MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 24, No. 2, 2022, pp. 791–809
http://pubsonline.informs.org/journal/msom
ISSN 1523-4614 (print), ISSN 1526-5498 (online)

Crowdsourcing Last-Mile Deliveries

Soraya Fatehi,[a] Michael R. Wagner[b,c]

[a] Naveen Jindal School of Management, University of Texas at Dallas, Richardson, Texas 75080; [b] Michael G. Foster School of Business, University of Washington, Seattle, Washington 98195; [c] Amazon.com, Seattle, Washington 98109

Contact: soraya.fatehi@utdallas.edu, https://orcid.org/0000-0002-3383-148X (SF); mrwagner@uw.edu, https://orcid.org/0000-0003-2077-3564 (MRW)

Received: March 8, 2020
Revised: October 26, 2020
Accepted: January 10, 2021
Published Online in Articles in Advance: May 24, 2021

https://doi.org/10.1287/msom.2021.0973

Abstract. ***Problem definition*:** Because of the emergence and development of e-commerce, customers demand faster and cheaper delivery services. However, many retailers find it challenging to efficiently provide fast and on-time delivery services to their customers. ***Academic/practical relevance*:** Amazon and Walmart are among the retailers that are relying on independent crowd drivers to cope with on-demand delivery expectations. ***Methodology*:** We propose a novel robust crowdsourcing optimization model to study labor planning and pricing for crowdsourced last-mile delivery systems that are utilized for satisfying on-demand orders with guaranteed delivery time windows. We develop our model by combining crowdsourcing, robust queueing, and robust routing theories. We show the value of the robust optimization approach by analytically studying how to provide fast and guaranteed delivery services utilizing independent crowd drivers under uncertainties in customer demands, crowd availability, service times, and traffic patterns; we also allow for trend and seasonality in these uncertainties. ***Results*:** For a given delivery time window and an on-time delivery guarantee level, our model allows us to analytically derive the optimal delivery assignments to available independent crowd drivers and their optimal hourly wage. Our results show that crowdsourcing can help firms decrease their delivery costs significantly while keeping the promise of on-time delivery to their customers. ***Managerial implications*:** We provide extensive managerial insights and guidelines for how such a system should be implemented in practice.

Funding: M. R. Wagner acknowledges the support of a Neal and Jan Dempsey Faculty Fellowship.
Supplemental Material: The e-companion is available at https://doi.org/10.1287/msom.2021.0973.

Keywords: crowdsourcing • on-demand deliveries • robust optimization • queueing theory

1. Introduction

Be it for a large business, such as Amazon or Walmart, or a smaller retailer, such as a local grocery store, last-mile deliveries are one of the most expensive aspects of delivery logistics for orders placed online, and managing them is becoming increasingly important as customers demand faster and cheaper delivery services. Implementing last-mile deliveries for fast orders can be challenging, especially during peak demand. According to Insider Intelligence (2017), 25% of customers do not proceed with their online orders if same-day delivery is not available. Nevertheless, only 20% of retailers offer same-day delivery, which is indicative of the challenges that are inherent in providing on-demand delivery services. To overcome these challenges, an increasing number of retailers, including Amazon and Walmart, are considering *crowdsourcing* last-mile deliveries.

Amazon is using independent crowdsourced drivers under a program called Amazon Flex (flex.amazon.com) and pays the drivers an hourly wage of $18–$25. In addition to its warehouses, Amazon is using vacant lands, such as outdoor parking lots, as delivery stations and is using crowdsourced drivers via the Flex program for making its Amazon Now and Amazon Fresh deliveries (Phillips 2018). Walmart similarly offers fast deliveries in all of its major markets and product categories under the crowdsourced Spark delivery program (www.drive4spark.com).

In this paper, we study last-mile deliveries via independent crowd drivers who decide whether to work and when to work. Specifically, we study crowdsourced last-mile delivery systems for on-demand orders, with guaranteed delivery time windows, where the time between placing an order and receiving it ranges from two hours to same day. According to Taylor (2018), crowd drivers' independence and customer delay sensitivity are integral to crowdsourced on-demand service platforms. Firms should intelligently decide on the crowd compensation and labor planning to have enough crowd drivers willing to participate, such that they are not too busy, to ensure on-time deliveries, nor too idle, where their effective earning rate is too low for participation. It is challenging for retailers to make pricing and labor planning decisions because there is uncertainty in customer demands (*when* and *where* will a customer place an order), crowd availability (*whether* and *when* independent crowd drivers are available for delivery services), and customer delivery service times (delivering to a house differs from delivering to a gated community), as well as varying road traffic conditions.

1.1. Literature Review

Our paper is related to three streams of research: crowdsourcing last-mile deliveries, the sharing economy and on-demand platforms, and queueing and routing via robust optimization. In particular, we are not aware of any paper that combines aspects of all three streams, as we do in our paper.

1.1.1. Crowdsourcing Last-Mile Deliveries. In this stream, Li et al. (2016) develop a heuristic algorithm for a problem where a set of taxi drivers is serving both people and parcels. Arslan et al. (2019) develop algorithms for the problem of crowdsourcing deliveries using a rolling horizon approach, which repeatedly matches drivers and deliveries every time a driver or a delivery arrives. Cao et al. (2018) study the use of ride-sharing platforms such as Uber and Lyft, along with an in-house van delivery system, for making last-mile deliveries and derive exact and asymptotic results for the expected number of packages that can be delivered during a time horizon. However, these papers focus on algorithmic solutions, whereas we focus on deriving *analytical* pricing and labor planning policies for on-demand crowdsourced delivery systems with guaranteed delivery time windows. The paper most relevant to ours is Qi et al. (2018), which studies crowdsourcing deliveries. One major difference between our study and the study by Qi et al. (2018) is that in our paper, packages must be delivered to customers within guaranteed delivery time windows, whereas Qi et al. (2018) do not consider such a constraint.

1.1.2. The Sharing Economy and On-Demand Platforms. There are a number of related papers that study peer-to-peer sharing. He et al. (2017) develop a distributionally robust optimization model for studying service region design for electric vehicle-sharing providers that offer one-way trips to customers. Benjaafar et al. (2019) characterize the equilibrium of peer-to-peer product sharing where owners rent their assets to nonowner consumers on an as needed basis to generate income from renting. There are also several papers studying the operation and pricing of on-demand service platforms. Gurvich et al. (2016) analyze an on-demand service provider that is using self-scheduling agents and wants to maximize its profit; they show that under self-scheduling, the firm offers lower service levels to its customers, which lower the firm's profit and are costly to customers. Cachon et al. (2017) show that the use of surge pricing on service platforms with self-scheduling providers benefits customers, providers, and the platform. Taylor (2018) examines pricing in on-demand service platforms where agents are independent and customers are delay sensitive and shows that uncertainties in agent opportunity costs and customer valuations impact the platform's per service price. Similarly, Bai et al. (2019) study the optimal price and wage rate on on-demand service platforms with price- and time-sensitive customers and earning-sensitive service providers. Benjaafar et al. (2018) analyze labor welfare in on-demand service platforms with independent workers. In contrast to these papers, in our work crowd drivers are utilized to deliver packages to customers within guaranteed delivery time windows.

1.1.3. Queueing and Routing via Robust Optimization. In this stream, Bandi et al. (2015) study the worst-case performances of single-server and multiserver queueing systems using robust optimization. Sungur et al. (2008) is one of the first papers that studies the capacitated vehicle routing problem with demand uncertainty using a robust optimization approach. Similarly, Gounaris et al. (2013) consider the capacitated vehicle routing problem under the case where demand is uncertain and belongs to a polyhedron uncertainty set, and the objective is to minimize total routing costs. Carlsson and Delage (2013) utilize a robust optimization framework to develop branch-and-bound algorithms for the stochastic multiple-vehicle routing problem, where the location of demand points and their distribution are unknown; their objective is to minimize the worst-case work load over all subregions, where each subregion is served by a single vehicle. Liu et al. (2020) utilize robust optimization for order assignments in last-mile delivery faced by food service providers. However, these papers apply robust optimization to classic routing problems; in contrast, we apply robust optimization to a novel application of on-demand deliveries via crowd drivers. Furthermore, no paper has combined robust queueing with robust routing, as we do.

1.2. Contributions

Our paper makes the following contributions.

1. This paper is the first, to our knowledge, that analytically studies crowdsourcing last-mile deliveries, with guaranteed delivery time windows (e.g., same-day or two-hour deliveries), under nonstationary uncertainties. We develop our optimization model by combining crowdsourcing, robust queueing, and robust routing theory. We derive closed form expressions for the delivery cost and system times of all customers in the crowd delivery system, which allows us to carefully analyze crowd compensation and labor planning. Our results can help firms in designing crowdsourced delivery systems, for their on-demand orders with guaranteed delivery times, to minimize their last-mile delivery costs. Additionally, our results help firms to decide on the on-time delivery guarantee level depending on their customers' sensitivities to delay and crowd drivers' opportunity costs.

2. Our results show that crowdsourcing can help firms significantly decrease their delivery costs while keeping the promise of fast on-time deliveries to customers. Our analysis shows that, under low levels of uncertainty or for low on-time guarantee levels, firms should assign fewer packages in each crowd delivery tour to create a stream of available delivery work, which encourages crowd drivers' participation, without having to offer them high wages. In contrast, under high levels of uncertainty or for high on-time guarantee levels, firms should assign more packages in each delivery tour but at the same time, offer higher hourly wages.

3. We design realistic numerical experiments that cover 17 Seattle ZIP codes that are served by an Amazon Flex depot, using real travel distances on road networks. Our experimental results show that the performance of our proposed robust solution, which is faster and easier to implement, performs significantly better than two benchmarks, a stochastic counterpart and a nonrandomized heuristic, in terms of the firm's cost savings and the number of on-time deliveries for customers' on-demand orders. Furthermore, our results show that a crowd delivery system, if designed optimally, is significantly cheaper than the exclusive use of a third-party logistics (3PL) firm, especially for customers' fast same-day and two-hour delivery orders.

2. A Robust Crowdsourcing Model

Challenges in crowdsourced delivery platforms come from uncertainties in both the supply of crowd drivers and the demand of customers. In these systems, the availability of independent crowd drivers is not guaranteed, as each driver has their own individual uncertain opportunity cost, which results in an uncertain supply of crowd drivers that depends on the offered wage rate. On the other hand, customer orders must be delivered within guaranteed delivery time windows, under uncertainty in the timing of customer orders, on-location service times, and delivery locations, possibly with trend and seasonality. These uncertainties make it challenging to deliver customer orders on time for a given service-level guarantee. We study how a firm should design an efficient crowdsourced last-mile delivery system, to ensure participation of a sufficient number of crowd drivers, for on-time delivery of all customer orders while still minimizing delivery costs.

In our paper, we develop our main model utilizing robust optimization, a modern approach to decision making under uncertainty, where uncertainty is captured using deterministic set membership (whose structures are motivated by various limit theorems of probability), rather than stochastic distributions. There are advantages in using robust optimization in studying crowdsourcing last-mile deliveries with guaranteed delivery time windows. First, unlike stochastic optimization, robust models revolve around worst-case scenarios, or in other words, they "under-promise but over-deliver" (Kuhn et al. 2019). This is important for our application because customers are sensitive to delay, and hence, firms need to make guaranteed on-time package deliveries; robust optimization allows a firm to promise on-time deliveries with high probability. Furthermore, the robust approach allows us to capture trend and seasonality in customer arrivals, service times, and traffic patterns, as well as heterogeneity in crowd drivers, over multiple service subregions with varying areas and populations; it is typically difficult to analytically capture all these patterns in a stochastic modeling approach (although simulation can be used, it generally does not provide analytical insights).

2.1. Problem Definition

In crowdsourced delivery platforms, as the firm should make sure a sufficient number of crowd drivers are available at the distribution center/pickup location for on-time delivery of customers' fast (one-hour, two-hour, or same-day) delivery orders, the firm asks for availability of crowd drivers beforehand. In particular, the firm announces the delivery work that it needs, along with associated details, including the start date and time. Each crowd driver who sees this announcement can choose to sign up for it and provide the requested delivery services or not, depending on what compensation the firm has offered compared with their opportunity costs. Hence, in our problem, the firm aims to make labor planning and crowd compensation decisions for a time horizon (e.g., half a day or a day) in advance of demand realization.

In our study, a firm has two options for last-mile delivery: The firm can (1) outsource all deliveries to a 3PL firm, such as Deliv or FedEx, or (2) assign all packages to crowd drivers to deliver. The main monetary difference between using 3PL companies for delivery and utilizing crowd drivers is the payment structure: 3PL companies typically charge firms per package delivery, whereas the crowd is paid hourly wages for value-added time. In our model, we also allow for hybrid solutions, where both a 3PL firm and the crowd are utilized, to add a layer of flexibility to a firm's delivery system. Hybrid solutions are motivated by practice: Amazon, in addition to using crowdsourced delivery services under the Flex program out of its warehouse stations and vacant lots (Phillips 2018), utilizes the 3PL companies UPS and FedEx for making its fast same-day deliveries (Kim 2019, Levy 2019). Similarly, Walmart has started using a crowd delivery program (Spark) while still using 3PL firms (such as FedEx). Finally, Cao et al. (2018) study a hybrid solution where the firm can use ride-sharing platforms such as Uber and Lyft along with in-house van delivery systems for making last-mile deliveries. However, this approach has not been successful for all firms:

Walmart recently retracted from using ride-sharing platforms (Reuters 2018).

3PL firms typically charge a fee per package for delivery, which is a function of the distance between the origin and the destination and the delivery time window; we let c_T denote this per package cost (the distance and delivery time window dependence is implicit), where the T subscript indicates a (delivery) truck. In contrast, in crowdsourced delivery applications drivers are typically paid an hourly wage w. As is the case in practice, we assume the driver is paid only for travel and servicing time (i.e., value-adding time) but not idle time.

Firms typically offer customers guarantees on delivery times. For instance, Amazon offers next-day, same-day, and even two-hour guaranteed delivery windows. We let α denote the desired delivery duration requested by the customer, and we assume that all customers have the same target delivery duration. We also refer to α as the *service level* that all customers are guaranteed to receive.

