Transportation Science

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

Enhancing Underutilized Bus Routes with Advance Reservations and Semiflexible Routing

Md Hishamur Rahman; , Ye Chen; , Shijie Chen; , Yanshuo Sun; , Muhammad Imran Younus Siddiqui, Matthew Mohebbi, Nikola Marković;

informs
https://pubsonline.informs.org/journal/trsc
TRANSPORTATION SCIENCE
Articles in Advance, pp. 1–23
ISSN 0041-1655 (print), ISSN 1526-5447 (online)

Enhancing Underutilized Bus Routes with Advance Reservations and Semiflexible Routing

Md Hishamur Rahman,[a] Ye Chen,[b] Shijie Chen,[c] Yanshuo Sun,[c] Muhammad Imran Younus Siddiqui,[d] Matthew Mohebbi,[d] Nikola Marković[a,*]

[a] Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, Utah 84112; [b] Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, Richmond, Virginia 23284; [c] Department of Industrial and Manufacturing Engineering, Florida A&M University–Florida State University College of Engineering, Florida State University, Tallahassee, Florida 32310; [d] IT Curves, Gaithersburg, Maryland 20879

*Corresponding author

Contact: hisham.rahman@utah.edu, https://orcid.org/0000-0001-7555-5703 (MHR); ychen24@vcu.edu, https://orcid.org/0000-0002-2906-4126 (YC); sc20hw@fsu.edu, https://orcid.org/0000-0003-4396-1357 (SC); y.sun@eng.famu.fsu.edu, https://orcid.org/0000-0003-2943-4323 (YS); myounus@itcurves.net (MIYS); mmohebbi@itcurves.net (MM); nikola.markovic@utah.edu, https://orcid.org/0000-0003-0883-2701 (NM)

Received: January 30, 2024
Revised: November 30, 2024; February 22, 2025
Accepted: April 17, 2025
Published Online in Articles in Advance: May 19, 2025

https://doi.org/10.1287/trsc.2024.0561

Abstract. This paper seeks to improve an underutilized conventional bus route by converting it into a semiflexible transit system where passengers provide advance notice of their intended stops, allowing buses to skip downstream stops without demand by taking shortcuts. This approach increases stop density, reduces walking distances to and from bus stops, and maintains operational efficiency. To design this system, we develop optimization models that maximize the number of stops while adhering to tour duration and arrival time constraints. A case study in Allegany County, Maryland, demonstrates significant enhancements for routes that were both underutilized (where the probability of a stop lacking demand exceeded 45%) and had layouts conducive to substantial shortcuts. In these instances, the number of stops can be increased by up to 160%, with the actual improvement depending on route configuration, passenger demand, and advance notice requirements.

Funding: Financial support from the the National Science Foundation [Grant 2055347] is gratefully acknowledged.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/trsc.2024.0561.

Keywords: transit • flexible routing • probabilistic analysis • online learning • approximate Bayesian inference

1. Introduction

Conventional public transit, with fixed routes and schedules, operates efficiently in densely populated areas but is often underutilized in low-density regions, where long walking distances to bus stops discourage potential users (Sun et al. 2017). In contrast, demand-responsive transit provides personalized curb-to-curb service, eliminating the need for walking. However, its high operating costs limit widespread availability, often restricting it to specific groups, such as seniors and people with disabilities (Davison et al. 2014, Rahman et al. 2023).

To bridge the gap between conventional and fully flexible transit services, we propose a semiflexible transit system to enhance underutilized bus routes. This system converts fixed stops into on-demand stops, which passengers must request in advance, and allows buses to skip stops without demand by taking shortcuts. By increasing stop density, this approach reduces walking distances while maintaining similar tour times to conventional routes. Next, we compare conventional transit with the proposed semiflexible system.

System 1 (Conventional Transit). A conventional bus route has a fixed number of stops and a predetermined schedule, requiring the bus to arrive at each stop at specific times. Although a bus may skip stops without demand to reduce dwell time, the time saved is often lost waiting at downstream stops to realign with the fixed schedule. This structure often leads to inefficiencies, such as stopping at locations without demand, idling at empty stops, and requiring passengers to walk long distances because of a limited number of stops.

System 2 (Proposed Semiflexible Transit). To improve transit efficiency, we propose transitioning to on-demand stops, which passengers request in advance. This allows the bus to skip stops without demand by taking shortcuts, reducing both travel and dwell times. This approach increases stop density, reduces walking distances, and maintains the same or even shorter bus tour durations as System 1 with high probability.

Requiring passengers to provide longer advance notice enables the bus to skip more downstream stops, allowing for a higher stop density and shorter walking distances. Figure 1 illustrates how stop density increases with longer notice periods while maintaining consistent tour durations with high probability and serving the same demand as conventional transit. However, this requires passengers to plan their trips further in advance, creating a trade-off between user flexibility and reduced walking distances.

This paper presents several contributions to the field of semiflexible transit system design:

- We propose a semiflexible transit system that combines the benefits of demand-responsive and fixed-route transit. Passengers provide advance notice of their intended stops, allowing buses to skip downstream stops without demand by taking shortcuts. This approach increases the number of stops, reduces walking distances, and maintains operational efficiency.
- To design the proposed semiflexible transit system, we develop both stylized and simulation-based models. Stylized models explore trade-offs in system design, whereas simulation-based models, integrated with an online Bayesian learning framework, optimize bus operations across diverse real-world scenarios. These models aim to increase the number of stops while ensuring, with high probability, that bus tour durations remain the same or even shorter than conventional transit systems and that passenger requests made with appropriate notice are honored.
- We applied our models to bus routes in Allegany County, Maryland, to evaluate the proposed semiflexible routing strategy. Significant enhancements were observed on routes that were both underutilized (where the probability of a stop lacking demand exceeded 45%) and had layouts conducive to substantial shortcuts. In these cases, the number of stops increased by up to 160%, with the actual improvement depending on route configuration, passenger demand, and advance notice requirements.

This paper is structured as follows: Section 2 reviews relevant literature. Sections 3 and 4 describe the stylized and simulation-based models for designing the semiflexible transit system. Section 5 analyzes trade-offs through numerical experiments, and Section 6 applies these strategies to a case study in Allegany County, Maryland. Finally, Section 7 concludes the paper.

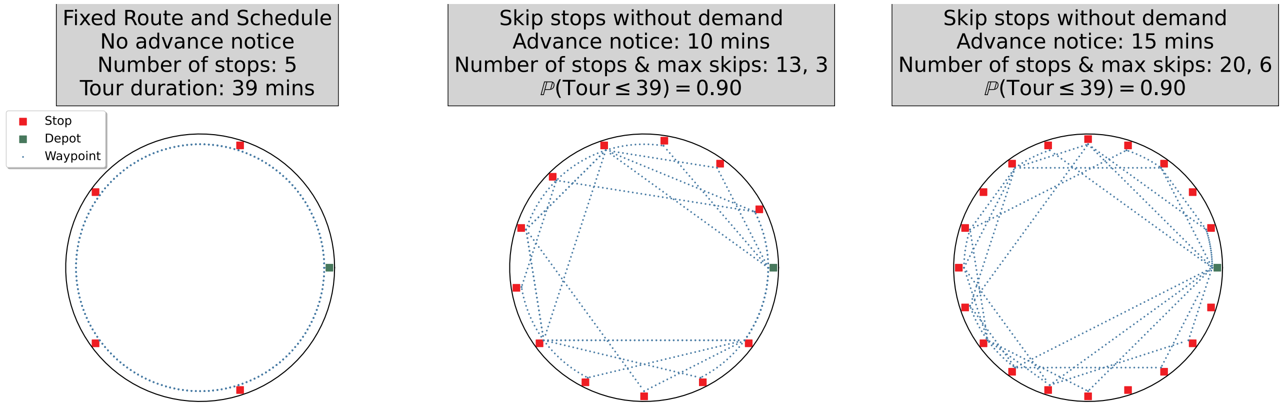
2. Literature Review

We begin by reviewing different types of flexible transit systems and contrasting them with the system proposed herein. Subsequently, we delve into various applications of stop skipping and explain how they differ from the objectives of our analysis.

2.1. Flexible Transit Systems

A flexible transit system typically provides door-to-door or stop-to-stop service without adhering to a fixed route. In contrast, a semiflexible transit system incorporates fixed stops, referred to as checkpoints, which vehicles are required to visit. Between these checkpoints, vehicles may deviate from the baseline route to serve optional on-demand stops, allowing for a degree of service flexibility. The transit cooperative research program identifies six main types of semiflexible transit services, illustrations of which are detailed in Online Appendix A. They include route deviation, point deviation, demand-responsive connector, request stops, flexible-route segments, and zone route (Koffman 2004). This literature review focuses on previous studies of semiflexible transit systems bearing resemblance to our proposed system.

Figure 1. (Color online) Illustrative Example of the Relationship Between Advance Notice and the Number of Stops in a Semiflexible Transit System, Using a Circular Fixed Route with Five Stops as the Baseline



Notes. The total number of passengers is kept constant, resulting in the probability of a stop not being used increasing with the number of stops. Bus trajectories (dotted lines) in multiple simulation cycles show the shortcuts (chords) taken in the semiflexible system by skipping stops without demand. The semiflexible system offers higher stop density, reducing walking distance to and from stops. It also maintains the same or even shorter tour duration with high probability. A simulation can be found at https://youtu.be/JeBA7_ToIR4.

2.1.1. Route Deviation. In this semiflexible transit system, vehicles follow a fixed schedule along a predetermined route, but may deviate to fulfill on-demand requests within a designated area around the route (Pratelli and Schoen 2001; Quadrifoglio, Hall, and Dessouky 2006). Studies suggest that route deviation performs well at lower demand levels (Pratelli et al. 2018). Various optimization models have been developed to determine the most efficient deviated routes and assess performance. For example, Pratelli and Schoen (2001) presented a model where passengers signal buses for deviated stops on-site using a push button. An enhanced model by Pratelli et al. (2018) integrates advance reservations via phone calls alongside on-site requests. Route deviation's primary advantage is its capability to serve additional passengers at deviated stops, but this often results in longer travel times for onboard passengers and increased waiting times for those at downstream stops (Pratelli and Schoen 2001). In contrast, our proposed system addresses on-demand requests by serving passengers exclusively at designated stops along the fixed route. Additionally, advance notice from passengers enables the bus to deviate from the fixed route for taking shortcuts to maintain operational efficiency.

2.1.2. Point Deviation. In point deviation systems, vehicles serve on-demand requests within a designated area, stopping at a limited number of fixed stops but not following a fixed route between them (Daganzo 1984). Previous studies have developed frameworks to assess the service quality of point deviation systems, finding that they can outperform conventional transit and route deviation systems, particularly at low demand levels (Daganzo 1984; Nourbakhsh and Ouyang 2012; Qiu et al. 2015; Pratelli et al. 2018; Zheng, Li, and Qiu 2018). Enhancements for point deviation systems, such as using predefined areas (Nourbakhsh and Ouyang 2012) and common meeting points (Li et al. 2023), have been suggested to improve efficiency. These models typically involve reservations via phone or online to schedule stops. Our proposed system is similar to point deviation, offering the flexibility of adaptive routing while ensuring that stops are served. Additionally, it maintains bus arrivals within a designated time window with high probability.

