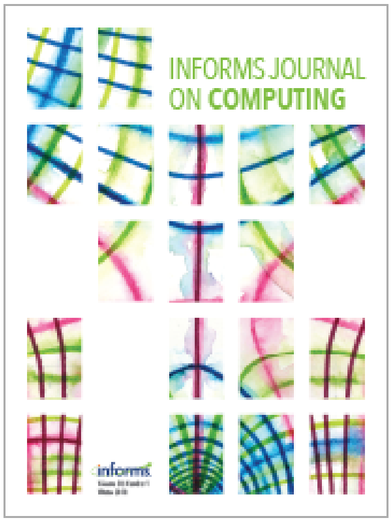


INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Managing Product Transitions: A Bilevel Programming Approach

Rahman Khorramfar, Osman Y. ÖzaltınKarl G. Kempf, Reha Uzsoy

informs
http://pubsonline.informs.org/journal/ijoc
INFORMS JOURNAL ON COMPUTING
Articles in Advance, pp. 1–17
ISSN 1091-9856 (print), ISSN 1526-5528 (online)

Managing Product Transitions: A Bilevel Programming Approach

Rahman Khorramfar,[a] Osman Y. Özaltın[b,*] Karl G. Kempf,[c] Reha Uzsoy[b]

[a] MIT Energy Initiative, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142; [b] Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, North Carolina 27695; [c] Intel Corporation, Chandler, Arizona 85226
*Corresponding author

Contact: khorram@mit.edu, https://orcid.org/0000-0002-7802-7460 (RK); oyozalti@ncsu.edu, https://orcid.org/0000-0002-0093-5645 (OYÖ); karl.g.kempf@intel.com, https://orcid.org/0000-0001-9860-5419 (KGK); ruzsoy@ncsu.edu, https://orcid.org/0000-0002-8068-550X (RU)

Received: August 12, 2021
Revised: March 5, 2022; April 22, 2022; May 16, 2022
Accepted: May 18, 2022
Published Online in Articles in Advance: June 30, 2022

https://doi.org/10.1287/ijoc.2022.1210

Abstract. We model the hierarchical and decentralized nature of product transitions using a mixed-integer bilevel program with two followers, a manufacturing unit and an engineering unit. The leader, corporate management, seeks to maximize revenue over a finite planning horizon. The manufacturing unit uses factory capacity to satisfy the demand for current products. The demand for new products, however, cannot be fulfilled until the engineering unit completes their development, which, in turn, requires factory capacity for prototype fabrication. We model this interdependency between the engineering and manufacturing units as a generalized Nash equilibrium game at the lower level of the proposed bilevel model. We present a reformulation where the interdependency between the followers is resolved through the leader's coordination, and we derive a solution method based on constraint and column generation. Our computational experiments show that the proposed approach can solve realistic instances to optimality in a reasonable time. We provide managerial insights into how the allocation of decision authority between corporate leadership and functional units affects the objective function performance. This paper presents the first exact solution algorithm to mixed-integer bilevel programs with interdependent followers, providing a flexible framework to study decentralized, hierarchical decision-making problems.

History: Accepted by Andrea Lodi, Area Editor for Design & Analysis of Algorithms—Discrete.
Funding: This research was supported by the National Science Foundation (NSF) [Grant CMMI-1824744]. Any opinions stated are those of the authors and do not necessarily reflect the position of NSF.
Supplemental Material: The online supplement is available at https://doi.org/10.1287/ijoc.2022.1210.

Keywords: product transitions • bilevel programming • constraint and column generation • generalized Nash equilibrium

1. Introduction

Product transitions involve the introduction of new products to replace those that are currently being sold to customers in response to ever-changing customer preferences (Klastorin and Tsai 2004), shorter product life cycles (Wu et al. 2009), and increasing global competition (Wu et al. 2005). Effective management of product transitions can bring significant competitive advantage to a firm, whereas poor decisions can have serious adverse business consequences. A central challenge during product transitions is the coordination of the product development process, which creates new products, and the supply chain that manufactures and distributes them. The product development process requires access to factory capacity to fabricate prototypes to assess their manufacturability and functionality, whereas the supply chain requires timely delivery of new product designs with enhanced features and reduced costs to stay competitive in the market. Although the challenges faced by firms in managing product transitions have been addressed in the literature (Lim and Tang 2006, Bilginer and Erhun 2010), most previous studies share several shortcomings:

- They treat the firm as a single decision entity with complete information about all aspects of its operations. However, as firms grow in size and diversify their product portfolio, operational control decisions become distributed among different functional groups, rendering detailed centralized planning impractical. The critical resource allocation decisions and domain knowledge are distributed among different functional groups within the firm, such as product divisions and manufacturing units, that are trying to reconcile their local, potentially conflicting, objectives with those of the corporate management (Bansal et al. 2020).
- They ignore the hierarchy of decision makers involved in managing product transitions. In practice, upper-level management addresses strategic/tactical

planning, resulting in directives to the functional groups that must make and implement operational decisions.

- They treat product transitions in isolation from routine operations, neglecting their impact on products that are not directly involved in the transition. However, firms with broad product portfolios, such as large semiconductor manufacturers, must often manage simultaneous product transitions in several markets supplied from shared manufacturing facilities. Hence new product introductions can adversely affect other products with which they share manufacturing, development, or marketing resources (Ulrich and Eppinger 2016).

Understanding the information flow and hierarchical relation between the decision-making units involved in product transitions is essential to developing more realistic models. Motivated by a large semiconductor manufacturing firm, we present a bilevel model that represents a decentralized, hierarchical environment in which product transitions take place. Although semiconductor manufacturing firms differ in the specifics of their manufacturing processes and product development activities, the knowledge and data required for the successful introduction of new products and retirement of older ones are usually distributed across three units: (1) corporate management (CORP), whose leadership role is to direct the company toward achieving its strategic business goals; (2) one or more product engineering (ENG) units, each of which is responsible for new product development activities in a specific market segment, including market research, design, implementation, and prototype testing; and (3) a manufacturing (MFG) organization that manufactures existing products for sale and prototypes for products under development by the ENG units.

CORP communicates desired production quantities to MFG based on demand. The ENG units determine specifications for new products through market research and then develop these product specifications into detailed design. The product development process alternates between periods of design activity and prototype fabrication. Once a product design has been tested as far as possible in software, the prototypes are fabricated to fully debug the new designs and assess their manufacturability. Several such cycles of design activity, prototype fabrication, and design testing and refinement may be required before the design is verified. Hence, the completion of product development may be delayed if ENG units do not have timely access to factory capacity for prototype fabrication or if too many design cycles are needed. The factory capacity used by the ENG units reduces the capacity available for MFG to meet current orders for revenue-generating products but is essential to maintain a profitable product pipeline in the future.

Optimizing corporate objectives by effectively coordinating the functional units involved in product transitions is extremely challenging. We propose a bilevel model where CORP, acting as a leader, seeks to maximize its revenue over a finite planning horizon subject to the decisions of the MFG and ENG units, which act as followers playing a generalized Nash equilibrium game in the lower level. The decision problem faced by the MFG unit is a linear program. The ENG solves a mixed-integer program that can be computationally difficult to solve even in a deterministic environment. Unlike traditional bilevel programs, the decisions of the two followers in our model are interdependent; the ENG and MFG units share factory capacity, and MFG is dependent on ENG for the development of new products to meet future demand.

The specific decision units considered in our model represent the decentralized, hierarchical organizational structure in which product transitions take place in many large corporations and particularly in the global semiconductor manufacturing firm motivating our work in this paper. Different organizational structures might lead to different decentralized decision settings in other firms, which could be studied using similar approaches. For example, the CORP and MFG can be merged into a single player that maximizes the difference between revenue and manufacturing costs. The main contributions of this paper are as follows:

- We propose a mixed-integer bilevel programming model to effectively coordinate product transitions in a large decentralized firm by considering the technological constraints and objectives of the functional units involved in product transitions. To the best of our knowledge, this is the first bilevel model developed for product transitions in the literature.
- The proposed bilevel programming model is also novel in considering two interdependent followers. To the best of our knowledge, this paper presents the first solution algorithm for mixed-integer bilevel programs with interdependent followers.
- We present two single-level reformulations of the bilevel model and develop an efficient solution approach based on these reformulations. We perform extensive computational experiments to evaluate the performance of the solution approach and provide key policy insights about the implications of decision hierarchy.

The remainder of this paper is organized as follows. We review the literature on product transitions and production planning in semiconductor manufacturing, as well as the state of the art in bilevel programming and generalized Nash equilibrium problem, in Section 2. We formulate a mixed-integer bilevel model for product transitions in Section 3. We develop a solution approach in Section 4 and present the results of computational experiments in Section 5. We conclude the paper and discuss future research directions in Section 6.

2. Literature Review

2.1. Product Transitions

A growing body of researchers has examined product transitions (Ferrer and Swaminathan 2006, Bilginer and Erhun 2010). Li et al. (2013) and Bhaskaran et al. (2015) study capacity planning for product transitions considering demand uncertainty, the effect of competition, and service-level requirements. The impact of initial investment on quality improvement and time-to-market are examined in Wu et al. (2009). A rich body of work examines the impact of customer behavior on product transition decisions such as the optimal timing of product introduction (Druehl et al. 2009, Liao and Seifert 2015) and product rollover strategies (Liang et al. 2014, Lobel et al. 2015). Klastorin and Tsai (2004) develop a game-theoretical model to capture the interactions among pricing, timing, and product design when entering a new market. Koca et al. (2010) analyze the impact of preannouncement and inventory decisions on the demand of new products.

Although this extensive literature provides useful insights into several aspects of the product transition problem, it fails to consider several important aspects of the problem addressed in this paper, particularly the hierarchy of decision makers, the interactions between them, and the complex technological constraints that an implementable solution must satisfy. It also fails to consider the impact of product transitions on other products that share capacity with the new products but are not in transition themselves. Thus our paper presents a novel direction for studying this complex and important business problem.

2.2. Production Planning in Semiconductor Manufacturing

A number of authors have addressed decentralized production planning in semiconductor manufacturing. Karabuk and Wu (2003) propose a multistage stochastic programming model under demand and yield uncertainty with a manufacturing unit and several product managers. In a subsequent paper, Karabuk and Wu (2005) consider corporate headquarters and product managers as the units involved in the capacity allocation process, and they propose a game-theoretic approach to elicit private information from the product managers to maximize expected corporate profit. More recently, Bansal et al. (2020) consider a simplified version of the product transition problem considered in this paper. They develop two iterative combinatorial auction schemes based on Lagrangian relaxation and column generation that seek to coordinate negotiations over factory capacity between MFG, acting as the auctioneer, and ENG units that bid for factory capacity. Bansal et al. (2020) do not enforce a strict decision hierarchy but instead employ a decentralized iterative combinatorial auction to coordinate the participants toward an implementable solution.

2.3. Generalized Nash Equilibrium Problem

The generalized Nash equilibrium problem (GNEP) is a noncooperative game in which each player's strategy set depends on the strategies of other players (Facchinei and Kanzow 2010). From an optimization perspective, GNEP is a decentralized model where decision makers share a set of *coupling constraints*. The outcome of a GNEP is a set of *equilibria*, a collection of strategies from which no decision maker has incentive to deviate unilaterally. GNEP has many applications, including pricing in telecommunication networks (Altman and Wynter 2004), electricity market analysis (Le Cadre et al. 2020), and transportation problems (Stein and Sudermann-Merx 2018, Sagratella et al. 2020). Despite its diverse applications, solution methods for GNEP are generally confined to cases where players control a set of continuous variables (Dreves et al. 2011, Facchinei et al. 2014, Aussel and Sagratella 2017).

