



Operations Research

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

In This Issue

To cite this article:

(2022) In This Issue. Operations Research 70(2):iii-viii. <https://doi.org/10.1287/opre.2022.2270>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2022, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

In This Issue

<https://doi.org/10.1287/opre.2022.2270>

Copyright: © 2022 INFORMS

A Framework to Run Personalized Promotions

The availability of individual-level transaction data allows retailers to implement personalized operational decisions. Although such decisions have been around for several years now in online platforms, recent technological developments open new opportunities to extend similar practices to bricks-and-mortar settings (e.g., by using electronic price tags to show different prices to different customers or by using beacon-based technology to send promotion offers to targeted customers). In “Personalized Retail Promotions Through a DAG-Based Representation of Customer Preferences,” Jagabathula, Mitrofanov, and Vulcano propose a back-to-back procedure for running customized promotions in retail operations contexts, from the construction of a nonparametric choice model where customer preferences are represented by directed acyclic graphs to the formulation of the promotion optimization problem. The empirical validation of their proposal on real supermarket data shows the promising performance of their approach over state-of-the-art benchmarks.

The Value of Transshipment in Offline Grocery Retailing

Transshipment in retailing is a practice where one outlet ships its excess inventory to another outlet with inventory shortages. By balancing inventories, transshipment can reduce waste and increase fill rate at the same time. In “Separation of Perishable Inventories in Offline Retailing Through Transshipment,” Li, Yu, and Du explore the idea of transshipping perishable goods with a fixed finite lifetime in offline grocery retailing. In the offline retailing of perishable goods, customers typically choose the newest items first, which can lead to substantial waste. They show that, in this context, transshipment plays two roles. One is inventory balancing, which is well known in the literature. The other is inventory separation, which is new to the literature. That is, transshipment allows a retailer to put newer inventory in one outlet and older inventory in the other. This makes it easier to sell older inventory and reduces waste as a result.

Modeling Crew Assignments for Urban Transport Services Using Differentiated Flows

Public transit agencies need to judiciously deploy their limited crew members to operate numerous daily scheduled services, while meeting duty and working time regulations for each crew member. Since crew costs account for a large portion of the organizations’ operating expenses, minimizing the total crew and transfer costs is very important. But, with hundreds of daily trips and millions of possible crew itineraries, optimizing trip-to-crew assignment decisions is challenging. In “Crew Assignment with Duty Time Limits for Transport Services: Tight Multicommodity Models,” Balakrishnan, Mirchandani, and Lin propose a novel integer optimization model that represents itineraries as multicommodity flows, differentiated by first trip and depot, to capture the duty time limits and incorporate

additional requirements such as selecting equitable schedules. The authors show that this compact model can be tighter than previous formulations, further strengthen the model, and propose a restricted optimization approach combined with an optimality test to generate near-optimal solutions quickly. Extensive computational tests using well-known and real-life problem instances show that the proposed model and solution approach can be very effective in practice.

Diversification vs. Diversity – How is the Efficiency of Markets Affected?

Prices aggregate information that is dispersed in the economy; they thereby facilitate the allocation of scarce resources. Inefficiencies may arise from deviations of market prices from their fundamental values. In “Market Efficient Portfolios in a Systemic Economy,” Awiszus, Capponi, and Weber investigate the impact of the trade-off between diversity and diversification on inefficiencies in secondary markets due to asset illiquidity and leverage constraints of financial institutions. The authors identify two key determining factors. These are the systemic significance of the banks and the statistical properties of the fundamental asset shocks. Systemic significance is driven by the banks’ target leverage, their trading strategies, and the illiquidity characteristics of the assets. The paper demonstrates that portfolio diversification is typically not efficient. In fact, efficient portfolio holdings may strongly deviate from this standard paradigm, especially if the banks have similar characteristics.