2.1.3. Demand-Responsive Connector and Zone Route. Demand-responsive connector systems and zone routes function as feeder services, gathering passengers within a responsive zone and transporting them to a transfer terminal linked to the main transit network (Chandra and Quadrifoglio 2013; Zheng, Li, and Qiu 2018). These services typically alternate between fixed-route operations during peak demand and on-demand service in low demand periods (Cayford and Yim 2004). Analytical models have been developed to determine the demand density required to switch between fixed-route and demand-responsive policies (Quadrifoglio and Li 2009, Li and Quadrifoglio 2010), and to optimize system design and operation, including factors like service cycle length and service area (Chang and Schonfeld 1991, Chandra and Quadrifoglio 2013). Although these models aim to enhance service quality by minimizing passenger and operator costs, they do not focus on increasing mobility by maximizing the number of stops, which is a key aspect of our study. Our proposed system's ability to maintain tour duration while accommodating on-demand requests makes it suitable for use as a demand-responsive connector system or a zone route.

2.1.4. Hybrid Systems. Hybrid systems integrate various types of semiflexible transit services to capitalize on their combined strengths. Systems that integrate fixed-route with point/route deviation services, for example, have been explored in several studies (Aldaihani and Dessouky 2003, Aldaihani et al. 2004, Crainic et al. 2005, Chen and Nie 2017, Sipetas and Gonzales 2021). These studies have developed analytical and mathematical programming models to optimize design, routing, and scheduling of hybrid systems. Other research has focused on broader modeling and simulation frameworks for planning and evaluating various semiflexible transit options (Cortés, Pagès, and Jayakrishnan 2005; Errico et al. 2013, 2021; Silva, Vinel, and Kirkici 2022). A common challenge in hybrid systems is the requirement for passenger transfers, often leading to inconvenience. Our proposed semiflexible system, however, is tailored to seamlessly integrate with fixed systems, thereby avoiding the need for transfers.

2.2. Stop Skipping

The concept of skipping stops has received significant attention in both scientific literature and practical applications. However, most existing studies focus on optimizing stop skipping to enhance operational efficiency and reduce costs, without aiming to increase the number of stops to minimize walking distances. In contrast, our study leverages stop skipping and shortcuts to introduce additional stops while maintaining service quality and efficiency. To contextualize our approach, we review several key studies that explore stop-skipping strategies with distinct objectives and methodologies.

For instance, Liu et al. (2013) investigated a bus stop-skipping scheme under random travel times, aiming to minimize total waiting time, in-vehicle travel time, and operating costs. They developed a genetic algorithm incorporating Monte Carlo simulation to optimize stop-skipping and deadheading strategies. Building on this, Chen et al. (2015) focused on integrating headway optimization with stop-skipping control in bus rapid transit

(BRT) systems. Their study, based on Beijing BRT Line 2, demonstrated that optimized headways with stop skipping could reduce total costs and in-vehicle time for passengers compared with traditional strategies. More recently, Zhang et al. (2021) developed an agent-based simulation model to optimize real-time bus stop-skipping and holding strategies. Their work accounted for random travel times and passenger arrivals to evaluate operational efficiency. Notably, the study found that real-time stop skipping combined with holding performed best under deterministic bus dispatch schedules, whereas stop skipping alone was more effective in cyclic dispatch systems where buses are reassigned to the same route after completing their first tour.

2.3. Stop Skipping with Shortcuts

In addition to skipping stops, a few studies have evaluated the impacts of taking shortcuts, but these efforts do not address the potential for increasing the number of stops—a central focus of our proposed semiflexible system. Instead, existing research primarily examines shortcuts as a means to improve operational efficiency and reduce travel times.

Hadas and Ceder (2008) introduced a novel method for optimizing public transit systems using multiagent systems. This approach integrates real-time data on vehicle locations, passenger demands, and travel times to enhance service reliability and efficiency. Key operational tactics such as holding, stop skipping, and shortcuts were employed to improve transfer synchronization and reduce travel times. Building on this work, Hadas and Ceder (2010) developed a dynamic programming model to further improve bus service efficiency by minimizing total travel time. Their model applied several real-time operational tactics, including holding, stop skipping, speed adjustments, and short-turn operations, and leveraged real-time data to dynamically adjust vehicle operations, reducing transfer times and increasing the likelihood of direct transfers.

Similarly, Osama et al. (2016) examined the effects of unscheduled stops and route deviations on bus service performance in Cairo. Using a prototype automatic vehicle location system, they collected data for a specific bus route and found that nearly 50% of trips involved shortcuts, where bus drivers deviated from the scheduled route. More recently, Aktaş, Sörensen, and Vansteenwegen (2023) introduced a variable neighborhood search (VNS) algorithm to optimize the performance of a single bus line during peak hours. Their demand-responsive system targeted periods with significantly imbalanced passenger demand in opposite directions, increasing service frequency toward the city center by allowing some buses to skip less crowded stops or sections of the route when traveling away from the center. The VNS algorithm determined which buses should visit all stops and which should take shortcuts based on expected demand.

2.4. Contributions to the Literature

Our study differs from previous research in several key ways. First, prior studies have primarily used stop skipping and shortcuts to improve operational efficiency and reduce costs. In contrast, our study strategically employs stop-skipping and shortcut strategies to increase the density of on-demand stops, bringing them closer to passenger origins and destination, reducing their walking distance to the nearest stop, and maintaining service quality and efficiency.

Second, existing semiflexible transit systems often deviate from their routes to pick up or drop off passengers, which is likely to increase overall travel time and reduce service predictability. In contrast, we maximize the number of stops while keeping the bus tour duration of the semiflexible system within that of the fixed transit system with high probability. Advance reservations for on-demand service are considered only at predetermined stops, aiming to provide a more predictable and efficient transit service. Finally, our advance reservation policy guarantees service to passengers with high probability, provided they request their intended stops with appropriate notice, which has not been considered before in the context of semiflexible transit systems.

3. Methodology: Stylized Model

In this section, we outline the methodology for designing the proposed semiflexible transit system by examining a stylized scenario with uniformly distributed demand across evenly spaced stops along the bus route, an example of which is given in Figure 1. The model tailored for this scenario is developed in Sections 3.1 to 3.4, whereas the problem of maximizing the number of stops is addressed in Sections 3.5 and 3.6. The theoretical analysis presented herein motivates the use of heuristics and the integration of derived constraints into a Bayesian learning framework to address more intricate scenarios discussed in Section 4.

3.1. Preliminaries

Let n denote the number of evenly spaced stops in System 2, and let p denote the probability of a stop having no demand. Given a fixed and uniformly spread demand, increasing the number of stops results in a higher probability that any given stop will not be used as either an origin or a destination. This intuitive result can be rigorously proved under the conditions stated below. Specifically, the monotonicity of p with respect to n is formally proved in Online Appendix B.

Assumption 1 (Demand Distribution). *Passengers' origins and destinations are uniformly distributed along the bus route.*

Assumption 2 (Passenger Behavior). *Each passenger uses the closest stops to their origin and destination, which are not the same. Moreover, the choices of origins and destinations among the passengers are independent.*

Proposition 1. *Consider System 2 with the passenger distribution and behavior stated in Assumptions 1 and 2. Given a fixed demand of Ψ passengers per round trip, the probability that a stop will not be used by any customer in a bus roundtrip is*

$$p = \left(1 - \frac{2}{n}\right)^{\Psi}, \tag{1}$$

which increases with the number of stops n. For large n, we have $p \sim e^{-\frac{2\Psi}{n}}$.

3.2. Bus Tour Duration

Let X_k denote the number of times the bus skips exactly k consecutive stops in its tour. Note that X_k is a discrete random variable with k taking integer values ranging from 1 to $n-1$; for example, $k = n-1$ means that the bus skips all stops except the depot. However, we limit the skipping of consecutive stops to no more than m at any given time (where $m \le n-1$). Accordingly, we model System 2 while conditioning on the following event:

$$A_{n,m} = \{X_k = 0, \ \ \forall\, k \in \{m+1, m+2, \ldots, n-1\}\},$$

which ensures the bus can skip at most m stops consecutively.

Let $T_{n,m}$ denote the bus tour duration serving n stops, with the bus skipping at most m consecutive stops. This duration includes both driving time and the time spent waiting at bus stops for passengers to board and alight. Let c represent the maximum driving time when all n stops are visited, and let s denote the time spent at each stop. Let $\zeta_{n,k}$ denote the amount of time saved by skipping k consecutive stops along the route. The bus tour duration is then given by

$$T_{n,m} = \underbrace{c + n \cdot s}_{\text{Full tour duration}} - \underbrace{\sum_{k=1}^{m} X_k(\zeta_{n,k} + k \cdot s)}_{\text{Time savings from skipping stops}}, \tag{2}$$

where the summand reflects the time savings from skipping k consecutive stops, both by shortening the driving distance by $\zeta_{n,k}$ and reducing dwell time by $k \cdot s$. Because $T_{n,m}$ depends on the random variables $X_1, \ldots, X_m$, which are *not* independent, it is itself a random variable.

To initiate the analysis of System 2, we need to derive the cumulative distribution function (CDF), expected value, and variance of $T_{n,m}$. We should keep in mind that the random variables $X_1, \ldots, X_m$ depend on p, the probability of skipping a stop, which is a function of n as given in (1). To proceed, we first derive the conditional probability mass function (PMF) of the random variables $X_1, \ldots, X_m$ given $A_{n,m}$ in Proposition 2. The formal proof of Proposition 2 is provided in Online Appendix C.

Proposition 2. *Suppose the probability of skipping a stop is p. Let $x_0 \triangleq n - \sum_{j=1}^{m}(j+1)x_j$ and $\mathcal{X}_{n,m} \triangleq \{(x_0, \ldots, x_m) : \sum_{j=0}^{m}(j+1)x_j = n, x_j \ge 0, \ \forall\, 0 \le j \le m\}$. Then, for $m \ge 1$,*

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m \,|\, A_{n,m}) = \frac{\frac{(\sum_{j=0}^{m} x_j)!}{\prod_{j=0}^{m}(x_j!)}\left(\frac{1-p}{p}\right)^{\sum_{j=0}^{m} x_j} 1_{\{(x_0,\ldots,x_m)\in\mathcal{X}_{n,m}\}}}{\sum_{(x_0,\ldots,x_m)\in\mathcal{X}_{n,m}} \frac{(\sum_{j=0}^{m} x_j)!}{\prod_{j=0}^{m}(x_j!)}\left(\frac{1-p}{p}\right)^{\sum_{j=0}^{m} x_j}},$$

where $1_{\{(x_0,\ldots,x_m)\in\mathcal{X}_{n,m}\}}$ is binary indicator that is equal to one if $(x_0, \ldots, x_m) \in \mathcal{X}_{n,m}$ and zero otherwise.

