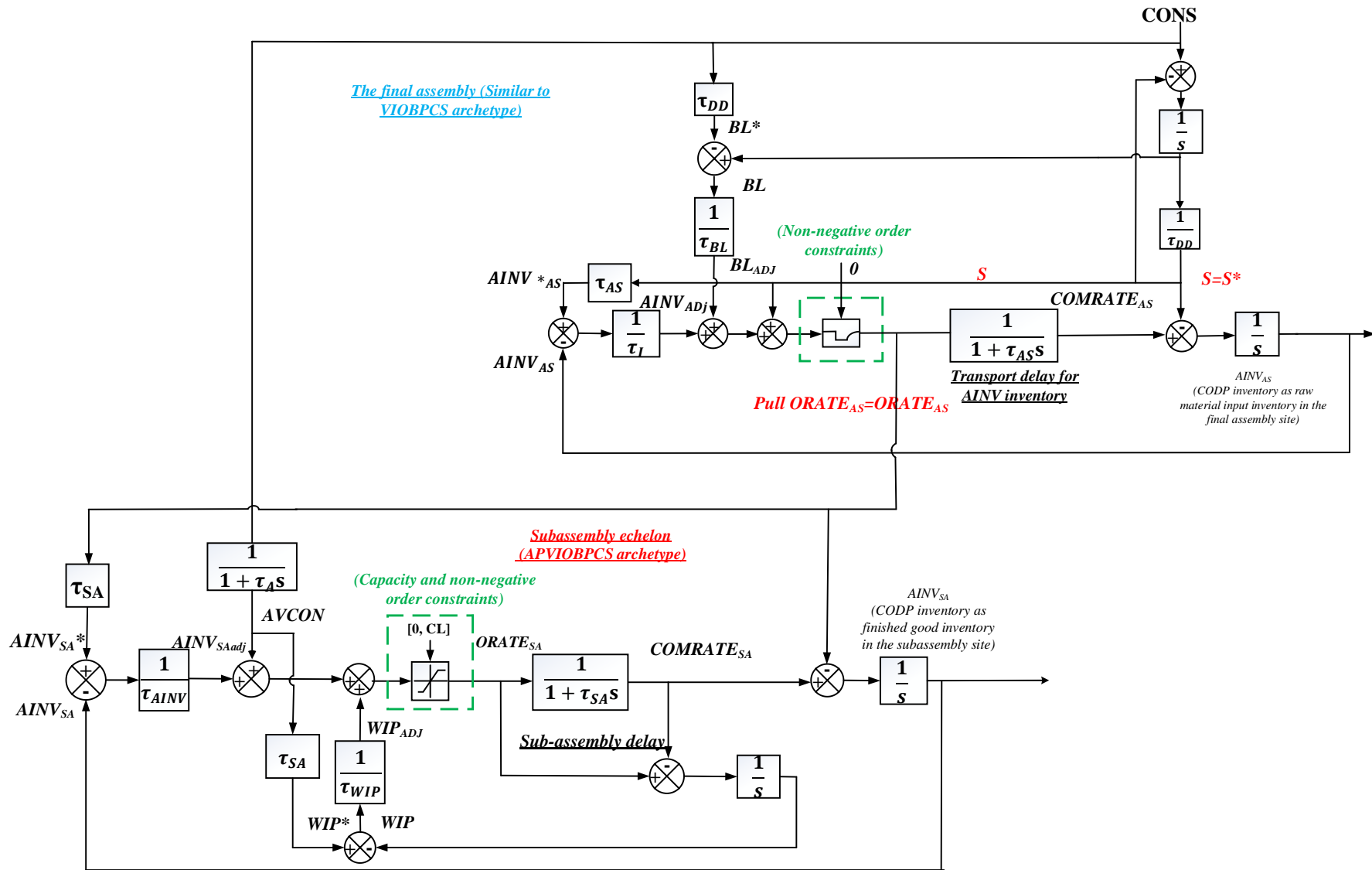
Figure 3. The block diagram of the simplified ATO system.

4.1. The impact of feedback and feedforward control on ATO system dynamics

By assuming that all nonlinearities are inactive (i.e. negative orders are permitted and no CL) it is possible to formulate the transfer functions of AINV$_{\text{AS}}$, AINV$_{\text{SA}}$ and ORATE$_{\text{SA}}$, i.e. two inventories and the supplier's capacity adjustment, in relation to CONS can be derived as follows:

$$\frac{AINV_{AS}}{CONS} = \frac{(-\tau_i\tau_{DD}^2 s^2 + \tau_{BL}s(\tau_i + \tau_{AS}) + \tau_{BL})(1 + \tau_{AS}s)}{(1 + \tau_i s + \tau_i\tau_{AS}s^2)(\tau_{BL} + \tau_{BL}\tau_{DD}s)} \quad (19)$$

$$\frac{ORATE_{SA}}{CONS} = \frac{(1 + \tau_{SA})\begin{pmatrix}(1 + \tau_A s)(1 + \tau_{AS}s)(\tau_{BL} + \tau_i\tau_{BL}s + \tau_{AS}\tau_{BL}s - \tau_i\tau_{DD}^2 s^2)(1 + \tau_{SA}s)\tau_{WIP} \\ + s\tau_{AINV}(1 + \tau_i s(1 + \tau_{AS}s))\tau_{BL}(1 + \tau_{DD}s)(\tau_{SA} + \tau_{WIP})\end{pmatrix}}{(1 + \tau_i s + \tau_i\tau_{AS}s^2)(\tau_{BL} + \tau_{BL}\tau_{DD}s)(1 + \tau_A s)} \quad (20)$$
$$(\tau_{WIP} + (\tau_{AINV}\tau_{SA} + \tau_{AINV}\tau_{WIP})s + \tau_{AINV}\tau_{SA}\tau_{WIP}s^2)$$

$$\frac{AINV_{SA}}{CONS} = \frac{\begin{pmatrix}\tau_{AINV}s(-\tau_{AS}(2 + \tau_{AS}s)\tau_{BL} + (1 + \tau_i s(1 + \tau_{AS}s))\tau_{BL}\tau_{DD} + \tau_i s(1 + \tau_{AS}s)\tau_{DD}^2)\tau_{SA} \\ + \begin{pmatrix}\tau_{AINV}s(-\tau_{AS}(2 + \tau_{AS}s)\tau_{BL} + (1 + \tau_i s(1 + \tau_{AS}s))\tau_{BL}\tau_{DD} + \tau_i s(1 + \tau_{AS}s)\tau_{DD}^2) \\ -(-1 + \tau_{AINV}s)(1 + \tau_{AS}s)((1 + (\tau_i + \tau_{AS})s)\tau_{BL} - \tau_i\tau_{DD}^2 s^2)\tau_{SA}\end{pmatrix}\tau_{WIP} \\ -\tau_A s(1 + \tau_{AS}s)((1 + (\tau_i + \tau_{AS})s)\tau_{BL} - \tau_i\tau_{DD}^2 s^2) \\ (\tau_{AINV}\tau_{SA} + (\tau_{AINV} + (-1 + \tau_{AINV}s)\tau_{SA})\tau_{WIP})\end{pmatrix}}{(1 + \tau_i s + \tau_i\tau_{AS}s^2)(\tau_{BL} + \tau_{BL}\tau_{DD}s)(1 + \tau_A s)} \quad (21)$$
$$(\tau_{WIP} + (\tau_{AINV}\tau_{SA} + \tau_{AINV}\tau_{WIP})s + \tau_{AINV}\tau_{SA}\tau_{WIP}s^2)$$

We exploit the Initial Value Theorem (IVT) and Final Value Theorem (FVT) to mathematically crosscheck the correctness of the transfer function, guide the appropriate initial condition required by a simulation and to understand the final steady state value of the dynamic response so as to help verify any simulation. Hence, the initial and final values of AINV$_{\text{AS}}$, AINV$_{\text{SA}}$, and ORATE$_{\text{SA}}$ in responding to a unit step input are obtained.

$$\lim_{s\to\infty} s\frac{AINV_{AS}}{CONS} = 0 \qquad \lim_{s\to 0} s\frac{AINV_{AS}}{D} = \tau_{AS}$$

$$\lim_{s\to\infty} s\frac{AINV_{SA}}{CONS} = 0 \qquad \lim_{s\to 0} s\frac{AINV_{SA}}{D} = \tau_{SA}$$

$$\lim_{s\to\infty} s\frac{ORATE_{SA}}{CONS} = 0 \qquad \lim_{s\to 0} s\frac{ORATE_{SA}}{CONS} = 1 \qquad (22)$$

As expected, the initial values of AINV$_{\text{AS}}$, AINV$_{\text{SA}}$ and ORATE$_{\text{SA}}$ are zero, similar to the results obtained by John, Naim and Towill (1994). The final value of the ORATE$_{\text{SA}}$ for the upstream sub-assembly system is, as expected, equal to demand, i.e. 1. The final value of the AINV$_{\text{SA}}$ and AINV$_{\text{AS}}$ are determined by the coefficient τ_{SA} and τ_{AS}, i.e. the steady state of two inventory in responding a step demand equal to desired inventory level. Based on Equation (19) to (21), the final assembly system is characterised as a third-order polynomial, while a sixth-order polynomial describes the sub-assembler's manufacturing system. Also, there is a third-order polynomial, $(1 + \tau_i s + \tau_i\tau_{AS}s^2)(\tau_{BL} + \tau_{BL}\tau_{DD}s)$, in both characteristic equations (CEs), which confirms that the dynamic behaviour of the final assembly system is not influenced by the sub-assembler manufacturing system, while the dynamic performance of the supplier manufacturing system can be partially manipulated by the final assembly control policies under the ATO system. We now assess the CEs of Equations (19) to (21) by obtaining the roots as follows:

$$R_{1\&2} = -\frac{1}{2\tau_{AS}} \pm \frac{\sqrt{\tau_i{}^2 - 4\tau_i\tau_{AS}}}{2\tau_i\tau_{AS}}, \qquad R_3 = -\frac{1}{\tau_A}, R_4 = -\frac{1}{\tau_{DD}}$$

$$R_{5\&6} = -\frac{1}{2}\left(\frac{1}{\tau_{SA}} + \frac{1}{\tau_{WIP}}\right) \pm \frac{\sqrt{-4\tau_{AINV}\tau_{SA}\tau_{WIP}^2 + (\tau_{AINV}\tau_{SA} + \tau_{AINV}\tau_{WIP})^2}}{2\tau_{AINV}\tau_{SA}\tau_{WIP}} \quad (23)$$

Inspecting Equation (23):

1. Given that the physical delays, τ_{SA} and τ_{AS}, are positive, the ATO state is permitted to be stable for any positive control policies, i.e. possible value of $\tau_A, \tau_{AINV}, \tau_{WIP}$ and τ_I. However, the system's response will be continuously oscillatory if $\tau_{SA} = -\tau_{WIP}$, that is, the $R_{5\&6}$ become purely imaginary with no real part.