As a firm needs to make crowd compensation and labor planning decisions before demand is realized, without using the package characteristics (e.g., location) of the realized demand, we develop a static optimization model that allows us to derive generic optimal policies for the crowdsourced last-mile delivery system. We adopt a simple randomized allocation policy, where a $P \in [0, 1]$ proportion of deliveries is randomly assigned to be delivered by the crowd, and the remaining $(1 - P)$ proportion is delivered using a 3PL firm. This randomized allocation policy allows us to derive generic analytical results and managerial insights under demand and supply uncertainties, without using the package characteristic of the realized demand. In Section 4, we propose a dynamic benchmark model that uses package characteristics (e.g., location) to determine which packages should be crowdsourced and which should be assigned to the 3PL firm, and we compare its performance with the randomized allocation policy.

In the next section, we develop a robust optimization model of crowdsourcing last-mile deliveries for on-demand orders with guaranteed delivery time windows. We aim to derive optimal policies for crowdsourced last-mile deliveries under supply and demand uncertainties and under the service-level guarantee that the firm offers its customers.

2.2. Models of Uncertainty

In this section, we detail our uncertainty sets. Note that the support of an uncertainty set does not need to coincide with the support of the associated random parameter; indeed, the discussion on pages 32–33 of Ben-Tal et al. (2009) shows that this disconnect can result in improved performance.

2.2.1. Customer Interarrival Times. Let the interarrival time of customer i be A_i, with mean and standard deviation of μ_i^a and σ_i^a, respectively. If customers arrive according to a Poisson process with rate λ, the random allocation policy splits this process into a Poisson process served by the crowd, of rate $P\lambda$, and another Poisson process of rate $(1 - P)\lambda$ that is served by a 3PL firm; in this analysis, we assume that $p > 0$. This observation motivates us to redefine the mean and standard deviation of the interarrival time of customer i, who is served by the crowd, to be μ_i^a/P and σ_i^a/P, respectively, because the mean interarrival time for a Poisson process is $1/\lambda$. Similar to the approaches in Bandi et al. (2015, 2018), we define the interarrival times uncertainty set for these customers to be

$$U_A(P) = \left\{ (A_1, \ldots, A_n) \geq 0 : \left| \frac{\sum_{i=k+1}^{j} (A_i - \mu_i^a/P)}{\sqrt{\sum_{i=k+1}^{j} (\sigma_i^a/P)^2}} \right| \leq \gamma^a, 0 \leq k < j \leq n \right\}, \tag{1}$$

which is motivated by the central limit theorem (CLT): $\lim_{j\to\infty} P\left(\frac{\Sigma_{i=k+1}^{j}(A_i - \mu_i^a/P)}{\Sigma_{i=k+1}^{j}(\sigma_i^a/P)^2} \leq x \right) = \Phi(x)$, where Φ is the standard normal distribution. The parameter γ^a can be chosen based on the probability with which the firm wants the interarrival time inequalities to be satisfied, such as $\gamma^a \in \{2, 3\}$, because if Z is a standard normal random variable, $P(|Z| \leq 2) \approx 0.954$ and $P(|Z| \leq 3) \approx 0.997$. If customers' delay sensitivity is high, then the firm should choose a higher γ^a value.

Our approach to modeling interarrival times captures trend and seasonality in customer arrivals, via proper selection of the μ_i^a and σ_i^a parameters, and is different from Bandi et al. (2015, 2018), whose approaches do not consider nonstationary arrivals. Our modeling of uncertainty sets with nonstationarity is motivated by Mamani et al. (2016) and Bandi et al. (2019).

2.2.2. Driver Route Durations. We assume there is a single depot in the service area, and each crowd driver, with capacity Q, serves a set of q customers at a time, between visits to the depot, where $q \in \{1, \ldots, Q\}$; these sets are formed and served using a first come, first served (FCFS) policy. This routing policy is motivated by the current practice of crowd delivery systems, where crowd drivers are assigned delivery sets of packages that fit within their vehicle capacity, and is similar to the $G/G/m$ heavy traffic policy and the tour-partitioning capacitated heavy traffic policy studied in Bertsimas and Van Ryzin (1993) with $n = q$ and the $q \in \{1, \ldots, Q\}$ being optimized in our optimization model. We assume each driver takes an optimal traveling salesman problem (TSP) tour through the q customer locations and not FCFS within a set. This assumption is motivated by observations from industry: The Amazon Flex application gives the driver directions for the optimal tour (logistics.amazon.com); similarly, Walmart also offers the optimal delivery tour to its drivers (Lore 2017). The sets themselves are served in the FCFS order.

We denote the TSP tour length that visits q city locations $x_1, \ldots, x_q \in \mathbb{R}^2$ as $L_q(x_1, \ldots, x_q)$; for simplicity, we write $L_q = L_q(x_1, \ldots, x_q)$. We assume that the number of crowdsourced customers n is a multiple of q: $n = mq$, for some positive integer $m \geq 1$; if we consider n to be large enough, this approximation will not change our results materially. Therefore, the drivers collectively serve m sets of size q.

We assume customers are located in M ZIP codes, each with area $\mathbb{A}_i$ and population p_i, $i = 1, \ldots, M$, where the probability density function that each customer is in a given ZIP code is proportional to the ZIP code's population, and, given a ZIP code, the customer's location is conditionally uniformly distributed over the ZIP code's area. The probability density function of a customer's location $x \in \mathbb{R}^2$, which is potentially located in one of M ZIP codes, each with area $\mathbb{A}_i$ and population p_i, $i = 1, \ldots, M$, is a generalization of a result in Carlsson and Delage (2013) and is given by

$$f(x) = \begin{cases} \dfrac{p_i}{\sum_{j=1}^{M} \mathbb{A}_j p_j}, & \text{if } x \text{ is in ZIP code } i, \ i = 1, \ldots, M \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

There exists a probabilistic limit theorem for the optimal TSP tour length, which is effectively a law of large numbers: Originally derived by Bearwood et al. (1959) and further developed by Steele (1981), there exists a constant $\tilde{\beta}$ such that $\lim_{q\to\infty} \frac{L_q}{\sqrt{q}} = \tilde{\beta} \iint_{\mathbb{R}^2} f(x)^{1/2} dx$, almost surely. From this asymptotic result and the probability density function $f(x)$, we can define

$$\begin{aligned} \lim_{q\to\infty} \frac{L_q}{\sqrt{q}} &= \tilde{\beta} \sum_{i=1}^{M} \iint_{x \text{ in ZIP code } i} \sqrt{\frac{p_i}{\sum_{j=1}^{M} \mathbb{A}_j p_j}} dx \\ &= \tilde{\beta} \sum_{i=1}^{M} \sqrt{\frac{p_i}{\sum_{j=1}^{M} \mathbb{A}_j p_j}} \mathbb{A}_i. \end{aligned} \tag{3}$$

The analysis has not accounted for the depot, where the driver picks up packages. As motivated by Amazon Flex and Walmart's Spark program, depot locations can vary; for instance, some fast Amazon Flex deliveries are food products from a Whole Foods Market location, of which there are many. Therefore, we also consider the depot location to be random. For simplicity, we assume that the depot has the same distribution as the customers' locations. Therefore, a driver's tour starts from the depot, visits q customers, and returns to the depot, which has length L_{q+1}. Note that the driver might have to return to the depot for returning undeliverable packages and/or customer-returned items, which motivates us to use a tour, rather than a path. Let L_{q+1}^j denote the optimal length of the jth tour through the depot, which serves the q customers $(j-1)q+1, (j-1)q+2, \ldots, jq$, and let v_j denote the average travel speed for the jth tour. Thus, L_{q+1}^j / v_j represents the optimal tour *duration* of the jth tour. Motivated by the Bearwood et al. (1959) limit theorem, we define

$$\begin{aligned} U_L^\ell = \Bigg\{ \frac{L_{q+1}^\ell}{v_\ell} : \beta_L \sum_{i=1}^{M} \sqrt{\frac{p_i}{\sum_{j=1}^{M} \mathbb{A}_j p_j}} \mathbb{A}_i \\ \leq \frac{L_{q+1}^\ell}{\sqrt{q+1}} \leq \beta_U \sum_{i=1}^{M} \sqrt{\frac{p_i}{\sum_{j=1}^{M} \mathbb{A}_j p_j}} \mathbb{A}_i \Bigg\} \end{aligned} \tag{4}$$

as the uncertainty set for the ℓth tour duration. The β_L and β_U parameters are lower and upper bounds, respectively, on $\tilde{\beta}$, which reflect the fact that the limit theorem does not hold exactly for a finite q. These parameters also control the level of conservatism in our model; as customers' delay sensitivity increases, the firm should choose a higher β_U value and a lower β_L value. A similar uncertainty set, motivated by the strong law of large numbers, was utilized in Wagner (2018). Note that, although the motivating TSP asymptotic results are based on the Euclidean distance metric, the robust uncertainty set does not require any specific distance metric; indeed, in our numerical experiments we utilize road distances measured on a real transportation network in the Seattle area. Similarly, Carlsson and Delage (2013), Cachon (2014), and Qi et al. (2018) utilize this limit theorem when in reality, the distances are calculated using the Manhattan metric. In the sequel, we show how we are modeling traffic patterns via the sets U_L^ℓ by appropriately setting the speeds v_ℓ.

2.2.3. Customer On-Site Service Times. The arrival of a customer entails a request for a package delivery. After the server travels to the customer's location, there is an on-site service time (e.g., parking, finding the appropriate apartment, etc.). Let X_i denote the on-site service time of the ith customer with mean and standard deviation μ_i^s and σ_i^s, respectively. We adopt and modify the service time uncertainty set, for multiple servers in a nonrouting queueing context, from Bandi et al. (2015). More specifically, we adapt their uncertainty set to be defined in terms of the service times of *sets* of customers, rather than individual customer service times. Furthermore, we define the uncertainty set for nonstationary service times. We again assume there are $n = mq$ crowdsourced customers, resulting in m tours of size q that are served by N available crowd drivers; in the sequel, we discuss how N is determined endogenously by self-selecting drivers. Let $\tau^m \triangleq \frac{m}{N}$, assumed to be integer, denote the average number of tours taken by each driver. We next assign the m tours of customers, in order of their arrivals, into the following partitions: $J_1 = \{1, \dots, N\}, J_2 = \{N+1, \dots, 2N\}, \dots, \ J_{\tau^m} = \{(\tau^m - 1)N + 1, \dots, m\}$. The uncertainty set U_X^N for service times in a multiple-driver scenario is defined as

$$U_X^N = \left\{ \begin{array}{c} (X_1, \dots, X_{mq}) \geq 0 : \left| \dfrac{\sum_{i \in I} \sum_{k=(j_i-1)q+1}^{j_i q} (X_k - \mu_k^s)}{\sqrt{\sum_{i \in I} \sum_{k=(j_i-1)q+1}^{j_i q} (\sigma_k^s)^2}} \right| \leq \gamma^s, \\ \forall j_i \in J_i, \quad \forall I \subseteq \left\{1, \dots, \tau^m\right\} \end{array} \right\}, \tag{5}$$

where γ^s plays the same role as γ^a and j_i is the index that refers to the members in set J_i. Simply stated, the uncertainty set U_X^N considers all different interset combinations of service times under all possible partitions for serving m sets by N available crowd drivers; we refer to Bandi et al. (2015) for a more in-depth discussion of this uncertainty set. Similar to the interarrival times uncertainty set, we can model nonstationarity in service times in the set U_X^N via the μ_i^s and σ_i^s parameters.

Finally, as in the robust optimization literature, our uncertainty sets are motivated by limit theorems of probability that typically assume independence of random variables. However, the independence assumption is not required for these uncertainty sets.

2.3 Our Crowdsourcing Model

2.3.1. Service Time Guarantees. We let S_i denote the *system time* of customer i in the crowdsourced delivery system, defined as the time that passes between a customer placing an order until receiving the package, which entails wait time, transportation time, and on-site time. The service levels of all customers are satisfied if $S_i \leq \alpha, \ i = 1, \dots, mq$. If time e is needed for processing and preparation of delivery orders, we should replace α with $\alpha - e$; however, for simplicity in the analysis we consider $e = 0$ throughout. Regarding 3PL delivery, we assume that all customers are guaranteed to receive their packages on time; to realistically capture this assumption, we let c_T be decreasing in α, as supported by FedEx delivery fees. Therefore, we only analyze the system times of customers in the crowdsourced delivery system.

2.3.2. Crowdsourced System Stability Condition. Implicit in our discussion is a queueing system. Customers must wait for their sets (of size q) to form, and customer sets must wait for one of the N available crowd drivers to be free to serve them; in particular, a queue of sets will form while all available drivers are busy serving other sets. Therefore, we require a stability condition for this set queueing system. The service time for this queueing system consists of tour travel time and on-site time. The expected on-site time for a set of q customers is $\frac{\sum_{i=1}^{mq} \mu_i^s}{m}$. From uncertainty set U_L^ℓ, defined in Equation (4), an upper bound on the average tour time for a set of q customers is $\frac{\sum_{i=1}^m \beta_U \left(\sum_{z=1}^M \sqrt{\frac{p_z}{\sum_{j=1}^M \mathbb{A}_j p_j}} \mathbb{A}_z \right) \sqrt{q+1}/v_i}{m}$; in the following analysis, for expository clarity, we replace $\beta_U \left(\sum_{z=1}^M \sqrt{\frac{p_z}{\sum_{j=1}^M \mathbb{A}_j p_j}} \mathbb{A}_z \right)$ with β. Therefore, $\frac{\sum_{i=1}^{mq} \mu_i^s}{m} + \frac{\sum_{i=1}^m \beta \sqrt{q+1}/v_i}{m}$ is an upper bound on the expected service time of each set. The expected interarrival time between sets of q customers, who are served via crowdsourced delivery, is $\frac{\sum_{i=1}^{mq} \mu_i^a / P}{m}$. As a result, $\rho \triangleq P \left(\frac{\sum_{i=1}^{mq} \mu_i^s}{m} + \frac{\sum_{i=1}^m \beta \sqrt{q+1}/v_i}{m} \right) \Big/ \left(N \frac{\sum_{i=1}^{mq} \mu_i^a}{m} \right)$ is an upper bound on the utilization of the queueing system serving sets of q customers. For queue stability, we require ρ to be strictly less than 1: $\rho < 1$. For simplicity in the analysis, we write $\rho < 1$ as $\rho \leq 1 - \varepsilon$, where ε is smaller than any other number.