In recent years, several researchers have studied GNEP with mixed-integer variables. Sagratella (2017) consider generalized potential Nash games, a special case of GNEP where players (unknowingly) optimize the same objective function over the aggregated feasible strategy set of all players. Huppmann and Siddiqui (2018) propose an exact reformulation to find Nash equilibria in a noncooperative game with binary decision variables by including compensation payments and incentive-compatibility constraints. Sagratella (2019) show that, under mild assumptions, the set of equilibrium points of a GNEP with mixed-integer variables and linear coupling constraints is finite and propose algorithms to generate all possible equilibria. More recently, Sagratella et al. (2020) propose a mixed-integer GNEP to address the noncooperative fixed charge transportation problem.

2.4. Bilevel Programming

Bilevel programming (BP) provides a powerful tool for modeling hierarchical decision-making problems in which the outcome of decisions by an upper-level authority (the leader) is affected by the response from a lower-level entity (follower) that seeks to optimize its own objective function (Bard 2013). From the perspective of game theory, BP models a static Stackelberg game (Dempe 2002, Bard 2013). An important feature of BP is that the feasible region of each level's decision problem may be impacted by variables controlled by the other level. This embedded hierarchy reflects the organizational structure in a firm and renders BP a suitable approach for the product transition problem studied in this paper.

BPs can be generally classified as bilevel linear programs (BLPs) with no integer variables and mixed-integer bilevel linear programs (MIBLPs) that have integer variables in at least one decision level. BLPs are nonconvex and strongly NP-hard (Colson et al. 2007), although they can be reformulated as a single-level model by enforcing the optimality of the follower's decisions through Karush-Kuhn-Tucker conditions (Bertsimas and Tsitsiklis 1997). The resulting single-level problem, which is usually nonlinear, can be solved using branch-and-bound (Bard and Moore 1990, Hansen et al. 1992), branch-and-cut (Audet et al. 2007), Bender's decomposition (Saharidis and Ierapetritou 2009, Nishi et al. 2011), or penalty function methods (Campelo et al. 2000).

Bard and Moore (1990; 1992), the first to consider discrete variables in the lower-level problem, apply a branch-and-bound algorithm and establish several properties of pure integer BPs. DeNegre and Ralphs (2009) extend this work to develop a branch-and-cut algorithm. Wang and Xu (2017) propose an algorithm for pure integer BPs that removes bilevel infeasible solutions by disjunctive cuts.

Xu and Wang (2014) consider a BP with a mixed-integer lower level and a pure integer upper level with bounded variables. They propose a branch-and-bound algorithm where branching occurs on the slack variables of the follower's constraints. Lozano and Smith (2017) extend this work by employing a single-level value function reformulation. Fischetti et al. (2017, 2018) propose a branch-and-cut algorithm for MIBLPs by generating intersection cuts. Yue et al. (2019) develop a projection-based approach for MIBLPs that first reformulates the problem into a single-level equivalent and then decomposes it by implicitly enumerating the follower's integer variables. MIBLPs have been applied in transportation (Arslan et al. 2018), robust multicommodity network design (Sun et al. 2018; 2019), hazardous materials transportation (Liu and Kwon 2020), homeland security and defense (Aksen and Aras 2012), electricity markets (Lavigne et al. 2000), product introduction (Hemmati and Smith 2016), healthcare (Özaltın et al. 2018), and supply chain (Yue and You 2017).

The limited research on BPs with multiple followers mostly considers BLPs. Shi et al. (2007) study a BLP with multiple followers who simultaneously determine a set of common variables. Calvete and Galé (2007) propose a method to transform a problem with multiple independent followers into a BLP with a single follower whose objective function is the sum of the objective functions of all followers. Calvete et al. (2019) consider a nonlinear BP with multiple followers in a rank pricing problem and present a nonlinear single-level reformulation. Tavaslıoğlu et al. (2019) develop a value function-based approach for MIBLPs with multiple independent followers. Their approach assumes that all followers have the same constraint matrix, and the leader's variables impacting the feasible regions of the followers can only take integer values. To the best of our knowledge, BPs with interdependent followers and discrete variables in both levels have not been considered in the literature.

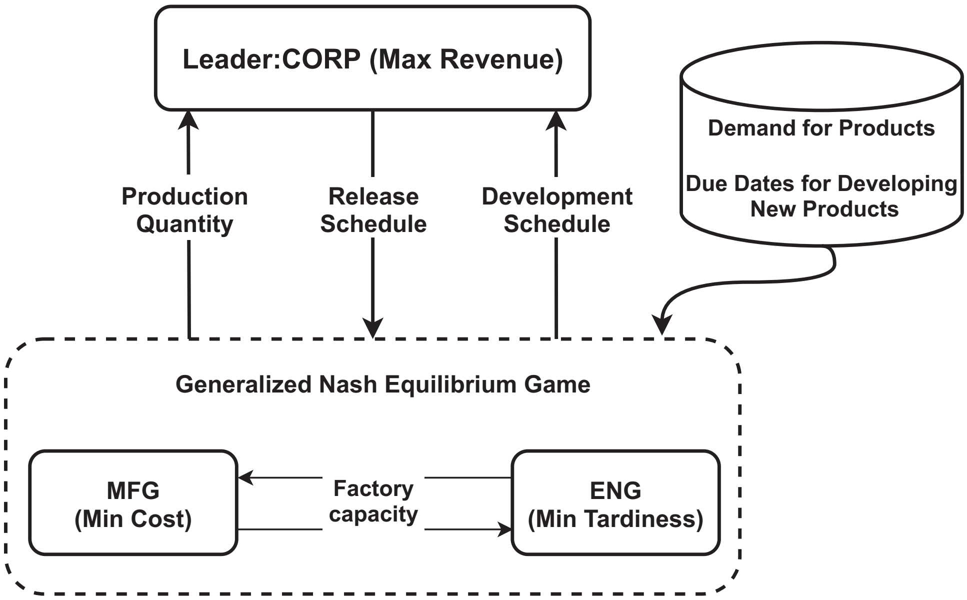
3. Model Formulation

We present a bilevel model of the product transition management problem to capture its decentralized, hierarchical nature. CORP acts as a leader, seeking to maximize its total revenue over a finite planning horizon $\mathcal{T} = \{1, \ldots, T\}$. Revenue is earned by selling the products manufactured by the MFG unit. The MFG and ENG units are two interdependent followers, responding to the decisions of the CORP in a manner that optimizes their local objective functions. They play a generalized Nash equilibrium game over the usage of factory capacity.

Figure 1 depicts the information exchange between the CORP, MFG, and ENG units. CORP determines when to release new products to MFG for sale in the market. However, new products cannot be released to MFG unless ENG completes their development, rendering MFG dependent on the decisions of ENG. On the other hand, ENG cannot complete the development of new products before their due dates without timely access to factory capacity for prototype fabrication, leading to competition between MFG and ENG over factory capacity. As indicated by the solid black arrows in Figure 1, once MFG and ENG reach an equilibrium over the usage of factory capacity, MFG communicates the production quantities to the CORP. We refer to this model as the bilevel product transition model with a generalized Nash equilibrium (BPTM-Nash).

Although we consider one MFG and one ENG unit in the lower level of our model, the MFG subproblem may represent the combined decisions of multiple manufacturing facilities, and the ENG subproblem may represent the integrated decisions of several product development units. Thus the factory capacity allocated to MFG in the optimal solution to BPTM-Nash may be shared by multiple manufacturing units. In the same vein, the factory capacity allocated to ENG in the optimal solution of BPTM-Nash may be shared by multiple product development units. Hence the proposed bilevel model with two followers is representative of the decisions encountered in practical instances of the product transition management problem. For the sake of exposition, we shall focus on the case with two followers, although we extend our approach to consider multiple followers in Appendix C of the online supplement and examine the impact of

Figure 1. Schematic View of the Proposed Bilevel Model with CORP as the Leader and MFG and ENG as Interdependent Followers Who Play a Generalized Nash Equilibrium Game over Factory Capacity



Notes. Arrows represent "decision variables" communicated between CORP and the followers. The demand estimates and due dates for the development of products are parameters given to the followers.

additional followers on the solution time in Section 5.3.

Let N denote the set of all products, and let $P \subset N$ denote the set of new products currently under development. Table 1 summarizes the model parameters, and Table 2 presents the upper- and lower-level decision variables. CORP specifies the value of the upper-level variable $Y_{pt} \in \{0,1\}$ to release new product $p \in P$ to MFG at the beginning of period t. Thus $Y \equiv \{Y_{p\in P, t\in \mathcal{T}}\}$ represents the new product *release schedule*. Given a product release schedule Y, we denote the generalized Nash equilibrium game between MFG and ENG in the lower level, as well as its solution set, by GNEP(Y). Let $\mathcal{X}$ denote the tuple (X, B, I, F, Z, V). We formulate the BPTM-Nash model as follows:

$$[\text{BPTM-Nash}] \qquad \max \sum_{n\in N}\sum_{t\in\mathcal{T}} \pi_{nt}(D_{nt} + B_{n,t-1} - B_{nt}) \tag{1a}$$

$$\text{subject to } Y_{p,t-1} \le Y_{pt} \qquad t \in \mathcal{T}\backslash\{1\}, \tag{1b}$$

$$Y_{pt} \le \sum_{\tau\le t} Z_{p\tau} \qquad p \in P, t \in \mathcal{T}, \tag{1c}$$

$$\mathcal{X} \in \text{GNEP}(Y), \tag{1d}$$

$$Y_{pt} \in \{0,1\} \qquad p \in P, t \in \mathcal{T}. \tag{1e}$$

The upper-level objective function (1a) maximizes the total revenue of the firm. The demand D_{nt} is a constant, independent of the decision variables, but we retain it to clarify that the objective function denotes the revenue earned by satisfying demand. Constraints (1b) ensure that once a new product is released to MFG in period t, it can be manufactured in all subsequent periods. Constraints (1c) ensure that the development of new product $p \in P$ is completed in period $\tau \le t$ (i.e., $Z_{p\tau} = 1$) before its release to MFG in period t. Constraint (1d) ensures that $\mathcal{X}$ is a solution to the generalized Nash equilibrium game between MFG and ENG in the lower level of the bilevel model. CORP controls the product release schedule through the Y_{pt} variables in BPTM-Nash. We follow an optimistic approach, allowing CORP to select the most advantageous equilibrium of GNEP(Y) if more than one exists.