2. Three feedback inventory loops, AINV$_{AS}$, AINV$_{SA}$ and WIP adjustment, may characterize oscillations of the ATO state if the square root part of $R_{1\&2}$ and $R_{5\&6}$ become negative, i.e. $\tau_I{}^2 - 4\tau_I\tau_{AS} < 0$ and $-4\tau_{AINV}\tau_{WIP}^2\tau_{SA} + (\tau_{AINV}\tau_{WIP} + \tau_{AINV}\tau_{SA})^2 < 0$. The corresponding CODP inventory-based control policies, τ_I, τ_{AINV} and τ_{WIP}, should be carefully adjusted to avoid the possible oscillatory system response.

3. Given the sub-assembler manufacturing delay, τ_{SA}, and associated inventory adjustment time (τ_{WIP}) are longer than downstream transport acquisition delay τ_{AS}, the real part of $R_{5\&6}$, $-\frac{1}{2}\left(\frac{1}{\tau_{SA}} + \frac{1}{\tau_{WIP}}\right)$, is smaller than the real part of $R_{1\&2}$, i.e. $-\frac{1}{2\tau_{AS}}$. In other word, $R_{5\&6}$ are located in a closer position to the origin s plane comparing the location of $R_{1\&2}$. As a result, the upstream inventory feedback loops, and forecasting loop may dominate the dynamic behaviour of the ATO state. Particularly, inventory loop-based control policies, τ_{AINV}, plays a key role in influencing the whole state's oscillatory behaviour.

To further understand the oscillation and system recovery properties, we derive the ω_n and ζ of two second order polynomials, $(1 + \tau_i s + \tau_i\tau_{AS}s^2)$ and $\left(1 + \frac{(\tau_{AINV}\tau_{SA} + \tau_{AINV}\tau_{WIP})}{\tau_{WIP}}s + \tau_{AINV}\tau_{SA}s^2\right)$:

$$\omega_{n1} = \sqrt{\frac{1}{\tau_{AS}\tau_I}} \qquad \zeta_1 = \frac{1}{2}\sqrt{\frac{\tau_I}{\tau_{AS}}} \qquad (24)$$

$$\omega_{n2} = \sqrt{\frac{1}{\tau_{AINV}\tau_{SA}}}, \quad \zeta_2 = \frac{(\tau_{SA} + \tau_{WIP})}{2\tau_{WIP}}\sqrt{\frac{\tau_{AINV}}{\tau_{SA}}}$$

For AINV$_{AS}$, both ω_{n1} and ζ_1 are determined by τ_I under physically fixed τ_{AS}, and τ_I has the reverse impact on nature frequency and damping ratio. The final assembly system's response and inventory recovery speed will be slower as the increase of τ_I, due to the decrease of ω_{n1}. However, the increase of τ_I will give the larger value of damping ratio and lead to the corresponding more 'damped' system with less oscillations. Also, ω_{n1} and ζ_1 could lead to such impact on the dynamic behaviour of AINV$_{SA}$ and ORATE$_{SA}$ at the subassembly site. Furthermore, τ_{AINV} and τ_{SA} negatively determine the value of natural frequency for the upstream supplier AINV$_{SA}$ feedback loop. The increase of their value will lead to slow system recovery speed to reach the steady state condition due to the decrease of value of ω_{n2}. Moreover, τ_{AINV} and τ_{WIP} have the reverse impact for ζ_2.

4.2. The impact of nonlinearities on ATO dynamic performance

When capacity and non-negative order constraints are active in the ATO system, the dynamic behaviour becomes more complex. We now explore the impact of two nonlinearities separately in responding sinusoidal demand by using describing function methods.

4.2.1. Linearization of capacity and non-negative order constraints at the subassembly echelon

In the linear system, upstream sub-assembly production capacity is assumed as unlimited and the order is permitted to be negative. This means that the sub-assembler can freely return the raw materials to their suppliers and any order rate received can be allocated for immediate production. These are unrealistic assumptions due to the expensive production line system, e.g. see Lin, Spiegler, and Naim (2017), and the forbidden return policy usually agreed between material suppliers and the sub-assembler manufacturers. So, both constraints should be considered when analysing the dynamics of the ATO system. We now linearize such nonlinearities before analysing their impact on the dynamic behaviour of the ATO system. Specifically, in an open-loop form of such nonlinearity, an input $DORATE_{SA}(t)$

$$DORATE_{SA}(t) = A \cdot \cos(wt) + B \qquad (25)$$

where A is the amplitude, B is the mean and w is the angular frequency ($w = \frac{2\pi}{t}$), will produce an output $ORATE_{SA}(t)$ with the same frequency but different mean and amplitude. Figure 4 reports the main characteristics of this nonlinearity. The output $ORATE_{SA}$ does not rely on the past value of the input $DORATE_{SA}$, but it varies depending on input's actual status based on the upper and lower limit.

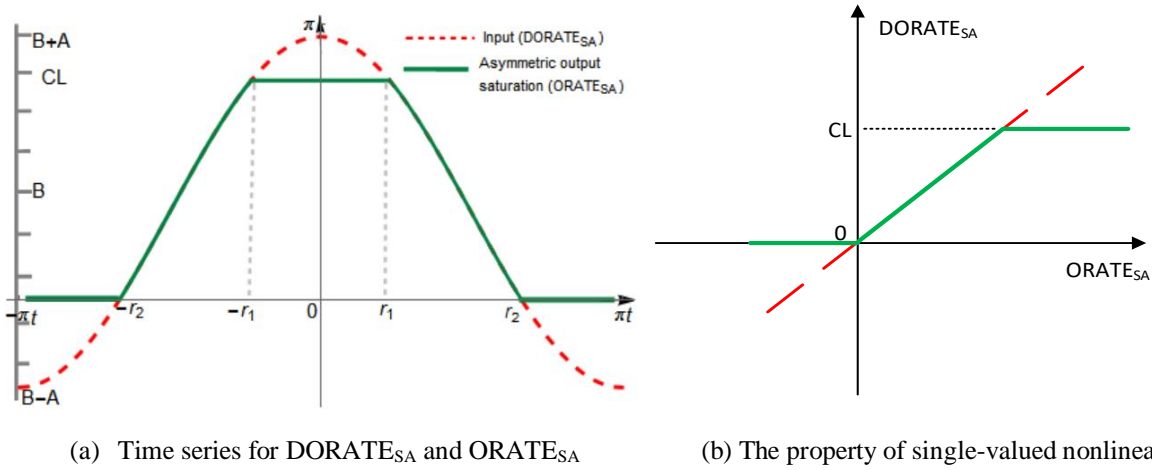(a) Time series for $DORATE_{SA}$ and $ORATE_{SA}$ (b) The property of single-valued nonlinearity

Figure 4. Asymmetric output saturation in relation to sinusoidal input $DORATE_{SA}$

Note that a fundamental requirement for the system is that CL must be at least larger than averaged demand due to the accumulative errors driven by the feedback integrator (1/s). In other word, the $DORATE_{SA}$ will increase exponentially if manufacturing capacity is less than the averaged demand rate and the system will become unstable. Under the assumption, the output function, $ORATE_{SA}$, can be represented by three linear piecewise equations as follow:

$$ORATE_{SA}(t) = \begin{cases} 0 & if \quad DORATE_{SA} < 0 \\ DORATE_{SA} & if \quad 0 < DORATE_{SA} < CL \\ CL & if \quad DORATE_{SA} > CL \end{cases} \quad (26)$$

To analyse the discontinuous nonlinearities in the ATO system, the describing function method can be applied (Spiegler et al. 2016; Spiegler and Naim 2017). This method is a quasi-linear representation for a nonlinear element subjected to specific input signal forms such as Bias, Sinusoid and Gaussian process and system's low-pass filter property (Vander and Wallace 1968). The principle advantage of using the describing function method is it enables the aid of analytically designing nonlinear systems. The basic idea is to replace the nonlinear component by a type of transfer function, or a gain derived from the effect of input (e.g. sinusoidal input). For an asymmetric saturation, as illustrated in Figure 5, $DORATE_{SA}$ is smaller than zero or greater than CL, at least two terms need to

be identified: one term describes the change in amplitude ($N_{A(CA)}$) in relation to the input amplitude and the other defines the change in mean ($N_{B(CA)}$) in relation to the input mean. Furthermore, output phase angle (φ) in relation to the input angle may also be changed.

Thereby given the input, i.e. Equation (25), the output ORATE$_{SA}$ can be approximated to:

$$ORATE_{SA}(t) = N_{A(CA)} \cdot A \cdot \text{Cos}(wt + \phi) + N_{B(CA)} \cdot B \qquad (26)$$

The Fourier series expansion can be applied to obtain $N_{A(CA)}$, $N_{B(CA)}$ and φ:

$$ORATE_{SA}(t) \approx b_0 + a_1 \cos(wt) + b_1 \sin(wt) + a_2 \cos(2wt) + b_2 \sin(2wt) + \cdots$$

$$\approx b_0 + \sum_{n=1}^{\infty}(a_n \cos(nwt) + b_n \sin(nwt)) \qquad (27)$$

Where the Fourier coefficient can be determined by:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} ORATE_{SA}(t) \cos(nwt)\, d_{wt} \qquad (28)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} ORATE_{SA}(t) \sin(nwt)\, d_{wt} \qquad (29)$$

$$b_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} ORATE_{SA}(t) d_{wt} \qquad (30)$$

and ORATE$_{SA}$ is the piecewise linear function (Figure 6a):

$$ORATE_{SA}(t) = \begin{cases} 0 & -\pi < wt < -r_2 \\ A \cdot \cos(wt) + B & -r_2 < wt < -r_1 \\ CL & -r_1 < wt < r_1 \\ A \cdot \cos(wt) + B & r_1 < wt < r_2 \\ 0 & r_2 < wt < \pi \end{cases} \quad (0 < r_1 < r_2 \leq \pi) \ (31)$$

To approximate periodic series, only the first, or fundamental harmonic is needed and thereby we need to find the first order coefficient of Fourier series expansion demonstrated in Equation (28)-(30). Note that such approximation is often useful for the symmetric system including only odd functions and thus high order harmonic can be effectively attenuated by the linear dynamic of the system, i.e. the property of low-pass filter. For the asymmetric system, as the focus of this study, the aid of simulation is recommended (Atherton, 1975) to verify the analytical results. We now obtain the first harmonic of the piecewise linear Equation (31):

$$ORATE_{SA}(t) = b_0 + a_1 \cdot \cos(wt) + b_1 \cdot \sin(wt) = b_0 + \sqrt{a_1^2 + b_1^2} \cos(wt + \phi) \qquad (32)$$