The PMF represents the probability of a specific bus tour, modeled as an n-digit binary sequence where one denotes a skipped stop and zero denotes a visited stop. The binary sequence must consist of blocks of the form $\underbrace{1\ldots1}_{j}0$, where $0 \le j \le m$, representing up to m consecutive skipped stops followed by a visited stop. The PMF is derived based on the likelihood of these blocks occurring, considering the probability p of skipping a stop, which is explicitly computed in Proposition 1. Each sequence configuration belongs to the set $\mathcal{X}_{n,m}$, where the number of blocks must satisfy the equation $\sum_{j=0}^{m}(j+1)x_j = n$, ensuring that the total number of stops is n. The last binary digit is always zero, ensuring that the bus returns to the depot. This derivation accounts for all possible valid bus tours, with the PMF providing the probability of any given tour.

By utilizing the derived PMF, we can compute the conditional expectation of $T_{n,m}$ given $A_{n,m}$ as follows:

$$\mathbb{E}(T_{n,m} \,|\, A_{n,m}) = \sum_{(x_0,\ldots,x_m)\in\mathcal{X}_{n,m}} \left(c + n \cdot s - \sum_{k=1}^{m} x_k(\zeta_{n,k} + k \cdot s)\right) \mathbb{P}(X_1 = x_1, \ldots, X_m = x_m \,|\, A_{n,m}). \tag{3}$$

Similarly, we use the derived PMF and the conditional expectation to compute the conditional variance of $T_{n,m}$ given $A_{n,m}$ as follows:

$$\mathrm{Var}(T_{n,m} \,|\, A_{n,m}) = \mathbb{E}(T_{n,m}^2 \,|\, A_{n,m}) - [\mathbb{E}(T_{n,m} \,|\, A_{n,m})]^2, \tag{4}$$

with detailed calculations presented in Online Appendix D.

Last, we derive the CDF of $T_{n,m}$ given $A_{n,m}$. Note that (2) implies that each realization of $X_1, X_2, \ldots, X_m$ results in a different realization of $T_{n,m}$; thus, we can compute the conditional CDF of $T_{n,m}$ by leveraging the conditional PMF of $(X_1, X_2, \ldots, X_m)$ specified in Proposition 2.

Define $\mathcal{T}(t, A_{n,m}) \triangleq \{(x_1, \ldots, x_m) \in A_{n,m} : c + n \cdot s - \sum_{k=1}^{m} x_k(\zeta_{n,k} + k \cdot s) \leq t\}$, that is, the set of all realizations of $X_1, X_2, \ldots, X_m$ that result in $T_{n,m} \leq t$. Then, we can express the conditional CDF as the sum of the probabilities of all such possible outcomes:

$$\mathbb{P}(T_{n,m} \leq t | A_{n,m}) = \sum_{(x_1, x_2, \ldots, x_m) \in \mathcal{T}(t, A_{n,m})} \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m | A_{n,m}). \tag{5}$$

We will utilize (5) to ensure that, with a high probability, the bus tour duration in System 2 does not exceed its counterpart in System 1. This allows for effective comparison and evaluation of the performance of the two systems.

3.3. Bus Arrivals

Let $I = \{1, 2, \ldots, n\}$ represent the set of stops. A bus departing from the depot at time 0 and arriving at stop $i \in I$ can skip at most m consecutive stops before reaching stop i. Denote by $G_{i,m}$ the arrival time of the bus at stop i. Let S_i be a binary indicator that equals one if stop i is skipped and zero otherwise. Note that, given $S_i = 0$, we must have $m \leq i - 1$ because the bus can skip at most $i - 1$ consecutive stops before arriving at stop i. Then, given $A_{n,m}$ and $S_i = 0$, we can derive the following expression for $G_{i,m}$ by treating stop i as the "depot" and following the computation for $T_{n,m}$ in (2):

$$G_{i,m} = \underbrace{c_i + (i-1)s}_{\text{Arrival time without skipping}} - \underbrace{\sum_{k=1}^{m} X_k^{(i)}(\zeta_{n,k} + k \cdot s)}_{\text{Time savings from skipping stops}},$$

where c_i represents the driving time from the depot to stop i when no stops are skipped, and the superscript on $X_k^{(i)}$ indicates that it counts the number of k consecutive skips on a tour terminating at the ith stop. Noting that c_n equals c from (2) and that $X_k^{(n)} = X_k$, we have that $T_{n,m} = G_{n,m} + s$; that is, the arrival time of the bus back at the depot (stop n) is offset by s from the bus tour duration. Thus, we have $\mathbb{E}(G_{n,m} | A_{n,m}, S_n = 0) = \mathbb{E}(T_{n,m} | A_{n,m}) - s$, and $\mathrm{Var}(G_{n,m} | A_{n,m}, S_n = 0) = \mathrm{Var}(T_{n,m} | A_{n,m})$.

Similarly, if we consider stop i as the depot and set $n = i$ in Proposition 2, then the resulting $\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m | A_{i,m})$ is exactly the conditional PMF of $X_1^{(i)}, X_2^{(i)}, \ldots, X_m^{(i)}$ given $A_{n,m}$ and $S_i = 0$. Then, similarly to (3) and (4), we can obtain $\mathbb{E}(G_{i,m} | A_{n,m}, S_i = 0)$ and $\mathrm{Var}(G_{i,m} | A_{n,m}, S_i = 0)$ as follows:

$$\mathbb{E}(G_{i,m} | A_{n,m}, S_i = 0) = \sum_{(x_0, \ldots, x_m) \in \mathcal{X}_{i,m}} \left(c_i + i \cdot s - \sum_{k=1}^{m} x_k(\zeta_{n,k} + k \cdot s) \right) \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m | A_{i,m}) - s, \tag{6}$$

$$\mathrm{Var}(G_{i,m} | A_{n,m}, S_i = 0) = \mathbb{E}(G_{i,m}^2 | A_{n,m}, S_i = 0) - [\mathbb{E}(G_{i,m} | A_{n,m}, S_i = 0)]^2, \tag{7}$$

with detailed calculations presented in Online Appendix D. Next, we will use (6) and (7) to ensure that a request to use stop i is met with a high probability.

3.4. Fulfilling Passenger Requests

Recall that System 2 should be designed to fulfill passenger requests with a high probability. To achieve this goal, we construct a time interval during which the bus is expected to arrive at stop i with a probability of at least $1 - \alpha$. Recall that $G_{i,m}$ given $A_{n,m}$ and $S_i = 0$ is a function of $(X_1^{(i)}, \ldots, X_m^{(i)})$, so we can explicitly characterize the conditional probability distribution of $G_{i,m}$. Therefore, we choose $\varepsilon_{i,1-\alpha/2}$ to be the smallest real number with two decimal places such that

$$\mathbb{P}(\mathbb{E}(G_{i,m} | A_{n,m}, S_i = 0) - \varepsilon_{i,1-\alpha/2} \sqrt{\mathrm{Var}(G_{i,m} | A_{n,m}, S_i = 0)} \leq G_{i,m} \leq \mathbb{E}(G_{i,m} | A_{n,m}, S_i = 0) + \varepsilon_{i,1-\alpha/2} \sqrt{\mathrm{Var}(G_{i,m} | A_{n,m}, S_i = 0)} | A_{n,m}, S_i = 0) \geq 1 - \alpha. \tag{8}$$

Accordingly, the $100(1 - \alpha)\%$ confidence interval for the bus arrival time at stop i is given by

$$[\mathbb{E}(G_{i,m} | A_{n,m}, S_i = 0) - \varepsilon_{i,1-\alpha/2} \sqrt{\mathrm{Var}(G_{i,m} | A_{n,m}, S_i = 0)}, \mathbb{E}(G_{i,m} | A_{n,m}, S_i = 0) + \varepsilon_{i,1-\alpha/2} \sqrt{\mathrm{Var}(G_{i,m} | A_{n,m}, S_i = 0)}].$$

Next, our goal is to ensure that if a passenger does request to use a particular stop sufficiently far in advance, the operator will serve that stop with high probability. To achieve this, we introduce the look-ahead period $\tau_{n,m}$ to quantify how far downstream the bus is allowed to skip. We recall that a bus is allowed to skip at most m consecutive stops, which can also be expressed as downstream stops that are at most $\tau_{n,m}$ time units away. The value of this parameter is computed as

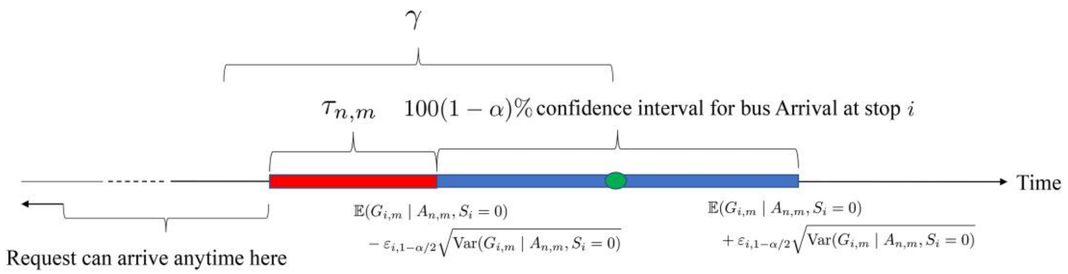
$$\tau_{n,m} = d_{n,m} + (m - 1)s,$$

where the first term represents the driving time needed to traverse m out of n stops, and the second term is the dwell time at stops (see Online Appendix E for an illustration). Therefore, a request to use stop i must be submitted at least $\tau_{n,m}$ time units before the lower bound of the interval to ensure that the bus does not skip the stop $100(1 - \alpha)\%$ of the time (see Figure 2 for an illustration). Accordingly, for each stop i, we impose the constraint

$$\varepsilon_{i,1-\alpha/2} \sqrt{\mathrm{Var}(G_{i,m} | A_{n,m}, S_i = 0)} + \tau_{n,m} \leq \gamma, \tag{9}$$

where the parameter γ represents the advance notice with which passengers must notify the operator of their intent to use a stop. This constraint ensures that if a passenger submits their request at least γ time units before

Figure 2. (Color online) Illustration of Passenger Reservation with Respect to Bus Arrival Time Window at Stop i



the bus's expected arrival at the desired stop, the stop will be visited $100(1-\alpha)\%$ of the time. Notably, a larger value of γ gives the transit operator more flexibility to skip stops without demand and, in turn, deploy additional stops (i.e., increase both m and n). However, a larger γ also inconveniences passengers, as requiring longer notice forces them to plan their trips further in advance to ensure their stops are not skipped. Thus, there is a trade-off between increasing the number of stops and the required advance notice. Transit operators must carefully consider this balance to improve coverage through higher stop density while limiting passenger inconvenience. Additionally, operational uncertainties arising from probabilistic constraints may prevent stop requests from being fulfilled $100 \cdot \alpha\%$ of the time. In such cases, operators can promptly notify the passengers through real-time communication if the stop cannot be served, and if necessary, alternative transportation (e.g., transportation network companies or taxi services) can be arranged to minimize passenger inconvenience.