2.3.3. Crowd Participation Constraint. We assume that there is a sufficiently large and heterogeneous supply of *potential* crowd drivers for the firm to tap into for its delivery work, with the population given by $\mathbb{N}$. The most important factor that motivates the participation of these crowd drivers is financial remuneration (Asdecker and Zirkelbach 2020). The manner in which crowd delivery platforms typically function is that the firm announces the delivery work that it needs along with associated details, including the start date and time, and then, each crowd driver who sees this announcement can choose to sign up for it and provide the requested delivery services, or not, depending on the compensation that the firm offers compared with their opportunity costs. Let K_i be crowd driver i's uncertain opportunity cost, with mean μ_i^k and standard deviation σ_i^k. Drivers have uncertainty in their opportunity costs, driven by different factors, such as the impact of weather or the different days of the week. Similar to the uncertainty set U_X^N, we define a CLT-motivated uncertainty set for the crowd's opportunity costs as

$$U_K = \left\{ (K_1, \dots, K_{\mathbb{N}}) \geq 0 : \left| \frac{\sum_{i \in I} K_i - \mu_i^k}{\sqrt{\sum_{i \in I} \left(\sigma_i^k\right)^2}} \right| \leq \gamma^k, \ \forall I \subseteq \{1, \dots, \mathbb{N}\} \right\}, \tag{6}$$

where γ^k controls the level of conservatism in crowd participation. If the μ_i^k and σ_i^k parameters are not known, the firm can use a single mean $\mu_i^k = \mu^k$ and standard deviation $\sigma_i^k = \sigma^k$, for all i.

During the time horizon that we are solving the problem for, let N denote the number of crowd drivers that the firm needs in order to deliver all packages to its customers on time. The firm can make sure to have at least N willing to participate crowd drivers by offering them a sufficiently high compensation (but not higher than necessary). Each of these N crowd driver's expected hourly income under the utilization of ρ and hourly wage w is ρw or equivalently, $wP\left(\frac{\Sigma_{i=1}^{mq}\mu_i^s}{m}+\frac{\Sigma_{i=1}^{m}\beta\sqrt{q+1}/v_i}{m}\right)\Big/\left(N\frac{\Sigma_{i=1}^{mq}\mu_i^a}{m}\right)$. As crowd drivers are compensated only for their busy times, and not their idle times, the firm should consider the hourly income ρw of crowd drivers and not just their hourly wage w. If the utilization ρ of crowd drivers is small, then the drivers are idle most of the time, which results in a small crowd hourly income, and the crowd is not as willing to participate in on-demand delivery services; as a result, the firm should offer a higher hourly wage to motivate the crowd to participate. However, if the utilization ρ of crowd drivers is high, then customers might not get their deliveries on time because of extended wait times. In the sequel, we analytically solve for the optimal hourly wage the firm should offer to the crowd to minimize its delivery cost while satisfying its customers' orders on time.

Motivated by Taylor (2018), if the expected hourly income of the crowd is at least its opportunity cost, then the crowd participates. We next analyze the opportunity costs of the crowd supply via the set U_K. Similar to Taylor (2018), we assume crowd drivers with higher opportunity costs participate if and only if crowd drivers with lower opportunity costs participate. Thus, if N crowd drivers participate, these are the N cheapest drivers from the supply of available crowd drivers in set U_K. To guarantee the availability of N crowd drivers, we analyze the worst-case opportunity costs of the cheapest N crowd drivers, for a given level of conservatism γ^k. This can be mathematically written as $\min_{S\subseteq\{1,\ldots,\mathbb{N}\},|S|=N}\max_{(K_1,\ldots,K_{\mathbb{N}})\in U_K}\frac{\Sigma_{i\in S}K_i}{N}$. In Lemma 1, we solve for these optimization problems. For simplicity in the analysis, we assume that μ_i^k depends on i but $\sigma_i^k=\sigma^k,\quad \forall i$ does not (as in, for instance, linear discriminant analysis).

Lemma 1.

$$\min_{\substack{S\subseteq\{1,\ldots,\mathbb{N}\},\\ |S|=N}}\ \max_{(K_1,\ldots,K_{\mathbb{N}})\in U_K}\frac{\sum_{i\in S}K_i}{N}=\frac{\sum_{i=1}^{N}\mu_{(i)}^k+\gamma^k\sqrt{N}\sigma^k}{N},$$

where $\mu_{(1)}^k\le\mu_{(2)}^k\le\cdots\le\mu_{(\mathbb{N})}^k$ are the order statistics of the mean parameters.

The term $\frac{\Sigma_{i=1}^{N}\mu_{(i)}^k+\gamma^k\sqrt{N}\sigma^k}{N}$ is the worst-case average opportunity cost of the cheapest crowd drivers of size N, with the level of conservatism set at γ^k. Motivated by Taylor (2018), the number of participating crowd drivers satisfies $wP\left(\frac{\Sigma_{i=1}^{mq}\mu_i^s}{m}+\frac{\Sigma_{i=1}^{m}\beta\sqrt{q+1}/v_i}{m}\right)\Big/\left(N\frac{\Sigma_{i=1}^{mq}\mu_i^a}{m}\right)\ge\frac{\Sigma_{i=1}^{N}\mu_{(i)}^k+\gamma^k\sqrt{N}\sigma^k}{N}$. This constraint guarantees that the participation constraint of the crowd drivers is satisfied with $P(|Z|\le\gamma^k)$ probability, where Z is a standard normal random variable (because of the CLT). The firm can influence the availability of crowd drivers through (a) the hourly wage and (b) the amount of on-demand delivery work it offers to the crowd. In Proposition 1, we show the inequality is tight at optimality, and we obtain the crowd supply function $w=\left(\Sigma_{i=1}^{N}\mu_{(i)}^k+\gamma^k\sqrt{N}\sigma^k\right)P\left(\frac{\Sigma_{i=1}^{mq}\mu_i^s}{m}+\frac{\Sigma_{i=1}^{m}\beta\sqrt{q+1}/v_i}{m}\right)\Big/\left(\frac{\Sigma_{i=1}^{mq}\mu_i^a}{m}\right)$ that *endogenously* determines the number of available crowd drivers. In this crowd supply function, the availability of the crowd, namely N, is an increasing function of the hourly wage offered by the firm, an increasing function of the amount of on-demand delivery work, and a decreasing function of the crowd opportunity costs. According to this supply function, as variability in the opportunity cost of the crowd increases, the firm should offer a higher hourly wage w to guarantee the availability of crowd drivers.

We next analyze the firm's objective function. The drivers' expected value-adding time for m sets of size q consists of traversing m tours, which takes $\Sigma_{j=1}^{m}L_{q+1}^{j}/v_j$ time, plus the on-site service time for mq customers, namely $\Sigma_{i=1}^{mq}X_i$. The sum of these expressions, divided by mq, is the value-adding time per package (e.g., hour per package). If we multiply this value-adding time per package by w, the driver's wage rate per time unit (e.g., dollars per hour), we obtain the crowdsourced delivery cost rate per package (e.g., dollars per package) as $w\frac{\Sigma_{j=1}^{m}L_{q+1}^{j}/v_j+\Sigma_{i=1}^{mq}X_i}{mq}$. As detailed previously, a P proportion of customers is served via crowdsourced delivery, and the remaining $(1-P)$ proportion is served via a 3PL firm at cost c_T, which we consider to be the average of the 3PL delivery costs for all customers with different distances to the depot. Thus, the expected delivery cost rate per package is $Pw\frac{\Sigma_{j=1}^{m}L_{q+1}^{j}/v_j+\Sigma_{i=1}^{mq}X_i}{mq}+(1-P)c_T$; this expression will serve as our objective, to be minimized by appropriately selecting P and w. We subsequently optimize over q numerically as well. To obtain a conservative estimate of the impact of crowdsourcing last-mile deliveries, we solve the following robust optimization model:

$$\begin{aligned}
\min_{P\in(0,1],w>0}\quad & \max_{\substack{(X_1,\ldots,X_{mq})\in U_X^N\\ L_{q+1}^j/v_j\in U_L^j, j=1,\ldots,m}} Pw\left(\frac{\sum_{j=1}^{m}L_{q+1}^{j}/v_j+\sum_{i=1}^{mq}X_i}{mq}\right)+(1-P)c_T\\
\text{s.t.}\quad & \max_{\substack{(A_1,\ldots,A_{mq})\in U_A(P)\\ (X_1,\ldots,X_{mq})\in U_X^N\\ L_{q+1}^j/v_j\in U_L^j, j=1,\ldots,m}} S_i\le\alpha,\ i=1,\ldots,mq\\
& P\left(\frac{\sum_{i=1}^{mq}\mu_i^s}{m}+\frac{\sum_{i=1}^{m}\beta\sqrt{q+1}/v_i}{m}\right)\Bigg/\left(N\frac{\sum_{i=1}^{mq}\mu_i^a}{m}\right)\le 1-\varepsilon\\
& wP\left(\frac{\sum_{i=1}^{mq}\mu_i^s}{m}+\frac{\sum_{i=1}^{m}\beta\sqrt{q+1}/v_i}{m}\right)\Bigg/\left(N\frac{\sum_{i=1}^{mq}\mu_i^a}{m}\right)\\
& \quad\ge\min_{\substack{S\subseteq\{1,\ldots,\mathbb{N}\},\\ |S|=N}}\ \max_{(K_1,\ldots,K_{\mathbb{N}})\in U_K}\frac{\sum_{i\in S}K_i}{N}.
\end{aligned}\tag{7}$$

The objective function is the maximum expected delivery cost rate per package, with respect to uncertainties that exist in the crowd delivery system. The left-hand side of the first constraint is the worst-case system time for customer i, with respect to uncertainties in the crowd delivery system. This constraint guarantees that the system time is at most the service level α; the index i refers to any customer who is served via crowdsourced delivery. The second constraint ensures the underlying queue is stable. The third constraint is the crowd participation constraint that endogenously determines the number N of participating independent crowd drivers. In the proposed robust queueing system, the number of service providers is not exogenous and is endogenously determined by the crowd compensation. Similarly, Dong and Ibrahim (2020) study queueing systems with a random number of servers for on-demand service providers using fluid and stochastic-fluid approximations, where the firm can use either full-time employees or flexible workers.

Problem (7) allows a firm to intelligently decide on crowd compensation and labor planning to minimize logistical costs while making sure that there are enough willing to participate crowd drivers that are (1) not too busy to risk late deliveries or (2) not too idle, where the drivers' effective earning rate becomes too low for participation. In the next section, we simplify Problem (7) by solving the inner maximization problems over the uncertainty sets $U_A(P)$, U_X^N, U_L^j, and U_K.

2.4. Solution Methodology

In Lemma 2, we analyze the objective function of Problem (7) and derive an attainable upper bound for it; in the remainder of our analysis of the robust model, we utilize this bound as our objective.

Lemma 2.

$$\max_{\substack{(X_1,\dots,X_{mq})\in U_X^N \\ L_{q+1}^j/v_j\in U_L^j, j=1,\dots,m}} Pw\left(\frac{\sum_{j=1}^m L_{q+1}^j/v_j+\sum_{i=1}^{mq}X_i}{mq}\right)+(1-P)c_T \le \frac{\sum_{j=1}^m \beta\sqrt{q+1}/v_j + \sum_{j=1}^N\left(\sum_{i=1}^{\tau^m}\sum_{k=(j_i-1)q+1}^{j_iq}\mu_k^s+\gamma^s\sqrt{\sum_{i=1}^{\tau^m}\sum_{k=(j_i-1)q+1}^{j_iq}\left(\sigma_k^s\right)^2}\right)}{mq}.$$

We next study the first constraint of Problem (7), and for this purpose, we analyze the worst-case system times of customers. Bandi et al. (2015) study the worst-case system time of the last customer in a queue formed by customers. However, in this paper, we are studying the worst-case system times of *all* customers in a queue formed by *sets* of customers, whose arrival process is a superposition of customer arrival processes, interarrival times and service times are nonstationary. There is also an additional routing layer whose duration depends on traffic patterns.

In order to analyze the customers' system times, we first let $D_{(j-1)q+k}$, $k=1,\dots,q$ denote the *departure time* of customer k in set $j=1,\dots,m$. We provide set-specific upper bounds T_j for the customers' departure times D_k. The first N sets of customers start getting served after they are formed, as there are N idle drivers in the system at time zero. Therefore, we have the following upper bounds for all customer departure times in the first N sets:

$$\text{(Set1)}\quad D_k\le T_1\triangleq\sum_{\ell=1}^q A_\ell+\sum_{\ell=1}^q X_\ell+L_{q+1}^1/v_1,\ k=1,\dots,q \tag{8}$$

$$\text{(Set2)}\quad D_{q+k}\le T_2\triangleq\sum_{\ell=1}^{2q} A_\ell+\sum_{\ell=q+1}^{2q} X_\ell+L_{q+1}^2/v_2,\ k=1,\dots,q \tag{9}$$

$$\dots$$

$$\text{(Set}\,N)\quad D_{(N-1)q+k}\le T_N\triangleq\sum_{\ell=1}^{Nq} A_\ell+\sum_{\ell=(N-1)q+1}^{Nq} X_\ell+L_{q+1}^N/v_N,\quad k=1,\dots,q. \tag{10}$$

For instance, the bound T_2 for set 2 consists of three components: (1) $\Sigma_{\ell=1}^{2q}A_\ell$ is the time when set 2 is formed and ready for service, (2) $\Sigma_{\ell=q+1}^{2q}X_\ell$ is the total on-site service time for all the customers in set 2, and (3) L_{q+1}^2/v_2 is the optimal tour duration to visit all customers in set 2.

A common difficulty in the analysis of multiserver queues is that overtaking can occur, in that set i arrives after set j, but is serviced and leaves the system first. This results in the order of arrivals being different than the order of departures, which creates technical difficulties in the analysis. We perform an analysis motivated by that in Bandi et al. (2015), which precludes overtaking from happening, and sets depart in the same order that they arrive to the crowdsourcing queue. We denote the policy that precludes overtaking as Θ, and the subsequent results are for this policy. Later, we show that the worst-case system time under policy Θ is equal to the worst-case system time under the FCFS policy.