Where $\phi = \arctan(\frac{b_1}{a_1})$

By comparing Equation (26) and (32), we have the gain of the describing function as follows:

$$N_{A(CA)} = \frac{\sqrt{a_1^2 + b_1^2}}{A} \text{ and } N_{B(CA)} = \frac{b_0}{B} \qquad (33)$$

Due to the property of such single-valued nonlinearity, there is no output phase shift in relation to input, that is, $b_1 = 0$ and $\phi = 0$. By calculating the Fourier coefficient a_1 and b_0 (Mathematica$^{\text{TM}}$), the describing function gains are obtained as follow:

$$N_{A(CA)} = \frac{A \cdot \text{Cos}(r_1) \cdot \text{Sin}(r_1) + \big(2B + A \cdot \text{Cos}(r_2)\big) \cdot \text{Sin}(r_2) - A \cdot r_1 + A \cdot r_2}{A \cdot \pi} \qquad (34)$$

$$N_{B(CA)} = \frac{A \cdot \big(\text{Sin}(r_2) - \text{Sin}(r_1)\big) + A \cdot r_1 \cdot \text{Cos}(r_1) + B \cdot r_2}{B \cdot \pi} \qquad (35)$$

Where $r_1 = \text{Cos}^{-1}(\frac{CL-B}{A})$ and $r_2 = \text{Cos}^{-1}(\frac{-B}{A})$ and Equation (34) and (35) can be further simplified as:

$$N_{A(CA)} = \frac{(CL - B) \cdot \sqrt{1 - \frac{(CL-B)^2}{A^2}} + B \cdot \sqrt{1 - \frac{B^2}{A^2}} - A \cdot \cos^{-1}\left(\frac{CL-B}{A}\right) + A \cdot \cos^{-1}\left(-\frac{B}{A}\right)}{A \cdot \pi} \qquad (36)$$

$$N_{B(CA)} = \frac{A \cdot \left(\sqrt{1 - \frac{B^2}{A^2}} - \sqrt{1 - \frac{(CL-B)^2}{A^2}}\right) + A \cdot (CL - B) \cos^{-1}\left(\frac{CL-B}{A}\right) + B \cdot \cos^{-1}\left(-\frac{B}{A}\right)}{B \cdot \pi} \qquad (37)$$

Figure 5 gives the density plot for the value of N_A as the increase of A from CL to 8CL, and the increase of B from 0.1 CL to CL. Depending on different value of A and B, $N_{A(CA)}$ ranges between 0 and 1. Specifically, for a fixed B, $N_{A(CA)}$ appears to be decreasing in A and this implies that only a fraction of DORATE$_{SA}$ will be manufactured due to the capacity and non-negative order constraints. However, the influence of B on amplitude gain depends on the relationship between A and CL. If A is larger than CL, B gives little influence on amplitude gain due to the dominant influence of A on the $N_{A(CA)}$. if A is located within 0 and CL, $N_{A(CA)}$ depends on both A and B. $N_{A(CA)}$ may equal to 1 (the system will behave as linear) if DORATE$_{SA}$ do not exceed the constraint range, i.e. [0, CL], while only a fraction of DORATE$_{SA}$ will be manufactured if $N_{A(CA)} <1$.
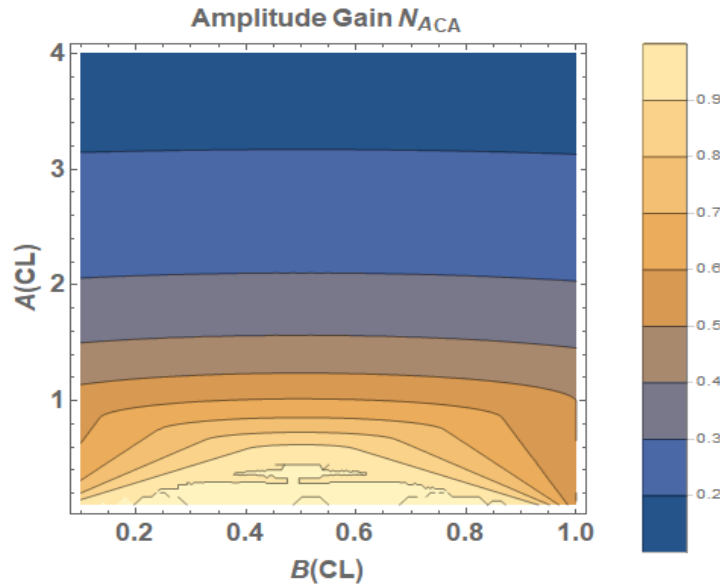

Figure 5. The density plot of $N_{A(CA)}$ based on A and B in relation to CL

Overall amplitude of DORATE$_{SA}$ play a major in influencing the value of $N_{A(CA)}$. This means the higher bullwhip, the less proportion of DORATE$_{SA}$ will be manufactured. To explore how the relationship of A, B and CL influences the output mean gain, $N_{B(CA)}$, we differentiate Equation (37) with respect to A and yield the following expression:

$$\frac{d_{N_{B(CA)}}}{d_A} = \frac{\sqrt{1 - \frac{B^2}{A^2}} - \sqrt{1 - \frac{(CL-B)^2}{A^2}}}{B\pi} \quad (38)$$

Equation (38) shows that the zero gradient can be achieved if $B = \frac{1}{2}CL$ and we obtain the corresponding value of $N_{B(CA)}$

$$N_{B(CA)}\big|_{B=\frac{1}{2}CL} = \frac{Cos^{-1}(-\frac{CL}{2A}) + Cos^{-1}(\frac{CL}{2A})}{\pi} = 1 \quad (39)$$

So output mean gain, $N_{B(CA)}$, equals 1 irrelevant of input amplitude A if averaged input mean is half of CL, due to the fact that system has a symmetric saturation in this case, i.e. equal influence of upper capacity and nonnegative order constraints. Also, the increase of A leads to the increase of $N_{B(CA)}$ if $B < \frac{1}{2}CL$, while $N_{B(CA)}$ is monotonically decreasing in A if $B > \frac{1}{2}CL$. This means that if averaged input demand is less than half of CL, the non-negative order constraint gives more impact on $N_{B(CA)}$ than the corresponding capacity constraint and thereby $N_{B(CA)}$ decreases by the increase of A due to order rate reaching zero more often than hitting CL. However, if averaged input demand is larger than the half of CL, $N_{B(CA)}$ is monotonically decreasing in A because the impact of capacity constraint dominates the output mean gain comparing the corresponding non-negative order constraint. Such findings are consistent with Spiegler et al. (2016a; 2016b)'s separate investigation of the effect of capacity constraint and non-negative order constraints on output mean gain. Figure 6 demonstrates two examples how $N_{B(CA)}$ varies as the increase of A related to AL when B=0.2CL and B =0.8CL.
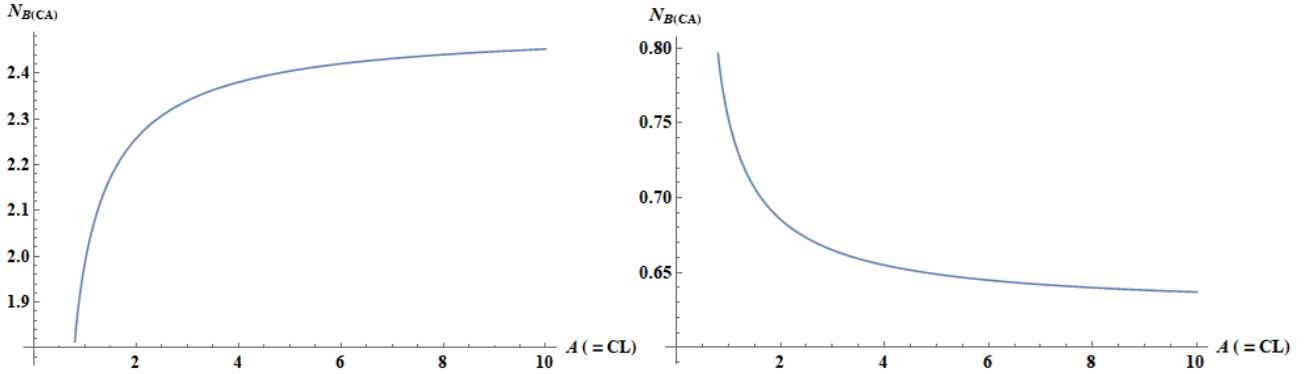

Figure 6. The change of $N_{B(CA)}$ as the increase of A in relation AL when B=0.2CL (Left) and B=0.8CL (Right).

4.2.2. Linearization of non-negative order constraints at the final assembly site

In the linear ATO system, $ORATE_{AS}$ at the final assembly echelon is permitted to take negative values. It means that excess PC components at the final assembly plant can be freely returned to the sub-assembler site. This is an unrealistic assumption due to long geographical distance and export/import policies between the final assembly and their PC parts subassemblies. As a result, the non-negative order constraint should be put into the model to prevent the free inventory return from final assembly site to the supplier site. The main characteristics of non-negative nonlinearity is reported in Figure 7 and Equation (39) shows the piece linear function of $ORATE_{AS}$:

$$ORATE_{AS}(t) = \begin{cases} 0 & if \quad Pull\ ORATE_{AS} < 0 \\ Pull\ ORATE_{SA} & if \quad Pull\ ORATE_{SA} > 0 \end{cases} \quad (39)$$

Where $Pull\ ORATE_{SA}(t) = A_1 \cdot \cos(wt) + B_1$ and $ORATE_{AS}(t)$ can be approximated by

$$ORATE_{AS}(t) \approx N_{A_{(NO)}}A_1 \cdot \cos(wt + \phi) + N_{B_{(NO)}}B_1 \qquad (40)$$
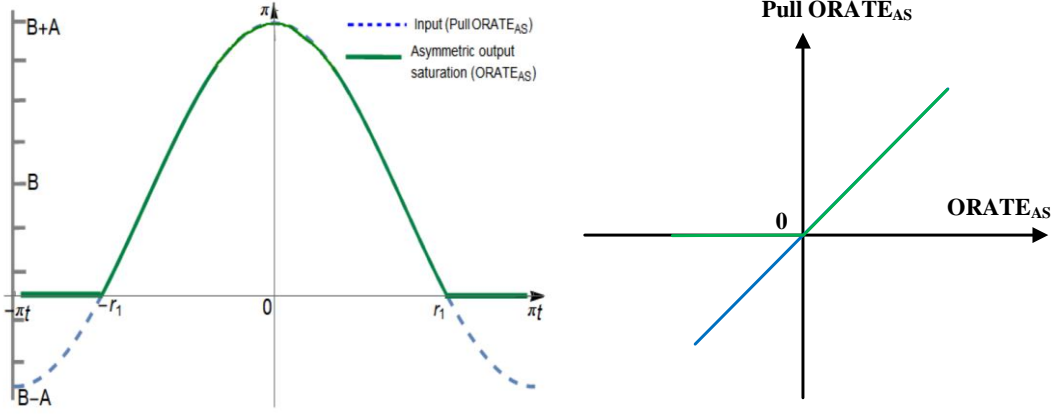

Figure 7. Main characteristics of non-negative order constraint at the downstream final assembly echelon.