Table 1. Parameters of the Bilevel Product Transition Model

Parameter	Description
D_{nt}	Demand for product n in period t
π_{nt}	Marginal revenue for product n in period t
r_{nt}	Marginal production cost of product n in period t
b_{nt}	Marginal backorder cost of product n in period t
H_{pt}	Capacity required for prototype fabrication of product $p \in P$ in period t
δ_p	Due date for completing the development of new product $p \in P$
w_p	Per-period tardiness cost for product $p \in P$
h_{nt}	Unit inventory holding cost for product n in period t
C_t	Factory capacity in period t

Table 2. Decision Variables of the Bilevel Product Transition Model

Unit	Variable	Description
CORP (leader)	Y_{pt}	1 if product p is released to MFG at period t and 0 otherwise (Y is referred to as the "product *release* schedule")
MFG (follower)	B_{nt}	Backorder of product n in period t
	X_{nt}	Fraction of factory capacity assigned to product n in period t
	I_{nt}	Inventory level of product n in period t
ENG (follower)	F_t	Fraction of factory capacity allocated for product development in period t
	V_p	Tardiness in the development of product p
	Z_{pt}	1 if product p is developed in period t and 0 otherwise (Z is referred to as the "product *development* schedule")

MFG's problem in GNEP(Y) can be written as the following linear program:

$$[\text{MFG}(F,Y)] \quad \min \sum_{n\in N}\sum_{t\in\mathcal{T}}(h_{nt}I_{nt} + r_{nt}X_{nt}C_t + b_{nt}B_{nt}) \quad (2a)$$

$$\text{subject to } F_t + \sum_{n\in N} X_{nt} = 1 \quad t\in\mathcal{T}, \quad (2b)$$

$$I_{nt} = I_{n,t-1} + X_{nt}C_t - (D_{nt} + B_{n,t-1} - B_{nt}) \quad n\in N, t\in\mathcal{T}, \quad (2c)$$

$$X_{pt} \le Y_{pt} \quad p\in P, t\in\mathcal{T}, \quad (2d)$$

$$\sum_{n\in N} X_{nt} \le 1 \quad t\in\mathcal{T}, \quad (2e)$$

$$X_{nt}, I_{nt}, B_{nt} \ge 0 \quad n\in N, t\in\mathcal{T}. \quad (2f)$$

MFG's objective function (2a) minimizes the sum of production, inventory holding, and backorder costs over the planning horizon. Coupling constraints (2b) ensure that the factory capacity allocated for product development (i.e., F_t) and that used by MFG to meet demand (i.e., $\sum_{n\in N} X_{nt}$) is equal to the available capacity in period t. The feasible region of the MFG unit, as a follower in the lower level of BPTM-Nash, depends on the decisions of both the leader CORP (who controls the Y variables) and those of the other follower ENG (who controls the F_t variables in (2b)). Constraints (2c) enforce inventory balance across time periods, whereas Constraints (2d), through the upper-level variable Y, ensure that new products cannot be manufactured unless they are released to MFG. Without loss of generality, we assume that $N \supset P$ (i.e., there is at least one current product) to ensure that MFG(F,Y) is feasible for any given (F, Y) tuple. If there is no such current product, we can create a dummy product with zero demand. Any factory capacity assigned to the dummy product would then represent unused factory capacity. Finally, Constraints (2e) impose the factory capacity limit in each period. For a given Y, we refer to Constraints (2c)–(2f) collectively as `MFG-primal-feasibility`(Y).

We formulate ENG's problem in GNEP(Y) as the following mixed-integer program:

$$[\text{ENG}(X)] \quad \min \sum_{p\in P} w_p V_P \quad (3a)$$

$$\text{subject to } F_t + \sum_{n\in N} X_{nt} = 1 \quad t\in\mathcal{T}, \quad (3b)$$

$$V_p \ge t\left(1 - \sum_{\tau<t} Z_{p\tau}\right) - \delta_p \quad p\in P, t\in\mathcal{T}, \quad (3c)$$

$$\sum_{p\in P} H_{pt}Z_{pt} \le C_tF_t \quad t\in\mathcal{T}, \quad (3d)$$

$$\sum_{t\in\mathcal{T}} Z_{pt} \le 1 \quad p\in P, \quad (3e)$$

$$V_p \ge 0, Z_{pt} \in \{0,1\} \quad p\in P, t\in\mathcal{T}. \quad (3f)$$

ENG's objective function (3a) minimizes the total weighted tardiness of new products in development as calculated in (3c). The coupling constraints (3b) also appear in MFG's problem. Constraints (3d) ensure that the factory capacity used for product development does not exceed F_t in each period t. These constraints allow for idle factory capacity, as the ENG does not have to use all of the allocated capacity. Finally, Constraints (3e) ensure that a new product is developed in a single period. This assumption can be relaxed by considering multiperiod product development; the proposed bilevel framework remains valid. For a given F, we refer to Constraints (3c)–(3f) as `ENG-feasibility`(F).

Tardiness is computed relative to the product development due dates in (3c). ENG may have to delay the completion of development activities depending on the amount of factory capacity it can obtain for prototype fabrication. If tardiness was not allowed, the ENG problem would not be necessary because constraints related to meeting product development deadlines could be enforced in the upper-level problem, leaving MFG as the only follower.

4. Solution Approach

We first show that there are multiple equilibrium solutions to the generalized Nash equilibrium game between MFG and ENG in the lower level for any given product release schedule by CORP.

Lemma 1. *For a given $\hat{Y}$, there is an equilibrium of* GNEP($\hat{Y}$) *for any nonnegative factory capacity allocation to ENG; that is,* $\hat{F}_t \geq 0,\ \forall t \in \mathcal{T}$.

Proof. Let $(\hat{X}, \hat{B}, \hat{I})$ be the optimal response of the MFG($\hat{F}, \hat{Y}$). Note that the ENG variable $F_t = 1 - \sum_{n\in N} \hat{X}_{nt} = \hat{F}_t,\ \forall t \in \mathcal{T}$ in any feasible strategy for ENG($\hat{X}$). Let $(\hat{F}, \hat{Z}, \hat{V})$ be the optimal response of ENG($\hat{X}$). Then $\hat{\mathcal{X}} = (\hat{X}, \hat{B}, \hat{I}, \hat{F}, \hat{Z}, \hat{V})$ is an equilibrium of GNEP($\hat{Y}$), completing the proof.

Given an upper-level product release schedule $\hat{Y}$, the decision problems of MFG and ENG can be merged into a single problem by taking a linear combination of their respective objective functions and enforcing all of their constraints simultaneously. The optimal solution to this problem would be an equilibrium of GNEP($\hat{Y}$) that does not necessarily optimize the revenue of the leader. However, as BPTM-Nash takes an optimistic approach, after deciding on the product release schedule Y, CORP should be able to select an equilibrium of the lower-level game to maximize its revenue. From Lemma 1, CORP can choose an equilibrium by setting the factory capacity allocation for product development (i.e., the F_t variable) in each time period $t \in \mathcal{T}$. We use this result to reformulate BPTM-Nash as a bilevel program with independent followers. Consider the following bilevel program:

$$[\text{BPTM}^*] \quad \max \sum_{n\in N}\sum_{t\in\mathcal{T}} \pi_{nt}(D_{nt} + B_{n,t-1} - B_{nt}) \tag{4a}$$

$$\text{subject to } Y_{pt} \leq \sum_{\tau \leq t} Z_{p\tau} \qquad p \in P,\ t \in \mathcal{T}, \tag{4b}$$

$$Y_{p,t-1} \leq Y_{pt} \quad p \in P, t \in \mathcal{T} \setminus \{1\}, \tag{4c}$$

$$(X, B, I) \in \arg\min \text{MFG}(F, Y), \tag{4d}$$

$$(Z, V) \in \arg\min \text{ENG}(F), \tag{4e}$$

$$F_t \geq 0 \qquad t \in \mathcal{T}, \tag{4f}$$

where ENG(F) is formulated as $\phi_E(F) = \min\{\sum_{p\in P} w_p V_P : \texttt{ENG-feasibility}(F)\}$. In this model, the factory capacity F_t allocated to product development in each period t is decided in the upper level and passed to the followers, decoupling the MFG(F, Y) and ENG(F) subproblems in the lower level of BTPM*.

Proposition 1. *BPTM* is equivalent to BPTM-Nash.*

Proof. We show that there is a feasible solution to BPTM* corresponding to any feasible solution to BPTM-Nash, and vice versa. Consider a bilevel feasible solution $(\hat{Y}, \hat{\mathcal{X}})$ to BPTM* satisfying the upper-level constraints (1b) and (1c) in BPTM-Nash. By (4d), it also satisfies the optimality conditions of the MFG problem. Note that $F_t = 1 - \sum_{n\in N} \hat{X}_{nt} = \hat{F}_t$ for $t \in \mathcal{T}$ in ENG($\hat{X}$). Therefore, if $(\hat{Z}, \hat{V}) \in \arg\min \text{ENG}(\hat{F})$, then $(\hat{Z}, \hat{V}, \hat{F}) \in \arg\min \text{ENG}(\hat{X})$. Thus $\hat{\mathcal{X}}$ is an equilibrium of GNEP($\hat{Y}$), and $(\hat{Y}, \hat{\mathcal{X}})$ is a bilevel feasible solution to BPTM-Nash.

Now consider a bilevel feasible solution $(Y', \mathcal{X}')$ to BPTM-Nash. This solution satisfies the upper-level constraints (4b) and (4c) in BPTM*. It also satisfies Constraints (4d) because MFG's response must be optimal in an equilibrium of GNEP(Y'). Note that $F'_t = 1 - \sum_{n\in N} X'_{nt}$ for $t \in \mathcal{T}$ in ENG(F'). Therefore, if $(Z', V', F') \in \arg\min \text{ENG}(X')$, then $(Z', V') \in \arg\min \text{ENG}(F')$. Thus, $(Y', \mathcal{X}')$ is a bilevel feasible solution to BPTM*.

The fact that BPTM* and BPTM-Nash share the same upper-level objective function completes the proof.

The argument in the proof of Proposition 1 can be extended to any bilevel problem with multiple followers who play a GNEP with exclusively equality coupling constraints. This property was exploited in Sagratella et al. (2020) for a noncooperative fixed charge transportation problem. They showed that any solution satisfying the coupling equality constraints can be used to find an equilibrium of the GNEP by solving each player's problem independently after fixing the variables in the coupling constraints to their values in that feasible solution.

Although Proposition 1 separates the lower-level problem into two independent followers' problems, solving BPTM* is still difficult because the continuous variables F_t are passed to both lower-level problems in each period. To alleviate this difficulty, we present a modified version of BPTM* where the factory capacity constraint (2b) of the MFG(F, Y) problem is enforced in the upper level:

$$[\text{BPTM}] \quad \max \sum_{n\in N}\sum_{t\in\mathcal{T}} \pi_{nt}(D_{nt} + B_{n,t-1} - B_{nt}) \tag{5a}$$

$$\text{subject to } Y_{pt} \leq \sum_{\tau \leq t} Z_{p\tau} \qquad p \in P,\ t \in \mathcal{T}, \tag{5b}$$

$$Y_{p,t-1} \leq Y_{pt} \qquad t \in \mathcal{T}\setminus\{1\}, \tag{5c}$$

$$F_t + \sum_{n\in N} X_{nt} = 1 \quad p \in P, t \in \mathcal{T}, \tag{5d}$$

$$(X, B, I) \in \arg\min \text{MFG}(Y), \tag{5e}$$

$$(Z, V) \in \arg\min \text{ENG}(F), \tag{5f}$$

where MFG(Y) is formulated as

$$\min\left\{\sum_{n\in N}\sum_{t\in\mathcal{T}} h_{nt} I_{nt} + r_{nt} X_{nt} C_t + b_{nt} B_{nt} : \texttt{MFG-primal-feasibility}(Y)\right\}.$$

Proposition 2. *Any optimal solution to BPTM, if it exists, is also optimal to BPTM*.*

Proof. Consider an optimal solution $(\mathcal{Y}^*, \mathcal{X}^*)$ to BPTM. This solution satisfies upper-level constraints (4b) and (4c) in BPTM*. By (5f), it is also optimal to ENG(F^*). The feasible region of MFG(Y^*) contains that of MFG(F^*, Y^*). As (X^*, B^*, I^*) is optimal to MFG(Y^*) and satisfies (2b), it is also optimal to MFG(F^*, Y^*). The fact that BPTM* and BPTM share the same upper-level objective function completes the proof.

In BPTM, the factory capacity F_t allocated to ENG in each period t is a continuous upper-level variable passed to the lower-level ENG problem, which is a mixed-integer model. In the bilevel programming literature, researchers have allowed only integer upper-level variables to appear in a mixed-integer lower-level problem in order to guarantee the existence of an optimal solution (Vicente et al. 1996, Köppe et al. 2010). In BPTM, however, the continuous F variables do not appear in the upper-level (CORP's) objective function; they only impact the lower-level ENG problem through the Z variables, whose values are constrained to be binary for any value of the F variables. Thus, an optimal solution is always attained when BPTM is feasible.