Motivated by Krivulin (1994), which generalized the recursive equations of Lindley (1952) to the multiserver $G/G/m$ queue, we define the following recursive set of upper bounds for the departure times of customers in the subsequent sets $N+1,N+2,\dots,m$, which might require waiting for a driver to become free:

$$D_{(j-1)q+k}^\Theta\le T_j^\Theta\triangleq\max\left\{\sum_{\ell=1}^{jq}A_\ell,T_{j-N}^\Theta\right\}+\sum_{\ell=(j-1)q+1}^{jq}X_\ell+L_{q+1}^j/v_j,\quad j=N+1,\dots,m,\ k=1,\dots,q, \tag{11}$$

where for $j=1,\dots,N$, $T_j^\Theta=T_j$ as defined in Equations (8)–(10), $\Sigma_{\ell=1}^{jq}A_\ell$ is the time when the jth set is formed and ready for service; $\max\left\{\Sigma_{\ell=1}^{jq}A_\ell,T_{j-N}^\Theta\right\}$ is the first time that set j begins service by a free driver, where T^{Θ}_{j-N} is the exact time that a driver becomes free from serving the $(j-N)$ th departed set.

Customers' system times can be calculated from the departure times as follows:

$$S_{(j-1)q+k} = D_{(j-1)q+k} - \sum_{\ell=1}^{(j-1)q+k} A_\ell,\ \ k=1,\dots,q,\ \ j=1,\dots,m, \tag{12}$$

where $\Sigma_{\ell=1}^{(j-1)q+k} A_\ell$ is the arrival time of customer k in the jth set to the system. In Lemma 3, we utilize the bounds on departure times to derive upper bounds on the system times S_i for customers $i=1,\dots,mq$.

Lemma 3. *System times of customers k in set j in a multiple-driver queue under policy* Θ *have the following upper bounds:*

$$S^{\Theta}_{(j-1)q+k} \le B^{\Theta}_{(j-1)q+k} \triangleq \sum_{\ell=(j-1)q+k+1}^{jq} A_\ell + \sum_{\ell=(j-1)q+1}^{jq} X_\ell + L^j_{q+1}/v_j,\ \ k=1,\dots,q,\ \ j=1,\dots N$$

$$S^{\Theta}_{(j-1)q+k} \le B^{\Theta}_{(j-1)q+k} \triangleq \max_{i=0,\dots a^j}\left\{\sum_{\ell=1}^{(j-iN)q} A_\ell + \sum_{\ell=0}^{i}\left(\sum_{p=(j-\ell N-1)q+1}^{(j-\ell N)q} X_p + \frac{L^{j-\ell N}_{q+1}}{v_{j-\ell N}}\right)\right\} - \sum_{\ell=1}^{(j-1)q+k} A_\ell,\ \ k=1,\dots,q,\ \ j\ge N+1,$$

where $a^j = \lceil \frac{j}{N} \rceil - 1$.

Lemma 3 defines B^{Θ}_i as the upper bound on S^{Θ}_i, for any i. Considering the bounds for the customers within an arbitrary set $j\in\{1,\dots,m\}$, we observe that $B^{\Theta}_{(j-1)q+1} \ge B^{\Theta}_{(j-1)q+2} \ge \cdots \ge B^{\Theta}_{jq}$; in other words, the upper bound for the ith customer assigned to the set is at least the upper bound of the $(i+1)$ th customer, and the first customer assigned to the set (in FCFS) has the largest upper bound (because of the largest wait time for the set to form). Therefore, we have the following implications: $B^{\Theta}_{(j-1)q+1} \le \alpha \Rightarrow B^{\Theta}_{(j-1)q+k} \le \alpha,\ k=1,\dots,q,\ j=1,\dots,m$, which further imply

$$\max_{\substack{(A_1,\dots,A_{mq})\in U_A(P)\\ (X_1,\dots,X_{mq})\in U_X^N\\ L^j_{q+1}/v_j\in U^j_L, j=1,\dots,m}} B^{\Theta}_{(j-1)q+1} \le \alpha,\ \ j=1,\dots,m \Rightarrow \max_{\substack{(A_1,\dots,A_{mq})\in U_A(P)\\ (X_1,\dots,X_{mq})\in U_X^N\\ L^j_{q+1}/v_j\in U^j_L, j=1,\dots,m}} S^{\Theta}_i \le \alpha,\ \ i=1,\dots,mq.$$

Consequently, we replace the $n=mq$ service constraints of Problem (7) with the m constraints in the first inequality. In the next set of results, we analyze this constraint, namely the maximum upper bound on the system time of the first customer in the jth set, $j\in\{1,\dots,m\}$. Letting $\mathbf{L}/\mathbf{v} = (L^1_{q+1}/v_1,\dots,L^m_{q+1}/v_m)$, $\mathbf{A} = (A_1,\dots,A_{mq})$, and $\mathbf{X} = (X_1,\dots,X_{mq})$, we solve the inner optimization problems, for $j=1,\dots,m$, of the first constraint sequentially:

$$\max_{\substack{\mathbf{A}\in U_A(P)\\ \mathbf{X}\in U_X^N\\ \mathbf{L}/\mathbf{v}\in \times_{j=1}^m U^j_L}} B^{\Theta}_{(j-1)q+1}(\mathbf{L}/\mathbf{v},\mathbf{A},\mathbf{X}) = \max_{\mathbf{L}/\mathbf{v}\in\times_{j=1}^m U^j_L}\left\{\max_{\mathbf{A}\in U_A(P)}\left\{\max_{\mathbf{X}\in U_X^N} B^{\Theta}_{(j-1)q+1}(\mathbf{L}/\mathbf{v},\mathbf{A},\mathbf{X})\right\}\right\}. \tag{13}$$

We first analyze the innermost optimization problem over U^N_X in the right-hand side of Equation (13), whose solution is presented in Lemma 4.

Lemma 4. *For* $j=1,\dots,m$, *we have that*

$$\max_{\mathbf{X}\in U_X^N} B^{\Theta}_{(j-1)q+1}\left(\frac{\mathbf{L}}{\mathbf{v}},\mathbf{A},\mathbf{X}\right) \le \begin{cases} \displaystyle\sum_{\ell=(j-1)q+2}^{jq} A_\ell + \sum_{\ell=(j-1)q+1}^{jq}\mu^s_\ell + \gamma^s\sqrt{\sum_{\ell=(j-1)q+1}^{jq}(\sigma^s_\ell)^2} + L^j_{q+1}/v_j,\ \ j=1,\dots,N \\ \displaystyle\max_{i=0,\dots a^j}\left\{\begin{array}{l}\displaystyle\sum_{\ell=1}^{(j-iN)q} A_\ell + \sum_{\ell=0}^{i}\sum_{p=(j-\ell N-1)q+1}^{(j-\ell N)q}\mu^s_p \\ \displaystyle+\gamma^s\sqrt{\sum_{\ell=0}^{i}\sum_{p=(j-\ell N-1)q+1}^{(j-\ell N)q}(\sigma^s_p)^2} + \sum_{\ell=0}^{i}\frac{L^{j-\ell N}_{q+1}}{v_{j-\ell N}} \\ \displaystyle-\sum_{\ell=1}^{(j-1)q+1} A_\ell\end{array}\right\},\ \ j=N+1,\dots,m.\end{cases}$$

In Lemma 5, we solve the second-tier optimization over $U_A(P)$ in the right-hand side of Equation (13).

Lemma 5. *For* $j=1,\dots,m$, *we have the following upper bounds:*

$$\max_{\mathbf{A}\in U_A(P)}\left\{\max_{\mathbf{X}\in U_X^N} B^{\Theta}_{(j-1)q+1}(\mathbf{L}/\mathbf{v},\mathbf{A},\mathbf{X})\right\} \le \begin{cases} g^N, & j=1,\dots,N \\ \max\left\{\displaystyle\max_{i=1,\dots,a^j} f^N_i, g^N\right\}, & j=N+1,\dots,m,\end{cases}$$

where $g^N = \Sigma_{\ell=(j-1)q+2}^{jq}\mu^a_\ell/P + \gamma^a\sqrt{\Sigma_{\ell=(j-1)q+2}^{jq}(\sigma^a_\ell/P)^2} + \Sigma_{\ell=(j-1)q+1}^{jq}\mu^s_\ell + \gamma^s\sqrt{\Sigma_{\ell=(j-1)q+1}^{jq}(\sigma^s_\ell)^2} + L^j_{q+1}/v_j$ *and* $f^N_i = -\Sigma_{\ell=(j-iN)q+1}^{(j-1)q+1}\mu^a_\ell/P + \gamma^a\sqrt{\Sigma_{\ell=(j-iN)q+1}^{(j-1)q+1}(\sigma^a_\ell/P)^2} + \Sigma_{\ell=0}^{i}\Sigma_{p=(j-\ell N-1)q+1}^{(j-\ell N)q}\mu^s_p + \gamma^s\sqrt{\Sigma_{\ell=0}^{i}\Sigma_{p=(j-\ell N-1)q+1}^{(j-\ell N)q}(\sigma^s_p)^2} + \Sigma_{\ell=0}^{i} L^{j-\ell N}_{q+1}/v_{j-\ell N}$.

In Lemma 6, we solve the final outer optimization over $\times_{j=1}^m U^j_L$ in the right-hand side of Equation (13).

Lemma 6. *For $j = 1, \dots, m$, we have the following upper bounds:*

$$\max_{\mathbf{L}/\mathbf{v} \in \times_{j=1}^m U_L^j} \left\{ \max_{\mathbf{A} \in U_A(P)} \left\{ \max_{\mathbf{X} \in U_X^N} B^{\Theta}_{(j-1)q+1}(\mathbf{L}/\mathbf{v}, \mathbf{A}, \mathbf{X}) \right\} \right\} \le \begin{cases} \hat{g}^N, & j = 1, \dots, N \\ \max\left\{ \max\limits_{i=1,\dots,a^j} \hat{f}_i^N, \hat{g}^N \right\}, & j = N+1, \dots, m, \end{cases}$$

where $\hat{g}^N = \Sigma_{\ell=(j-1)q+2}^{jq} \mu_\ell^a / P + \gamma^a \sqrt{\Sigma_{\ell=(j-1)q+2}^{jq} (\sigma_\ell^a/P)^2} + \Sigma_{\ell=(j-1)q+1}^{jq} \mu_\ell^s + \gamma^s \sqrt{\Sigma_{\ell=(j-1)q+1}^{jq} (\sigma_\ell^s)^2} + \beta\sqrt{q+1}/v_j$ *and* $\hat{f}_i^N = -\Sigma_{\ell=(j-iN)q+1}^{(j-1)q+1} \mu_\ell^a / P + \gamma^a \sqrt{\Sigma_{\ell=(j-iN)q+1}^{(j-1)q+1} (\sigma_\ell^a/P)^2} + \Sigma_{\ell=0}^{i} \Sigma_{p=(j-\ell N-1)q+1}^{(j-\ell N)q} \mu_p^s + \gamma^s \sqrt{\Sigma_{\ell=0}^{i} \Sigma_{p=(j-\ell N-1)q+1}^{(j-\ell N)q} (\sigma_p^s)^2} + \Sigma_{\ell=0}^{i} \beta\sqrt{q+1}/v_{j-\ell N}$.

The next result shows that the bounds B_i^{Θ} are also applicable to the FCFS policy.

Lemma 7. *The system time of customer i, $i = 1, \dots, mq$, under the FCFS policy S_i^{FCFS} is at most the bound B_i^{Θ} defined in Lemma 3 for the no-overtaking policy Θ:*

$$S_i^{\text{FCFS}}(\mathbf{L}/\mathbf{v}, \mathbf{A}, \mathbf{X}) \le B_i^{\Theta}(\mathbf{L}/\mathbf{v}, \mathbf{A}, \mathbf{X}), \quad \forall \mathbf{X} \in U_X^N, \quad \forall \mathbf{A} \in U_A(P), \quad \forall \mathbf{L}/\mathbf{v} \in \bigtimes_{j=1}^{m} U_L^j.$$

As a result of Lemma 7, all of the analyses of the Θ policy are applicable and equivalent to the FCFS policy; we therefore replace, in the sequel, the Θ superscript with *FCFS*. Thus, Lemmas 1–6 allow us to define the following approximation to Problem (7):

$$\begin{aligned} z = \min_{P \in (0,1],\, w>0} \; & Pw \left(\frac{\sum_{j=1}^{m} \beta\sqrt{q+1}/v_j + \sum_{j=1}^{N} \left(\sum_{i=1}^{\tau^m} \sum_{k=(j_i-1)q+1}^{j_i q} \mu_k^s + \gamma^s \sqrt{\sum_{i=1}^{\tau^m} \sum_{k=(j_i-1)q+1}^{j_i q} (\sigma_k^s)^2} \right)}{mq} \right) + (1-P)c_T \\ \text{s.t.} \; & \max_{i=1,\dots,a^j} \left\{ -\sum_{\ell=(j-iN)q+1}^{(j-1)q+1} \mu_\ell^a \Big/ P + \gamma^a \sqrt{\sum_{\ell=(j-iN)q+1}^{(j-1)q+1} (\sigma_\ell^a/P)^2} + \sum_{\ell=0}^{i} \sum_{p=(j-\ell N-1)q+1}^{(j-\ell N)q} \mu_p^s + \gamma^s \sqrt{\sum_{\ell=0}^{i} \sum_{p=(j-\ell N-1)q+1}^{(j-\ell N)q} \left(\sigma_p^s\right)^2} + \sum_{\ell=0}^{i} \frac{\beta\sqrt{q+1}}{v_{j-\ell N}} \right\} \le \alpha, \; j = N+1, \dots, m \\ & \sum_{\ell=(j-1)q+2}^{jq} \mu_\ell^a/P + \gamma^a \sqrt{\sum_{\ell=(j-1)q+2}^{jq} (\sigma_\ell^a/P)^2} + \sum_{\ell=(j-1)q+1}^{jq} \mu_\ell^s + \gamma^s \sqrt{\sum_{\ell=(j-1)q+1}^{jq} (\sigma_\ell^s)^2} + \frac{\beta\sqrt{q+1}}{v_j} \le \alpha, j = 1, \dots, m \\ & P\left(\frac{\sum_{i=1}^{mq} \mu_i^s}{m} + \frac{\sum_{i=1}^{m} \beta\sqrt{q+1}/v_i}{m} \right) \Big/ \left(N \frac{\sum_{i=1}^{mq} \mu_i^q}{m} \right) \le 1 - \varepsilon \\ & wP\left(\frac{\sum_{i=1}^{mq} \mu_i^s}{m} + \frac{\sum_{i=1}^{m} \beta\sqrt{q+1}/v_i}{m} \right) \Big/ \left(N \frac{\sum_{i=1}^{mq} \mu_i^a}{m} \right) \ge \frac{\sum_{i=1}^{N} \mu_{(i)}^k + \gamma^k \sqrt{N} \sigma^k}{N}. \end{aligned} \tag{14}$$