How to Design Good Policies for Resource Allocation Problem

A key challenge in the resource allocation problem is to find near-optimal policies to serve different customers with random demands/revenues, using a fixed pool of capacity (properly configured). Three classes of allocation policies, responsive (with perfect hindsight), adaptive (with information updates), and anticipative (with forecast information) policies, are widely used in practice. In “Stochastic Knapsack Revisited: The Service Level Perspective,” Lyu, Chou, Teo, Zheng, and Zhong analyze and compare the performances of these policies for both capacity minimization and revenue maximization models. In both models, the performance gaps between optimal anticipative policies and adaptive policies are shown to be bounded when the demand and revenue of each item are independently generated. In contrast, the gaps between the optimal adaptive policies and responsive policies can be arbitrarily large. More importantly, they show that the techniques developed, and the persistence values obtained from the optimal responsive policies can be used to design good adaptive and anticipative policies for the other two variants of resource allocation problems.

Will Dynamic Pricing Stabilize or Destabilize Grocery Supply Chains?

Suppose that technology reduces price-adjustment costs (e.g., the costs of printing and changing price tags), and as a result prices at grocery stores change more dynamically. Will this change mean less stability or more stability for grocery supply chains? In other words, will more dynamic pricing downstream mean more last-minute purchases, more overtime work, and more empty space in trucks and warehouses? Or will it mean more regular and more standardized orders, smoother schedules, and less waste? To answer this question, the authors fit a pricing and ordering model to data from a large Chinese supermarket chain (daily prices, sales, inventories, and shipments from products from seven categories at 78 stores for 3.5 years) and then simulate a counterfactual grocery chain in which the estimated price-adjustment costs are set to zero. In “Menu Costs and the Bullwhip Effect: Supply Chain Implications of Dynamic Pricing,” Bray and Stamatopoulos find that the removal of price-adjustment costs stabilizes the supply chain, reducing both its shipment volatility, its sales volatility, and its bullwhip (the difference between the shipment and sales volatility).

Relaxing the Fixed Economies of Scale Assumption in Hub Network Design

The economies of scale in hub location are usually modeled by a constant parameter, which captures the benefits companies obtain through consolidation. In “Single Allocation Hub Location with Heterogeneous Economies of Scale,” Rostami, Chit-saz, Arsian, Laporte, and Lodi relax this assumption and consider hub-hub connection costs as piecewise linear functions of the flow amounts. This spoils the triangular inequality property of the distance matrix, making the classical flow-based model invalid and further complicates the problem. The authors tackle the challenge by building a mixed-integer quadratically constrained program and by developing a methodology based on constructing Lagrangian function, linear dual functions, and specialized polynomial-time algorithms to generate enhanced cuts. The developed method offers a new strategy in Benders-type decomposition through relaxing a set of complicating constraints in subproblems when such relaxation is tight. The results confirm the efficacy of the solution methods in solving large-scale problem instances.

Constrained Assortment Optimization in the PCL Model

Assortment optimization involves selecting a subset of products to offer to customers in order to maximize revenue. Often, the selected subset must also satisfy some constraints, such as capacity or space usage. Two key aspects in assortment optimization are (1) modeling customer behavior and (2) computing optimal or near-optimal assortments efficiently. The paired combinatorial logit (PCL) model is a generic customer choice model that allows for arbitrary correlations in the utilities of different products. The PCL model has greater modeling power than other choice models, such as multinomial-logit and nested-logit. In “Constrained Assortment Optimization Under the Paired Combinatorial Logit Model,” Ghuge, Kwon, Nagarajan, and Sharma provide efficient algorithms that find provably near-optimal solutions for PCL assortment optimization under several types of constraints. These include the basic unconstrained problem (which is already intractable to solve exactly), multidimensional space constraints, and partition constraints. The authors also demonstrate via extensive experiments that their algorithms typically achieve over 95% of the optimal revenues.

A Monge Sequence based Approach to Characterize the Competitive Newsvendor Problem

Replicating cash flows of multiple agents in game-theoretic settings tends to be a challenging task. In “Technical Note—A Monge Sequence-Based Approach to Characterize the Competitive Newsvendor Problem,” Bansal and Nagarajan consider the competitive newsvendor game where multiple newsvendors choose inventory levels before demand arrival and the unmet demand of each newsvendor spills over to multiple other newsvendors. They show that this spillover behavior and the resulting cash flows of each newsvendor can be replicated within a transportation problem after assigning artificial costs on spillover behavior. This replication provides an opportunity to study structural properties of the problem, as well as determine the equilibrium of the game. This paradigm of using artificial costs within an optimization framework to replicate agents’ cash flows can be used in many other games as well.