We propose a solution approach that generates upper and lower bounds in each iteration based on a single-level value function reformulation of BPTM. The main idea of this reformulation is to enforce the optimality of MFG(Y) and ENG(F) using constraints in the upper-level problem. Let ζ_{nt}, θ_{pt}, and ψ_t be dual variables associated with Constraints (2c), (2d), and (2e), respectively. The dual of the MFG(Y) problem is then given by

$$\max -\sum_{n\in N}\sum_{t\in\mathcal{T}}\zeta_{nt}D_{nt} + \sum_{p\in P}\sum_{t\in\mathcal{T}}\theta_{pt}Y_{pt} + \sum_{t\in\mathcal{T}}\psi_t \tag{6a}$$

$$\text{subject to } \zeta_{nt} - \zeta_{n,t+1} \le h_{nt} \qquad n\in N,\ t\in\mathcal{T}, \tag{6b}$$

$$-C_t\zeta_{pt} + \psi_t + \theta_{pt} \le r_{pt}C_t \qquad p\in P,\ t\in\mathcal{T}, \tag{6c}$$

$$-C_t\zeta_{nt} + \psi_t \le r_{nt}C_t \qquad n\in N\backslash P,\ t\in\mathcal{T}, \tag{6d}$$

$$-\zeta_{nt} + \zeta_{n,t+1} \le b_{nt} \qquad n\in N,\ t\in\mathcal{T}, \tag{6e}$$

$$\zeta_{nt}\in\mathbb{R}, \psi_t, \theta_{pt} \le 0 \qquad p\in P,\ t\in\mathcal{T}. \tag{6f}$$

We refer to Constraints (6b)–(6f) as `MFG-dual-feasibility`. Because MFG(Y) is a linear program, we can enforce its optimality through the primal feasibility, dual feasibility, and strong duality conditions (Labbé et al. 1998, Zeng and An 2014). The strong duality condition requires the equality of primal and dual objectives at optimality, and it can be stated as

$$\sum_{n\in N}\sum_{t\in\mathcal{T}} h_{nt}I_{nt} + r_{nt}X_{nt}C_t + b_{nt}B_{nt} = -\sum_{n\in N}\sum_{t\in\mathcal{T}}\zeta_{nt}D_{nt} + \sum_{p\in P}\sum_{t\in\mathcal{T}}\theta_{pt}Y_{pt} + \sum_{t\in\mathcal{T}}\psi_t. \tag{7}$$

We define an auxiliary variable ω_{pt} to reformulate the bilinear term $\theta_{pt}Y_{pt}$ through McCormick (1976) inequalities: $0\le\omega_{pt}\le MY_{pt}$, $\omega_{pt}\le\theta_{pt}$, $\omega_{pt}\ge\theta_{pt} - M(1-Y_{pt})$, where M is a large positive constant. We collectively refer to `MFG-primal-feasibility(Y)`, `MFG-dual-feasibility`, and the strong duality condition (7) as `MFG-optimality(Y)` and to Constraints (5b)–(5d) as `Upper-level-feasibility`. We can now restate BPTM as

$$[\text{EPTM}] \quad \max \sum_{n\in N}\sum_{t\in\mathcal{T}}\pi_{nt}(D_{nt} + B_{n,t-1} - B_{nt})$$
$$\text{subject to } \texttt{Upper-level-feasibility},\ \texttt{MFG-optimality}(Y),\ (Z,V)\in\arg\min \text{ENG}(F).$$

Although `MFG-optimality(Y)` is a set of linear constraints, the extended product transition model (EPTM) remains a bilevel program because of the constraint $(Z,V)\in\arg\min\text{ENG}(F)$. We now present two exact single-level reformulations of EPTM that enforce the optimality of ENG with a finite set of constraints and variables.

4.1. Single-Level Reformulations of the EPTM

Let $\mathcal{Z}$ denote the set of all possible distinct product development schedules for ENG, and let J denote the index set of $\mathcal{Z}$. The size of $\mathcal{Z}$ increases exponentially with the number of time periods T and the number of new products $|P|$, but it is finite and bounded above by $2^{T\times|P|}$. A given product development schedule $j\in J$ will have total weighted tardiness $\sum_{p\in P}w_p\hat{V}_p^j$ and factory capacity requirement $\hat{H}_t^j = \sum_{p\in P}H_{pt}\hat{Z}_{pt}^j$ in each time period $t\in\mathcal{T}$. We use this fact to enforce the optimality of ENG using a finite set of constraints and variables.

4.1.1. Exact Reformulation 1.

Without loss of generality, we assume that the factory capacity requirements for developing new products are positive integers (i.e., $H_{pt}\in\mathbb{Z}_+$). Therefore, the factory capacity requirement $\hat{H}_t^j$ for product development in period t under development plan j is also a positive integer. We define binary variable $\gamma_t^j, j\in J, t\in\mathcal{T}$, which takes the value of 0 if ENG has been allocated sufficient factory capacity for product development plan j in period t and 1 otherwise. Consider the following single-level reformulation of the EPTM:

$$[\text{MP}_1(J)] \quad \phi^* = \max\sum_{n\in N}\sum_{t\in\mathcal{T}}\pi_{nt}(D_{nt}+B_{n,t-1}-B_{nt})$$
$$\text{subject to } \texttt{Upper-level-feasibility},\ \texttt{MFG-optimality}(Y),\ \texttt{ENG-feasibility}(F),$$

$$C_tF_t \ge \hat{H}_t^j - \hat{H}_t^j\gamma_t^j \qquad t\in\mathcal{T}, j\in J, \tag{9a}$$

$$C_tF_t \le \hat{H}_t^j - \gamma_t^j + C_t(1-\gamma_t^j) \qquad t\in\mathcal{T}, j\in J, \tag{9b}$$

$$\sum_{p\in P}w_pV_P \le M\sum_{t\in\mathcal{T}}\gamma_t^j + \sum_{p\in P}w_p\hat{V}_P^j \qquad j\in J, \tag{9c}$$

$$\gamma_t^j\in\{0,1\} \qquad t\in\mathcal{T}, j\in J. \tag{9d}$$

The mixed-integer program MP$_1(J)$ determines the values of the ENG variables Z and V. Their feasibility to the ENG problem is enforced by the `ENG-feasibility`(F) constraints and their optimality by (9a)–(9d). Constraints (9a) and (9b) ensure that sufficient factory capacity is available to ENG to implement product development plan j in period t (i.e., $C_tF_t \geq \hat{H}_t^j$) if $\gamma_t^j = 0$. Otherwise, if $\gamma_t^j = 1$, (9a) is not binding, and (9b) ensures that the factory capacity allocated for product development in period t is strictly less than $\hat{H}_t^j$ (i.e., $C_tF_t \leq \hat{H}_t^j - 1$), and hence product development plan j cannot be implemented in period t because of insufficient factory capacity. Note that if product development plan j is not feasible, it might still be the case that $C_tF_t \geq \hat{H}_t^j$ for some periods (i.e., Constraints (9a) are enforced as $\gamma_t^j = 0$), and $C_{t'}F_{t'} \leq \hat{H}_{t'}^j - 1$ for at least one period t' (i.e., Constraints (9b) are enforced as $\gamma_{t'}^j = 1$).

Finally, Constraints (9c) ensure that if sufficient factory capacity is available in every period for product development plan j (i.e., $\sum_{t\in\mathcal{T}}\gamma_t^j = 0$), then the weighted tardiness of the new products (i.e., the ENG objective function) should not exceed that under product development plan j (i.e., $\sum_{p\in P} w_p V_P \leq \sum_{p\in P} w_p \hat{V}_P^j$). Note that if $\sum_{t\in\mathcal{T}}\gamma_t^j = 0$, product development plan j is feasible, but it may not be selected by ENG for implementation if there are other feasible product development plans with lower total tardiness. The largest possible value for the left-hand side of Constraints (9c) is attained when no new products are developed until the last time period, so we can set $M = \sum_{p\in P}(T - \theta_p)w_p$.

Proposition 3. *$MP_1(J)$ is equivalent to the EPTM.*

Proof. We show that there is a feasible solution to $\text{MP}_1(J)$ corresponding to any feasible solution to EPTM, and vice versa. Consider a bilevel feasible solution $(\bar{Y}, \bar{\mathcal{X}})$ to EPTM, and for every $j \in J$ and $t \in \mathcal{T}$, define

$$\bar{\gamma}_t^j = \begin{cases} 1 & \text{if } C_t\bar{F}_t < \hat{H}_t^j, \\ 0 & \text{if } C_t\bar{F}_t \geq \hat{H}_t^j. \end{cases}$$

Note that $(\bar{\mathcal{X}}, \bar{\gamma})$ is feasible to Constraints (9a) and (9b) by construction. It is also feasible to (9c) because $(\bar{Z}, \bar{V}) \in \arg\min \text{ENG}(\bar{F})$. Thus, $(\bar{Y}, \bar{\mathcal{X}}, \bar{\gamma})$ is a feasible solution to $\text{MP}_1(J)$. We now consider a feasible solution $(Y^*, \mathcal{X}^*, \gamma^*)$ to $\text{MP}_1(J)$. This solution satisfies `upper-level-feasibility`, `MFG-optimality`(Y^*), and `ENG-feasibility`(F^*). The index set of all feasible $\text{ENG}(F^*)$ solutions is given by $J^* = \{j \in J | \hat{H}_t^j \leq C_tF_t^*, t \in \mathcal{T}\}$. Note that $\gamma_t^{*j} = 0$ for every $j \in J^*$ and $t \in \mathcal{T}$ because of Constraints (9a) and (9b). Thus, $\sum_{t\in\mathcal{T}}\gamma_t^{*j} = 0$ and Constraints (9c) ensure that

$$\sum_{p\in P} w_p V_P^* \leq \sum_{p\in P} w_p \hat{V}_P^j, \quad j \in J^*.$$

As a result, (Z^*, V^*) is optimal to the $\text{ENG}(F^*)$ problem, and $(Y^*, \mathcal{X}^*)$ is feasible to EPTM. The fact that $\text{MP}_1(J)$ and EPTM have the same objective completes the proof.

4.1.2. Exact Reformulation 2.

To obtain an alternative single-level reformulation of EPTM, we define $\mathcal{B}_t(j) = \{k \in J | \hat{H}_t^k \leq \hat{H}_t^j\}$ as the set of all product development plans k whose factory capacity requirement does not exceed that of development plan j in time period t. Thus, if the factory capacity allocated to product development in period t is not sufficient to execute plan k (i.e., strictly less than $\hat{H}_t^k$), it will not be sufficient for plan j either. We define an auxiliary binary variable $\upsilon_t^j \in \{0,1\}$. If $\upsilon_t^j = 1$, the factory capacity allocation for product development in period t (i.e., C_tF_t) is enforced to be less than $\hat{H}_t^j$. If $\upsilon_t^j = 0$, on the other hand, no requirement is imposed on C_tF_t. The main difference between the υ_t^j and γ_t^j variables is that if $\gamma_t^j = 0$, sufficient factory capacity is allocated to execute plan j in period t, whereas if $\upsilon_t^j = 0$, the product development plan j is feasible in period t if $\sum_{k\in\mathcal{B}_t(j)}\upsilon_t^k = 0$ and infeasible if $\sum_{k\in\mathcal{B}_t(j)}\upsilon_t^k = 1$. We can now reformulate EPTM as follows:

$$[\text{MP}_2(J)] \qquad \phi^* = \max \sum_{n\in N}\sum_{t\in\mathcal{T}} \pi_{nt}(D_{nt} + B_{n,t-1} - B_{nt})$$

subject to `Upper-level-feasibility`, `MFG-optimality`(Y), `ENG-feasibility`(F),

$$C_tF_t \leq C_t - \sum_{j\in J}(C_t + 1 - \hat{H}_t^j)\upsilon_{tj} \qquad t \in \mathcal{T}, \qquad (10a)$$

$$\sum_{p\in P} w_p V_P \leq M \sum_{t\in\mathcal{T}}\sum_{k\in\mathcal{B}_t(j)} \upsilon_t^k + \sum_{p\in P} w_p \hat{V}_P^j \qquad j \in J, \qquad (10b)$$

$$\upsilon_t^j \in \{0,1\} \qquad t \in \mathcal{T}, j \in J. \qquad (10c)$$

Constraints (10a)–(10c) enforce the optimality of the $\text{ENG}(F)$ problem. In particular, Constraints (10a) state that υ_t^j can be equal to 1 for at most one $j \in J$ such that $C_tF_t \leq \hat{H}_t^j - 1$ in each time period t. Constraints (10b) ensure that if there is no infeasible development plan in $\mathcal{B}_t(j)$ in any time period, the weighted tardiness of the ENG problem cannot exceed that of product development plan j. Lozano and Smith (2017) presented a similar single-level reformulation for general mixed-integer bilevel programs. As in the first reformulation, we set $M = \sum_{p\in P}(T - \theta_p)w_p$ in Constraints (10b).