3.5. Maximizing the Number of Stops

Our objective is to maximize the value of n, which represents the number of stops, while ensuring that the duration of the bus tour in the flexible system remains below the duration of the same tour in System 1 with high probability. Additionally, we aim to serve reservations submitted within the designated time windows with high probability. To achieve this objective, we formulate the following optimization problem:

$$\max_{n\in\mathbb{Z}_{++},\, m\in\mathbb{Z}_{++}} \quad n \tag{10}$$

$$\text{s.t.} \quad \mathbb{P}(T_{n,m} \le b|A_{n,m}) \ge 1-\beta, \tag{11}$$

$$\varepsilon_{i,1-\alpha/2}\sqrt{\mathrm{Var}(G_{i,m}|A_{n,m},S_i=0)}+\tau_{n,m} \le \gamma, \quad \forall i \in I, \tag{12}$$

$$m \le M. \tag{13}$$

The constraint in (11) is formulated using the PMF given in (5). This constraint ensures that, with a probability of $1-\beta$, the bus in System 2 returns to the depot before its counterpart in System 1 with $\overline{n}$ stops, thus enabling it to wait until time b to start the next tour, where $b = c + \overline{n} \cdot s$. This ensures that the variability in bus tour duration does not carry over to the subsequent tour and enables a fair comparison between the two systems. The set of constraints in (12) is based on (9). These constraints utilize the conditional variance as specified in (7), along with the parameter $\varepsilon_{i,1-\alpha/2}$ defined in (8). By incorporating these constraints, we can provide assurance that passengers' requests will be accommodated, and their stops will be reliably served. Constraint (13) restricts the number of consecutive skips to a maximum of M. This enables the operator to directly regulate the variability in the bus route.

It may be intuitively expected that the time windows for bus arrivals should expand at downstream stops. This could suggest that the constraint in (12) for the final stop, n, would dominate the same constraints for $i \in \{1, 2, \ldots, n-1\}$. Consequently, one might think that relaxing the constraints in (12) for $i \in \{1, 2, \ldots, n-1\}$ would not affect the overall problem. However, as we will demonstrate in Remark 1 (see proof in Online Appendix F), this is not the case.

Remark 1. The term $\varepsilon_{i,1-\alpha/2}\sqrt{\mathrm{Var}(G_{i,m}|A_{n,m}, S_i=0)} + \tau_{n,m}$ is not monotonically increasing with respect to i. Consequently, the constraints in (12) cannot be relaxed for $i \in \{1, 2, \ldots, n-1\}$.

3.6. Solution Approach

To devise an efficient solution method for (10)–(13), it is essential to understand the bounds imposed by the constraints. As n increases while m is fixed, the bus must visit more locations, extending the tour duration; thus, n cannot increase indefinitely without violating Constraint (11). Conversely, as n decreases while m remains fixed, the bus can skip a larger proportion of stops, leading to higher variance in arrival times and thereby increasing the risk of violating Constraint (12). Therefore, Constraint (11) establishes the upper bound for n, whereas (12) defines the lower bound. Feasible points are confined to the region defined by these bounds, though they may not be uniformly distributed within this region because of the nonmonotonic behavior of $\mathbb{P}(T_{n,m} \le b|A_{n,m})$ and $\varepsilon_{i,1-\alpha/2}\sqrt{\mathrm{Var}(G_{i,m}|A_{n,m}, S_i=0)} + \tau_{n,m}$ with respect to n.

We tackle Problem (10)–(13) using a custom grid search (Algorithm 1), which primarily relies on the upper bound to identify the optimal solution (n^*, m^*) that maximizes the number of stops. The algorithm starts with the maximum possible number of consecutive skips, $m = M$, and the corresponding upper bound for n. It then decreases n either until the first feasible point is found or until confirming that no feasible n exists. Afterward, it decrements m by one and repeats the process. Throughout, Algorithm 1 avoids evaluating solutions that are dominated by the best feasible solution found so far.

Algorithm 1 (A Custom Grid Search to Determine Optimal Solution for (10)–(13))

Input: γ; β; s; v; c; b; M; $\overline{n}$
Output: Optimal solution (n^*, m^*).

procedure Determining optimal solution based on custom grid search
  $m \leftarrow M; m^* \leftarrow 1; n^* \leftarrow \overline{n}$
  $n^* \leftarrow \overline{n}$ ▷ Set initial best solution
  while $m > 0$ **do**
    $\overline{N}_m = \inf\left\{n \in \mathbb{Z}^+ : \zeta_{n,m} < s, n \geq \frac{\overline{n}\cdot s\cdot(m+1)}{s-\zeta_{n,m}}\right\}$ ▷ Set upper bound
    $n \leftarrow \overline{N}_m$
    while $n > n^*$ **do**
      if $\mathbb{P}(T_{n,m} \leq b | A_{n,m}) \geq 1 - \beta$ **then**
        $\tau_{n,m} = d_{n,m} + (m-1)s$
        if $\varepsilon_{i,1-\alpha/2}\sqrt{\mathrm{Var}(G_{i,m}|A_{n,m}, S_i = 0)} + \tau_{n,m} \leq \gamma$, $\forall i \in I$ **then**
          if $m = M$ **then**
            $m^* \leftarrow m; n^* \leftarrow n$ ▷ Update best feasible solution
          else
            if $n > n^*$ **then**
              $m^* \leftarrow m; n^* \leftarrow n$ ▷ Update best feasible solution
            end if
          end if
        end if
      end if
      $n \leftarrow n - 1$
    end while
    $m \leftarrow m - 1$
  end while
  return (n^*, m^*) ▷ Return the optimal solution
end procedure

Next, we focus on the upper bound used in Algorithm 1. Specifically, we show that Constraint (11) is violated for any n greater than the established upper bound in Proposition 3 (proved in Online Appendix J). To motivate this proposition, we assume two intuitive properties of $\zeta_{n,k}$. First, for any given n, $\zeta_{n,k}$ increases with k because skipping more stops results in greater time savings. Second, for any fixed k, $\zeta_{n,k}$ decreases as n increases. Some of the subsequent analysis will be established based on these two properties.

Assumption 3. *The function $\zeta_{n,k}$, which represents the time saved by skipping k consecutive and evenly spaced stops along the route, has the following properties:*

- *For fixed n, $\zeta_{n,k}$ increases with k .*
- *For fixed k, $\zeta_{n,k}$ decreases with n, and $\lim_{n\to\infty}\zeta_{n,k} = 0$.*

Proposition 3. *Consider any fixed m, and define*

$$\overline{N}_m \triangleq \inf\left\{n \in \mathbb{Z}^+ : \zeta_{n,m} < s, n > \frac{\overline{n}\cdot s\cdot(m+1)}{s-\zeta_{n,m}}\right\}. \quad (14)$$

Suppose that Assumption 3 holds, which guarantees that the set in (14) is nonempty. Then, for all $n \geq \overline{N}_m$, it holds that $\mathbb{P}(T_{n,m} \leq b | A_{n,m}) = 0$.

Note that the upper bound $\overline{N}_m$ can be obtained by starting at $n = \overline{n}$ and incrementing it until the two conditions are met. Computing $\overline{N}_m$ in this manner is sufficient for Algorithm 1 to subsequently find an optimal solution. However, if we consider a specific bus route configuration in which an explicit expression of $\zeta_{n,k}$ is available, then we can further determine a slightly looser but closed-form upper bound $\breve{N}_m$ and use it to obtain $\overline{N}_m$ more efficiently as follows. For example, if we consider the circular route in Figure 1, then $\zeta_{n,k}$ represents the time difference between traveling along the circumference and the chord bypassing k stops:

$$\zeta_{n,k} = \frac{2r\pi(k+1)}{nv} - \frac{2r\sin\left(\frac{\pi(k+1)}{n}\right)}{v}, \quad (15)$$

where r is the radius of the circular route, and v is the speed of the vehicle (see Figure 16 in Online Appendix H for an illustration). As shown in Proposition 5, our explicit characterization of $\zeta_{n,k}$ allows us to derive a closed-form upper bound $\breve{N}_m$ for $\overline{N}_m$, thereby bracketing $\overline{N}_m$ in the finite interval $[\overline{n}(m+1), \breve{N}_m]$. Armed with this bracket, we can efficiently determine $\overline{N}_m$ by using a standard search method such as binary search, until the conditions in (14) are satisfied. Before deriving this upper bound (proved in Online Appendix I), we show in Proposition 4 that $\zeta_{n,k}$ under the circular route model satisfies the properties of Assumption 3. The formal proof of this proposition is provided in Online Appendix H.

Proposition 4. *For the circular route model, $\zeta_{n,k}$, as given in (15), satisfies the properties of Assumption 3.*

Proposition 5. *Consider any fixed m, and define*

$$\breve{N}_m = \left\lceil \frac{2r(m+1)}{sv}\left(\pi - \sin\left(\frac{\pi}{\overline{n}}\right)\overline{n}\right) + \overline{n}\cdot(m+1)\right\rceil.$$

Then, under the circular route model with $\zeta_{n,k}$ given in (15), it holds that $\overline{N}_m \leq \breve{N}_m$.

We can interpret the bound $\breve{N}_m$ as follows. The second term, $\overline{n}(m+1)$, indicates that if we allow skipping up to m consecutive stops, then for every $m+1$ stops, we can always skip m of them and visit the last one. This approach is feasible because we still visit $\overline{n}$ stops as in System 1, although without considering shortcuts. Thus, $\overline{n}(m+1)$ would serve as an upper bound if the bus were constrained to travel only along the circle. Furthermore, the first term, $\frac{2r(m+1)}{sv}\left(\pi - \sin\left(\frac{\pi}{\overline{n}}\right)\overline{n}\right)$, accounts for the time-saving benefit of taking shortcuts. Together, these two terms in the definition of $\breve{N}_m$ illustrate the maximum number of stops that can be possibly added if no probability constraint on bus tour duration is considered.

Next, we turn our attention to Algorithm 1 and its computational complexity. Note that the upper bound $\overline{N}_m$ is applicable to all feasible values of n for a given fixed m. Furthermore, because $\zeta_{n,m}$ is increasing in m (Assumption 3), we can see that $\overline{N}_m$ is also increasing in m. Thus, we have $\sup_{m\le M}\overline{N}_m = \overline{N}_M$, which implies that Algorithm 1 examines at most $M(\overline{N}_M - \overline{n})$ points to return an optimal solution (n^*, m^*). However, for each (n,m), we need to compute the set $\mathcal{X}_{n,m}$ defined in Proposition 2, whose cardinality increases exponentially with n and m. Specifically, the upper bound on its cardinality is $\min\left\{\frac{(2n+m(m+3))^m}{m!\cdot(m+1)!\cdot 2^m}, \binom{m+\lfloor\frac{n}{2}\rfloor}{m}\right\}$, resulting in an exponential run time for Algorithm 1. The computational complexity of Algorithm 1 is presented in Proposition 6.