Risk Management and Mean Field Type Control

A risk management version of the classical investment-consumption problem known as Merton’s problem in the finance literature is proposed. Risk is measured by variance, which introduces a nonlinear function of the expected value into the control problem. Standard stochastic theory cannot properly handle this type of nonlinear stochastic optimization problem. Therefore, in “A Risk Extended Version of Merton’s Optimal Consumption and Portfolio Selection,” Bensoussan, Hoe, Kim, and Yan study this time-inconsistent problem within the mean field-type control framework. They derive the sufficient condition of optimality and solve the problem completely. Numerical results illustrating the effect of risk on optimal policies are also presented. Applications can be numerous, including all kinds of investment decisions or operations decisions.

Scheduling Portfolio Execution in the Era of Passive Investment

Over the past decade, there has been a significant rise in assets managed under passive and systematic strategies. Such strategies hold and trade portfolios in a coordinated manner, often concentrating trading around the end of the trading session. Simultaneously, there has been a rise in activity from market participants that act as liquidity providers, themselves trading along portfolio directions. In “Cross-Sectional Variation of Intraday Liquidity, Cross-Impact, and Their Effect on Portfolio Execution,” Min, Maglaras, and Moallemi investigate the implications of these two observations, specifically exploring how the phenomenon of portfolio liquidity provision leads to cross-security impact and influences the optimal execution schedules of risk-neutral traders that seek to minimize their expected execution costs. They show that the optimized schedules deviate from the naïve approach that trades each security separately and instead, couple the trading intensity across stocks so as to benefit from the liquidity provided along attractive portfolio trading directions. Empirical analysis demonstrates that coupled optimized schedules could lower costs by as much as 15% relative to the naïve approach.

Pricing a New Product with Data

Before a product launch (or even after a launch), firms often have little demand information and often do not know critical information such as the market size, the willingness-to-pay distribution, or the adoption speed. The lack of information makes pricing a new product challenging, and insufficient data makes demand forecasting very difficult. This is particularly costly for

new products because the current price not only affects the current revenue, but also the number of adopters who can influence future demand. In “Data-Driven Pricing for a New Product,” Zhang, Ahn, and Uichanco consider a setting in which a firm can learn by observing early sales data at (different) prices over time. They propose a simple and computationally tractable pricing policy that guides price changes after introducing the product. Using mathematical proofs and computational study, the authors show that their method substantially outperforms existing methods even with a very few price changes during a selling season.

Contiguity in Political Districting

For nearly 60 years, operations research techniques have assisted in the creation of political districting plans, beginning with an integer programming model. This model, which seeks compactness as its objective, tends to generate districts that are contiguous, or nearly so, but provides no guarantee of contiguity. In “Imposing Contiguity Constraints in Political Districting Models” Validi, Buchanan, and Lykhovyd consider and analyze four different contiguity models (two old and two new). Their computer implementation can handle redistricting instances as large as Indiana (1,511 census tracts). Their fastest approach uses a branch-and-cut algorithm, where contiguity constraints are added in a callback. Critically, many variables can be fixed to zero a priori by Lagrangian arguments. All test instances and source code are publicly available.

Capacitated Assortment Optimization: Hardness and Approximation

Assortment optimization is an important problem arising in various applications. In many practical settings, the assortment is subject to a capacity constraint. In “Technical Note—Capacitated Assortment Optimization: Hardness and Approximation,” Désir, Goyal, and Zhang study the capacitated assortment optimization problem. The authors first show that adding a general capacity constraint makes the problem NP-hard even for the simple multinomial logit model. They also show that under the mixture of multinomial logit model, even the unconstrained problem is hard to approximate within any reasonable factor when the number of mixtures is not constant. In view of these hardness results, the authors present near-optimal algorithms for a large class of parametric choice models including the mixture of multinomial logit, Markov chain, nested logit, and d-level nested logit choice models. In fact, their approach extends to a large class of objective functions that depend only on a small number of linear functions.