Proposition 4. *If $MP_2(J)$ is feasible, there exists an optimal solution $(Y^*, \mathcal{X}^*, \upsilon^*)$ to $MP_2(J)$ such that in each period $t \in \mathcal{T}$, $\upsilon_t^{*j_t} = 1$ for $j_t \in \text{argmin}\{\hat{H}_t^k | C_tF_t^* \leq \hat{H}_t^k - 1, k \in J\}$ and $\upsilon_t^{*j} = 0$ for all $j \in J \setminus \{j_t\}$.*

Proof. Let $(Y^*, \mathcal{X}^*, v^*)$ be an optimal solution to $\text{MP}_2(J)$. Constraints (10a) ensure that $v_t^{*j} = 1$ for at most one $j \in \hat{J}_t = \{k | C_t F_t^* \leq \hat{H}_t^k - 1, k \in J\}$ and $v_t^{*j} = 0$ for $j \in J \setminus \hat{J}_t$ in each $t \in \mathcal{T}$.

Case 1. There exists a $t \in \mathcal{T}$ such that $v_t^{*j} = 0$ for all $j \in \hat{J}_t$. In this case, we can construct $(Y^*, \mathcal{X}^*, \bar{v})$ such that $\bar{v} = v^*$ except $\bar{v}_t^{j_t} = 1$. Then, $\bar{v}_t$ satisfies (10a) because $C_t F_t^* \leq \hat{H}^{j_t} - 1$. This solution also satisfies Constraints (10b) because $\bar{v}_t^j \geq v_t^{*j}$ for all $j \in J$ and $t \in \mathcal{T}$. As a result, $(Y^*, \mathcal{X}^*, \bar{v})$ is a feasible solution.

Case 2. There exists a $t \in \mathcal{T}$ such that $v_t^{*\ell} = 1$ for $\ell \in \hat{J}_t$, but $\ell \neq j_t$. In this case, we can construct $(Y^*, \mathcal{X}^*, \bar{v})$ such that $\bar{v} = v^*$ except $\bar{v}_t^{j_t} = 1$ and $\bar{v}_t^{\ell} = 0$. Then, $\bar{v}_t$ satisfies (10a) because $C_t F_t^* \leq \hat{H}^{j_t} - 1$. Note that $\hat{H}_t^{j_t} \leq \hat{H}_t^{\ell}$ by definition of j_t. This solution also satisfies Constraints (10b) because for any $j \in J$ such that $\hat{H}_t^{j_t} \leq \hat{H}_t^j < \hat{H}_t^{\ell}$, we have $\sum_{k \in \mathcal{B}_t(j)} \bar{v}_t^k = 1 \geq \sum_{k \in \mathcal{B}_t(j)} v_t^{*k} = 0$. As a result, $(Y^*, \mathcal{X}^*, \bar{v})$ is a feasible solution.

In either case, the constructed feasible solution has the same objective function and thus is also optimal, completing the proof.

Proposition 5. *$MP_2(J)$ is equivalent to EPTM.*

Proof. We show that there is a feasible solution to $\text{MP}_2(J)$ corresponding to any feasible solution to EPTM, and vice versa. Consider a bilevel feasible solution $(\bar{Y}, \bar{\mathcal{X}})$ to EPTM. Based on Proposition 4, a solution $j_t \in J$ such that $j_t \in \arg\min\{\hat{H}_t^k | C_t \bar{F}_t \leq \hat{H}_t^k - 1, k \in J\}$ for every $t \in \mathcal{T}$ can be identified. Now set $\bar{v}_t^{j_t} = 1$ and $\bar{v}_t^j = 0$ for all $j \in J \setminus \{j_t\}$. By construction, $(\bar{Y}, \bar{\mathcal{X}}, \bar{v})$ is feasible to Constraints (10a). Suppose, by contradiction, that $(\bar{Y}, \bar{\mathcal{X}}, \bar{v})$ violates at least one of the Constraints (10b). Then there is a solution $j' \in J$ such that $\sum_{p \in P} w_p \bar{V}_P > \sum_{p \in P} w_p \hat{V}_P^{j'}$ and $\bar{v}_t^k = 0$ for every $t \in \mathcal{T}$ and $k \in \mathcal{B}_t(j')$. This means that $(\hat{Z}^{j'}, \hat{V}^{j'})$ is a feasible solution to $\text{ENG}(\bar{F})$, and its corresponding objective value is less than $\sum_{p \in P} w_p \bar{V}_P$. This contradicts the fact that $(\bar{Z}, \bar{V}) \in \arg\min \text{ENG}(\bar{F})$. Thus, $(\bar{Y}, \bar{\mathcal{X}}, \bar{v})$ is a feasible solution to $\text{MP}_2(J)$.

Now consider a feasible solution $(Y^*, \mathcal{X}^*, v^*)$ to $\text{MP}_2(J)$. This solution satisfies `Upper-level-feasibility`, `MFG-optimality`(Y^*), and `ENG-feasibility`(F^*). The index set of all feasible $\text{ENG}(F^*)$ solutions is given by $J^* = \{j \in J | C_t F_t^* \geq \hat{H}_t^j, t \in \mathcal{T}\}$. Note that $v_t^{*k} = 0$ for every $k \in \mathcal{B}_t(j)$, $j \in J^*$ and $t \in \mathcal{T}$, because otherwise, $C_t F_t^* \leq \hat{H}_t^k - 1$ as a result of Constraints (10a), contradicting $C_t F_t^* \geq \hat{H}_t^j$, as $\hat{H}_t^j \geq \hat{H}_t^k$ by definition of $\mathcal{B}_t(j)$. Thus, $\sum_{t \in \mathcal{T}} \sum_{k \in \mathcal{B}_t(j)} v_t^{*k} = 0$ for $j \in J^*$, and Constraints (10b) ensure that

$$\sum_{p \in P} w_p V_P^* \leq \sum_{p \in P} w_p \hat{V}_P^j, \quad j \in J^*.$$

As a result, (Z^*, V^*) is optimal to the $\text{ENG}(F^*)$ problem, and $(Y^*, \mathcal{X}^*)$ is feasible to EPTM. The fact that $\text{MP}_2(J)$ and EPTM have the same objective completes the proof.

4.2. Constraint and Column Generation Algorithm

The number of product development schedules in set J grows exponentially with the number of new products $|P|$ and the number of time periods T, rendering the solution time of $\text{MP}_1(J)$ or $\text{MP}_2(J)$ using off-the-shelf mixed-integer programming solvers prohibitively time consuming even for moderate-sized instances. We thus propose a constraint and column generation (CCG) method where Constraints (9a)–(9d) (in the first reformulation, $MP_1(J)$) or Constraints (10a)–(10c) (in the second reformulation, $MP_2(J)$) are initially relaxed and then enforced through an iterative process. Because the main steps of the CCG method are similar for $MP_1(J)$ and $MP_2(J)$, we refer to both reformulations simply as MP(J) in the rest of this section. A restricted master problem MP(J^k) is solved in each iteration k where $J^k \subseteq J$ and $J^0 = \emptyset$. The optimal objective function value ϕ^k of MP(J^k) is an upper bound on the optimal objective value ϕ^* of the EPTM (i.e., $\phi^k \geq \phi^*$) because the feasible region of MP(J^k) contains that of MP(J).

Let $(Y^k, \mathcal{X}^k)$ be an optimal solution of MP(J^k) at iteration k. We first solve $\text{ENG}(F^k)$ to determine $\phi_E(F^k)$, the optimal objective value of ENG under the current factory capacity allocation F^k. We then solve the following subproblem to check the feasibility of ENG's product development schedule for the product release schedule Y^k:

$$[\text{ENG}'(F^k, Y^k)] \qquad \phi'_E(F^k, Y^k) = \min \sum_{p \in P} w_p V_p$$

subject to `ENG-feasibility`(F^k),

$$\sum_{\tau \leq t} Z_{p\tau} \geq Y_{pt}^k \quad p \in P, t \in \mathcal{T}. \tag{11}$$

The only difference between subproblems $\text{ENG}(F^k)$ and $\text{ENG}'(F^k, Y^k)$ is Constraint (11), which forces ENG to meet the product release schedule Y^k. Let (Z', V') be the optimal solution to $\text{ENG}'(F^k, Y^k)$, and let $\phi'_E(F^k, Y^k)$ be the optimal objective value. If $\phi'_E(F^k, Y^k) = \phi_E(F^k)$, then Z' is an optimal product development schedule for ENG in response to the factory capacity allocation F^k, and it is feasible to the product release schedule Y^k. Thus, $(Y^k, \mathcal{X}') = (Y^k, F^k, X^k, B^k, I^k, Z', V')$ is bilevel feasible. Note that $(Y^k, \mathcal{X}')$ is also bilevel optimal because the feasible region of the restricted master problem MP(J^k) contains the feasible region of MP(J). If, on the other hand, $\phi'_E(F^k, Y^k) > \phi_E(F^k)$, we generate new columns and constraints to add to MP(J^k). Algorithm 1 presents the CCG algorithm.

Algorithm 1 (CCG Algorithm)

1 **Initialization:** set $k = 0$, $J^k = \emptyset$
2 **while** $J^k \subseteq J$ **do**
3 Solve MP(J^k).
4 **if** MP(J^k) *is infeasible or* $J^k = J$ **then**
5 MP(J) is infeasible, STOP.
6 **end**
7 Let $(Y^k, \mathcal{X}^k) = (Y^k, F^k, X^k, B^k, I^k, Z^k, V^k)$ be the optimal solution to MP(J^k).
8 Solve ENG(F^k) to get the optimal solution Z^* and the objective value $\phi_E(F^k)$.
9 **if** $\sum_{p\in P} w_p V_p^k \le \phi_E(F^k)$ **then**
10 Return the optimal bilevel solution $(Y^k, \mathcal{X}^k)$.
11 **end**
12 Solve ENG′(F^k, Y^k) to get the optimal solution (Z', V') and the objective value $\phi'_E(F^k, Y^k)$.
13 **if** $\phi_E(F^k) = \phi'_E(F^k, Y^k)$ **then**
14 Return the bilevel optimal solution, $(Y^k, \mathcal{X}') = (Y^k, F^k, X^k, B^k, I^k, Z', V')$.
15 **end**
16 Add Constraints (9a)–(9d) to $\text{MP}_1(J^k)$ (or Constraints (10a)–(10c) to $\text{MP}_2(J^k)$).
17 Set $J^k = J^k \cup \{Z^*\}$, $k = k + 1$.
18 **end**

Proposition 6. *Algorithm 1 terminates in a finite number of iterations.*

Proof. A solution to ENG is generated in line 1. If the same solution has been generated in previous iterations, the stopping condition in line 9 will be satisfied. Thus, in each iteration, either a new product development schedule is added to J^k or the algorithm terminates. The result follows because the set of distinct feasible product development schedules J is finite.