For each customer set j, $i = 1, \dots, a^j$ indexes the customer sets formed before set j that are also served by the same crowd driver who serves set j. For analyzing the system times in Problem (14), let $E_i^j \triangleq \Sigma_{\ell=(j-iN)q+1}^{(j-1)q+1} \mu_\ell^a - \gamma^a \sqrt{\Sigma_{\ell=(j-iN)q+1}^{(j-1)q+1} (\sigma_\ell^a)^2}$ denote the best-case interarrival time between set j and the ith preceding set, all served by the same driver. Let $F_i^j \triangleq \Sigma_{\ell=0}^{i} \Sigma_{p=(j-\ell N-1)q+1}^{(j-\ell N)q} \mu_p^s + \gamma^s \sqrt{\Sigma_{\ell=0}^{i} \Sigma_{p=(j-\ell N-1)q+1}^{(j-\ell N)q} (\sigma_p^s)^2} + \Sigma_{\ell=0}^{i} \frac{\beta\sqrt{q+1}}{v_{j-\ell N}}$ denote the cumulative worst-case system times (including the on-site service times and tour times) of set j and the i preceding sets, all of which served by the same driver. Finally, let $G_j \triangleq \Sigma_{\ell=(j-1)q+2}^{jq} \mu_\ell^a + \gamma^a \sqrt{\Sigma_{\ell=(j-1)q+2}^{jq} (\sigma_\ell^a)^2}$ denote the worst-case waiting time for the first customer in set j until her set forms. Using these terms, the closed form solutions to Problem (14) are provided in Proposition 1.

Proposition 1. *If $\alpha > \max_{j=1,\dots,m}\left\{F_{i=0}^j\right\}$, then Problem (14) is feasible if and only if $\hat{P} \le \min\left\{\bar{P}^N, \tilde{P}, 1\right\}$; if feasible, the optimal solution is $P^* = \min\left\{\bar{P}^N, \tilde{P}, 1\right\}$, with the crowd supply function* $w^* = \dfrac{\frac{\Sigma_{i=1}^N \mu_{(i)}^k + \gamma^k \sqrt{N} \sigma^k}{N} \left(\frac{\Sigma_{i=1}^{mq} \mu_i^a}{m} \right)}{P^* \left(\frac{\Sigma_{i=1}^{mq} \mu_i^s}{m} + \frac{\Sigma_{i=1}^{m} \beta\sqrt{q+1}/v_i}{m} \right)}$, *and the minimized worst-case delivery cost rate* $z(N) = P^* \left(w^* \frac{\Sigma_{j=1}^m \beta\sqrt{q+1}/v_j + \Sigma_{j=1}^N \left(\Sigma_{i=1}^{\tau^m} \Sigma_{k=(j_i-1)q+1}^{j_i q} \mu_k^s + \gamma^s \sqrt{\Sigma_{i=1}^{\tau^m} \Sigma_{k=(j_i-1)q+1}^{j_i q} (\sigma_k^s)^2} \right)}{mq} - c_T \right) + c_T$,

where $\bar{P}^N = \min_{j=N+1,\dots,m} \left\{ \min_{i \in s_2^j} \left\{ \frac{E_i^j}{F_i^j - \alpha} \right\} \right\}, \hat{P} = \max_{j=1,\dots,m} \left\{ \frac{G_j}{\alpha - F_{i=0}^j} \right\}$,

$\tilde{P} = \dfrac{(1-\varepsilon) N \frac{\Sigma_{i=1}^{mq} \mu_i^a}{m}}{\frac{\Sigma_{i=1}^{mq} \mu_i^s}{m} + \frac{\Sigma_{i=1}^m \beta\sqrt{q+1}/v_i}{m}}$, $s_2^j = \left\{ i = 1, \dots, a^j : \alpha - F_i^j < 0 \right\}$, $j = N+1, \dots, m$.

According to Proposition 1, the optimal hourly wage that the firm should offer to crowd drivers changes with the delivery time window and the different uncertainties that exist in the delivery system, including the uncertainty in customer purchasing patterns, delivery service times, and crowd opportunity costs. In particular, the hourly wage that the firm should offer to crowd drivers is higher as the uncertainty in their opportunity costs (i.e., σ^k) increases; this is to ensure their participation. The precise nature of the dependence of the optimal crowd hourly wage on the other various uncertainties in the system is not straightforward (i.e., nonlinear) because of the underlying queue; we study this dependence in Section 5.1 via a numerical analysis. Finally, if the firm wants to ensure crowd participation with an even higher probability (i.e., increasing γ^k), the hourly wage that the firm should offer is even higher.

Our analysis shows that there exists an upper bound on P, denoted by $\bar{P}^N$, that, if exceeded, results in the on-time delivery of packages being at risk. $\tilde{P}$ is another upper bound on P, which ensures that the underlying queue in the crowd delivery system is stable. $\hat{P}$ is a lower bound on P that guarantees that sufficient delivery orders are assigned to the available crowd drivers such that customers do not wait too long for the formation of their delivery tours and thus, can receive their orders within α. Respecting these bounds means that the number of packages assigned to crowd drivers is large enough to help with the reduction of logistical cost but at the same time, small enough to not jeopardize the promise of on-time delivery. In other words, if firms want to use a crowdsourced system for their last-mile deliveries, they cannot assign too little nor too much work to the crowd—it has to be a moderate workload.

3. Benchmark 1: A Stochastic Crowdsourcing Model

In this section, we propose and solve a stochastic counterpart to our robust model, where we replace the worst-case expressions by expected values; we aim to compare our worst-case analysis approach against an expected value analysis approach. The comparison of these two models, in Section 5, will evaluate the benefit of using robust optimization over stochastic optimization. Furthermore, although trend and seasonality can be rather easily incorporated into the robust model, we found them to be analytically intractable to incorporate into our stochastic model, which is based on the analysis in Bertsimas and Van Ryzin (1993); consequently, we develop the stochastic model for stationary random variables.

We assume customers arrive according to a Poisson Process with rate λ; similar to the robust crowdsourcing model, a proportion P of customers is randomly assigned to crowdsourced delivery, and the remaining $(1-P)$ proportion is served via a 3PL firm. The mean interarrival time between crowdsourced customers is $\frac{1}{P\lambda}$. We again form sets of q customers based on an FCFS policy, and these sets are served by the first available crowd driver, who then visits the customers within a set using the optimal TSP tour. The time for a set to form has an Erlang distribution with mean $\frac{q}{P\lambda}$ and variance $\frac{q}{P^2\lambda^2}$. The on-site service times have a general distribution, with mean μ^s and variance σ_s^2. Customer locations are distributed according to the density $f(x)$, and L_{q+1} again denotes the shortest tour through q customer locations; according to Bertsimas and Van Ryzin (1993), $E[L_{q+1}]=\hat{\beta}\sqrt{q+1}$, where from Equation (3), we set $\hat{\beta}=\tilde{\beta}\Sigma_{i=1}^{M}\sqrt{\frac{p_i}{\Sigma_{j=1}^{M}\mathbb{A}_j p_j}}\mathbb{A}_i$. We denote $\hat{\beta}\sqrt{q+1}/v$ and $\sigma_{L_{q+1}}^2/v^2$ as the mean and variance of the tour duration L_{q+1}/v, respectively, where v is the average travel speed for all crowd delivery tours.

We can model the queue formed by the sets as a $G/G/m$ queue. Assuming N crowd drivers are willing to provide on-demand delivery services for the firm, the utilization of the crowd delivery system is $\frac{P\lambda}{qN}\left(\hat{\beta}\sqrt{q+1}/v+q\mu^s\right)$. We assume that the opportunity costs of available crowd drivers are distributed according to a general distribution with mean μ^k and standard deviation σ^k. Similar to the approach in the robust model, the number of participating crowd drivers satisfies $w\frac{P\lambda}{qN}\left(\hat{\beta}\sqrt{q+1}/v+q\mu^s\right)\geq\mu^k$; in Proposition 2, we show this inequality is tight, and we obtain the supply function $w=N\mu^k\frac{q}{P\lambda(\hat{\beta}\sqrt{q+1}/v+q\mu^s)}$ that endogenously determines the number of participating crowd drivers, as a function of the crowd hourly wage and the delivery work load.

Using a similar analysis to that in Bertsimas and Van Ryzin (1993), we calculate the expected system times of customers. The expected time a customer must wait until her set forms is $W_{\text{expected-time-to-form-a-set}}=\frac{1}{q}\left(0+\frac{1}{\lambda}+\frac{2}{\lambda}+\cdots+\frac{q-1}{\lambda}\right)=\frac{(q-1)}{2\lambda}$. Using the heavy traffic limit introduced by Kingman (1974), we write an approximation for the expected waiting time in queue (for a set of size q) as $W_{\text{expected-waiting-time-for-sets}}\approx\frac{\frac{P\lambda}{q}\left(\frac{q}{P^2\lambda^2}+\frac{1}{N^2}\left(\sigma_L{}^2/v^2+q\sigma_s{}^2\right)\right)}{2\left(1-\frac{P\lambda}{qN}\left(\hat{\beta}\sqrt{q+1}/v+q\mu^s\right)\right)}$. This approximation becomes more accurate as the system utilization increases toward one. In our simulation study, covering the Seattle area, the utilization of crowd drivers under the stochastic policy is at least 0.8, which is high enough for this heavy traffic approximation (Whitt 1993). Hence, we assume the quality of the Kingman (1974) approximation for our stochastic analysis is practically reasonable.

Finally, the expected wait time for a customer to be served, after the driver starts serving the customer's set, is at most $W_{\text{expected-service-time-of-sets}}=\hat{\beta}\sqrt{q+1}/v+\frac{1}{q}(\mu^s+2\mu^s+\cdots+q\mu^s)=\hat{\beta}\sqrt{q+1}/v+\frac{(q+1)\mu^s}{2}$. As a result, the expected system time of a customer in the stable crowdsourced delivery system (i.e., $\frac{P\lambda}{qN}(\hat{\beta}\sqrt{q+1}/v+q\mu^s)\leq 1-\varepsilon$) is at most $\frac{(q-1)}{2\lambda}+\frac{\frac{P\lambda}{q}\left(\frac{q}{P^2\lambda^2}+\frac{1}{N^2}\left(\sigma_L{}^2/v^2+q\sigma_s{}^2\right)\right)}{2\left(1-\frac{P\lambda}{qN}\left(\hat{\beta}\sqrt{q+1}/v+q\mu^s\right)\right)}+\hat{\beta}\sqrt{q+1}/v+\frac{(q+1)\mu^s}{2}$. As a result of the analysis, a stochastic optimization variant of Model (7) can be written as the following model, with its solution provided in the subsequent proposition:

$$
\begin{aligned}
z_s=\min_{P\in(0,1],w>0}\quad & Pw\left(\frac{\hat{\beta}\sqrt{q+1}/v}{q}+\mu^s\right)+(1-P)c_T\\
\text{s.t.}\quad & \frac{(q-1)}{2\lambda}+\frac{\frac{P\lambda}{q}\left(\frac{q}{P^2\lambda^2}+\frac{1}{N^2}\left(\sigma_L{}^2/v^2+q\sigma_s{}^2\right)\right)}{2\left(1-\frac{P\lambda}{qN}\left(\hat{\beta}\sqrt{q+1}/v+q\mu^s\right)\right)}\\
& +\hat{\beta}\sqrt{q+1}/v+\frac{(q+1)\mu^s}{2}\leq\alpha\\
& \frac{P\lambda}{qN}\left(\hat{\beta}\sqrt{q+1}/v+q\mu^s\right)\leq 1-\varepsilon\\
& w\frac{P\lambda}{qN}\left(\hat{\beta}\sqrt{q+1}/v+q\mu^s\right)\geq\mu^k.
\end{aligned}
\tag{15}
$$

Proposition 2. *If* $\frac{(q-1)}{2\lambda}+\hat{\beta}\sqrt{q+1}/v+\frac{(q+1)\mu^s}{2}\leq\alpha$ *and* $\Delta\geq 0$, *then Problem (15) is feasible if and only if*

$P_1 \le \min\{P_2, P_3, 1\}$; *if feasible, the optimal solution is* $P^* = \min\{P_2, P_3, 1\}$ *with the crowd supply function* $w^* = \frac{\mu^k qN}{P^* \lambda(\hat{\beta}\sqrt{q+1}/v + q\mu^s)}$, *and the minimized delivery cost rate* $z_s(N) = \frac{\mu^k N}{\lambda} + (1 - P^*)c_T$, *where* $U = \frac{\lambda}{qN^2}\left(\sigma_L{}^2/v^2 + q\sigma_s{}^2\right) + \frac{2\lambda}{qN}\left(\alpha - \frac{(q-1)}{2\lambda} - \hat{\beta}\sqrt{q+1}/v - \frac{(q+1)\mu^s}{2}\right)\left(\hat{\beta}\sqrt{q+1}/v + q\mu^s\right)$, $\Delta = \left(\alpha - \frac{(q-1)}{2\lambda} - \hat{\beta}\sqrt{q+1}/v - \frac{(q+1)\mu^s}{2}\right)^2 - 4U/\lambda$, $P_1 = \frac{\alpha - \frac{(q-1)}{2\lambda} - \hat{\beta}\sqrt{q+1}/v - \frac{(q+1)\mu^s}{2} - \sqrt{\Delta}}{2U}$, $P_2 = \frac{\alpha - \frac{(q-1)}{2\lambda} - \hat{\beta}\sqrt{q+1}/v - \frac{(q+1)\mu^s}{2} + \sqrt{\Delta}}{2U}$, *and* $P_3 = \frac{(1-\varepsilon)qN}{\lambda(\hat{\beta}\sqrt{q+1}/v + q\mu^s)}$.