Optimal Pricing under Multiple-Discrete Customer Choices and Diminishing Return of Consumption

Customers make purchase decisions based on the attributes of the products offered and their prices. While the customer selects only one unit of a product in some settings, she purchases multiple products and even possibly multiple units of each product in other settings. Although several studies in the literature have addressed the former case, little attention has been paid to the latter case, this paper’s subject. In this paper, the authors consider the customer’s problem and show that the set of products she purchases is one of the ordered sets based on the product attribute and prices. In “Technical Note—Optimal Pricing Under Multiple-Discrete Customer Choices and Diminishing Return of Consumption,” Huh and Li show that the firm’s optimal pricing problem can be solved efficiently based on another ordering among the products.

What’s the Best Product Variant to Bring to Market?

When companies develop new products, there are often competing designs from which to choose to take to market. How to decide? Traditional methods, such as focus groups, do not scale to the modern marketplace in which tastes evolve rapidly. In “Robust Learning of Consumer Preferences,” Feng, Caldentey, and Ryan develop a data-driven approach to deciding which design to produce by displaying a sequence of subsets of possible designs to potential customers. Their framework finds solutions that are robust to any model of consumer choice within a broad family that includes common choice models studied in the literature as special cases. Previous research focuses on algorithms whose performances are tied to a given choice model. Their algorithm is shown to be asymptotically optimal in a worst-case sense. The promising practical performance of the algorithm is demonstrated through a comprehensive numerical study using real data.

Multiplicative Pacing Equilibria in Auction Markets

Budgets play a significant role in ad markets that implement sequential auctions such as those hosted by internet companies. In “Multiplicative Pacing Equilibria in Auction Markets,” Conitzer, Kroer, Sodomka, and Stier-Moses look at pacing in an ad marketplace using the lens of game theory. The goal is understanding how bids must be shaded to maximize advertiser welfare, at equilibrium. Motivated by the real-world auction mechanism, they construct a game where advertisers in the auctions choose a multiplicative factor not larger than 1 to possibly reduce their bids and best respond to the other advertisers. The authors study the theoretical properties of the game such as existence and uniqueness of equilibria, offers an exact algorithm to compute them, connects the game to well-known abstractions such as Fisher markets, and performs a computational study with real-world-inspired instances. The main insights are that the solutions to the studied game can be used to improve the outcomes achieved by a closer-to-reality dynamic pacing algorithm and that buyers do not have an incentive to misreport bids or budgets when there are enough participants in the auction.

How to Assess the Benefits of Coordination in Managing Hospital Resources

In providing patient care, hospitals rely on multiple types of resources, such as operating rooms, recovery beds, labs, and diagnostic equipment, that are often controlled and managed as separate entities and by different decision makers. In “Surgical Case-Mix and Discharge Decisions: Does Within-Hospital Coordination Matter?” Bavafa, Örmeci, Savin, and Virudachalam focus on the interaction between “front-end” resources, such as operating rooms, and “backroom” resources, such as recovery beds, and compare hospital profitability under the fully coordinated, optimal approach to hospital resource management and under alternative decentralized approaches often encountered in practice. The authors identify settings in which the benefits of coordination are likely to be high as well as settings in which those benefits are at best moderate. In a given hospital, only hospital managers are in a position to estimate with any degree of certainty potential costs of coordinated management of hospital resources, and the article’s analysis of the benefits of coordination empowers hospital managers to make informed decisions on the desirability of replacing the often decentralized “status quo” by centralized resource management.