Proposition 7. *Algorithm 1 either returns the optimal solution to* MP(J) *or identifies infeasibility.*

Proof. In any iteration k, the feasible region of MP(J^k) contains the feasible region of MP(J). Therefore, if MP(J^k) is infeasible, MP(J) is also infeasible, and the algorithm will terminate in line 5. If the condition in line 9 is satisfied, the current MP(J^k) solution satisfies the optimality conditions for ENG and is thus bilevel optimal. If the condition in line 13 is satisfied, (Z', V') is optimal to ENG(F^k) and feasible to the product release schedule Y^k. Thus, $(Y^k, \mathcal{X}') = (Y^k, F^k, X^k, B^k, I^k, Z', V')$ is a bilevel feasible solution. It is also bilevel optimal because MP(J^k) is a relaxation of MP(J) whose objective does not include variables of the ENG. If none of the three stopping conditions within the while loop is satisfied, the algorithm will stop when $J^k = J$—that is, when all ENG solutions have been generated. At that point, it can be concluded that there is no bilevel feasible solution to MP(J), completing the proof.

The CCG algorithm solves the restricted master problem MP(J^k) in each iteration. Our computational experiments showed that MP(J^k) is a difficult mixed-integer program, and a high quality initial feasible solution can speed up its solution considerably. We develop a greedy heuristic to generate a feasible solution to MP(J^k) in Appendix A of the online supplement.

4.3. Lower Bound for BPTM

We propose a heuristic to generate a bilevel feasible solution to BPTM that will provide a lower bound. Consider the following mixed-integer program, which we refer to as the extended ENG (EENG) problem:

$$[\text{EENG}] \qquad \eta^* = \min \sum_{p\in P} w_p V_p,$$
$$\text{subject to } \texttt{Upper-level-feasibility}, \quad \texttt{MFG-optimality}(Y), \quad \texttt{ENG-feasibility}(F).$$

Let η^* denote the optimal objective value of the EENG, and let F^* denote the factory capacity allocation to ENG in the corresponding optimal solution. Note that $\eta^* \ge \phi_E(F^*)$ because each (Z^*, V^*) solution with $F = F^*$ in the feasible solution set of EENG is feasible to ENG(F^*), and these two problems have the same objective function. Therefore, if $\eta^* = \phi_E(F^*)$, an optimal solution to EENG is a bilevel feasible solution to BPTM, because it simultaneously satisfies the optimality conditions for both MFG and ENG.

Our preliminary computational experiments, however, showed that off-the-shelf solvers struggle to solve even moderate-size instances of EENG in a reasonable time. Therefore, we have developed a warm-start method to generate an initial feasible solution to EENG. In particular, we solve a multiple knapsack problem (MKP) to generate product development schedules yielding solutions for the Z variables. Each time period is treated as a knapsack with a certain amount of factory capacity allocated to product development. New products under development represent items to be packed into these knapsacks.

The "weight" of each new product (item) $p \in P$ is set to the factory capacity H_{pt} required for its development. Recall that δ_p denotes the due date for developing product $p \in P$. In the MKP we define the "value" of developing product p in time period $t \in \mathcal{T}$ as

$$S_{pt} = \begin{cases} -\infty & \text{if } t < T^0, \\ T - (\delta_p - t) & \text{if } t \le \delta_p, \\ (\delta_p - t) w_p & \text{if } t > \delta_p, \end{cases}$$

where T^0 is the earliest time period a new product can be developed. The definition of S_{pt} favors developing new products close to their due dates. This property allows the warm-start method to generate product development schedules that are not too focused on minimizing the weighted tardiness objective of the ENG.

We define a coefficient f to control the factory capacity allocated to product development in each time period $t \in \mathcal{T}$ as a fraction of the total demand for the current products (i.e., $\sum_{n \in N \setminus P} D_{nt}$). This is equivalent to specifying a minimum fill rate for current demand that MFG must meet. The formulation of the MKP for a given f is as follows:

$$\text{MKP}(f) \quad \max \sum_{p \in P} \sum_{t \in \mathcal{T}} S_{pt} Z_{pt} \tag{13a}$$

$$\sum_{p \in P} H_{pt} Z_{pt} \le C_t - f \sum_{n \in N \setminus P} D_{nt} \quad t \in \mathcal{T}, \tag{13b}$$

$$Z_{pt} \in \{0,1\} \quad p \in P, t \in \mathcal{T}. \tag{13c}$$

We solve MKP(f) with $f \in \{1.0, 1.2, 1.4\}$ in the numerical experiments. The resulting product development schedules (i.e., Z variables) are fixed in EENG to generate feasible solutions. These solutions are then provided to the solver as initial feasible solutions when optimizing the EENG. We henceforth refer to this algorithm as the lower bounding algorithm (LBA). Note that if EENG is infeasible, then BPTM is also infeasible because the feasible region of BPTM is a subset of the feasible region of EENG. Thus the LBA can identify the infeasibility of a BPTM instance.

5. Numerical Experiments

We run computational experiments to evaluate the performance of the proposed solution algorithm and model formulations. We then present an extension of our model with multiple ENG units and analyze scalability of the proposed solution method. Finally, we provide managerial insights about the impact of product mix, the cost of decentralization, and the value of CORP's leadership.

5.1. Instance Generation

Our test instances represent a realistic demand pattern observed during product transition periods in the semiconductor industry: demand for current products diminishes to 0 over time, whereas the demand for new products rises over time and then stabilizes. Specifically, in our test instances, the current products have positive demand in the first 60% of the planning horizon, and new products have positive demand in the last 60%. Thus there is a time interval with positive demand for both current and new products. The demand for current products diminishes in the last quarter of their lifetime, whereas demand for new products increases in the first quarter of their lifetime. Because semiconductor companies operate in different markets such as mobile devices, memory, and servers (Rash and Kempf 2012), we consider four discrete uniform distributions to model demand. Each product is randomly assigned to a market.

The due date for the development of each new product is selected randomly with equal probability from the periods with positive demand for that product. The factory capacity C_t in period t is randomly generated from a discrete uniform distribution between $\lceil 0.7 \sum_{n \in N} D_{nt} \rceil$ and $\lceil 1.2 \sum_{n \in N} D_{nt} \rceil$. Once the factory capacity is determined, the factory capacity H_{pt} required for the development of new product p in period t is generated randomly from a discrete uniform distribution between $\lceil 0.2 C_t \rceil$ and $\lceil 0.6 C_t \rceil$. The values of other model parameters are given in Table 3.

5.2. Computational Performance

We analyze the performance of our algorithm using 24 problem classes with number of periods $T \in \{12, 18, 24\}$, number of products $|N| \in \{12, 14\}$, and number of new products $|P| \in \{4, 5, 6, 7\}$. We report the size of each problem class in Appendix B of the online supplement. We generate six instances per problem class for a total of $24 \times 6 = 144$ instances. The LBA identified four infeasible instances that were replaced with new instances. We set the time limit to 7,200 seconds for the overall solution method, 3,600 seconds for solving the master problem in each iteration of the CCG algorithm, and 300 seconds for the LBA in the first iteration. We do not stop an iteration if it is in progress when the time limit is exceeded. We run experiments on a computer with two Intel Xeon 2.99 GHz processors and 32 GB RAM. We implement our algorithms in C++ and solve the mixed-integer programs using CPLEX 12.9. All the code developed for this paper is available in a GitHub repository found at https://github.com/RahmanKhorramfar91/IJOC2022-Managing-Product-Transitions.

Table 4 reports the performance of the solution method for each reformulation. The `t-max(min)` column reports the maximum (minimum) solution time in seconds, and the `t-avg` column is the average solution time for optimally solved instances. The `iter` column reports the maximum (minimum) number of iterations for optimally solved instances, and the `slv` column reports the number of instances (out of six instances in each

Table 3. Parameter Values

Parameter	Value	Parameter	Value
Demand in Market 1	U[600, 1000]	h (holding cost)	0.5
Demand in Market 2	U[800, 1200]	r (production cost)	$2h$
Demand in Market 3	U[1200, 1600]	b (backorder cost)	$10h$
Demand in Market 4	U[1400, 1800]	$w \in \mathbb{Z}^P$ (tardiness cost)	U[5, 200]
T^0	$0.4T$	$\pi \in \mathbb{Z}^N$ (revenue)	U[25, 30]

Table 4. The Performance of the Solution Algorithm with Exact Reformulations 1 and 2

	Exact reformulation 1				Exact reformulation 2				
Class	`t-max(min)`	`t-avg`	`iter`	`slv`	`t-max(min)`	`t-avg`	`iter`	`slv`	`LBA`
C1	9 (1)	5	4 (1)	6	7 (1)	3	4 (1)	6	6
C2	7 (2)	5	2 (1)	6	95 (3)	20	76 (1)	6	6
C3	120 (5)	35	8 (1)	6	167 (9)	70	10 (1)	6	6
C4	4,401 (14)	910	10 (1)	6	1,082 (13)	297	3 (1)	6	6
C5	7 (1)	4	7 (1)	6	5 (1)	3	4 (1)	6	6
C6	12 (2)	6	5 (1)	6	21 (2)	7	9 (1)	6	6
C7	84 (27)	42	14 (1)	6	133 (12)	57	10 (1)	6	6
C8	633 (66)	307	14 (2)	6	775 (35)	305	9 (2)	6	6
C9	13 (3)	8	3 (1)	6	14 (4)	9	4 (1)	6	6
C10	284 (14)	119	5 (1)	6	296 (7)	109	20 (1)	6	6
C11	2,489 (464)	1,384	17 (1)	6	2,907 (454)	1,044	16 (1)	6	0
C12	8,545 (257)	3,452	5 (1)	5	3,322 (825)	1,818	4 (1)	5	1
C13	28 (4)	12	4 (1)	6	25 (4)	11	3 (1)	6	6
C14	315 (4)	85	3 (1)	6	312 (4)	86	3 (1)	6	5
C15	984 (11)	354	5 (1)	6	791 (10)	314	3 (1)	6	2
C16	7,236 (491)	3,220	8 (1)	6	3,433 (216)	1,790	5 (1)	6	4
C17	19 (3)	8	2 (1)	6	39 (3)	17	3 (1)	6	6
C18	403 (11)	83	3 (1)	6	443 (13)	113	3 (1)	6	5
C19	4,031 (190)	1,586	5 (1)	6	3,413 (183)	1,946	5 (1)	6	4
C20	—	—	—	0	3,191 (3,191)	3,191	1 (1)	1	3
C21	40 (6)	18	1 (1)	6	135 (7)	35	3 (1)	6	6
C22	312 (26)	144	10 (1)	6	312 (19)	103	9 (1)	6	5
C23	1,736 (96)	646	18 (1)	6	3,027 (95)	1,039	26 (1)	6	4
C24	—	—	—	0	—	—	—	0	2
Avg.	1,535 (152)	692			1,095 (267)	588			

problem class) that are solved optimally within the time limit. The `LBA` column gives the number of instances for which the LBA was able to generate a lower bound within the time limit. The last row of the table reports the average time over all instances.