Where $N_{A(NO)}$ is the change in amplitude in relation to the input amplitude and $N_{B(NO)}$ is the change in mean in relation to the input mean under non-negative order constraint policy. Similar to the linearization of capacity and non-negative constraints for the sub-assembler production, the describing function method can be applied for linearizing such nonlinearity. The corresponding describing function gain can be derived as follows:

$$N_{A_{(NO)}} = \frac{\dfrac{B \cdot \sqrt{1 - \dfrac{B^2}{A^2}}}{A} + \text{Cos}^{-1}(-\dfrac{B}{A})}{\pi} \qquad (41)$$

$$N_{B_{(NO)}} = \frac{A \cdot \sqrt{1 - \dfrac{B^2}{A^2}} + B \cdot \text{Cos}^{-1}(-\dfrac{B}{A})}{B \cdot \pi} \qquad (42)$$
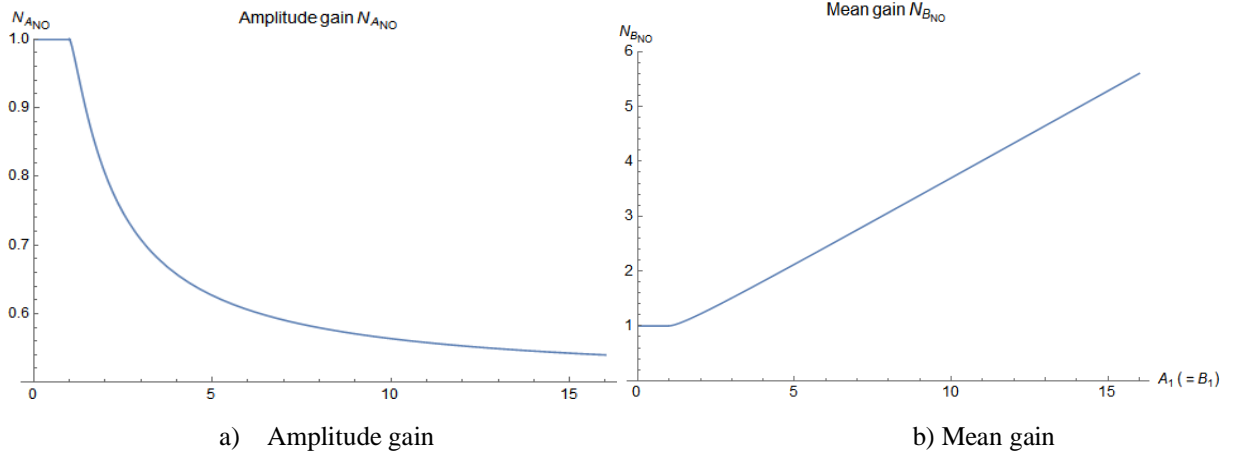

a) Amplitude gain

b) Mean gain

Figure 8. Terms of describing function for the non-negativity constraint.

Figure 8 illustrates how the coefficients of the describing function vary as A_1 increases for any $B_1 > 0$. For values of A_1 lower than B_1, the system behaves as linear and output o(t) will be equal to the input do(t) corresponding to $N_{A(NO)} = 1$ (Figure 8a). However, when A_1 increases then only a fraction of this rate will actually be ordered corresponding to $N_{A(NO)} < 1$. By inspecting Equation (41), we find that as Ado approaches infinity, N_{A1} approaches 0.5. So, the amplitude gain of the describing function can only vary from 0.5 to 1. On the other hand, the value of N_{B1} rises as A_1 increases because the limit value of the order rate is at its minimum (Figure 8b).

4.2.3. Predicting the system's dynamic behaviour

Although two nonlinearities in the ATO system have different features, they both decrease their corresponding output amplitude gains ($N_{A(CA)}$ and $N_{A(NO)}$) as the increase of input amplitude. Root locus techniques (Spiegler et al. 2016a; Spiegler and Naim 2017) can be used to predict how these nonlinearities affect the system responses. By replacing the ⊞ and ⊔ with the corresponding amplitude gains, $N_{A(CA)}$ and $N_{A(NO)}$, respectively, and using block diagram algebra, we obtain the new CEs and compare with CEs in linear ATO state based on Equation (19)-(21).

$$CE_{OEM} \colon (\tau_{BL} + \tau_{DD}\tau_{BL}s)\big(N_{A(NO)} + \tau_i s + \tau_i \tau_{AS}s^2\big) \quad (43)$$

$$CE_{supplier} \colon \frac{(1 + \tau_A s)(\tau_{BL} + \tau_{DD}\tau_{BL}s)\big(N_{A(NO)} + \tau_i s + \tau_i \tau_{AS}s^2\big)}{\big(N_{A(CA)}\tau_{WIP} + (\tau_{AINV}\tau_{WIP} + \tau_{AINV}\tau_{SA}N_{A(CA)})s + \tau_{AINV}\tau_{WIP}\tau_{SA}s^2\big)} \quad (44)$$

The new ω and ζ can be derived as follow:

$$\omega_{n1} = \sqrt{\frac{N_{A(NO)}}{\tau_{AS}\tau_I}} \qquad \zeta_1 = \frac{1}{2}\sqrt{\frac{\tau_I}{\tau_{AS}N_{A(NO)}}} \qquad (45)$$

$$\omega_{n2} = \sqrt{\frac{N_{A(CA)}}{\tau_{AINV}\tau_{SA}}}, \qquad \zeta_2 = \frac{(N_{A(CA)}\tau_{SA} + \tau_{WIP})}{2\tau_{WIP}}\sqrt{\frac{\tau_{AINV}}{\tau_{SA}N_{A(CA)}}}$$

Regarding the downstream final assembly system, the incorporation of $N_{A(NO)}$ (ranging between 0.5 - 1) will result in a reverse impact on ω_{n1} and ζ_1, that is, the decrease of ω_{n1} but increase ζ_1 as the decrease of $N_{A(NO)}$. This means the incorporation of non-negative order constraint at the final assembly site leads to a 'more damped' system with less oscillations at the expense of slow system recovery speed. Also, as indicated by the Section 4.2.2, the $N_{B(NO)}$ will increase as the increase of input demand amplitude. The dynamic response of upstream sub-assembler variables, however, are influenced by both nonlinearities. The decrease of output amplitude gain, $N_{A(CA)}$, resulted from the capacity and non-negative order constraints, leads to the decrease of ω_{n2} and ζ_2. This gives both slower and more oscillatory dynamic response of the ATO system. Note that depending on the relationship between mean of input demand and the half of capacity constraint (i.e. the dominant zone), the increase of demand amplitude may lead to the increase or decrease of $N_{B(NO)}$.

5. Numerical study

In this section numerical simulation is conducted to test whether the analytical results derived from the linearized model (Section 4) hold under the nonlinear case. Specifically, the nonlinear hybrid ATO model (Figure 3), including capacity and non-negative order constraints, is used as the base simulation model and the numerical study is conducted via Simulink^{TM} (Matlab). We assume that the lead times ratio between τ_{SA} and τ_{AS} is 1:2 (i.e. 4 and 8 for transportation and component manufacturing delay). This assumption represents the long-term upstream subassembly manufacturing time, and relatively short time for component acquisition delay between supplier and the final assembly echelon.(Kumar and Craig; Katariya et al. 2014):

$$\tau_{SA} = 2\tau_{AS} = 2\tau_I = 8\tau_{DD} = 8, \tau_{WIP} = 16\ \tau_{AINV} = 8$$

5.1 Feedback and feedforward control policy

To test the impact of feedback and feedforward control loops ($\tau_I, \tau_A, \tau_{AINV}$) on bullwhip and CODP inventory variance analytically derived in Section 3.1, a unit step demand increase as the input is used due to its advantage of offering rich information for the dynamic behaviour of the system (John, Naim and Towill 1994). The recommended settings of both VIOBPCS and APVIOBPCS will be used as the initial design as illustrated above in Section 5, although we vary different control policies to understand the impact of each control policy on dynamic performance in the nonlinear environment (i.e. capacity limit is set as 2). All results are shown in Figure 9.
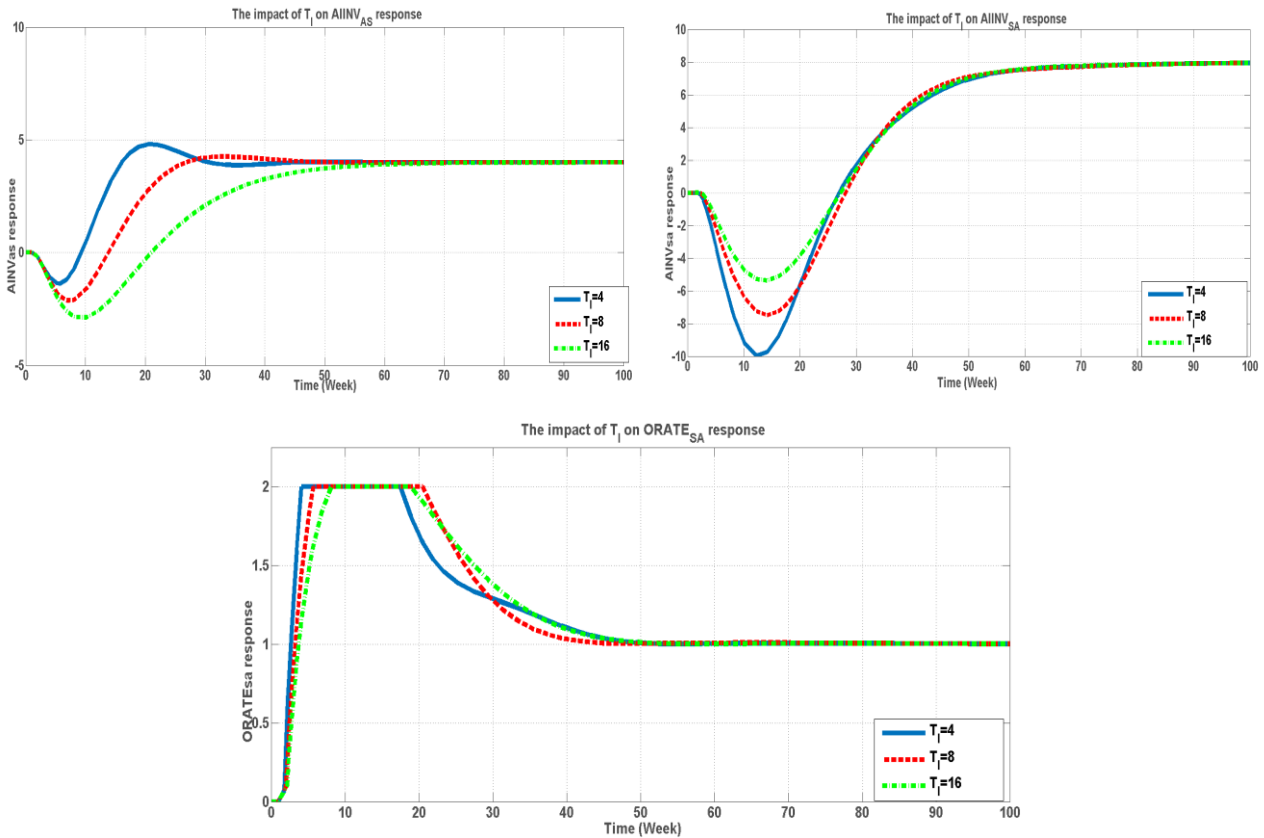