Protecting Networks Against Strategic Attacks

Ensuring the security of critical infrastructures is crucial for the society's welfare and prosperity. However, these infrastructure networks are inherently vulnerable to both intentional and unintentional threats. In "Network Inspection for Detecting Strategic Attacks," Dahan, Sela, and Amin study a strategic network inspection problem, formulated as a large-scale bilevel optimization problem, in which a utility seeks to determine an inspection strategy with minimum number of smart detectors that ensures a desirable expected detection performance under worst-case attacks. The authors derive structural properties of optimal solutions and show that the problem can be solved using Nash equilibria of a large-scale zero-sum game. Their analysis leads to a computationally tractable and operationally feasible solution approach with theoretical guarantees based on combinatorial objects that capture the nature of equilibrium inspection and attack strategies. Their computational study indicates that utilities can achieve a high level of protection in large-scale networks by strategically positioning a small number of detectors.

Better Matching for Reliable and Flexible Ridesharing

Ridesharing platforms have radically changed the way people get around in urban areas, but there remain challenges undercutting the mission of "making transportation as reliable as running water." A particular concern is that drivers strategize: calling riders to find out their destinations and canceling trips that are not worthwhile, declining trips and chasing surge prices in neighboring areas, and going off-line before large events will end in anticipation of a price increase. In "Spatio-Temporal Pricing for Ridesharing Platforms," Ma, Fang, and Parkes show that such strategic behaviors are symptoms of inefficiencies in the pricing and dispatching rules governing today's platforms. They propose the Spatio-Temporal Pricing mechanism, which solves for the welfare-optimal matching of drivers to trips, and sets prices that are appropriately smooth in both space and time such that the best thing for drivers to do is accept any proposed trip dispatch. This demonstrates that ridesharing platforms can succeed in optimally orchestrating trips and providing reliable transportation for riders, while still leaving drivers with the flexibility to decide how to work.

New Model for Strategic Workforce Planning in Organization

In "Strategic Workforce Planning Under Uncertainty," Jaillet, Loke, and Sim propose a data-driven model for conducting strategic workforce planning in organizations. The model optimizes for recruitment and promotions by balancing the risks of not meeting headcount, budget, and productivity constraints, while keeping within a prescribed organizational structure. Analysis using the model indicates that there are increased workforce risks faced by organizations that are not in a state of growth or organizations that face limitations to organizational renewal (such as bureaucracies).

School Choice in Chile

In "School Choice in Chile," Correa, N. Epstein, R. Epstein, Escobar, Rios, Aramayo, Bahamondes, Bonet, Castillo, Cristi, B. Epstein, and Subiabre describe the design and implementation of the new school admissions system in Chile. The design, based on the celebrated work of Gale, Shapley, and Roth, involves several challenges to comply with the Chilean legislation. For instance, the system includes different priorities and quotas for different groups of students. Moreover, the system

operates nationwide and in all grade levels. As a result of the latter, one of the primary goals was to favor the joint assignment of siblings to the same schools. To accomplish this, the authors propose a heuristic approach that dynamically updates preferences, and breaks ties at the family level to increase the probability that siblings are assigned to the same school. The system, introduced in 2016 and still in use today, serves more than half a million students each year.

Pricing with Samples

In "Pricing with Samples," Allouah, Bahamou, and Besbes study the value of data for pricing purposes. Although pricing is central across industries, little is known about the minimal amount of data needed to achieve good pricing decisions. The authors propose a novel approach to quantify the informational content of data, through the introduction of a new class of robust data-driven policies and the development of factor-revealing dynamic programs. Studying the prototypical case of data coming in the form of samples from the willingness to pay of customers, they show that even a few samples (as few as 10) go a very long way in uncovering "good" prices. For example, quite strikingly, against a general class of distributions (monotone increasing hazard rate distributions), a single observation guarantees 64% of the performance an oracle with full knowledge of the distribution would achieve, two samples suffice to ensure 71%, and 10 samples guarantee 80% of such performance.