Similar to the robust analysis in Proposition 1, P_1 is a lower bound on P that guarantees that enough delivery orders are assigned to the available crowd such that customers do not wait too long for their crowd delivery set to form. P_2 is an upper bound on P that guarantees that the number of assigned delivery orders to the available crowd drivers is small enough so that deliveries can be made on time, and P_3 is another upper bound on P to guarantee a stable crowd delivery queueing system. Similar to the robust approach, even though the number of crowd drivers N is determined endogenously, the supply function allows us to indirectly optimize over w by using N.

4. Benchmark 2: A Nonrandomized Heuristic Crowdsourcing Model

In this section, we propose and analyze a capacitated vehicle routing problem (CVRP)-based heuristic benchmark model that uses package characteristics (i.e., delivery location) to determine which packages should be crowdsourced and which should be assigned to the 3PL firm. In Section 5, we compare the quality of this dynamic and heuristic-based benchmark approach, which utilizes package characteristics, with that of the static and randomized allocation policy of Section 2. This benchmark utilizes and modifies the savings algorithm, originally proposed by Clarke and Wright (1964), which is a heuristic method for solving the static deterministic CVRP, and it is used in many commercial vehicle routing packages (Simchi-Levi et al. 2014).

In contrast to the static deterministic CVRP, where all customer locations are known a priori, if customers arrive (revealing their location and demand) dynamically over time, it is a dynamic CVRP. In order to solve the dynamic CVRP, we apply the savings algorithm in a rolling horizon framework. In particular, whenever q customers arrive—a batch of size q—we reapply the algorithm.

We next modify the algorithm to accommodate the case where each demand point can be served by either an available crowd driver or a 3PL firm. In our modified savings algorithm (Algorithm 1), as the initial solution, we assign each customer to a 3PL firm at cost c_T^i, where the superscript i indicates customer i, whose distance to the depot influences this cost (i.e., if customer i is located farther from the depot, c_T^i increases); in our subsequent numerical analysis, we use FedEx delivery prices with volume discounts to parameterize c_T^i. The cost of this initial solution is $\Sigma_{i=1}^n c_T^i$.

We next apply the basic idea of the savings algorithm and calculate the savings for (1) serving customer i in a single crowdsourced delivery tour, instead of the 3PL firm, and (2) combining customers i and j and serving them using a single crowdsourced delivery tour. The savings s_{ii} we obtain from reassigning customer i from the 3PL firm to a single crowdsourced delivery tour are $s_{ii} = c_T^i - 2wd_i$, where c_T^i is the delivery cost via the 3PL, d_i is the distance from customer i's location to the depot, and $2wd_i$ is the delivery cost via a crowd driver with hourly wage w. Note that, for simplicity, we consider d_i to be the travel time, between customer i and the depot, in hours. The savings we obtain from reassigning customers i and j from the 3PL firm to a crowdsourced delivery tour, which serves both customers i and j, is $s_{ij} = c_T^i + c_T^j - w(d_i + d_j + d_{ij})$, where $c_T^i + c_T^j$ is the delivery cost for customers i and j if they are both served via the 3PL, d_{ij} is the distance between customers i and j, and $w(d_i + d_j + d_{ij})$ is the delivery cost if customers i and j are served in a single tour via a crowd driver. Note that a crowd driver also experiences an on-site service time X_i when delivering a package to customer i (e.g., parking, finding the appropriate apartment, etc.); we include these on-site service times in the savings matrix as $s_{ii} = c_T^i - w(2d_i + X_i), s_{ij} = c_T^i + c_T^j - w(d_i + d_j + d_{ij} + X_i + X_j)$.

Algorithm 1 outlines the steps of our proposed modified savings algorithm, based on the savings algorithm in Larson and Odoni (1981) and Simchi-Levi et al. (2014), for every batch arrival of size q. Loosely speaking, the proposed algorithm outsources deliveries, where either their location is not close to other deliveries or they require large on-site service times, to the 3PL. Note that, as guaranteed by Step 4, the service-level constraints of all customers are satisfied under this nonrandomized allocation policy.

Algorithm 1 (The Modified Savings Algorithm)

Step 1. Assign each customer to a 3PL firm.

Step 2. Calculate the savings s_{ij}, $\forall i, j$.

Step 3. List the savings in descending order of magnitude.

Step 4. Choose the first feasible link (i, j) from the savings, where

1. i and j are not assigned to the same route,
2. i and j are traveled first or last in their routes,
3. there is an available crowd driver for serving the newly formed route from merging the routes of i and j,
4. merging the routes of i and j does not violate the service-level constraints.

Add the new route to the solution, delete the merged routes consisting of i and j, and delete s_{ij} from the savings list.

Step 5. Repeat Step 4 until there is no feasible link.

5. Numerical Analysis

In this section, we evaluate the performances of the optimal policies from the robust crowdsourcing model and the two benchmark models, via a real-world simulation analysis covering 17 ZIP codes in Seattle, Washington. These ZIP codes include the downtown area as well as sparsely populated areas, served by one active Amazon Flex depot (Amazon Flex Info 2018). Our simulation experiment is designed as follows. We randomly generated 1,000 customer locations according to the probability density function $f(x)$ in Equation (2). Figure 1 shows the 17 ZIP codes, the 1,000 randomly generated customers (green diamonds), and the Amazon Flex depot (black star). The boundaries of the ZIP codes are from the 2016 U.S. ZIP codes via ArcGIS.

We assume orders arrive from the 1,000 customers in Figure 1 via a Poisson process, where customer on-site service times are exponentially distributed with a mean of three minutes. We consider mean customer interarrival times that follow the pattern in the left panel in Figure 2, where μ_L^a and μ_H^a are the mean interarrival times during busy and normal hours, respectively, which captures heavier buying patterns before/after normal business hours as well as the lunch hour. As a baseline, we consider $\mu_H^a = 2\mu_L^a = 5$ minutes. The center panel of Figure 2 shows a heterogeneous supply of crowd drivers of size $\mathbb{N} = 50$, with two different expected opportunity costs, split evenly among the $\mathbb{N}$ drivers. According to Cook et al. (2018), on average, the weekly earnings of Uber drivers from January 2015 to March 2017 was \$376.38 for 17.06 hours covering 29.83 trips, where there was an average wait time of approximately 8 minutes between trips. This is equivalent to working for 21.04 hours and earning \$376.38 weekly, which results in earnings of \$17.9 per hour. In our numerical analysis, we assume that crowd drivers' opportunity costs are normally distributed; the crowd drivers with lower opportunity costs have an expected opportunity cost of $\mu_L^k = \$18$ an hour, and the ones with higher opportunity costs have an expected opportunity cost of $\mu_H^k = \$22$ an hour. The baseline standard deviation of the opportunity costs is $\sigma^k = \$3$ an hour. Finally, we consider a crowd driver's mean vehicle speed that follows the pattern in the right panel in Figure 2, where v_L and v_H are delivery speeds during busy and normal hours, respectively, which captures rush hour traffic patterns. As a baseline, we consider $v_H = 2v_L = 30$ miles per hour (mph), which is motivated by the speed limits in the city of Seattle (Seattle Department of Transportation 2018).

In order to calculate the distances between each customer and the depot as well as that between customers, we use *real distances* (i.e., on a street network) via the origin destination cost matrix implemented in the Network Analyst tool in ArcGIS. We calculate the TSP tours of each instance using the nearest neighbor algorithm (Johnson and McGeoch 1997); we also tested our results using the two-opt local search algorithm (Croes 1958), and our results remain unchanged.

We set the drivers' capacity $Q = 50$ as the baseline value; this number is supported by Amazon Flex drivers' YouTube videos (Ducklow 2016, Ivy 2016). Amazon's Prime Now program (primenow.amazon.com) offers two-hour deliveries. Additionally, Amazon offers same-day delivery services: If customers order by noon, they can receive their orders by 9 p.m. the same day. Therefore, we consider $\alpha \in \{2, 4, 9\}$ hours as the range of customer service levels.

We also need to have c_T^i estimates. The delivery fees of USPS, UPS, and FedEx vary significantly based on the distance between origin and destination and the delivery time window. Therefore, c_T^i, the 3PL delivery cost for package i, should be a function of the distance between a customer's location and the depot, as well as the service level α. We use distance-dependent FedEx delivery fees for same-day, four-hour, and two-hour deliveries (FedEx Economy Delivery 2016, FedEx Priority Delivery 2016, FedEx Standard Delivery 2016), for packages weighing between 0 and 50 pounds (87% of Amazon packages weigh less than 5 pounds: Wells and Stevens 2016). According to FedEx (2019), depending on a firm's annual shipping volume, the firm can get discounts of up to 30%. Thus, we consider the FedEx delivery prices, discounted by $\eta = 30\%$, as representative pricing for a 3PL firm.

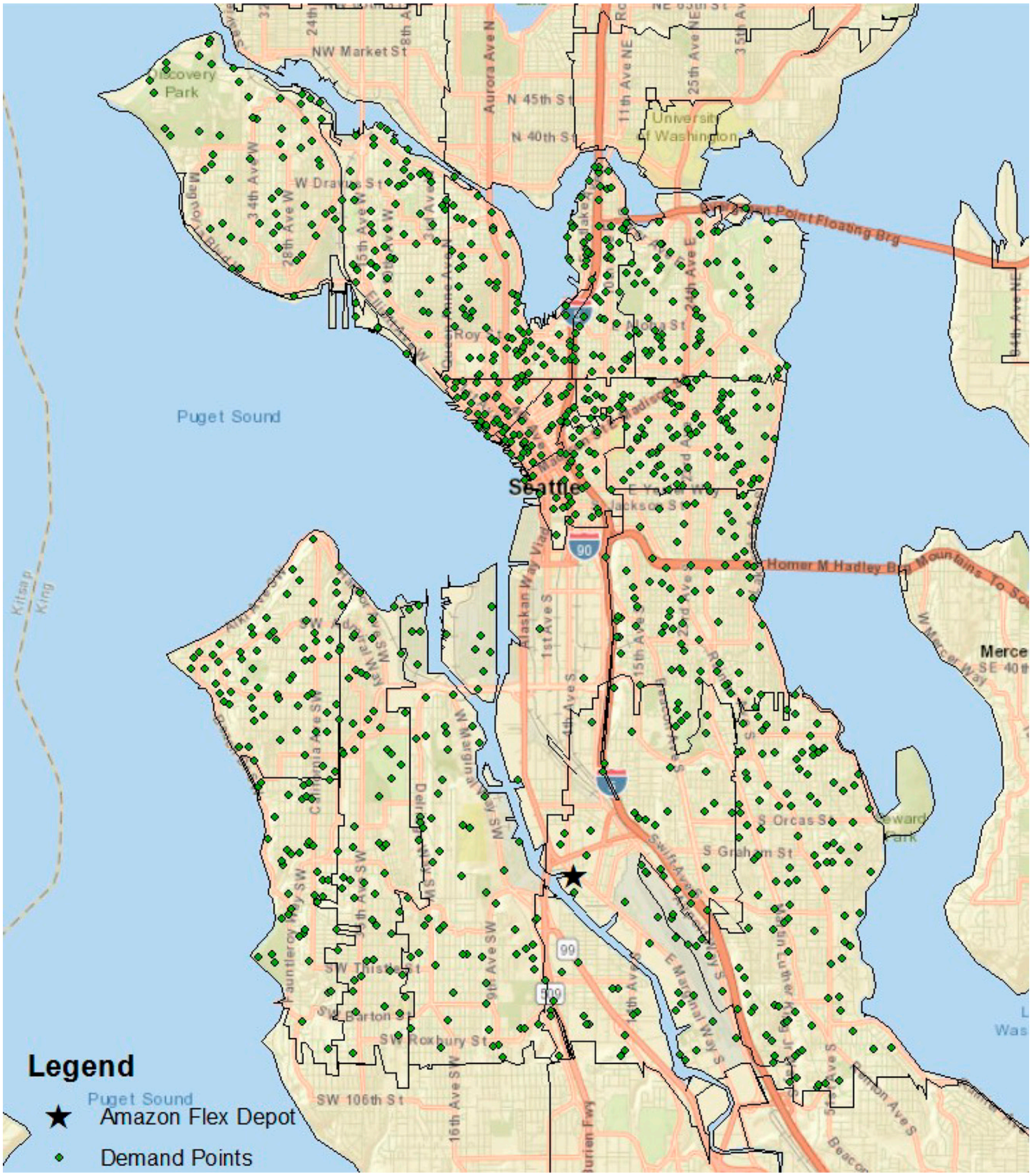
In our robust and stochastic models, we use a single α-dependent value of c_T (the modified savings algorithm, Algorithm 1, uses all c_T^i values). We set $c_T = \Sigma_{i=1}^n c_T^i / n$, where c_T^i is the FedEx delivery cost for customer i who is located d_i miles from the depot. Our results are as follows: The average c_T values, after an $\eta = 30\%$ discount, for all $n = 1{,}000$ customers in all 17 ZIP codes are $c_T = \$9.55$ for $\alpha = 9$ hours, $c_T = \$12.04$ for $\alpha = 4$ hours, and $c_T = \$14.51$ for $\alpha = 2$ hours.

In the robust model, we estimate an upper bound on tour lengths for each tour size $q \in \{1, \ldots, Q\}$, starting and ending at the Amazon Flex depot, by $\beta_U \Sigma_{i=1}^M \sqrt{\frac{p_i}{\Sigma_{j=1}^M \mathbb{A}_j p_j}} \mathbb{A}_i \sqrt{q+1}$. Using the population and area of each of the 17 Seattle ZIP codes in Figure 1, we calculate the multiplier $\Sigma_{i=1}^M \sqrt{\frac{p_i}{\Sigma_{j=1}^M \mathbb{A}_j p_j}} \mathbb{A}_i = 6.94$. Additionally, we consider $\beta_U = 2\tilde{\beta} \approx 1.4$, in which $\tilde{\beta}$ is the Bearwood–Halton–Hammersley constant for Euclidean distances on a plane (Steele 1997), and the two multiplier reflects a finite q in the limit theorem for the robust model.