Figure 9a. The impact of τ_I on AINV$_{AS}$, AINV$_{SA}$ and ORATE$_{SA}$ dynamic response.
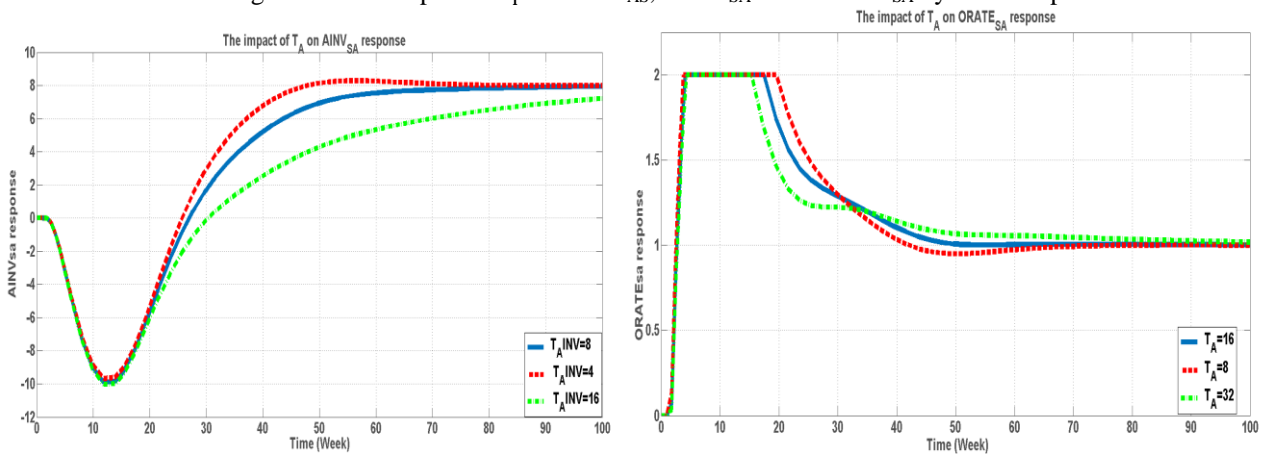

Figure 9b. The impact of τ_A on AINV$_{SA}$ and ORATE$_{SA}$ dynamic response.
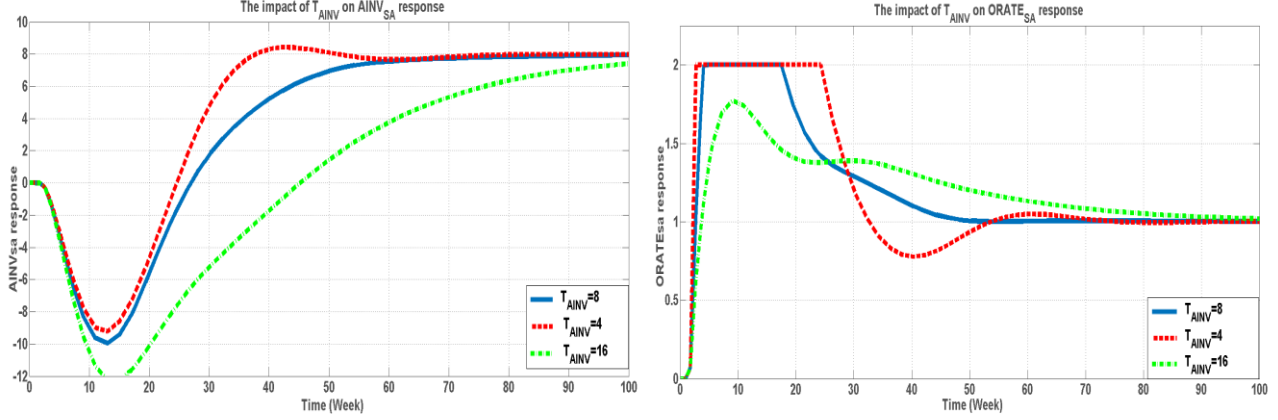
Figure 9c. The impact of τ_{AINV} on $AINV_{SA}$ and $ORATE_{SA}$ dynamic response.

In general, the simulation results support the analytical insights. The increase of τ_I lead to less oscillatory system response due to the increase of ζ_1, at the expense of slower response of the ATO system (e.g. slow recovery of $AINV_{AS}$, Figure 9a) driven by the decrease of ω_{n1}. As a result, τ_I should be carefully adjusted due to the availability of $AINV_{AS}$ that directly relate to the customer service level, i.e. whether all incoming customized orders can be immediately final-assembled and shipped out. Similar to the effect of τ_I, the increase of τ_{AINV} benefit to the more 'damped' system compromised by slow system recovery due to the decrease of ω_{n2} and increase of ζ_2. Furthermore, the simulation result supports the analytical result that τ_{AINV} significantly influences the dynamic behaviour of the ATO system, comparing the influence of τ_I and τ_A (Compare Figure 9a, 9b and 9c). Quick adjustment of τ_{AINV} leads to high bullwhip and significant oscillations, while long-term adjustment causes slow system recovery to reach the steady state condition. Thus, the upstream subassembly echelon should carefully tune their inventory policy to benefit the ATO system performance by reducing the cost of supply chain dynamics due to bullwhip and inventory variance. Note that compared to other control policies, the forecasting policy (τ_A) has less impact on the dynamics of the ATO system.

5.2. The impact of nonlinearities on ATO dynamic performance

To test whether the analytical results of nonlinearities derived from the linearized model (Section 4.2) hold under the nonlinear model, the asymmetrical capacity and non-negative constraint zone is set as $[0, 1]$. i.e. the minimum value will not be less than 0 and the CL is 1. Specifically, as analytically derived from Section 4.2.1, the sinusoidal input amplitude directly influences the describing function gain, $N_{A(CA)}$ and $N_{B(CA)}$, of the sinusoidal output response at the sub-assembly plant under capacity and non-negative order constraints. Table 4 presents the comparison between analytical and simulation results of $N_{A(CA)}$. Input amplitudes ranging between 0.3 to 4 with 0.1rad/week frequency are used to examine the output amplitude gain change. Within reasonable error range, the simulation results support the analytical insights.

$N_{A(CA)}$ simulation (analytical) results	A=0.3	A=1	A=2	A=4
B=0.2	0.833 (0.890)	0.500 (0.574)	0.250 (0.311)	0.165 (0.158)
B=0.5	1 (1)	0.500 (0.608)	0.250 (0.314)	0.165 (0.158)
B=0.8	0.833 (0.890)	0.500 (0.574)	0.250 (0.311)	0.165 (0.158)

Table 4. Comparison between simulation and analytical results.

Also, as highlighted by the analytical findings that the impact of input amplitude on mean gain ($N_{B(CA)}$) depends on the relationship between the average of input and the half of capacity limit, numerical simulation is implemented for test the analytical result shown in Table 5. The mean value of input demand is set 0.2, 0.5 (half) and 0.8 units to represent the different nonlinear dominated zones (non-negative order or capacity constraints). Input amplitudes ranging between CL to 5CL with 0.1rad/week frequency are used to examine the output mean gain change. It can be concluded that the simulation results support the analytical insights.

$N_{B(CA)}$ simulation results	A=0.3	A=1	A=2	A=3	A=4	Summary
B=0.2	1.08	1.660	1.770	1.845	1.590	$N_{B(CA)}$ is larger than 1 and is monotonically increasing in A
B=0.5	1.021	0.986	0.996	1.002	0.1004	$N_{B(CA)}$=1 within a reasonable error range
B=0.8	0.931	0.806	0.791	0.783	0.763	$N_{B(CA)}$ is monotonically decreasing in A and less than 1

Table 5. Numerical simulation result for $N_{B(CA)}$ based on different input amplitude and mean.

Furthermore, simulation is conducted to test analytical insights derived by Root locus techniques (Spiegler et al. 2016a; Spiegler and Naim 2017) regarding the prediction of the impact of nonlinearities on the system responses. Due to input frequency does not impact on the output gains of nonlinearities (the property of single-value discontinuous nonlinearities), CL=3 and sinusoidal demand pattern with mean=1, frequency = 3 rad/week and amplitudes = 5 is implemented for a better visualization. All results are shown in Figure 10. It should be noted that a mix of step increase and sinusoidal demand patterns are adopted with zero initial condition, which has the advantage of visualizing the impact of nonlinearities on ATO dynamics in responding to both types of patterns simultaneously.
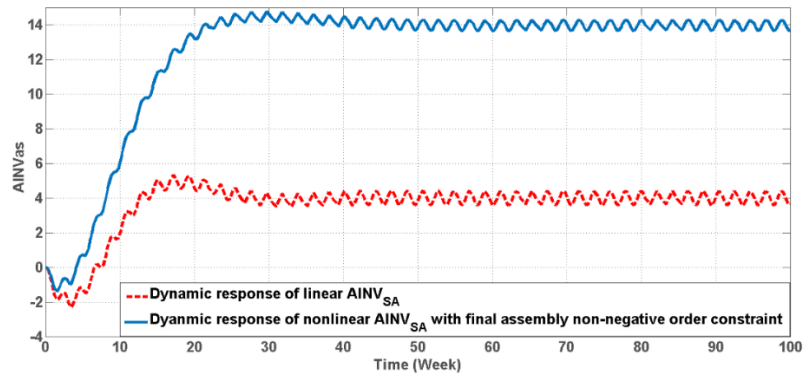