Proposition 6. *The time complexity of Algorithm 1 for finding the optimal solution to Problem (10)–(13) is*

$$\mathcal{O}\left((M^2\overline{N}_M^2 + M\overline{N}_M^3)\cdot \min\left\{\frac{(2\overline{N}_M + M(M+3))^M}{M!\cdot(M+1)!\cdot 2^M}, \binom{M+\left\lfloor \frac{\overline{N}_M}{2}\right\rfloor}{M}\right\}\right).$$

In particular, for the circular route model, the time complexity of Algorithm 1 is $\mathcal{O}(M^2(\frac{e^2}{2})^M)$.

Algorithm 2 (Dynamic Programming Solution to Determine N_m^*)

Input: $m, \overline{N}_m, s, b, c$
Output: Tightened upper bound N_m^*
$n \leftarrow \overline{N}_m$
Initialize $t_{n,m}^* \leftarrow c + n\cdot s$ ▷ Initialize $t_{n,m}^*$ at the largest possible value
while $t_{n,m}^* > b$ **do**
  $n \leftarrow n - 1$
  Calculate $\zeta_{n,k}$ for each $k \in \{1,2,\ldots,m\}$
  Define $D \leftarrow \{d_i = 0 \mid i \in \{1,2,\ldots,n+1\}\}$ ▷ Initialize dynamic programming tabulation
  for $i \in [1,n]$ **do**
    for $k \in [1,m]$ **do**
      if $i \ge k+1$ **then**
        $d_i \leftarrow \max\{d_i, d_{i-(k+1)} + (\zeta_{n,k} + k\cdot s)\}$
      end if
    end for
  end for
  Initialize $x_k \leftarrow 0$ for each $k \in \{1,2,\ldots,m\}$
  $i \leftarrow n$
  while $i > 1$ **do**
    for $k \leftarrow [1,m]$ **do**
      if $i \ge k+1$ and $d_i = d_{i-(k+1)} + (\zeta_{n,k} + k\cdot s)$ **then**
        $x_k \leftarrow x_k + 1$
        $i \leftarrow i - (k+1)$
        break
      end if
    end for
  end while
  $t_{n,m}^* \leftarrow c + n\cdot s - \sum_{x_k=1}^{m} x_k(\zeta_{n,k} + k\cdot s)$
end while
Return $N_m^* \leftarrow n$

Since checking the feasibility of each point (n,m) is computationally intensive because of the size of the set $\mathcal{X}_{n,m}$, we further tighten the upper bound $\overline{N}_m$ in Proposition 7 (see proof in Online Appendix K) by repeatedly applying dynamic programming (Algorithm 2) to solve the unbounded knapsack problem. This additional effort to reduce the feasible region is well justified, especially for large m. Specifically, to eliminate a point (n,m), Algorithm 2 applies dynamic programming with a pseudopolynomial time complexity of $\mathcal{O}(nm)$. This approach is more efficient than constructing the set $\mathcal{X}_{n,m}$ and checking the feasibility of constraints because the lower bound on the cardinality of $\mathcal{X}_{n,m}$ is $\max\left\{\frac{n^m}{m!\cdot(m+1)!}, \binom{m+\lfloor\frac{n}{2}\rfloor}{m}\right\}$ (see the proof of Proposition 6 in Online Appendix J). Thus, the benefit of applying Algorithm 2 to reduce the feasible region outweighs the cost of the additional computations required to tighten the upper bound $\overline{N}_m$.

Proposition 7. *The output of Algorithm 2, denoted by* N_m^*, *provides a tighter upper bound on* n *than* $\overline{N}_m$; *that is,* $N_m^* \le \overline{N}_m$.

4. Methodology: Simulation-Based Models

Building on our previous discussion of a stylized bus route model with uniformly spaced stops, we now turn to more complex scenarios where stops are nonuniformly distributed. To capture this realism, we propose two distinct simulation-based models tailored to different demand scenarios. In the first scenario, where demand is assumed to remain unchanged despite the addition of new stops, our model places stops to minimize walking distances for passengers. Conversely, in the second scenario, which assumes an increase in demand due to new stops, our model deploys stops in a manner that maximizes demand coverage.

4.1. Minimizing Walking Distance for Fixed Demand

Consider a conventional bus route with a set of current stops, denoted by N. Passengers travel on foot from their starting locations to the nearest bus stop in N, and from their alighting bus stop in N to their final destinations. To enhance this route, we define C as a set of candidate locations for new bus stops, as identified by a transit authority. We consider the integration of a subset of these candidate stops, denoted by $S \subseteq C$, into the existing route. In this context, $\xi(N \cup S)$ represents the stochastic demand at stops, which includes both the existing stops and the new stops.

The function $f(N \cup S, \xi(N \cup S))$ quantifies the total walking distance for all passengers, taking into account the distances from each passenger's starting location to the nearest stop within $N \cup S$ and from their alighting stop to their final destination. Integration of candidate stops S into the existing route also considers the variability in bus tour durations. The duration of the bus tour and the arrival times at the stops, given the new configuration of stops and their demand, are denoted by $T(N \cup S, \xi(N \cup S))$ and $G_i(N \cup S, \xi(N \cup S))$, respectively. For notational convenience, we abbreviate them to T and G_i in the following context. These metrics provide insights into the system's performance, influenced by both the locations of the stops and passenger demand.

Our objective is to select a subset $S \subseteq C$ for deploying new bus stops to minimize the expected walking distance, ensuring compliance with constraints analogous to (11) and (12). This can be formulated as

$$\begin{aligned} \min_{S \subseteq C,\, \tau \geq 0} \quad & \mathbb{E}[f(N \cup S, \xi(N \cup S))] \\ \text{s.t.} \quad & \mathbb{P}(T(N \cup S, \xi(N \cup S)) \leq b \,|\, \pi(\tau)) \geq 1 - \beta, \\ & \varepsilon_{i,1-\alpha/2} \sqrt{\mathrm{Var}(G_i(N \cup S, \xi(N \cup S)) \,|\, \pi(\tau))} + \tau \leq \gamma, \\ & \qquad \forall i \in N \cup S, \end{aligned} \tag{16}$$

where the variance in the bus's arrival times at individual stops, as well as the overall duration of the bus tour, is conditioned on the policy $\pi(\tau)$. This policy establishes restrictions on the number of stops a bus can consecutively skip. Specifically, under $\pi(\tau)$, a bus is not permitted to bypass any downstream stop that lies beyond a threshold of τ time units, where τ is the look-ahead period.

4.2. Greedy Solution Approach

To address Problem (16), we devise a greedy heuristic, which is coupled with a Bayesian framework for the evaluation of constraints, as detailed in Algorithm 3. This algorithm iteratively evaluates each remaining candidate stop one at a time, selecting the feasible stop that provides the highest improvement in the objective value. This process is repeated until no further improvements can be made. We estimate the objective function in (16), representing the expected walking distance, through the application of the sample average approximation (SAA) technique (Shapiro, Dentcheva, and Ruszczynski 2021). Let $\mathcal{W}$ represents a set of scenarios for stochastic demand, where each scenario $w \in \mathcal{W}$ corresponds to a specific realization of demand patterns. Similarly, $\xi_w(N \cup S)$ denotes the demand realization under scenario w, for the combined set of existing stops N and new stops S. Specifically, for a given set S and policy $\pi(\tau)$, we simulate bus operations under each demand scenario w and approximate the objective function in (16) as follows:

$$\frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} f(N \cup S,\, \xi_w(N \cup S)),$$

where the summation over $w \in \mathcal{W}$ allows us to average the simulated walking distances across all demand scenarios, providing a comprehensive evaluation of the expected walking distance.

Concurrently, we employ a Bayesian framework to analyze the simulation experiment outputs and assess violation of constraints in (16) with as few simulation runs as possible. This approach holds a distinct advantage over traditional frequentist methods, which generally necessitate a large predetermined number of simulations. In contrast, we will sequentially conduct simulation experiments to collect observations about the arrival times and the bus tour duration. Then, we will establish a Bayesian online learning model for estimating the unknown parameters, which can be compactly updated using the sequential observations. Based on this Bayesian framework, we are able to efficiently evaluate the feasibility of S and $\pi(\tau)$ by invoking the constraints in (16). In particular, we establish a stopping rule for determining the termination of the sequential experiments, which helps reduce the number of simulation experiments needed for estimation while still maintaining desirable solution quality.

In addition to reducing the number of simulation experiments, we also aim to decrease the duration of individual simulation runs by aborting them immediately once the bus tour duration surpasses the upper bound b. This operation results in two types of observations from our simulations: (i) complete observations, which occur when the bus tour duration remains below the upper bound b, allowing this simulation run to complete, and (ii) censored observations, which occur when the bus tour duration exceeds b, leading to early termination of this simulation run and resulting in a binary signal that only indicates the upper bound was exceeded. Obviously, such early termination reduces the duration of our simulation runs whenever it occurs, but the resulting censored observations only carry incomplete information about the simulated bus tours. Nonetheless, we are able to develop a compact online learning model via an approximate Bayesian approach (Chen et al. 2022), which effectively learns from both complete and censored observations and thus achieves the same level of learning efficiency as if only complete observations were collected.

4.2.1. Conjugate Bayesian Approach for Arrival Times. Denote by $\mathfrak{G}_{i,r}$ the arrival time at stop i in the rth simulation experiment and assume $\mathfrak{G}_{i,1}, \mathfrak{G}_{i,2}, \ldots$ are independent and identically distributed (i.i.d.) samples of G_i. Because G_i must be positive, we model G_i with a log-normal distribution, that is, $\log(G_i) \sim \mathcal{N}(\Omega_i, \Sigma_i^2)$, where both Ω_i and Σ_i^2 are unknown parameters. We specifically use the log-normal distribution because it not only has a positive support but also allows flexible parameterization that can fit a variety of density curve shapes. Then, to efficiently learn these unknown parameters from the sequential observation $\mathfrak{G}_{i,r}$, we establish a Bayesian learning model by assuming a normal-inverse-gamma (NIG) prior distribution for (Ω_i, Σ_i^2), that is, $\Sigma_i^2 \sim \text{Inverse-Gamma}(\chi_{i,0}, \Upsilon_{i,0})$, and $\Omega_i | \Sigma_i^2 \sim \mathcal{N}(\Theta_{i,0}, \Sigma_i^2/\Lambda_{i,0})$, where $\chi_{i,0}$, and $\Upsilon_{i,0}$ are the parameters of the inverse-gamma distribution that represent our prior belief about the variance Σ_i^2, and $\Theta_{i,0}$ is our prior belief about the mean Ω_i with $\Lambda_{i,0}$ indicating how confident we are about our prior belief. Because of the conjugate property of the NIG prior (DeGroot 1970, Koch 2007), the posterior distribution of (Ω_i, Σ_i^2) still belongs to the NIG distribution family. Therefore, whenever we collect a new observation $\mathfrak{G}_{i,r}$ from the simulation experiment, we can compactly and efficiently update the posterior distribution of (Ω_i, Σ_i^2) by simply updating the parameters of the prior distribution at that time stage. To be specific, let $(\chi_{i,r}, \Upsilon_{i,r}, \Theta_{i,r}, \Lambda_{i,r})$ denote the prior parameters of (Ω_i, Σ_i^2) at the rth time stage where the observations from the first r simulation experiments, that is, $\mathfrak{G}_{i,1}, \ldots, \mathfrak{G}_{i,r}$, have been collected and the $(r+1)$st experiment is to be conducted. Then, when $\mathfrak{G}_{i,r+1}$ is collected from the $(r+1)$st experiment, the posterior parameters $(\chi_{i,r+1}, \Upsilon_{i,r+1}, \Theta_{i,r+1}, \Lambda_{i,r+1})$ can be obtained by

$$\begin{aligned}
\chi_{i,r+1} &= \chi_{i,r} + \frac{1}{2}, \\
\Upsilon_{i,r+1} &= \Upsilon_{i,r} + \frac{\Lambda_{i,r}(\log(\mathfrak{G}_{i,r+1}) - \Theta_{i,r})^2}{2(\Lambda_{i,r}+1)}, \\
\Theta_{i,r+1} &= \frac{\Lambda_{i,r}\Theta_{i,r} + \log(\mathfrak{G}_{i,r+1})}{\Lambda_{i,r}+1}, \\
\Lambda_{i,r+1} &= 1 + \Lambda_{i,r}. \qquad (17)
\end{aligned}$$

Then, $(\chi_{i,r+1}, \Upsilon_{i,r+1}, \Theta_{i,r+1}, \Lambda_{i,r+1})$ will become the prior parameters at the $(r+1)$st time stage. Therefore, we can recursively update these parameters using (17) at each time stage when a new $\mathfrak{G}_{i,r}$ is collected from the sequential simulation experiments.