We also perform an additional layer of numerical optimization for the robust model to determine the

Manufacturing & Service Operations Management, 2022, vol. 24, no. 2, pp. 791–809, © 2021 INFORMS

Figure 1. (Color online) One Thousand Randomly Generated Customers in 17 Seattle ZIP Codes

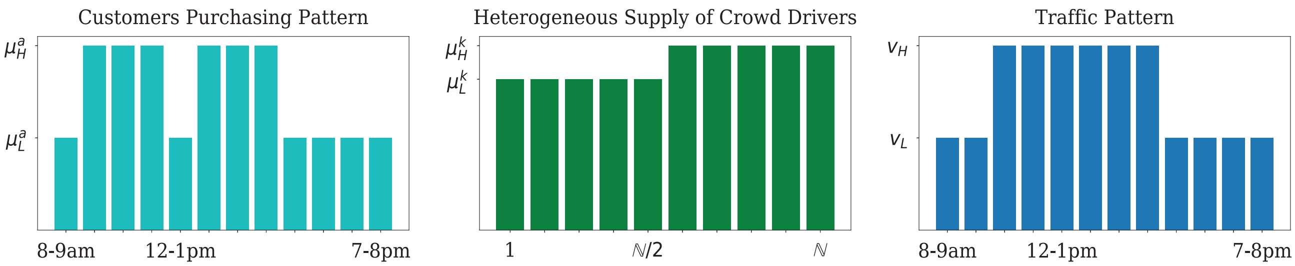


optimal value of q: We insert the solutions from Proposition 1 into Problem (14), which results in an objective function that appears to be convex in q (our conclusion is from extensive numerical analysis as we could not prove this analytically), and this is easily solvable numerically.

5.1. Analysis of the Robust Solution

Figure 3 shows the worst-case cost of the robust optimal policies as a function of demand uncertainty σ^a, for different levels of conservatism in the robust model. The results show that, as there is more demand uncertainty, the delivery cost increases according to a step

Figure 2. (Color online) Patterns of Random Variables over a Day



function. Each of the sharp increases is because of having an extra crowd driver in the crowdsourced delivery system. An increase in the number of crowd drivers decreases the utilization of the system, and hence, it enables the firm to better respond to demand uncertainties. Although in each of the plateaus, the number of crowd drivers does not change, crowd compensation and delivery assignments change as we move across the plateaus, as indicated for $\gamma^r = 3$ and 4 crowd drivers. In this example, on the left-hand side of the plateau, which corresponds to lower levels of demand uncertainty, the firm should assign a smaller number of packages ($q^* = 9$) to crowd drivers for the delivery tours. This allows the firm to create a stream of available delivery work for these independent crowd drivers, which helps with their participation, without having to offer them higher hourly wages. However, as we move to the other end of the plateau, which corresponds to higher demand uncertainty, the firm should assign a larger number of packages ($q^* = 18$) to the crowd drivers for the delivery tours. Assigning a larger number of packages means that the interarrival times between delivery tours increases or equivalently, the utilization of crowd drivers decreases, which helps the crowd delivery system to deal with higher levels of demand uncertainty to satisfy customer demands on time. However, this decrease in utilization of the system results in lower income for the crowd drivers, which can adversely impact their participation, and therefore, the firm should increase the crowd hourly wage; in this example, on the left-hand side, the hourly wage is \$23, and on the right-hand side, the hourly wage is \$30. In addition, the optimal value of P^* across each plateau is one. The changes of the optimal policy with respect to uncertainty in service times are qualitatively similar.

Figure 3. (Color online) Robust Optimal Policy as a Function of Demand Uncertainty

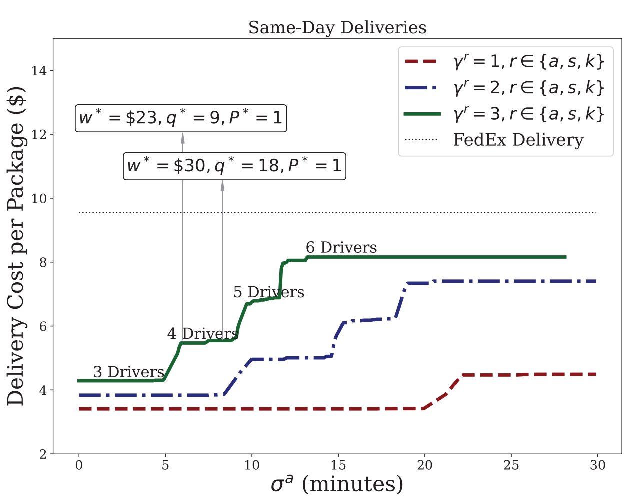
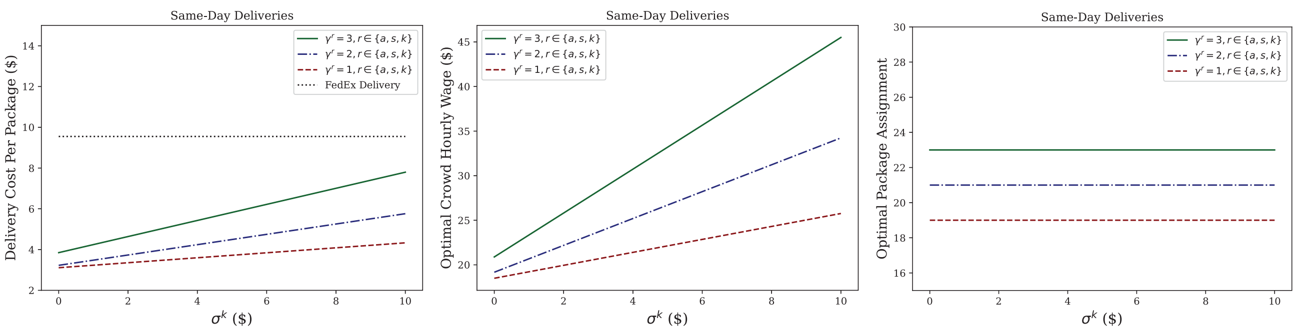


Figure 4 shows the robust optimal policies as a function of variability in crowd opportunity costs. As variability in crowd opportunity costs increases, the firm should not change the labor planning and crowd delivery assignment but rather, simply offer a higher hourly wage to guarantee the availability of independent crowd drivers. Increasing the service guarantee level from 95% to 99% (i.e., increasing the γ^r parameters in the robust uncertainty sets from two to three) is more expensive than increasing it from 68% to 95%, and this is because of a significant decrease in crowd utilization ρ. This results in the firm offering a higher hourly wage w for guaranteeing crowd participation. Thus, the firm, by offering satisfactory levels of an on-time delivery guarantee (e.g., 95%), can significantly reduce its delivery cost via crowdsourcing; we subsequently confirm this via the real simulation study over Seattle ZIP codes.

Similar to the results in each of the plateaus in Figure 3, for lower levels of conservatism γ^r, the firm assigns a smaller number of deliveries to each crowd delivery tour to create a stream of available delivery work for the independent crowd drivers to make them willing to participate without having to offer them high hourly wages. However, for higher levels of conservatism γ^r, the firm wants to keep utilization low to guarantee on-time deliveries. Thus, the firm assigns larger set sizes q to crowd drivers to increase the interarrival times between delivery tours and as a result, decrease crowd utilization; instead, they compensate them via higher hourly wages. Currently, Amazon (Amazon Flex 2019) and Postmates (Postmates 2018) offer crowd hourly wages of \$18–\$25, which is

Figure 4. (Color online) Robust Optimal Policy as a Function of Uncertainty in Crowd Opportunity Costs for $\mu_H^a = 3\mu_L^a = 10$ Minutes



consistent with our findings for service-level guarantees with 68%–95% probabilities for same-day deliveries.

Finally, we can apply our robust model on smaller time horizons (e.g., three-, four-, or six-hour durations, as motivated by Amazon Flex's delivery blocks) by defining our uncertainty sets for the time horizon of interest. Hence, the optimal policy can result in a nonstationary number of drivers and different package assignments across different time horizons. Under the trend and seasonality parameters of Figure 2 with the benchmark settings and $\mu_L^k = \mu_H^k = \$18$, our model for 4-hour deliveries results in the following policies for 12- and 4-hour time horizons, respectively:

$w^* = \$34.30,\ q^* = 7,\ P^* = 1,\ z = \$9.25$

8a.m. 9a.m. 10a.m. 11a.m. 12p.m. 1p.m. 2p.m. 3p.m. 4p.m. 5p.m. 6p.m. 7p.m. 8p.m.

and

$w^* = \$27.58,\ q^* = 9$	$w^* = \$24.40,\ q^* = 6$	$w^* = \$22.52,\ q^* = 8$
$P^* = 1,\ z = \$7.07$	$P^* = 1,\ z = \$5.46$	$P^* = 0.96,\ z = \$7.59$

8a.m. 9a.m. 10a.m. 11a.m. 12p.m. 1p.m. 2p.m. 3p.m. 4p.m. 5p.m. 6p.m. 7p.m. 8p.m.

These optimal policies show that considering smaller time horizons results in more flexibility for the firm, and hence, lower cost, because the firm can adjust its crowd compensation and labor planning in a more flexible manner.

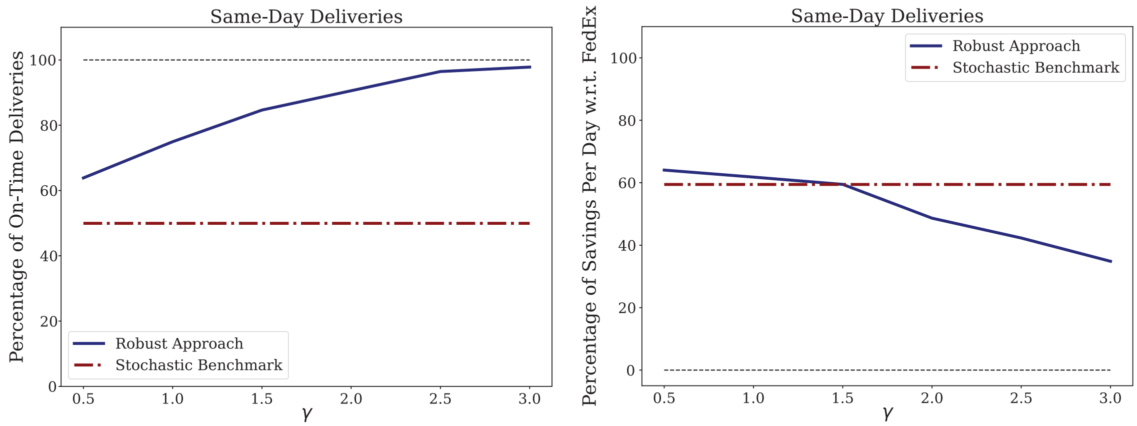
5.2. Comparison of the Robust and Benchmark Solutions

For our simulation studies over the 17 Seattle ZIP codes, for each parameter set, we generate 1,000 simulation trials. We simulate the customer arrivals to each of the ZIP codes during 12-hour time intervals, from 8 a.m. to 8 p.m., as a function of the daily pattern of mean interarrival times.

5.2.1. Comparison of the Robust and Stochastic Solutions. The plots in Figure 5 summarize the comparisons between the robust and stochastic optimal solutions for different levels of conservatism in the robust model, as a function of $\gamma^a = \gamma^s = \gamma^k = \gamma$. The goal of this section is to see how our robust analysis compares against an expected value analysis approach. Note that, because our stochastic model cannot handle nonstationary random variables, we first consider stationary arrivals, service times, delivery speeds, and crowd opportunity costs for the comparison. We set $v_H = v_L = 15$ mph and $\mu_H^k = \mu_L^k = \$18$, as supported by the average speed and average hourly wage of Uber drivers (Cook et al. 2018), respectively, and interarrival mean $\mu_H^a = \mu_L^a = 5$ minutes, where interarrival times of customer orders are exponentially distributed. The left panel in Figure 5 provides the percentage of on-time deliveries, and the right panel in Figure 5 shows the cost-savings percentage per day with respect to the discounted delivery costs that FedEx offers to retailers.

The left panel in Figure 5 shows that under the robust optimal policy for $\gamma = 3$, almost all customers receive their same-day delivery orders on time. The robust policy for $\gamma = 3$ is $(w^*, q^*, P^*) = (\$24.06, 12, 1)$, which, as motivated by the CLT, captures approximately 99% of variability, and that results in the on-time delivery of almost all orders. In contrast, under the stochastic optimal policy, which is $(w^*, q^*, P^*) = (\$18.24, 20, 1)$, approximately half of the customer orders were not delivered on time. The stochastic policy suggests that the firm should assign a relatively large number of deliveries and offer crowd drivers a relatively low hourly wage, with respect to the robust policy with $\gamma = 3$; this results in a significant number of customers not receiving their orders on time under the stochastic policy because of not having enough crowd drivers participating and the associated long wait time of customers until they get served. The right panel in Figure 5 shows that for $\gamma \le 1$, the cost savings under the robust optimal policy are higher than those of the stochastic policy. However, as γ increases further, the robust model becomes more conservative, which results in lower cost savings but higher on-time deliveries; in other words, the stochastic model can

Figure 5. (Color online) Comparison of the Performances of the Robust and Stochastic Optimal Solutions



reduce costs more than the robust model but at the expense of late deliveries.

However, the comparison is not complete, as the robust model is easily able to analytically handle nonstationary uncertainties. In contrast, our stochastic benchmark, based on the analysis of Bertsimas and Van Ryzin (1993), does not capture trend and seasonality in problem parameters. Figure 6 evaluates the performances of the robust and stochastic policies under our benchmark settings and nonstationary interarrival times. As the stochastic benchmark cannot capture nonstationary problem parameters, we utilized the weighted average of the means of the nonstationary interarrival times. The results in Figure 6 show that the robust model outperforms the stochastic benchmark for both cost savings and on-time deliveries.