Neither of the exact reformulations has consistently better computational performance than the other. The solution times for some instances are significantly shorter with exact reformulation 2 (e.g., the C4 and C16 problem classes) and those for others with exact reformulation 1 (e.g., the C23 problem class). It is interesting that some instances are solved optimally even when the LBA fails to generate an initial feasible solution—for example, all six instances in C11 are solved optimally although the LBA failed to generate an initial feasible solution for any of them. The warm-start strategy reduces the solution times significantly for both reformulations and allows larger instances to be solved. We also tried several acceleration strategies proposed in Lozano and Smith (2017), but none of them had substantial impact on the performance.

The number of new products has the greatest impact on the solution time. Solving the bilevel model becomes more difficult as the number of new products increases because MFG and ENG have to be coordinated on the development of more products. Although the problem size increases with number of periods T, a longer planning period provides scheduling flexibility for product development. However, in our experiments the impact of larger instance sizes is more significant than that of product development flexibility. In Table 4, the solution time increases with the number of periods to the point that neither formulation is able to solve instances with $|P| = 7$ new products and $T = 24$ periods.

5.3. Multiple Engineering Units

Because the proposed bilevel model is motivated by the need to coordinate the activities of a leader and two followers, a natural question that arises is how the procedure can be scaled to more than two followers. The semiconductor firm motivating this work has a single, global manufacturing organization; therefore we conduct experiments by varying the number of ENG units. Our formulation and the solution approach can be extended to consider multiple ENG units in a straightforward manner, as presented in Appendix C of the online supplement. We disable the LBA as well as the greedy heuristic presented in Appendix A of the online supplement because these procedures are developed for the case with one ENG unit. We also increase the time limit to six hours to get a better sense of how well the proposed solution approach scales to multiple ENG units.

We consider two problem classes in Table 5 with six and eight new products. We vary the number of ENG units and solve six randomly generated instances of each problem. As seen in Table 5, the solution time increases with the number of ENG units in each

Table 5. The Performance of the Solution Method with Varying Numbers of ENG Units

	$\|P\|$	#ENGs	t-max(min)	t-avg	iter	slv
$\|P\| = 6$	6	1	60 (1)	31	16 (1)	6
	3	2	187 (2)	48	11 (2)	6
	2	3	241 (1)	83	11 (1)	6
	1	6	11,016 (1)	2,783	21 (1)	4
$\|P\| = 8$	8	1	2,639 (1)	1,682	8 (1)	5
	4	2	5,425 (443)	2,521	8 (1)	6
	2	4	8,371 (304)	2,534	5 (1)	4
	1	8	1,017 (1,017)	1,017	1 (1)	1

Notes. The number of periods is $T = 12$, and the number of products is $|N| = 14$ for each instance. The number of new products per ENG unit is denoted by $|P|$. The time limit is set to six hours.

problem class. This increase is quite significant when considering more than four ENG units. However, the number of new products has more impact on the scalability of the solution method than the number of ENG units when there are fewer than four ENG units. For example, increasing the number of new products per ENG from three to four when there are two ENG units increases the average solution time from 48 seconds to 2,521 seconds; increasing the number of ENG units from two to three when the total number of new products is fixed at six increases the average solution time less significantly, from 48 seconds to 83 seconds.

These results suggest that our procedure scales well with the number of ENG units up to a certain limit when the total number of new products is fixed, and it is this latter quantity that drives the computational effort. If there is a need to increase the number of new products, then assigning more products to existing ENG units rather than forming new ENG units would be a better strategy to facilitate the allocation of factory capacity.

5.4. Impact of Product Mix

In this section we explore the impact of product mix (i.e., the number of current and new products in set N) on the optimal objective values of the CORP, MFG, and ENG problems. We consider instances with $T = 18$ periods and $|N| = 14$ products. We analyze the change in the optimal objective value of each decision entity by increasing the number of new products $|P|$ from 4 to 7 while keeping the total number of products $|N| = 14$. We randomly generate 6 instances with $|P| = 4, 5, 6$ and 7, for a total of 24 instances.

Table 6 shows that CORP's revenue decreases, MFG's cost increases, and ENG's tardiness decreases with the number of new products. Developing more new products during a fixed planning horizon requires that ENG be allocated more factory capacity. This, in turn, results in more backorder cost and less revenue, as demand cannot be fully satisfied by the MFG unit. Although tardiness may also increase when there are more new products to be developed, the ENG unit manages to use the increased factory capacity allocation to improve its objective in all instances in Table 6.

5.5. Cost of Decentralization

In this section we evaluate the impact of bilevel decision hierarchy on the objective function performance. Although a single-level model of the considered product transition problem is not realistic because of the decentralized organizational structure across the CORP, MFG, and ENG units, we consider a hypothetical case where all manufacturing and product design decisions are made by CORP as a central planner. The formulation of this single-level problem, referred to as the integrated model (IM), is given by

$$\text{(IM)} \quad \max \sum_{n\in N}\sum_{t\in T}\pi_{nt}(D_{nt}+B_{n,t-1}-B_{nt}) - \left(\sum_{n\in N}\sum_{t\in T}h_{nt}I_{nt}+r_{nt}X_{nt}C_t+b_{nt}B_{nt}\right) \tag{14a}$$

$$\text{subject to } Y_{pt} \le \sum_{\tau\le t} Z_{p\tau} \qquad p\in P,\ t\in\mathcal{T}, \tag{14b}$$

$$F_t + \sum_{n\in N} X_{nt} = 1 \qquad t\in\mathcal{T}, \tag{14c}$$

$$\texttt{MFG-primal-feasibility}(Y),$$
$$\texttt{ENG-feasibility}(F).$$

The objective function (14a) maximizes the profit given by the difference between revenue (i.e., CORP objective) and manufacturing costs (i.e., MFG objective). The constraints include all of the upper- and lower-level constraints in BPTM-Nash. We solve this single-level mixed-integer model using CPLEX 12.9. Table 7 shows the change in the optimal objective function values of CORP, MFG, and ENG, as well as the change in profit compared with bilevel model BPTM-Nash for six instances in problem class C3.

As seen in the last row of Table 7, the IM results in higher profit in all test instances. As expected, the decentralized nature of the problem leads to loss of profit. However, in a firm consisting of multiple autonomous functional units, as commonly seen in today's global high-technology firms, centralized decision models are hard to implement because of their incompatibility with the organizational structure,

Table 6. Average Change in the Optimal Objective Values of CORP, MFG, and ENG as the Number of New Products $|P|$ Increases

	$\|P\| = 4$ to 5	$\|P\| = 5$ to 6	$\|P\| = 6$ to 7
CORP (max revenue) (%)	−2.7	−3.4	−3.6
ENG (min tardiness) (%)	−52.8	−37.5	−11.7
MFG (min cost) (%)	6.6	8.5	12.7
Backorder cost	59,838	77,570	120,500
Inventory cost	−2,855	3,391	5,510
Production cost	−2,895	−5,538	−3,500

Table 7. Percent Change in the Optimal Objective Values of the CORP, MFG, ENG, and Profit in the Integrated Model Compared with BPTM-Nash (Reported Instances Are in Problem Class C3)

	Instance 1	Instance 2	Instance 3	Instance 4	Instance 5	Instance 6
CORP (max revenue) (%)	1.0	1.5	1.4	5.1	4.1	4.3
MFG (min cost) (%)	−26.9	−11.2	−13.9	−5.3	−8.4	−16.5
ENG (min tardiness) (%)	65.0	0.0	70.4	−21.8	237.6	0.5
Profit[a] (%)	7.9	4.9	5.4	9.8	9.4	12.1

[a]Profit = CORP objective − MFG objective.

management incentives, and the high cost of collecting and maintaining data for a large, centralized model (Bansal et al. 2020).

Table 7 displays a consistent pattern of relatively small increases in revenue (the upper-level objective in the bilevel problem) and large reductions in MFG cost. Clearly, the decisions made by CORP based on revenue allow MFG limited scope to reduce its cost while meeting demand in the bilevel model. Increases in the ENG objective (tardiness) suggest that the integrated model prefers to exploit some products by meeting large portions of their demand while delaying the introductions of others. Instance 4 is of particular interest because all three units are better off with the integrated solution, illustrating how the bilevel approach can lead to a Pareto-dominated solution.

5.6. The Value of CORP's Leadership

The aim of this section is to assess the value of CORP's leadership. To this end, we consider an alternative bilevel model in which MFG manages the factory capacity allocation acting as the leader. ENG in the lower level is the only follower. The formulation of this model, referred to as MFG-Leader, is given by

$$\text{(MFG-Leader)} \qquad \min \sum_{n \in N} \sum_{t \in T} h_{nt} I_{nt} + r_{nt} X_{nt} C_t + b_{nt} B_{nt} \tag{15a}$$

$$\text{subject to } Y_{pt} \le \sum_{\tau \le t} Z_{p\tau} \qquad p \in P,\ t \in \mathcal{T}, \tag{15b}$$

$$F_t + \sum_{n \in N} X_{nt} = 1 \qquad t \in \mathcal{T}, \tag{15c}$$

$$\texttt{MFG-primal-feasibility}(Y),$$

$$(Z, V) \in \arg\min \text{ENG}(F).$$

The last constraint in MFG-Leader formulates the decision problem of the ENG. Our proposed solution approach can be adapted to solve this model by modifying the restricted master problem. Specifically, we replace CORP's objective with the MFG objective and change the `MFG-optimality` (Y) constraints to `MFG-primal-feasibility`(Y) constraints in the restricted master problem MP(J^k).

Table 8 shows the difference in the optimal objective values of the CORP, MFG, and ENG in MFG-Leader compared with BPTM-Nash for six instances in problem class C3. As a leader, MFG can start producing new products as soon as it is ready without waiting for CORP's product release decision. Thus, the MFG objective function improves (i.e., manufacturing cost decreases) compared with BPTM-Nash in all instances. On the other hand, the impact on CORP and ENG objectives is mixed. CORP's revenue can decrease because CORP has no control over the decisions of MFG and ENG in MFG-Leader. However, it can also benefit from the new setting, as revenue depends on backorder levels that can improve under MFG's leadership. It is interesting to observe that all decision makers improve their objective values in instance 4 when MFG acts as the leader. Another interesting observation pertains to the "stability" of the bilevel solution obtained from BPTM-Nash with the leadership of CORP. If MFG and ENG realize that they can both improve their objective functions without the CORP's control, as in instances 3 and 4 in Table 8, they might collude to implement a solution that is not always in the best interest of CORP.

6. Conclusion

We formulate a mixed-integer bilevel model with interdependent followers for capacity coordination during product transitions in a firm with decentralized product development and manufacturing units. We propose two single-level reformulations and derive a solution algorithm based on constraint and column generation. We perform extensive computational experiments to

Table 8. Percent Change in the Optimal Objective Values of the CORP, MFG, and ENG in the MFG-Leader Model Compared with BPTM-Nash (Reported Instances Are in Problem Class C3)

	Instance 1	Instance 2	Instance 3	Instance 4	Instance 5	Instance 6
CORP (max revenue) (%)	−1.9	−0.5	−1.7	0.5	1.6	−1.4
MFG (min cost) (%)	−38.1	−14.8	−42.6	−8.6	−18	−37.4
ENG (min tardiness) (%)	105	59	−14.2	−17.6	198.9	0

examine the performance of the proposed solution approach and provide managerial insights.

To the best of our knowledge, this paper presents the first exact solution algorithm for mixed-integer bilevel programs with interdependent followers. The proposed bilevel model and solution approach are quite flexible and allow us to study the impact of different settings for the distribution of decision authority among the considered decision units. Therefore, this paper provides a useful tool to study decentralized, hierarchical decision problems in organizational design. Our computational experiments provide practical insights, but the reported numerical values should be interpreted with caution because our test instances are randomly generated. In practice, some problem parameters may be inherently related to each other. For example, tardiness weight associated with the development of a new product may depend on the demand, backordering cost, and unit revenue of that product. The decision makers should consider such relations when calibrating the proposed model.