Bayesian Exploration: Incentivizing Exploration in Bayesian Games

In a wide range of recommendation systems, self-interested individuals ("agents") make decisions over time, using information revealed by other agents in the past, and producing information that may help agents in the future. Each agent would like to exploit the best action given the current information but would prefer the previous agents to explore various alternatives to collect information. A social planner, by means of a well-designed recommendation policy, can incentivize the agents to balance exploration and exploitation in order to maximize social welfare or some other objective. The recommendation policy can be modeled as a multiarmed bandit algorithm under Bayesian incentive compatibility (BIC) constraints. This line of work has received considerable attention in the "economics and computation" community. Although in prior work, the planner interacts with a single agent at a time, the present paper allows the agents to affect one another directly in a shared environment. The agents now face two sources of uncertainty: what is the environment, and what would the other agents do? In "Bayesian Exploration: Incentivizing Exploration in Bayesian Games," Mansour, Slivkins, Syrgkanis, and Wu focus on "explorable" actions: those that can be recommended by some BIC policy. They show how the principal can identify and explore all such actions.

Budget Allocation for Nested Simulation

Nested simulation (also referred to as two-level simulation) finds a variety of applications such as financial risk measurement, and a central issue of nested simulation is how to allocate a finite amount of simulation budget to achieve the highest accuracy. In "Technical Note—Bootstrap-based Budget Allocation for Nested Simulation," Zhang, Liu, and Wang propose a bootstrap-based rule for simulation budget allocation for nested simulation. By utilizing the asymptotically optimal inner- and outer-level sample sizes that are typically unknown, the proposed method employs bootstrap sampling on a small amount

of initial samples to estimate the unknown optimal sample sizes, thus providing a reasonably good allocation rule for the main simulation. An allocation rule to ensure the asymptotic validity of confidence intervals is also given.

A Simple and Near-Optimal Static Pricing Policy for Reusable Resources

Reusable resources have become increasingly important because of the rise of cloud computing and ride-sharing vehicles in addition to more traditional examples, such as hotel rooms, airline seats, and rotatable spare parts. Optimal pricing policies for reusable resources involve the use of dynamic pricing, by which prices increase as resources are being used and decrease as resources free up. In “Technical Note—Static Pricing: Universal Guarantees for Reusable Resources,” Besbes, Elmachtoub, and Sun consider a foundational model for reusable resources and prove that a static price policy guarantees a constant fraction of the profit, market share, and service level of an optimal dynamic pricing policy. These results hold even when the number of units is small, and the service and arrival rates are arbitrary.

Improved Performance Guarantees for Sequential Batch-Testing

In “A Polynomial-Time Approximation Scheme for Sequential Batch Testing of Series Systems,” Segev and Shaposhnik study a recently introduced generalization of the classic sequential testing problem for series systems, consisting of multiple stochastic components. The conventional assumption in such settings is that the overall system state can be expressed as an AND function, defined with respect to the states of individual components. However, unlike the classic setting, rather than testing components separately, one after the other, we allow aggregating multiple tests to be conducted simultaneously, while incurring an additional set-up cost. This feature is present in numerous practical applications, where decision makers are incentivized to exploit economy of scale by testing subsets of components in batches. The main contribution of this paper is to devise a polynomial-time approximation scheme for the sequential batch-testing problem, by leveraging a number of techniques in approximate dynamic programming, based on a synthesis of ideas related to efficient enumeration methods, state-space collapse, and charging schemes.

Learning and Adaptive Control in Dynamic Service Platforms for Payoff Maximization

Many online service platforms have dedicated algorithms to match their available resources to incoming clients to maximize client satisfaction. One of the key challenges is to balance the generation of higher payoffs from existing clients and exploration of new clients’ unknown characteristics while at the same time satisfy the resource capacity constraints. In “Integrated Online Learning and Adaptive Control in Queueing Systems with Uncertain Payoffs,” Hsu, Xu, Lin, and Bell show that traditional approaches such as maximizing instantaneous payoffs with current knowledge or using queue-length based controls guided by “shadow prices,” would lead to suboptimal long-term payoffs. Instead, they propose a novel utility-guided assignment algorithm that seamlessly integrates online learning and adaptive control to provide high system payoffs with performance guarantees. The theoretical performance bound also lends system design insights into the impact of uncertain client dynamics, payoff learning, and backlogged clients. They further

develop a decentralized version of the algorithm, which is applicable to large systems and performs well even when the service rates are random.