5.2.2. Comparison of the Robust Solution and the Heuristic Benchmark Algorithm. The proposed robust model is a static model that uses a randomized allocation policy, whereas the heuristic dynamic benchmark algorithm, a modified savings algorithm (Algorithm 1), uses package characteristics (e.g., customer locations) to assign packages to the crowd or the 3PL provider. In this section, we evaluate how our robust model performs against the nonrandomized heuristic benchmark. For this comparison, we use the baseline parameter values, where we assume customer orders arrive according to a Poisson process with rate λ. Table 1 shows the cost savings of the robust model and the nonrandomized heuristic benchmark as a function of α and λ, respectively. The baseline for these cost saving calculations is the discounted FedEx delivery cost. The results show that, despite being a simpler and faster strategy, the robust model results in higher cost savings for the firm.

According to Table 1, as the customer order arrival rate λ decreases, the percentage of cost savings decreases until it becomes zero. The reason is that, as λ decreases, the interarrival time between orders increases, and consequently, the utilization of crowd drivers decreases. This decreased utilization translates to decreased income for crowd drivers, which adversely impacts their participation. Thus, the firm should offer a higher hourly wage to crowd drivers to ensure

Figure 6. (Color online) Performances of the Robust and Stochastic Optimal Solutions for $\mu_H^a = 3\mu_L^a = 15$ Minutes

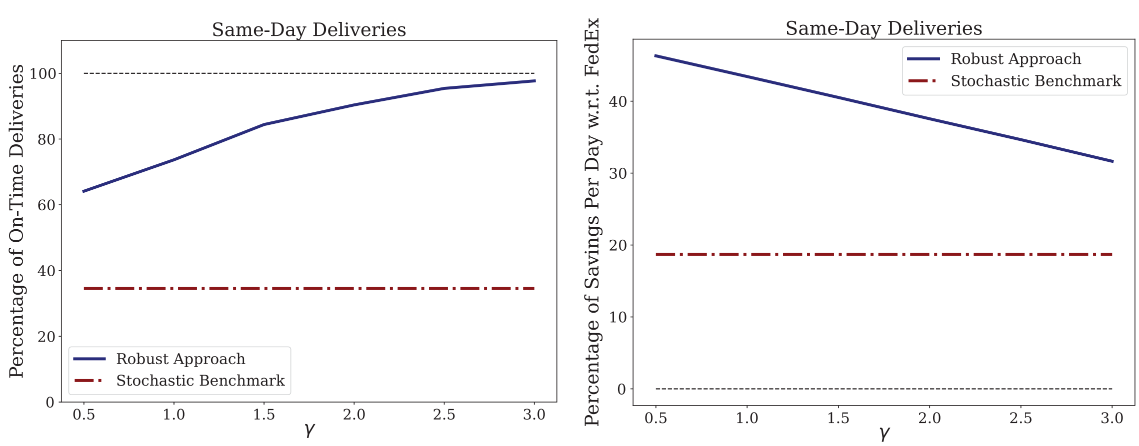


Table 1. Percentage of Cost Savings Under the Robust Optimal Policy and the Nonrandomized Heuristic

α	$\lambda = 24$	$\lambda = 12$	$\lambda = 6$	$\lambda = 3$
Same day, %	(70.1, 11.3)	(62.5, 12.4)	(49.3, 13.3)	(0, 0)
4 hour, %	(51.3, 20.1)	(45.0, 15.2)	(26.8, 8.8)	(0, 0)
2 hour, %	(18.2, 6.9)	(0, 0)	(0, 0)	(0, 0)

their participation. As λ decreases further, utilizing crowd drivers becomes more expensive than outsourcing to 3PL companies. At that point, the firm should outsource all package deliveries to a 3PL firm, and the cost savings from crowdsourcing become zero.

6. Conclusion

In this paper, we analytically studied crowdsourcing last-mile deliveries, with guaranteed delivery time windows, under nonstationary uncertainties. We developed our optimization model by combining crowdsourcing, robust queueing, and robust routing theory. We analytically solved the proposed robust crowdsourcing model and provided closed form solutions, which allowed us to derive managerial insights on how the operations of crowdsourced last-mile delivery systems for on-demand orders should be designed. Our results show that crowdsourcing helps firms significantly decrease their delivery costs while keeping the promise of on-time delivery to their customers for their on-demand orders. We validated the performance of the robust policy via a realistic simulation study over the 17 Seattle ZIP codes that modeled a real transportation network and varying purchasing and traffic patterns. Finally, the robust model outperformed the two proposed benchmark models in terms of the firm's delivery cost savings as well as the percentage of on-time deliveries.

Acknowledgments

The authors thank the associate editor and two reviewers for very insightful comments and suggestions that greatly improved the exposition of the paper.

References

Amazon Flex (2019) What is Amazon Flex? Accessed March 21, 2019, https://flex.amazon.com.

Amazon Flex Info (2018) All Amazon flex warehouse & delivery locations. Accessed August 9, 2018, https://www.moneypixels.com/rideshare/locations.

Arslan AM, Agatz N, Kroon LG, Zuidwijk R (2019) Crowdsourced delivery: A dynamic pickup and delivery problem with ad-hoc drivers. *Transportation Sci.* 53(1):222–235.

Asdecker B, Zirkelbach F (2020) What drives the drivers? A qualitative perspective on what motivates the crowd delivery workforce. *Proc. 53rd Hawaii Internat. Conf. System Sci., Maui, HI*, 4011–4020.

Bai J, So KC, Tang CS, Chen X, Wang H (2019) Coordinating supply and demand on an on-demand service platform with impatient customers. *Manufacturing Service Oper. Management* 21(3):556–570.

Bandi C, Bertsimas D, Youssef N (2015) Robust queueing theory. *Oper. Res.* 63(3):676–700.

Bandi C, Bertsimas D, Youssef N (2018) Robust transient analysis of multi-server queueing systems and feed-forward networks. *Queueing Systems* 89(3):351–413.

Bandi C, Trichakis N, Vayanos P (2019) Robust multiclass queuing theory for wait time estimation in resource allocation systems. *Management Sci.* 65(1):152–187.

Bearwood J, Halton JH, Hammersley JM (1959) The shortest path through many points. *Math. Proc. Cambridge Philos. Soc.* 55(4): 299–327.

Benjaafar S, Ding JY, Kong G, Taylor T (2018) Labor welfare in on-demand service platforms. Preprint, submitted January 17, http://dx.doi.org/10.2139/ssrn.3102736.

Benjaafar S, Kong G, Li X, Courcoubetis C (2019) Peer-to-peer product sharing: Implications for ownership, usage, and social welfare in the sharing economy. *Management Sci.* 65(2):477–493.

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) *Robust Optimization* (Princeton University Press, Princeton, NJ).

Bertsimas DJ, Van Ryzin G (1993) Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles. *Oper. Res.* 41(1):60–76.

Cachon GP (2014) Retail store density and the cost of greenhouse gas emissions. *Management Sci.* 60(8):1907–1925.

Cachon GP, Daniels KM, Lobel R (2017) The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing Service Oper. Management* 19(3):368–384.

Cao J, Olvera-Cravioto M, Shen ZJM (2018) Last-mile shared delivery: A discrete sequential packing approach. Preprint, submitted May 14, https://arxiv.org/abs/1805.05012.

Carlsson JG, Delage E (2013) Robust partitioning for stochastic multivehicle routing. *Oper. Res.* 61(3):727–744.

Clarke G, Wright JW (1964) Scheduling of vehicles from a central depot to a number of delivery points. *Oper. Res.* 12(4):568–581.

Cook C, Diamond R, Hall J, List JA, Oyer P (2018) The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers. National Bureau of Economic Research working paper series, Cambridge, MA.

Croes GA (1958) A method for solving traveling salesman problems. *Oper. Res.* 6(6):791–812.

Dong J, Ibrahim R (2020) Flexible workers or full-time employees? on staffing systems with a blended workforce. *Oper. Res.* 68(4): 1238–1264.

Ducklow L (2016) Amazon flex—how many packages are in a shift? Accessed November 3, 2017, https://www.youtube.com/watch?v=gvV1WQS99pk.

FedEx (2019) Increase your savings with FedEx. Accessed January 15, 2019, http://www.fedex.com/us/oadr/en/discounts/fa7b330664/shipping-discounts-programs.html.

FedEx Economy Delivery (2016) FedEx SameDay city economy. Accessed July 12, 2017, https://www.fedexsameday.com/docs/SameDayCity_RatesEconomy_Flyer_US.pdf.

FedEx Priority Delivery (2016) FedEx SameDay city priority delivery service. Accessed July 12, 2017, https://www.fedexsameday.com/docs/SameDayCity_RatesPriority_Lvbh_US.pdf.

FedEx Standard Delivery (2016) FedEx SameDay city standard delivery service. Accessed July 12, 2017, https://www.fedexsameday.com/docs/SameDayCity_RatesStandard_Flyer_US.pdf.

Gounaris CE, Wiesemann W, Floudas CA (2013) The robust capacitated vehicle routing problem under demand uncertainty. *Oper. Res.* 61(3):677–693.

Gurvich I, Lariviere M, Moreno A (2016) Operations in the on-demand economy: Staffing services with self-scheduling capacity. Preprint, submitted June 26, http://dx.doi.org/10.2139/ssrn.2336514.

He L, Mak HY, Rong Y, Shen ZJM (2017) Service region design for urban electric vehicle sharing systems. *Manufacturing Service Oper. Management* 19(2):309–327.

Insider Intelligence (2017) Same-day shipping is key to fighting Amazon. *Business Insider* (February 13), https://www.businessinsider.com/same-day-shipping-key-to-fighting-amazon-2017-2.

Ivy C (2016) Amazon flex - Denver - 1st day. Accessed November 3, 2017, https://www.youtube.com/watch?v=6U78VkZoErkt=767s.

Johnson DS, McGeoch LA (1997) The traveling salesman problem: A case study in local optimization. A case study in local optimization. Aarts EHL, Lenstra JK, eds. *Local Search in Combinatorial Optimization* (Wiley, New York), 215–310.

Kim E (2019) FedEx will no longer provide express shipping for Amazon in the US. *CNBC* (June 7). https://www.cnbc.com/2019/06/07/amazon-fedex-end-shipping-contract-in-us.html.

Kingman JFC (1974) The heavy traffic approximation in the theory of queues. Smith WL, Wilkinson RI, eds. *Proc. Sympos. Congestion Theory* (University of North Carolina, Chapel Hill, NC), 137–169.

Krivulin NK (1994) A recursive equations based representation for the $G/G/m$ queue. *Appl. Math. Lett.* 7(3):73–77.

Kuhn D, Esfahani PM, Nguyen VA, Shafieezadeh-Abadeh S (2019) Wasserstein distributionally robust optimization: Theory and applications in machine learning. *Operations Research & Management Science in the Age of Analytics*, TutORials in Operations Research (INFORMS, Catonsville, MD), 130–166.

Larson RC, Odoni AR (1981) 6.4.12 Single-Depot VRP. Accessed October 10, 2017, http://web.mit.edu/urban_or_book/www/book/chapter6/6.4.12.html.

Levy A (2019) FedEx Dumps Amazon Before Amazon Dumps It. *The Motley Fool* (June 10), https://www.fool.com/investing/2019/06/10/fedex-dumps-amazon-before-amazon-dumps-it.aspx.

Li B, Krushinsky D, Woensel TV, Reijers HA (2016) An adaptive large neighborhood search heuristic for the share-a-ride problem. *Comput. Oper. Res.* 66(2016):170–180.

Lindley DV (1952) The theory of queues with a single server. *Math. Proc. Cambridge Philos. Soc.* 48(2):277–289.

Liu S, He L, Shen Z-JM (2020) On-time last mile delivery: Order assignment with travel time predictors. *Management Sci.*, ePub ahead of print November 3, https://doi.org/10.1287/mnsc.2020.3741.

Lore M (2017) Serving customers in new ways: Walmart begins testing associate delivery. Accessed November 9, 2017, https://blog.walmart.com/innovation/20170601/serving-customers-in-new-ways-walmart-begins-testing-associate-delivery.

Mamani H, Nassiri S, Wagner M (2016) Closed-form solutions for robust inventory management. *Management Sci.* 63(5):1625–1643.

Phillips EE (2018) e-Commerce companies get creative in quest for last mile space. *Wall Street Journal* (December 9), https://www.wsj.com/articles/e-commerce-companies-get-creative-in-quest-for-last-mile-space-1544364000.

Postmates (2018) Postmates Courier Driver hourly salaries in the United States. Accessed August 8, 2018, https://www.indeed.com/salaries/Courier-Driver-Salaries-at-Postmates.

Qi W, Li L, Liu S, Shen ZJM (2018) Shared mobility for last-mile delivery: Design, operational prescriptions, and environmental impact. *Manufacturing Service Oper. Management* 20(4):737–751.

Reuters (2018) Walmart ends its grocery delivery partnerships with Uber and Lyft after they fail to take off. *CNBC* (May 8), https://www.cnbc.com/2018/05/08/walmarts-grocery-delivery-partnerships-with-uber-lyft-fail-to-take-off.html.

Seattle Department of Transportation (2018) Speed limits. Accessed March 14, 2019, https://www.seattle.gov/transportation/projects-and-programs/safety-first/vision-zero/speedlimits.

Simchi-Levi D, Chen X, Bramel J (2014) *The Logic of Logistics*, 3rd ed. (Springer, Berlin).

Steele JM (1981) Subadditive Euclidean functionals and nonlinear growth in geometric probability. *Ann. Probab.* 9(3):365–376.

Steele JM (1997) *Probability Theory and Combinatorial Optimization* (Society for Industrial and Applied Mathematics, Philadelphia).

Sungur I, Ordóñez F, Dessouky M (2008) A robust optimization approach for the capacitated vehicle routing problem with demand uncertainty. *IIE Trans.* 40(5):509–523.

Taylor T (2018) On-demand service platforms. *Manufacturing Service Oper. Management* 20(4):704–720.

Wagner M (2018) Robust inventory management: An optimal control approach. *Oper. Res.* 66(2):426–477.

Wells G, Stevens L (2016) Amazon conducts first commercial drone delivery. *Wall Street Journal* (December 14), https://www.wsj.com/articles/amazon-conducts-first-commercial-drone-delivery-1481725956.

Whitt W (1993) Approximations for the gi/g/m queue. *Production Oper. Management* 2(2):114–161.