A number of directions for future research emerge from this work. Despite the algorithmic enhancements we have implemented, such as the warm-start and lower bounding techniques, the computational burden of the solution approach remains substantial, requiring significant improvements to be able to address larger instances with multiple followers and more complex information flows among the followers. Exploring the relationship between the number of decision entities, the distribution of decision authority among them, and the number of feasible bilevel solutions is also of great practical interest. It may well be that under certain decision hierarchies there exist very few bilevel feasible solutions, but small adjustments to problem parameters such as capacity levels may make a larger number of such solutions available. Finally, exploring the relationship between iterative combinatorial auction-based approaches to product transitions and the bilevel approach proposed here is of both theoretical and practical interest.

References

Aksen D, Aras N (2012) A bilevel fixed charge location model for facilities under imminent attack. *Comput. Oper. Res.* 39(7):1364–1381.

Altman E, Wynter L (2004) Equilibrium, games, and pricing in transportation and telecommunication networks. *Networks Spatial Econom.* 4(1):7–21.

Arslan O, Jabali O, Laporte G (2018) Exact solution of the evasive flow capturing problem. *Oper. Res.* 66(6):1625–1640.

Audet C, Savard G, Zghal W (2007) New branch-and-cut algorithm for bilevel linear programming. *J. Optim. Theory Appl.* 134(2): 353–370.

Aussel D, Sagratella S (2017) Sufficient conditions to compute any solution of a quasivariational inequality via a variational inequality. *Math. Methods Oper. Res.* 85(1):3–18.

Bansal A, Uzsoy R, Kempf K (2020) Iterative combinatorial auctions for managing product transitions in semiconductor manufacturing. *IISE Trans.* 52(4):413–431.

Bard JF (2013) *Practical Bilevel Optimization: Algorithms and Applications, Nonconvex Optimization and Its Applications*, Vol. 30 (Springer Science & Business Media, Dordrecht, Netherlands).

Bard JF, Moore JT (1990) A branch and bound algorithm for the bilevel programming problem. *SIAM J. Sci. Statist. Comput.* 11(2):281–292.

Bard JF, Moore JT (1992) An algorithm for the discrete bilevel programming problem. *Naval Res. Logist.* 39(3):419–435.

Bertsimas D, Tsitsiklis JN (1997) *Introduction to Linear Optimization* (Athena Scientific, Belmont, MA).

Bhaskaran SR, Goel A, Ramachandran K (2015) Managing product transitions under technology uncertainty. Preprint, submitted October 29, https://ssrn.com/abstract=1775430.

Bilginer Ö, Erhun F (2010) Managing product introductions and transitions Cochran JJ, Cox LA Jr., Keskinocak P, Kharoufeh JP, Smith JC, eds. *Wiley Encyclopedia of Operations Research and Management Science*, online ed. (John Wiley & Sons, Oxford, UK), https://doi.org/10.1002/9780470400531.eorms0489.

Calvete HI, Galé C (2007) Linear bilevel multi-follower programming with independent followers. *J. Global Optim.* 39(3):409–417.

Calvete HI, Domínguez C, Galé C, Labbé M, Marin A (2019) The rank pricing problem: Models and branch-and-cut algorithms. *Comput. Oper. Res.* 105(May):12–31.

Campelo M, Dantas S, Scheimberg S (2000) A note on a penalty function approach for solving bilevel linear programs. *J. Global Optim.* 16(3):245–255.

Colson B, Marcotte P, Savard G (2007) An overview of bilevel optimization. *Ann. Oper. Res.* 153(1):235–256.

Dempe S (2002) *Foundations of Bilevel Programming* (Kluwer Academic Publishers, Dordrecht, Netherlands).

DeNegre ST, Ralphs TK (2009) A branch-and-cut algorithm for integer bilevel linear programs. Chinneck JW, Kristjansson B, Saltzman MJ, eds. *Operations Research and Cyber-Infrastructure* (Springer, New York), 65–78.

Dreves A, Facchinei F, Kanzow C, Sagratella S (2011) On the solution of the KKT conditions of generalized Nash equilibrium problems. *SIAM J. Optim.* 21(3):1082–1108.

Druehl CT, Schmidt GM, Souza GC (2009) The optimal pace of product updates. *Eur. J. Oper. Res.* 192(2):621–633.

Facchinei F, Kanzow C (2010) Generalized Nash equilibrium problems. *Ann. Oper. Res.* 175(1):177–211.

Facchinei F, Kanzow C, Sagratella S (2014) Solving quasi-variational inequalities via their KKT conditions. *Math. Programming* 144(1):369–412.

Ferrer G, Swaminathan JM (2006) Managing new and remanufactured products. *Management Sci.* 52(1):15–26.

Fischetti M, Ljubić I, Monaci M, Sinnl M (2017) A new general-purpose algorithm for mixed-integer bilevel linear programs. *Oper. Res.* 65(6):1615–1637.

Fischetti M, Ljubić I, Monaci M, Sinnl M (2018) On the use of intersection cuts for bilevel optimization. *Math. Programming* 172(1–2):77–103.

Hansen P, Jaumard B, Savard G (1992) New branch-and-bound rules for linear bilevel programming. *SIAM J. Sci. Statist. Comput.* 13(5):1194–1217.

Hemmati M, Smith JC (2016) A mixed-integer bilevel programming approach for a competitive prioritized set covering problem. *Discrete Optim.* 20(May):105–134.

Huppmann D, Siddiqui S (2018) An exact solution method for binary equilibrium problems with compensation and the power market uplift problem. *Eur. J. Oper. Res.* 266(2):622–638.

Karabuk S, Wu SD (2003) Coordinating strategic capacity planning in the semiconductor industry. *Oper. Res.* 51(6):839–849.

Karabuk S, Wu SD (2005) Incentive schemes for semiconductor capacity allocation: A game theoretic analysis. *Production Oper. Management* 14(2):175–188.

Klastorin T, Tsai W (2004) New product introduction: Timing, design, and pricing. *Manufacturing Service Oper. Management* 6(4):302–320.

Koca E, Souza GC, Druehl CT (2010) Managing product rollovers. *Decision Sci.* 41(2):403–423.

Köppe M, Queyranne M, Ryan CT (2010) Parametric integer programming algorithm for bilevel mixed integer programs. *J. Optim. Theory Appl.* 146(1):137–150.

Labbé M, Marcotte P, Savard G (1998) A bilevel model of taxation and its application to optimal highway pricing. Management Sci. 44(12, Part 1):1608–1622.

Lavigne D, Loulou R, Savard G (2000) Pure competition, regulated and Stackelberg equilibria: Application to the energy system of Quebec. *Eur. J. Oper. Res.* 125(1):1–17.

Le Cadre H, Jacquot P, Wan C, Alasseur C (2020) Peer-to-peer electricity market analysis: From variational to generalized Nash equilibrium. *Eur. J. Oper. Res.* 282(2):753–771.

Li H, Graves SC, Huh WT (2013) Optimal capacity conversion for product transitions under high service requirements. *Manufacturing Service Oper. Management* 16(1):46–60.

Liang C, Çakanyıldırım M, Sethi SP (2014) Analysis of product rollover strategies in the presence of strategic customers. *Management Sci.* 60(4):1033–1056.

Liao S, Seifert RW (2015) On the optimal frequency of multiple generation product introductions. *Eur. J. Oper. Res.* 245(3): 805–814.

Lim WS, Tang CS (2006) Optimal product rollover strategies. *Eur. J. Oper. Res.* 174(2):905–922.

Liu X, Kwon C (2020) Exact robust solutions for the combined facility location and network design problem in hazardous materials transportation. *IISE Trans.* 52(10):1156–1172.

Lobel I, Patel J, Vulcano G, Zhang J (2015) Optimizing product launches in the presence of strategic consumers. *Management Sci.* 62(6):1778–1799.

Lozano L, Smith JC (2017) A value-function-based exact approach for the bilevel mixed-integer programming problem. *Oper. Res.* 65(3):768–786.

McCormick GP (1976) Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. *Math. Programming* 10(1):147–175.

Nishi T, Hiranaka Y, Grossmann IE (2011) A bilevel decomposition algorithm for simultaneous production scheduling and conflict-free routing for automated guided vehicles. *Comput. Oper. Res.* 38(5):876–888.

Özaltın OY, Prokopyev OA, Schaefer AJ (2018) Optimal design of the seasonal influenza vaccine with manufacturing autonomy. *INFORMS J. Comput.* 30(2):371–387.

Rash E, Kempf K (2012) Product line design and scheduling at Intel. *Interfaces* 42(5):425–436.

Sagratella S (2017) Algorithms for generalized potential games with mixed-integer variables. *Comput. Optim. Appl.* 68(3):689–717.

Sagratella S (2019) On generalized Nash equilibrium problems with linear coupling constraints and mixed-integer variables. *Optimization* 68(1):197–226.

Sagratella S, Schmidt M, Sudermann-Merx N (2020) The noncooperative fixed charge transportation problem. *Eur. J. Oper. Res.* 284(1):373–382.

Saharidis GK, Ierapetritou MG (2009) Resolution method for mixed integer bi-level linear problems based on decomposition technique. *J. Global Optim.* 44(1):29–51.

Shi C, Zhou H, Lu J, Zhang G, Zhang Z (2007) The *K*th-best approach for linear bilevel multifollower programming with partial shared variables among followers. *Appl. Math. Comput.* 188(2):1686–1698.

Stein O, Sudermann-Merx N (2018) The noncooperative transportation problem and linear generalized Nash games. *Eur. J. Oper. Res.* 266(2):543–553.

Sun L, Karwan MH, Kwon C (2018) Generalized bounded rationality and robust multicommodity network design. *Oper. Res.* 66(1):42–57.

Sun L, Karwan MH, Kwon C (2019) Path-based approaches to robust network design problems considering boundedly rational network users. *Transportation Res. Record* 2673(3):637–645.

Tavaslıoğlu O, Prokopyev OA, Schaefer AJ (2019) Solving stochastic and bilevel mixed-integer programs via a generalized value function. Oper. Res. 67(6):1659–1677.

Ulrich KT, Eppinger SD (2016) *Product Design and Development* (McGraw-Hill, New York).

Vicente L, Savard G, Judice J (1996) Discrete linear bilevel programming problem. *J. Optim. Theory Appl.* 89(3):597–614.

Wang L, Xu P (2017) The watermelon algorithm for the bilevel integer linear programming problem. *SIAM J. Optim.* 27(3): 1403–1430.

Wu L, De Matta R, Lowe TJ (2009) Updating a modular product: How to set time to market and component quality. *IEEE Trans. Engrg. Management* 56(2):298–311.

Wu SD, Erkoc M, Karabuk S (2005) Managing capacity in the high-tech industry: A review of literature. *Engrg. Econom.* 50(2):125–158.

Xu P, Wang L (2014) An exact algorithm for the bilevel mixed integer linear programming problem under three simplifying assumptions. *Comput. Oper. Res.* 41(January):309–318.

Yue D, You F (2017) Stackelberg-game-based modeling and optimization for supply chain design and operations: A mixed integer bilevel programming framework. *Comput. Chem. Engrg.* 102(July):81–95.

Yue D, Gao J, Zeng B, You F (2019) A projection-based reformulation and decomposition algorithm for global optimization of a class of mixed integer bilevel linear programs. *J. Global Optim.* 73(1):27–57.

Zeng B, An Y (2014) Solving bilevel mixed integer program by reformulations and decomposition. Working paper, University of South Florida, Tampa. http://www.optimization-online.org/DB_FILE/2014/07/4455.pdf.