4.2.2. Approximate Bayesian Approach for Tour Duration. Denote by $\mathfrak{T}_r$ the bus tour duration in the rth simulation experiment and assume $\mathfrak{T}_1, \mathfrak{T}_2, \ldots$ are i.i.d. samples of T. Similarly as in Section 4.2.1, we model T with a log-normal distribution, that is, $\log(T) \sim \mathcal{N}(\rho, \sigma^2)$, where ρ is the unknown mean of $\log(T)$ with σ^2 being the variance. For notational convenience and model tidiness, we assume the variance σ^2 is known. In practice, we can estimate σ^2 using the sample variance, as will be illustrated later in Section 4.2.3.

To efficiently learn ρ using the sequential outputs, we again establish a Bayesian learning model by assuming it has a normal prior, that is, $\rho \sim \mathcal{N}(\theta_0, \upsilon_0)$. Should all $\mathfrak{T}_r$ be fully observed, we could then devise a compact Bayesian learning model utilizing conjugacy, as described in Section 4.2.1. However, given that any simulation run exceeding the bus tour duration threshold b is terminated immediately, not all $\mathfrak{T}_r$ will be observable. Thus, we define a binary signal ψ_r to indicate whether $\mathfrak{T}_r$ is observed in the rth simulation. That is, $\psi_r = 1$ if $\mathfrak{T}_r \le b$, and $\mathfrak{T}_r$ is observed, and $\psi_r = 0$ if $\mathfrak{T}_r > b$, and the exact value of $\mathfrak{T}_r$ will not be available because the simulation experiment is not finished because of early termination. Consequently, we will not be able to construct a conjugate Bayesian online learning model for ρ as for (Ω_i, Σ_i^2) in Section 4.2.1, because the conjugacy property no longer exists because of the incomplete observations. Therefore, we will adopt an approximate Bayesian approach to learn ρ, which can achieve the same level of learning efficiency as a conjugate Bayesian learning model while theoretically guaranteeing the consistency of the estimators (Chen and Ryzhov 2020). To proceed, denote by θ_r and υ_r the prior mean and variance of ρ at the rth time stage where the observations from the first r simulation experiments have been collected. Then, if $\psi_{r+1} = 1$ and $\mathfrak{T}_{r+1}$ is observed from the $(r+1)$st experiment, we can simply update the prior parameters based on the conjugacy property as in Section 4.2.1 using

$$\begin{aligned}
\theta_{r+1} &= \theta_r + \frac{\log(\mathfrak{T}_{r+1}) - \theta_r}{\sigma^2 + \upsilon_r}\upsilon_r, \\
\upsilon_{r+1} &= \upsilon_r - \frac{\upsilon_r^2}{\sigma^2 + \upsilon_r}. \qquad (18)
\end{aligned}$$

Then, $\mathcal{N}(\theta_{r+1}, \upsilon_{r+1})$ will become the prior distribution at the $(r+1)$st time stage.

However, if $\psi_{r+1} = 0$, then the conjugacy property is no longer available because $\mathfrak{T}_{r+1}$ will not be available. In such a case, Chen et al. (2022) propose an alternative for constructing closed-form updates for θ_r and υ_r based on approximate Bayesian inference. Specifically, they approximate the exact posterior distribution of ρ given $\psi_{r+1}=0$ with an artificial normal distribution $\mathcal{N}(\theta_{r+1}, \upsilon_{r+1})$ through matching the first two moments of these two distributions, that is, $\theta_{r+1}=\mathbb{E}(\rho|\psi_{r+1}=0)$ and $\upsilon_{r+1}=Var(\rho|\psi_{r+1}=0)$. Solving these moment-matching equations, we can obtain the following closed-form updates for θ_r and υ_r:

$$\theta_{r+1}=\theta_r+\frac{\phi(\eta_r)}{(1-\Phi(\eta_r))\sqrt{\sigma^2+\upsilon_r}}\upsilon_r, \tag{19}$$

$$\upsilon_{r+1}=\upsilon_r-\frac{\upsilon_r}{\sigma^2+\upsilon_r}+\frac{\upsilon_r^2}{\sigma^2+\upsilon_r}\cdot\left(1-\eta_r\frac{\phi(\eta_r)}{\Phi(\eta_r)}-\frac{\phi(\eta_r)^2}{\Phi(\eta_r)^2}\right)^2, \tag{20}$$

where $\eta_r=\frac{\log b-\theta_r}{\sqrt{\sigma^2+\upsilon_r}}$, and ϕ and Φ are the probability density function and CDF of the standard normal distribution respectively. Once θ_r and υ_r are updated using (19) and (20), we will discard the exact posterior distribution of ρ given $\psi_{r+1}=0$, and $\mathcal{N}(\theta_{r+1}, \upsilon_{r+1})$ will become the prior distribution at the $(r+1)$st time stage. Additionally, although the variance reductions in (18) and (20) are different, Chen et al. (2022) note that the impact of the additional variance reduction term in (20) is usually minimal because the variance υ_r becomes small when r gets sufficiently large, making (18) and (20) approximately identical. Therefore, they suggest using (18) uniformly to update υ_r. Thus, we follow their suggestion and summarize the recursive updates for θ_r and υ_r at each time r as follows:

$$\theta_{r+1}=\theta_r+\psi_{r+1}\frac{\log(\mathfrak{T}_{r+1})-\theta_r}{\sigma^2+\upsilon_r}\upsilon_r + (1-\psi_{r+1})\frac{\phi(\eta_r)}{(1-\Phi(\eta_r))\sqrt{\sigma^2+\upsilon_r}}\upsilon_r,$$

$$\upsilon_{r+1}=\upsilon_r-\frac{\upsilon_r^2}{\sigma^2+\upsilon_r}. \tag{21}$$

4.2.3. Bayesian Evaluation of Constraints. With the posterior distributions recursively specified by (17) and (21), we can sequentially evaluate the constraints in (16). At each time r, because $T|\rho$ is normally distributed, we can compute the probability that the bus tour duration remains within a desired threshold b given ρ is equal to its prior mean:

$$\mathbb{P}(T\le b|\rho=\theta_r)=\mathbb{P}(\log(T)\le\log b|\rho=\theta_r) = \Phi\left(\frac{\log b-\theta_r}{\sigma}\right). \tag{22}$$

Similarly, because $G_i|(\Omega_i,\Sigma_i^2)$ follows a log-normal distribution, we can specify the uncertainty of the arrival time at stop i given that Ω_i and Σ_i^2 are equal to their prior means as follows:

$$\mathrm{Var}\left(G_i|\Omega_i=\Theta_{i,r},\Sigma_i^2=\frac{\Upsilon_{i,r}}{\chi_{i,r}-1}\right) = \left[\exp\left(\frac{\Upsilon_{i,r}}{\chi_{i,r}-1}\right)-1\right]\exp\left(2\Theta_{i,r}+\frac{\Upsilon_{i,r}}{\chi_{i,r}-1}\right). \tag{23}$$

Note that the constraints in (16) are imposed such that the bus tour duration and the arrival times can be controlled within desired thresholds with certain confidence levels. Therefore, combining (22) and (23) with the constraints in (16), we propose a stopping rule to determine the termination of the sequential simulation experiments by checking the following at the beginning of each time stage r:

$$\Phi\left(\frac{\log b-\theta_r}{\sigma}\right)\ge 1-\beta, \tag{24}$$

$$\varepsilon_{i,1-\alpha/2}\cdot\sqrt{\left[\exp\left(\frac{\Upsilon_{i,r}}{\chi_{i,r}-1}\right)-1\right]\exp\left(2\Theta_{i,r}+\frac{\Upsilon_{i,r}}{\chi_{i,r}-1}\right)} + \tau\le\gamma, \quad \forall i\in N\cup S, \tag{25}$$

where $\varepsilon_{i,1-\alpha/2}$ is adaptively determined as at the beginning of Section 3.4. Specifically, for any potential inclusion of stops in S with policy $\pi(\tau)$, we will run at least r_{min} simulation experiments for the Bayesian learning model to warm up. Additionally, in practice, if σ^2 is unknown, we can estimate it sequentially with $\hat{\sigma}_r^2$, the sample variance of $\log(\mathfrak{T}_1),\ldots,\log(\mathfrak{T}_r)$, for each time stage $r\le r_{min}$, because the first r_{min} simulation experiments are complete with no early termination. Then, for all $r>r_{min}$, we can simply replace σ^2 with $\hat{\sigma}^2_{r_{min}}$ in the updating equations. Meanwhile, we will also run at most r_{max} simulation experiments to prevent excessive computation. Then, starting from the $(r_{min}+1)$st experiment, we will immediately terminate the simulation experiment and deem the inclusion S with policy $\pi(\tau)$ infeasible if Constraints (24) and (25) are violated in R consecutive experiments. Otherwise, if we can finish all r_{max} experiments with no R consecutive violations of (24)–(25), then we will deem the inclusion S with policy $\pi(\tau)$ feasible. Note that we impose R consecutive violations for concluding infeasibility to account for the numerical error due to the convergence of the Bayesian learning model. We repeat the abovementioned process by incrementing τ. Because τ determines the number of downstream stops that a bus can skip, using a precision of less than one minute is not meaningful from a practical perspective as it would not impact the number of stops a bus can skip. Therefore, we treat τ as an integer expressed in minutes and increment it by one in the algorithm. The entire procedure is summarized below in Algorithm 3.