Figure 10a. Linear and nonlinear AINV$_{AS}$ response (Final assembly non-negativity constraint only).
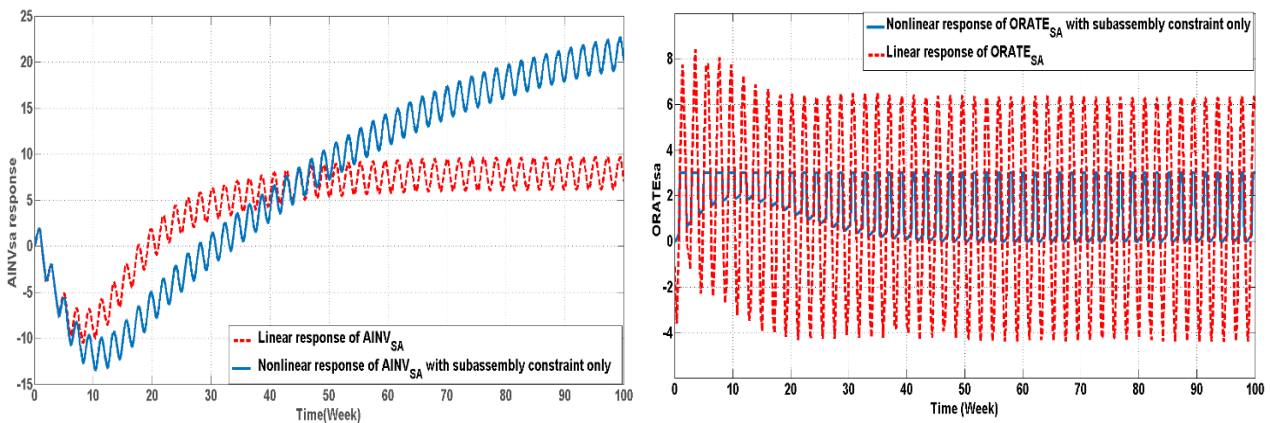
Figure 10b. Linear and nonlinear AINV$_{SA}$ response (sub-assembler's nonlinearity only)

Overall, the results support the analytical findings regarding transient behaviour. The incorporation of non-negative constraints at the final assemble echelon leads to less oscillations (with an increase in ζ_1) but slow recovery speed (due to a decrease in ω_{1n}). But it should be noted that the incorporation of such a nonlinearity increases the mean level of AINV$_{AS}$. This may improve the dynamic performance of the sub-assembler internal system by reducing AINV$_{SA}$ but contradicts the final assembly member's general objective, i.e. minimize inventory to reduce the risk of technological redundancy with ever shorter product life cycles of products entering the market. The sub-assembler's constraints for both capacity and non-negative order, verified by simulation (Figure 10b), reduce the bullwhip (ORATE$_{SA}$), at the expense of slowing AINV$_{SA}$ recovery speed as well as increasing its mean level, driven by the decrease in $N_{A(NO)}$ and an increase in $N_{B(NO)}$. This finding is well-recognized in the literature. e.g. see Cannella, Ciancimino, and Marquez (2008); Nepal, Murat, and Chinnam (2012); Ponte et al. (2017).

5.3. Sensitivity analysis

In the dynamic analysis above, one of the fundamental assumptions is that there is no loss of product quality or assembly line efficiency, which is not realistic in a real-world ATO system. For example, in the semiconductor industry, the unit yield (the percentage of good chips for each assembly die), assembly line yield rate (the percentage of good wafers per total) and the line yield (the percentage of good die per fabricated wafers) are important quality and efficiency related parameters (Gonçalves, Hines, and Sterman 2015; Mönch, Fowler, and Mason 2013) in influencing the dynamic behaviour of the system. By undertaking a sensitivity analysis, it is possible to check on the dynamic performance due to possible changes in quality and efficiency, that is, the physical parameters that the control policy designer cannot directly influence or change.

Specifically, we incorporate two general parameters related to the quality and efficiency, Y_F (final assembly line efficiency, the percentage of shippable goods for each final assembly line) and Y_S (subassembly quality yield rate), into the original nonlinear ATO model (Figure 3), as presented in Figure 11. The perfect quality and efficiency values ($Y_F = Y_S = 1$) are used as the baseline setting, and we vary the two parameters between 0.6 and 1. A step demand input is introduced, and all nonlinearities are temporarily removed to visualize the key dynamic propertiessuch as peak order overshoot (equivalent to bullwhip) and inventory variance. All results are presented in Figure 12.

The simulation results show that the quality yield rate and line efficiency have a negative impact on the dynamics of the ATO system. Decreases in Y_F and Y_S significantly increase the bullwhip of ORATE$_{SA}$, while comparing the significant impact of Y_F on inventory variance, Y_S has much less influence on AINV$_{SA}$, due to the safety stock setting of AINV$_{SA}$, i.e. AINV$^*_{SA}$ is only driven by Y_F. To be more specific, the decrease of final assembly line efficiency, Y_F, indicates the requirement of higher level of finished AINV$_{SA}$ to satisfy the end customized orders, which result the increase of the safety stock needed in the subassembly site, AINV$^*_{SA}$. This implies the importance of maintaining high final assembly line efficiency to not only ensure the customer service level, but also improve the dynamic performance of the whole ATO system to reduce supply chain dynamics related cost. Furthermore, as expected, the decrease of quality yield and efficiency may lead to the increase in final value of AINV$_{SA}$ and ORATE$_{SA}$ to ensure the same customer service level (the final value of AINV$_{SA}$ depends solely on Y_F). This leads to excess inventory and hence corresponding increases in inventory holding costs. For example, the final value of ORATE$_{SA}$ in responding to a unit step demand increase under $Y_S = 0.6$ approximately equal to 1.67, i.e. $FV_{ORATE_{SA}} = \frac{1}{0.6} \approx 1.67$.
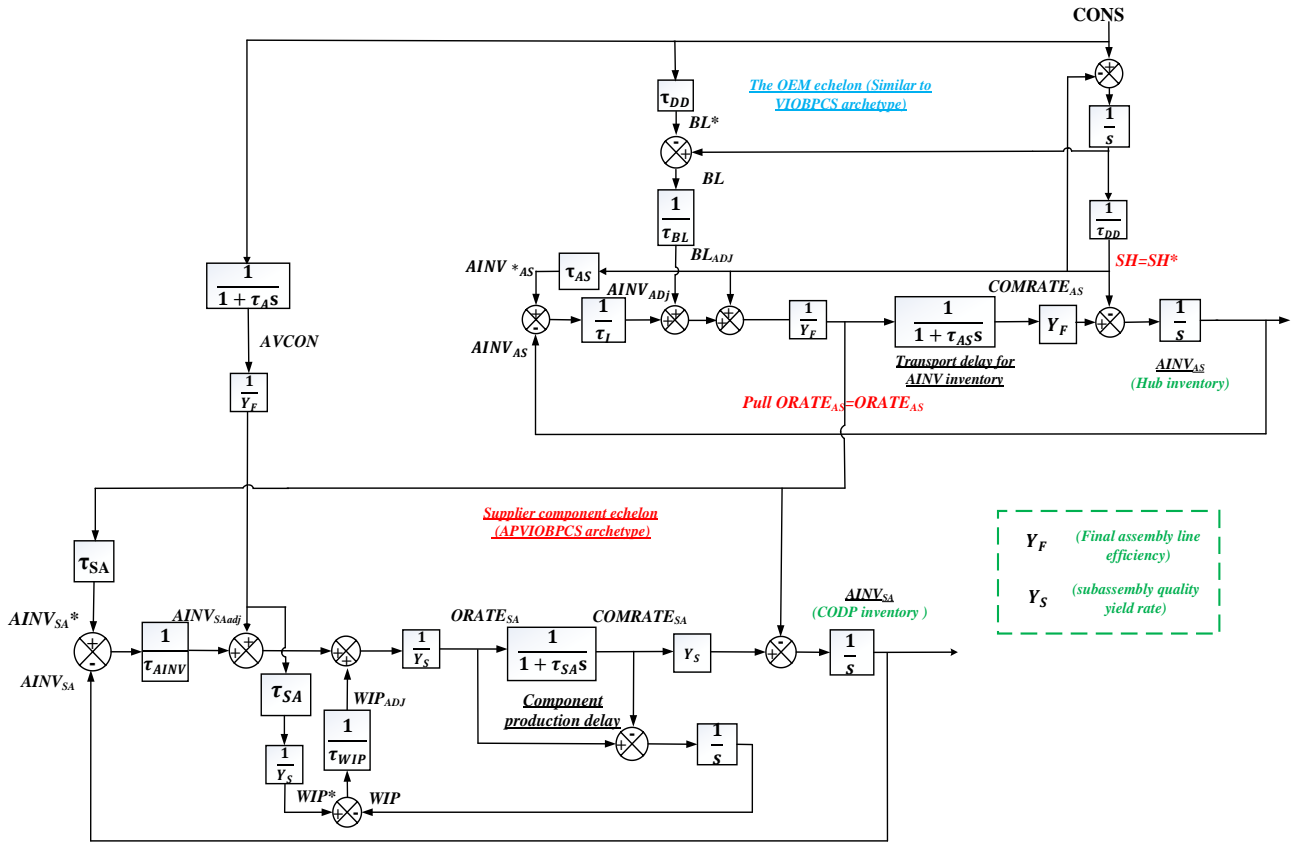
Figure 11. The incorporation of quality and efficiency parameters in the hybrid ATO state.
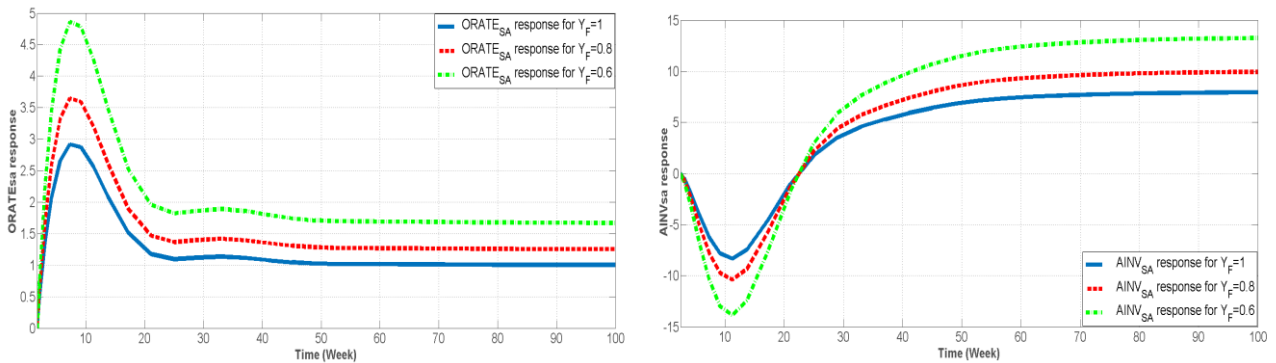

Figure 12a. The impact of final assembly line efficiency (Y_L) parameters on the dynamics of the ATO system.
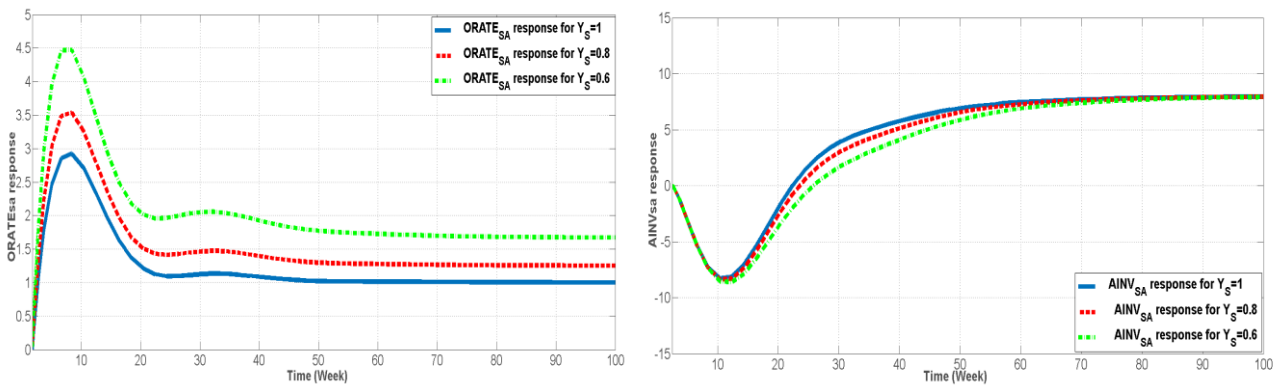

Figure 12b. The impact of subassembly quality yield rate (Y_S) parameters on the dynamics of the ATO system.