Dynamics of Parallel Service Systems

In parallel service systems, customers of various types are served simultaneously by several servers that have some different and some overlapping skills or capacities. Examples include customer routing in call centers, scheduling of hospital operating rooms, and driver assignment in ride-hailing services. First come first served (FCFS) is a natural and widespread resource allocation policy, yet its performance both during the transient stage and in the long run is often difficult to analyze. In particular, matching rates, which indicate what fraction from all services are of a specific server customer type, are difficult to evaluate under FCFS. In “Fluid Models of Parallel Service Systems under FCFS,” Nov, Weiss, and Zhang formulate a stochastic queueing model for such systems and study its fluid approximation. The authors demonstrate the usefulness and the limitations of exploring fluid performance in evaluating system behavior in terms of stability, responsiveness, and utilization.

Nonconvex Low-Rank Tensor Completion from Noisy Data

In “Nonconvex Low-Rank Tensor Completion from Noisy Data,” Cai, Li, Poor, and Chen investigate a problem of broad practical interest, namely, the reconstruction of a large-dimensional low-rank tensor from highly incomplete and randomly corrupted observations of its entries. Although a number of papers have been dedicated to this tensor completion problem, prior algorithms either are computationally too expensive for large-scale applications or come with suboptimal statistical performance. Motivated by this, the authors propose a fast two-stage nonconvex algorithm—a gradient method following a rough initialization—that achieves the best of both worlds: optimal statistical accuracy and computational efficiency. Specifically, the proposed algorithm provably completes the tensor and retrieves all low-rank factors within nearly linear time, while at the same time enjoying near-optimal statistical guarantees (i.e., minimal sample complexity and optimal estimation accuracy). The insights conveyed through our analysis of nonconvex optimization might have implications for a broader family of tensor reconstruction problems beyond tensor completion.

Efficiency Analysis for Multicomponent Production Processes

Conventional models concerned with efficiency analysis of organizations typically consider a single production process, or technology, in which all inputs are used in the production of all outputs. This approach does not account well for the situations in which the organizations are involved in several component production processes whose inputs and outputs may be shared by different processes. The main difficulty in modeling such technologies is the fact that we often do not know the exact allocation of the shared inputs and outputs to individual processes. In “Variable and Constant Returns-to-Scale Production Technologies with Component Processes,” Podinovski shows how this problem can be overcome by the consideration of the worst-case scenario for the allocation of the shared inputs and outputs to different components of the technology. This approach leads to the development of multicomponent variants of two well-established nonparametric models. An application involving universities in England demonstrates the usefulness

and improved discriminating power of the new models compared with their conventional analogues.

Assessing the Quality of a Family of Linear Programming Relaxations

The strength of an integer programming (IP) formulation is typically measured by the quality of its linear programming (LP) relaxation. Using this measure, it can be difficult to rank IP formulations across multiple right-hand sides. In “The Gap Function: Evaluating Integer Programming Models Over Multiple Right-Hand Sides,” Ajayi, Thomas, and Schaefer define the gap function of an IP formulation as the difference between its LP relaxation and optimal objective. The authors leverage superadditive duality to formulate linear programs that compute various metrics for relaxation quality, such as the expectation or extrema of the absolute or relative gap functions.

Spatial Capacity Planning

In “Spatial Capacity Planning,” Besbes, Castro, and Lobel study the relationship between capacity and performance for a service firm with spatial operations such as a ride-hailing platform in which each customer arrives in the system with the need to travel from an origin to a destination. In a classical M/M/n queueing model, the square root safety (SRS) staffing rule is known to balance server utilization and customer wait times. By contrast, the authors find that the SRS rule does not lead to such a balance in spatial systems. In a spatial environment, pick-up times increase the load in the system; furthermore, they are an endogenous source of extra workload that leads the system to only operate efficiently if there is sufficient imbalance between supply and demand. In particular, to obtain a balance of utilization and wait times, the service firm should use a higher safety factor, proportional to the offered load to the power of $2/3$.