Algorithm 3 (Greedy Heuristic and Approximate Bayesian Updating for Stop Allocation)

Input: $r_{min}, r_{max}, R, b, \gamma, \alpha, \beta, \mathcal{W}, C,$
Output: S^*: selected stops $S^* \subseteq C$; τ^*: optimized lookup period

$S^* \leftarrow \emptyset; J_{\tau-1} \leftarrow \emptyset; \tau \leftarrow 0; \sigma \leftarrow 1; \varepsilon_{i,1-\alpha/2} \leftarrow 0$
Initialize Bayesian parameters $\theta_0, \upsilon_0, \Theta_{i,0}, \Lambda_{i,0}, \chi_{i,0}, \Upsilon_{i,0}$
while $\varepsilon_{i,1-\alpha/2} \cdot \sqrt{\left[\exp\left(\frac{\Upsilon_{i,r}}{\chi_{i,r}-1}\right) - 1\right] \exp\left(2\Theta_{i,r} + \frac{\Upsilon_{i,r}}{\chi_{i,r}-1}\right)} + \tau \leq \gamma$, $\forall i \in I$ **do**
  $\tau \leftarrow \tau + 1$
  Generate passengers and determine set of stops skipped J_τ
  if $J_\tau = J_{\tau-1}$ **then** continue
  else
    $J_{\tau-1} \leftarrow J_\tau; S \leftarrow \emptyset; r \leftarrow 0$
    $\theta_r \leftarrow \theta_0; \upsilon_r \leftarrow \upsilon_0;\ \Theta_{i,r} \leftarrow \Theta_{i,0}; \Lambda_{i,r} \leftarrow \Lambda_{i,0}; \chi_{i,r} \leftarrow \chi_{i,0}; \Upsilon_{i,r} \leftarrow \Upsilon_{i,0}$
    repeat
      $r \leftarrow r + 1$;
      if $r \leq r_{min}$ **then**
        $\sigma^2 \leftarrow \hat{\sigma}_r^2$; Update $\theta_r, \upsilon_r, \Theta_{i,r}, \Lambda_{i,r}, \chi_{i,r}, \Upsilon_{i,r}$ using (17) and (21)
      end if
      if $r > r_{min}$ **then**
        $\sigma^2 \leftarrow \hat{\sigma}_{r_{min}}^2$; Update $\theta_r, \upsilon_r, \Theta_{i,r}, \Lambda_{i,r}, \chi_{i,r}, \Upsilon_{i,r}$ using (17) and (21)
        Calculate $\varepsilon_{i,1-\alpha/2}$ and check constraints using (24) and (25)
        if Constraints are not satisfied for R consecutive experiments **then** break
        end if
      end if
    until $r = r_{max}$
    if $r = r_{max}$ **then**
      $i^* \leftarrow \arg\min_{i \in C \setminus S} \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} f(N \cup S \cup \{i\}, \xi_w(N \cup S \cup \{i\}))$,
      $S \leftarrow S \cup \{i^*\}$
    end if
    if $|S| \geq |S^*|$ **then**
      $\tau^* \leftarrow \tau; S^* \leftarrow S$
    end if
  end if
end while
return S^*, τ^*

4.3. Computational Savings from Bayesian Approaches

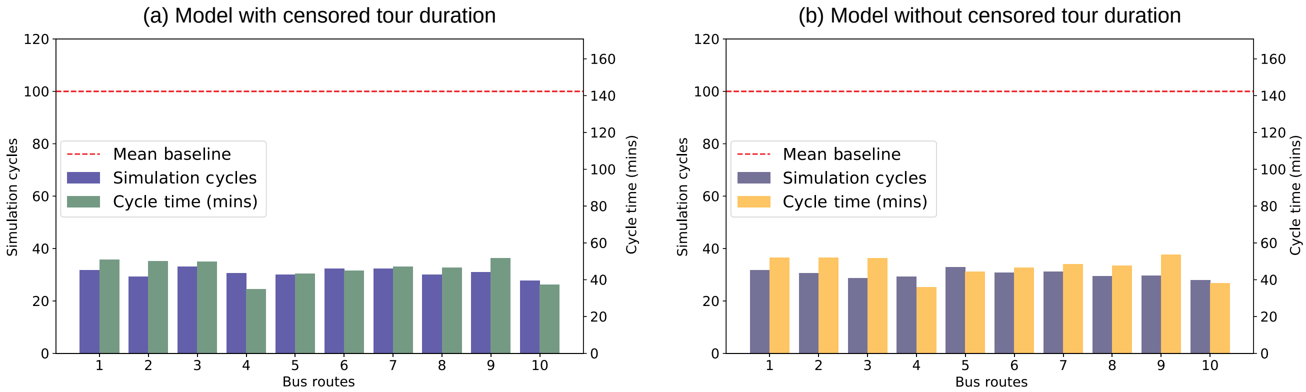
We begin by demonstrating the computational efficiencies of our proposed approximate Bayesian approach as compared with traditional frequentist methods. To highlight the efficiency improvement, we conduct simulations involving 10 bus route operations, with details to be elaborated upon later in the text. The results for each route are compared against those obtained using a frequentist approach, which consistently runs 100 simulations for each candidate solution and estimates the objective function with SAA. Our Bayesian approach, employing compact sequential updates to incorporate the information of each new observation, rapidly converge to an accurate depiction of the underlying distribution. This efficiency translates to significant computational savings, as demonstrated in Figure 3. Regarding the number of simulation cycles, the Bayesian method takes only 30 simulation cycles on average, achieving a 69% reduction compared with the frequentist approach. Meanwhile, the frequentist method requires an average of 142 minutes in simulation time, whereas the Bayesian method achieves similar accuracy using only 47 minutes on average, representing a 67.92% reduction in simulation time. The solution quality of the approximate Bayesian method for these instances is reported in Table 1. The difference between the walking distances from the Bayesian and frequentist methods averages around only 1.93%, which suggests that Bayesian method achieves a level of estimation accuracy equivalent to that of the frequentist method. However, the simulation time required by the Bayesian method is significantly reduced (67.92% shorter on average) in contrast to the frequentist method, which demonstrates the efficiency of our sequential approximate Bayesian approach.

Furthermore, in addition to measuring the reduction in the number of simulation experiments relative to a frequentist approach, we also assess the efficiency improvement of our proposed approximate Bayesian framework in which early termination will be applied to the simulation runs if their bus tour exceeds the threshold b. To do so, we design a benchmark in which all simulation runs are complete without early termination, and thus all complete $\mathfrak{T}_r$ can be observed with no censorship. Note that because all complete $\mathfrak{T}_r$ are available in this benchmark, we no longer need an approximate Bayesian approach and can simply build a conjugate Bayesian learning model for T exactly the same way as for G_i in Section 4.2.1. This head-to-head comparison between the results from panels (a) and (b) of Figure 3 shows that, although both approaches required the same number of simulations to meet the termination criterion, our approximate Bayesian framework with early termination was able to achieve a reduction of 2.1%–3.8% in the run time for individual simulations.

4.4. Maximizing Population Coverage for Increasing Demand

The previous discussion assumed a constant demand for transit, irrespective of the addition of new stops. However, in reality, the addition of new stops is likely to attract new passengers, which can affect the probability of stops being skipped. As the number of passengers increases, the probability of skipping stops decreases,

Figure 3. (Color online) Performance of Bayesian Approaches for 10 Route Instances



Note. A paired *t*-test comparing the simulation cycle reduction achieved by the model with censored tour duration (69%) and the model without censored tour duration (69%) showed no statistically significant difference in mean values.

which in turn reduces the number of new stops that can be added without exceeding the desired roundtrip duration. Therefore, it is important to consider the dynamic nature of demand and its impact on the feasibility of adding new stops while maintaining an acceptable bus tour duration. To this end, we consider another practical scenario in which additional demand is covered with each added stop. We adopt a simulation-based approach akin to that in Section 4.1, aiming to maximize population coverage by adding stops to an existing bus route. The same greedy solution method from Section 4.2 is used for this increasing demand case. The impacts and benefits from analyzing increasing demand are discussed in Online Appendix L.

5. Numerical Experiments of Analytical Derivations

In this section, we conduct empirical validation and sensitivity analysis of our analytical derivations for the stylized semiflexible bus transit system. Initially, we validate the derivations through a comprehensive simulation model. Subsequently, we optimize the number of stops and conduct a sensitivity analysis.

5.1. Validation via Simulation

We validate the analytical derivations in (2)–(9) using a simulation model with the same parameters. This model simulated a bus navigating a circular route, addressing passenger demands at various stops, with the number of stops n ranging from 7 to 25 in the flexible system. The purpose of this simulation was to confirm the accuracy of the stylized model by comparing key metrics such as driving time, waiting time, tour duration, and the time savings from shortcuts.

In the simulation, a bus was allowed to complete 1,000 cycles along the circular route, serving passenger demands that are uniformly distributed along the route. For validation, we systematically compared these key metrics between the analytical derivation and the simulation across the specified range of stops. The results, which underscore the consistency between the simulation outcomes and the derivations, are detailed in Online Appendix M.

5.2. Sensitivity Analysis

We optimized the number of stops and consecutive skips in the semiflexible system by solving the model (10)–(13) using Algorithm 1 for varying parameters. For

Table 1. Solution Quality Comparison for 10 Instances

	Walking (ft/passenger)		Time (mins)		Gap (%)	
Instance	Frequentist	Bayesian	Frequentist	Bayesian	Walking	Time
1	2,730.21	2,782.21	146.06	50.75	1.90	65.25
2	2,736.05	2,785.78	132.49	49.96	1.82	62.29
3	2,751.91	2,793.58	137.71	49.77	1.51	63.86
4	2,730.61	2,770.37	135.56	34.89	1.46	74.27
5	2,732.56	2,776.98	148.34	43.33	1.63	70.79
6	2,732.66	2,787.91	143.51	45.01	2.02	68.64
7	2,759.06	2,822.88	145.47	47.03	2.31	67.67
8	2,758.87	2,807.58	148.72	46.45	1.77	68.76
9	2,747.04	2,802.09	143.28	51.61	2.00	63.97
10	2,740.86	2,792.36	141.87	37.27	1.88	73.74

this purpose, we considered a fixed bus transit system (System 1) running on a 15-mile circular route (equivalent to a radius of 2.4 miles) with $\overline{n} = 5$ stops as our baseline. Assuming a speed of 25 mph and a dwell time of 30 seconds at each stop, the bus tour duration for the fixed system was kept deterministic at 38.7 minutes, the maximum possible tour duration when all stops are visited. The passenger demand in the fixed system is uniformly distributed along the route.

Considering the impact of the passenger demand on the probability of skipping stops, we explored how varying the probability of a stop in System 1 not being used ($\overline{p}$) affects the optimal solutions obtained from the optimization model (10)–(13), while keeping other parameters constant. The results of this analysis are depicted in Figure 4(a). Clearly, a decrease in $\overline{p}$ leads to more frequent use of existing stops and a lower likelihood of skipping them. This diminished scope for skipping stops consequently restricts the number of new stops that can be added, resulting in smaller values of (n^*, m^*). This observation highlights that the benefits of the proposed semiflexible system become more pronounced in scenarios with lower demand routes.