6. Discussion and conclusion.

In this paper we investigate the dynamic performance of the ATO system by using combined control engineering and system dynamic methods. Using a PC supply chain empirically reported by Berry, Towill and Wadsley (1994), Naylor, Naim and Berry (1999), Kapuscinski et al. (2004), Huang and Li (2010) and Katariya et al. (2014), as an example, a system dynamics model of ATO is developed and the IOBPCS family are used as the benchmark models. We explore the impact of major feedback and feedforward control loops, as well as nonlinearities present in the ATO system. The system dynamic simulation is adopted for testing and providing some further insights of the ATO dynamic property. All main findings and corresponding managerial implications are summarized in Table 6.

We contribute to the analysis of the ATO system structure from the system dynamics perspective. We reveal the impact of nonlinearities on the dynamic performance of an ATO system. The describing function is used to linearize nonlinearities present in the ATO system so that analytical insights can be obtained. We highlight the fact that, depending on the mean and amplitude of the demand, the non-negative order and capacity constraints in the ATO system may occur and their significant impact on system dynamics performance should be carefully considered. Failing to monitor nonlinearities, as traditionally assumed by the linear studies (Lin et al. 2017), may result in unwanted dynamic performance and thus dramatically increase the operational cost. For instance, final assemblers may underestimate the mean level of inventory and overestimate the inventory recovery speed if the non-negative order constraint is ignored. Sub-assemblers, analogously, may suffer increased inventory cost (i.e. the consequence of increasing/decreasing in inventory level and recovery speed) if capacity and non-negative order constraints are not considered at their production site. These analytical results, verified by simulation, offer robust insights for practitioners to monitor and control nonlinearities present in their ATO system to improve system dynamics behavior.

Furthermore, downstream final assembly inventory control policy impacts on the dynamic performance of both final assemblers and sub-assemblers, e.g. the quick recovery of inventory at the final assembly may benefits the customer service level for final assemblers but increase supply chain dynamics associated cost for the sub-assemblers due to exceed inventory variance and bullwhip. This highlights the importance of trade-off design and control in managing supply chain dynamics. Moreover, we found the forecasting policy may no longer play an important role in influencing dynamic behavior of the system, contradicting previous literature that assume linearity, such as a linear order-up-to system (Dejonckheere et al. 2002; Li, Disney and Gaalman, 2014). Note that quality yield and final assembly line efficiency also plays a substantial role in influencing the dynamic behavior of the ATO system.

This study, however, is limited to the analysis of a hybrid ATO system and ignored the possible switch between different states due to insufficient CODP inventory. The investigation of corresponding delivery lead times dynamics resulted from the switch can be an extension of this study. Furthermore, due to the importance of maintaining ATO structures to ensure customer service level, further control policy trade-off design between capacity and CODP inventory should be considered to minimize the corresponding operational cost within the context of the PC sector.

ATO system structure			Analytical and simulation results	Corresponding managerial implications
Control loops	Feedback loops		1. The ATO is stable for any positive value of $\tau_A, \tau_{AINV}, \tau_{WIP}$ and τ_I 2. ω_{n1} and ζ_1 are inversely proportional to τ_I 3. ω_{n2} and ζ_2 are inversely proportional to τ_{AINV} 4. τ_{AINV} plays a dominant role in influencing the whole state's oscillatory behaviour	1. There is a need to consider the inventory policy of the downstream echelon to avoid excessive bullwhip and inventory variance and associated costs. Managers need to avoid too quick an inventory adjustment, defined by τ_I. 2. There is a trade-off in the sub-assembler between capacity and CODP inventory variance defined by τ_{AINV}. This policy parameter needs to be carefully selected due to its dominant influence on the dynamic behavior of the ATO system. 3. The forecasting policy plays a substantively smaller role in influencing the dynamic performance of the ATO system in comparison to the other policies in the system, contrasting to previous studies that assumed linearity (Dejonckheere et al. 2002; Li, Disney and Gaalman, 2014).
	Feedforward loops		An increase in τ_A leads to a reduction in bullwhip (ORATE$_{SA}$ variance) at the expense of increasing AINV-$_{SA}$ variance, although the effect of τ_A is limited comparing feedback control loops	
Nonlinearities	Linearization	$N_{A(NO)}$ and $N_{B(NO)}$	1. The occurrence of non-negative order constraints at the final assembly site lead to a change in $N_{A(NO)}$ ranging between 0.5 and 1 depending on the amplitude of input demand. 2. $N_{B(NO)}$ increases as of demand amplitude increases	1.Being aware of the impact of the system's nonlinearities and constraints is very important for final assemblers. Depending on the demand amplitude, the non-negative order constraint at the final assembly plant may occur, such that $N_{B(NO)}$ will increase with demand amplitude, and this could lead to a significant increase in average inventory level, which increases total costs. 2. Production managers at the subassembly site should carefully consider capacity utilization, i.e. should the mean of the orders received from the downstream final assembly exceed half of the maximum capacity, then the dominant impact on CODP inventory dynamics will be the capacity constraint rather than the non-negative order low boundary. Under such condition, $N_{B(CA)}$ will increase with demand amplitude, leading to the decrease in average inventory level. In contrast, if the mean of the orders received is less than half of the maximum capacity then the non-negative order boundary dominates. This lead to the increase in average CODP inventory level at sub-assemblers. Alternatively, if the mean of the orders received equals half of the maximum capacity then nonlinearities do not have impact on the averaged inventory level.
		$N_{A(CA)}$ and $N_{B(CA)}$	1. $N_{A(CA)}$ decreases and approaches 0 as input demand amplitude increases, triggered by the occurrence of capacity and non-negative order constraints at the sub-assembler site. 2. The change of $N_{B(CA)}$, however, depends on the relationship between mean of input (B) and $\frac{1}{2}$CL. N$_{B(CA)}$ equals to 1 irrespective of input amplitude A if B=$\frac{1}{2}$CL. If B $< \frac{1}{2}$CL, an increase in A leads to an increase in N$_{B(CA)}$, while N$_B$ is monotonically decreasing in A if B $> \frac{1}{2}$CL	
	The impact of $N_{A(NO)}$ and $N_{B(NO)}$	ω_{n1} and ζ_1	A decrease in $N_{A(NO)}$ will result in a decrease in ω_{n1} and an increase in ζ_1.	1. An increase in demand amplitude, which influences $N_{A(NO)}$, will yield a system with lower bullwhip and inventory variance, although at the expense of a slower inventory recovery speed at the final assembly. The

	The impact of $N_{A(CA)}$ and $N_{B(CA)}$	ω_{n2} and ζ_2	The decrease of output amplitude gain, $N_{A(CA)}$, resulting from the capacity and non-negative order constraints at the sub-assembler site, will lead to a decrease in ω_{n2} and ζ_2	latter suggests a decrease in customer service level due to the increased probability of stock-out, in particular when the system's steady state condition is disturbed by a sudden but a sustained demand increase.

2. An increase in demand amplitude, which influences $N_{A(CA)}$, will decrease CODP inventory recovery speed at the subassembly, which also directly increases the stock-out probability of CODP inventory at the final assembly site. |
| **Quality and efficiency** | The impact of Y_F an Y_S | | The decrease of Y_F and Y_S significantly increases bullwhip, or $ORATE_{SA}$ variance, and Y_F also plays a key role in influencing the variance of $AINV_{SA}$ | Final assembler should pay attention to their final assembly line efficiency, defined by Y_F, and the sub-assembler needs to consider yield losses, given by Y_S, since they not only directly relate to the customer service level, i.e. whether the total orders can be delivered within the quoted lead times, but also increase supply chain dynamics costs of the upstream supplier in the ATO system. |

Table 6. The summary of findings and managerial implications in this study.

Reference

- Atan, Zümbül, Taher Ahmadi, Clara Stegehuis, Ton de Kok, and Ivo Adan. 2017 "Assemble-to-order systems: A review." *European Journal of Operational Research* 261 (3): 866-879.
- Atherton, Derek P.1975. "*Nonlinear control engineering.*", New York: Van Nostrand Reinhold.
- Berry, Danny, and Denis R. Towill. 1992. "Material flow in electronic product based supply chains." *International Journal of Logistics Management* 3 (2): 77-94.
- Berry, Danny, Denis R. Towill, and Nick Wadsley. 1994. "Supply chain management in the electronics products industry." *International Journal of Physical Distribution and Logistics Management* 24 (10): 20-32.
- Cannella, Salvatore, Elena Ciancimino, and Adolfo Crespo Marquez. 2008. "Capacity constrained supply chains: a simulation study." *International Journal of Simulation and Process Modelling* 4 (2): 139-147.
- Cannella, Salvatore, Elena Ciancimino., Jose M. Framinan., 2011. "Inventory policies and information sharing in multi-echelon supply chains". *Production Planning and Control* 22 (7), 649–659
- Cannella, Salvatore, Jalal Ashayeri, Pablo A. Miranda, and Manfredi Bruccoleri. 2014. "Current economic downturn and supply chain: the significance of demand and inventory smoothing". *International Journal of Computing Integrated Manufacturing*, 27 (3), 201–212.
- Cannella, Salvatore, Roberto Dominguez, Borja Ponte, and Jose M. Framinan. 2018. "Capacity restrictions and supply chain performance: Modelling and analysing load-dependent lead times." *International Journal of Production Economics*. In press: https://doi.org/10.1016/j.ijpe.2018.08.008.
- Chatfield, Dean C., and Alan M. Pritchard. 2013. "Returns and the bullwhip effect." *Transportation Research Part E: Logistics and Transportation Review* 49, 1: 159-175.
- Cheng, Feng, Markus Ettl, Yingdong Lu, and David D. Yao. 2012 "A production–inventory model for a push–pull manufacturing system with capacity and service level constraints." *Production and Operations Management* 21 (4): 668-681.
- Choi, Kanghwa, Ram Narasimhan, and Soo Wook Kim. 2012. "Postponement strategy for international transfer of products in a global supply chain: A system dynamics examination." *Journal of Operations Management* 30 (3): 167-179.
- Christopher, Martin, and Helen Peck. 2004. "Building the resilient supply chain." *The International Journal of Logistics Management* 15 (2): 1-14.
- Coyle, R.G.1977. *Management system dynamics*. Chichester: John Wiley & Sons Australia, Limited.
- Dejonckheere, Jeroen, Stephen M. Disney, Marc R. Lambrecht, and Denis R. Towill. 2002. "Transfer function analysis of forecasting induced bullwhip in supply chains." *International Journal of Production Economics* 78 (2): 133-144.
- Dominguez, Roberto, Salvatore Cannella, and Jose M. Framinan. 2015. "On returns and network configuration in supply chain dynamics." *Transportation Research Part E: Logistics and Transportation Review* 73: 152-167.
- Edghill, J., and D. R. Towill. 1990. "Assessing manufacturing system performance: frequency response revisited." *Engineering Costs and Production Economics* 19 (1-3): 319-326.
- Forrester, J. W. (1958). "Industrial dynamics: A breakthrough for decision makers." *Harvard Business Review* July-August: 37–66.
- Forrester, J. W. (1961). *Industrial dynamics*. Cambridge: MIT Press
- Gonçalves, Paulo, Jim Hines, and John Sterman. 2005. "The impact of endogenous demand on push–pull production systems." *System dynamics review* 21 (3): 187-216.
- Harrison, Terry P., Hau L. Lee, and John J. Neale. 2005. *The practice of supply chain management: where theory and application converge*. Springer Science & Business Media.
- Hedenstierna, Philip, and Amos HC Ng. 2011. "Dynamic implications of customer order decoupling point positioning." *Journal of Manufacturing Technology Management* 22, no. (8): 1032-1042.

- Huang, Yu-Ying, and Shyh-Jane Li. 2010. "How to achieve leagility: A case study of a personal computer original equipment manufacturer in Taiwan." *Journal of Manufacturing Systems* 29 (2-3): 63-70.

- Hussain, M., Khan, M., Sabir, H., 2015. Analysis of capacity constraints on the backlog bullwhip effect in the two-tier supply chain: a Taguchi approach. *International Journal of Logistics Research and Application*. 19(1): 1–21.

- Jeong, Sanghwa, Yonghun Oh, and Sangsuk Kim. 2000. "Robust Control of multi-echelon production-distribution systems with limited decision policy (II)." *KSME International Journal* 14 (4): 380-392.

- Simon, John, Mohamed Mohamed Naim, and Denis Royston Towill. 1994. "Dynamic analysis of a WIP compensated decision support system." *International Journal of Manufacturing System Design* 1 (4): 283-297.

- Katariya, Abhilasha Prakash, Sıla Çetinkaya, and Eylem Tekin. 2014. "Cyclic consumption and replenishment decisions for vendor-managed inventory of multisourced parts in Dell's supply chain." *Interfaces* 44 (3): 300-316.

- Karabuk, Suleyman, and S. David Wu. 2003. "Coordinating strategic capacity planning in the semiconductor industry." *Operations Research* 51 (6): 839-849.

- Kumar, Sameer, and Sarah Craig, 2007. "Dell, Inc.'s closed loop supply chain for computer assembly plants." *Information Knowledge Systems Management* 6, no. 3: 197-214.

- Lee, Hau L., Venkata Padmanabhan, and Seungjin Whang. 1997. "Information distortion in a supply chain: The bullwhip effect." *Management Science* 43 (4): 546-558.

- Lee, Hau L., and Christopher S. Tang. 1997. "Modelling the costs and benefits of delayed product differentiation." *Management Science* 43 (1): 40-53.

- Li, Qinyun, Stephen M. Disney, and Gerard Gaalman. 2014. "Avoiding the bullwhip effect using Damped Trend forecasting and the Order-Up-To replenishment policy." *International Journal of Production Economics* 149: 3-16.

- Lin, Junyi, Mohamed M. Naim, Laura Purvis, and Jonathan Gosling. 2017. "The extension and exploitation of the inventory and order based production control system archetype from 1982 to 2015." *International Journal of Production Economics* 194: 135-152.

- Lin, Junyi, Virginia LM Spiegler, and Mohamed M. Naim. 2017. "Dynamic analysis and design of a semiconductor supply chain: a control engineering approach." *International Journal of Production Research*. 56:13, 4585-4611.

- Mönch, Lars, John W. Fowler, and Scott J. Mason. 2013. *Production Planning and Control for Semiconductor Wafer Fabrication Facilities*. Berlin: Springer-Verlag New York.

- Nagatani, Takashi, and Dirk Helbing. 2004. "Stability analysis and stabilization strategies for linear supply chains." *Physica A: Statistical Mechanics and its Applications* 335 (3-4): 644-660.

- Naylor, J. Ben, Mohamed M. Naim, and Danny Berry. 1999. "Leagility: Integrating the lean and agile manufacturing paradigms in the total supply chain." *International Journal of Production Economics* 62 (1-2): 107-118.

- Nepal, Bimal, Alper Murat, and Ratna B. Chinnam. 2012. "The bullwhip effect in capacitated supply chains with consideration for product life-cycle aspects." *International Journal of Production Economics* 136(2): 318-331.

- Ponte, Borja, Xun Wang, David de la Fuente, and Stephen M. Disney. 2017. "Exploring nonlinear supply chains: the dynamics of capacity constraints." *International Journal of Production Research* 55, (14): 4053-4067.

- Sargent, R. G. 2013. "Verification and Validation of Simulation Models." *Journal of Simulation* 7 (1): 12–24.

- Sarimveis, Haralambos, Panagiotis Patrinos, Chris D. Tarantilis, and Chris T. Kiranoudis. 2008 "Dynamic modeling and control of supply chain systems: A review." *Computers & Operations Research* 35 (11): 3530-3561.
- Simon, Herbert A. 1952. "On the application of servomechanism theory in the study of production control." *Econometrica: Journal of the Econometric Society*: 247-268.
- Sipahi, Rifat, and Ismail Ilker Delice. 2010. "Stability of inventory dynamics in supply chains with three delays." *International Journal of Production Economics* 123 (1): 107-117.
- Spiegler, Virginia LM, Mohamed M. Naim, and Joakim Wikner. 2012. "A control engineering approach to the assessment of supply chain resilience." *International Journal of Production Research* 50 (21): 6162-6187.
- Spiegler, Virginia LM, Andrew Thomas Potter, M. M. Naim, and Denis Royston Towill. 2016a. "The value of nonlinear control theory in investigating the underlying dynamics and resilience of a grocery supply chain." *International Journal of Production Research* 54 (1): 265-286.
- Spiegler, Virginia LM, Mohamed M. Naim, Denis R. Towill, and Joakim Wikner. 2016b "A technique to develop simplified and linearised models of complex dynamic supply chain systems." *European Journal of Operational Research* 251(3): 888-903.
- Spiegler, Virginia LM, and Mohamed M. Naim. 2017. "Investigating sustained oscillations in nonlinear production and inventory control models." *European Journal of Operational Research* 261(2): 572-583.
- Sterman, John D. 1989. "Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment." *Management Science* 35 (3): 321-339.
- Syntetos, Aris A., Nicholas C. Georgantzas, John E. Boylan, and Brian C. Dangerfield. 2011 "Judgement and supply chain dynamics." *Journal of the Operational Research Society* 62 (6): 1138-1158.
- Towill, D. R. "Exponential smoothing of learning curve data. 1977. " *International Journal of Production Research* 15 (1): 1-15.
- Towill, Denis R. 1982. "Dynamic analysis of an inventory and order based production control system." *International Journal of Production Research* 20 (6): 671-687.
- Towill, Denis R., Li Zhou, and Stephen M. Disney. 2007. "Reducing the bullwhip effect: Looking through the appropriate lens." *International Journal of Production Economics* 108, (1-2): 444-453.
- Vander Velde, Wallace E. 1968. *Multiple-input describing functions and nonlinear system design*. McGraw-Hill, New York.
- Wang, Gang, and Angappa Gunasekaran. 2017. "Modeling and analysis of sustainable supply chain dynamics." *Annals of Operations Research* 250 (2): 521-536.
- Wang, Zhaodong, Xin Wang, and Yanfeng Ouyang. 2015. "Bounded growth of the bullwhip effect under a class of nonlinear ordering policies." *European Journal of Operational Research* 247 (1): 72-82.
- Warburton, Roger DH, Stephen M. Disney, Denis R. Towill, and Jonathan PE Hodgson. 2004. "Further insights into 'the stability of supply chains'." *International Journal of Production Research* 42 (3): 639-648.
- Wang, Xun, Stephen M. Disney, and Jing Wang. 2012. "Stability analysis of constrained inventory systems with transportation delay." *European Journal of Operational Research* 223 (1): 86-95.
- Wang, Xun, and Stephen M. Disney. 2016. "The bullwhip effect: Progress, trends and directions." *European Journal of Operational Research* 250 (3): 691-701.
- Wang, Xun, Stephen M. Disney, and Jing Wang. 2014. "Exploring the oscillatory dynamics of a forbidden returns inventory system." *International Journal of Production Economics* 147: 3-12.
- Wikner, Joakim. 2003. "Continuous-time dynamic modelling of variable lead times." *International Journal of Production Research* 41(12): 2787-2798.

- Wikner, Joakim, Mohamed M. Naim, Virginia LM Spiegler, and Junyi Lin. 2017. "IOBPCS based models and decoupling thinking." *International Journal of Production Economics* 194:153-166.
- Xiao, Yongbo, Jian Chen, and Chung-Yee Lee. 2012. "Single-Period Two-Product Assemble-to-Order Systems with a Common Component and Uncertain Demand Patterns." *Production and Operations Management* 19 (2): 216-232.
- Zhou, Li, Mohamed M. Naim, and Stephen M. Disney. 2017. "The impact of product returns and remanufacturing uncertainties on the dynamic performance of a multi-echelon closed-loop supply chain." *International Journal of Production Economics* 183: 487-502.