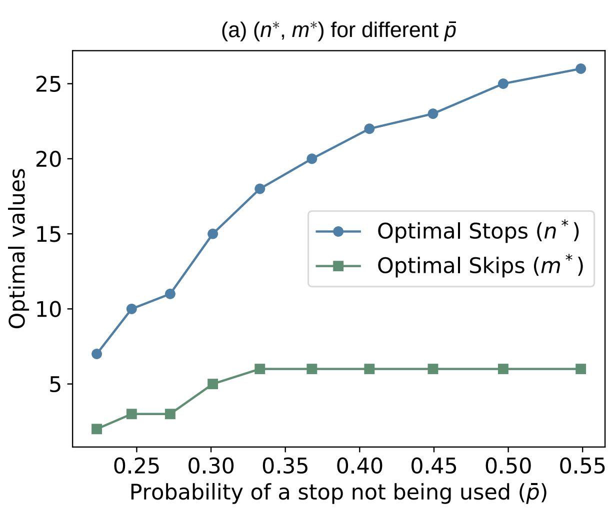
We also investigated the impact of the advance notice (γ) in (12) on the optimal solutions of the optimization model (10)–(13), while keeping other parameters constant. As illustrated in Figure 4(b), longer advance notice allows for the implementation of more stops, leading to higher values of n^* and m^*. This is logical, as larger values of γ provide operators with more flexibility to skip stops, though this also requires passengers to plan their trips further in advance. Detailed representations of the feasible solution space and the optimal solutions for a range of γ values are provided in Online Appendix N.

6. Allegany County Case Study: Simulation-Based Models

Allegany County is a sparsely populated county located in the western part of Maryland in the United States. With a population of around 70,000 people, the county has a relatively low population density, which is approximately 125 people per square mile (48 people per square kilometer). The county's transit network includes 10 bus routes with different configurations (see Online Appendix O). The route schedule and stop locations were extracted from General Transit Feed Schedule (GTFS) data (https://transitfeeds.com/). Additionally, automatic passenger counting (APC) data were mined to estimate average passenger demand at each stop, and the shortcuts between stops were queried from the GraphHopper routing application programming interface (https://graphhopper.com/).

In each route, we maintained all existing stops and considered 60 additional candidate locations for potential new stops in our simulation-based optimization. Initially, candidate stop locations were identified at the geographic midpoint between existing stops. The process then proceeded by identifying additional candidate stops at the midpoints between these newly established candidate stops until a total of 60 candidate stops were included. Although the candidate stops are selected along the bus routes, their locations are not restricted to the bus route itself. Any location deemed suitable by the transit authority is a viable option for stop placement. After applying simulation-based optimization, improvements were noted in the Blue, Red, and Yellow Lines. Other routes did not show improvement because of high utilization of existing stops (Figure 5(a)) or route configurations that do not allow for significant time savings from shortcuts when stops are skipped (Figure 5(b)). The improved routes, which have lower demand rates as shown in Figure 5(a), allow for more frequent skipping of stops. Additionally, the layout of the improved routes (Figure 6) allows shortcuts that yield greater savings compared with routes with a straight configuration, which limit the potential benefits of bypassing stops.

Figure 4. (Color online) Sensitivity Analysis of the Optimal Solution for Different Values of $\overline{p}$ and γ



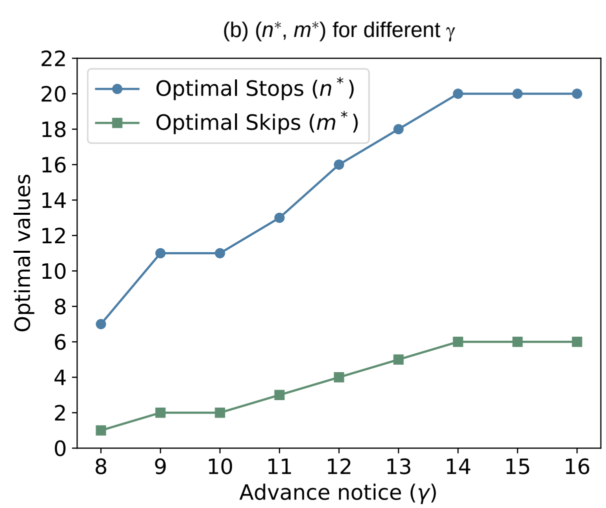
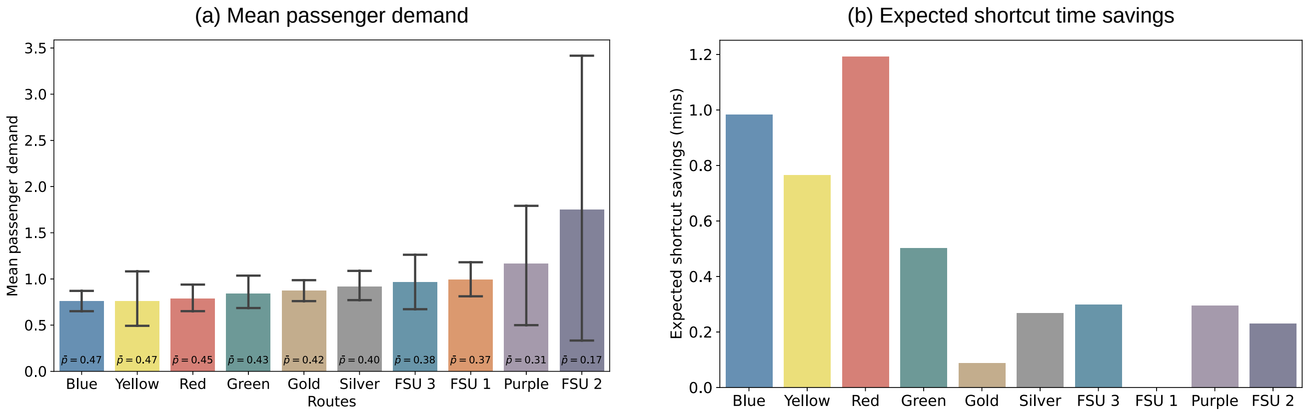


Figure 5. (Color online) Mean Passenger Demand and Shortcut Time Savings Distributions for Allegany County Routes



Note. FSU, Frostburg State University.

The parameters used in our simulation experiments are detailed in Table 2. The simulations are run for at least 20 warm-up cycles before starting the Bayesian evaluation of constraints. After completing the warm-up cycles, if the Bayesian constraints are satisfied, the simulation is run up to 35 cycles before adding a new stop. The fixed system's bus tour duration is determined based on the published schedule. To mirror real-world

Figure 6. (Color online) Candidate Stops and Possible Shortcuts in the Improved Routes

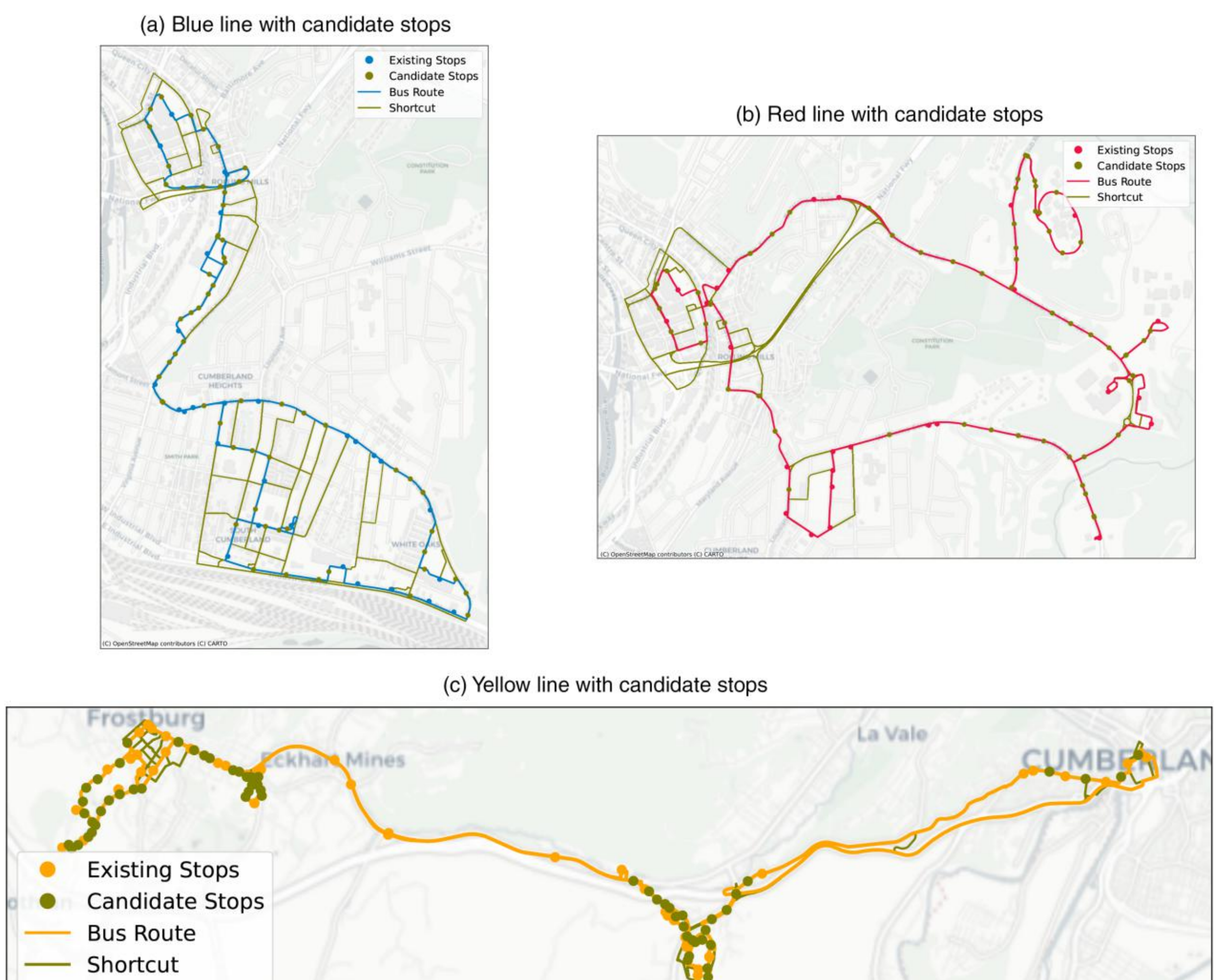


Table 2. Simulation Parameters and Their Values

Simulation parameter	Blue Line	Red Line	Yellow Line
Number of fixed system stops (N)	29	26	41
Fixed system tour duration (b) in min	100.75	119.43	112.50
Advance notice (γ) in min	10–30	10–30	10–30
Probability threshold ($1-\beta$)	0.80	0.80	0.80
Warm-up simulation cycles ($r_{\min}$)	20	20	20
Maximum simulation cycles ($r_{\max}$)	35	35	35
Number of candidate stops (C)	60	60	60

scenarios, we calibrated the passenger demand at these stops to match with APC data. This ensured that the probability of skipping a stop in the simulation model closely resembles that in the APC data (see Figures 7–9).

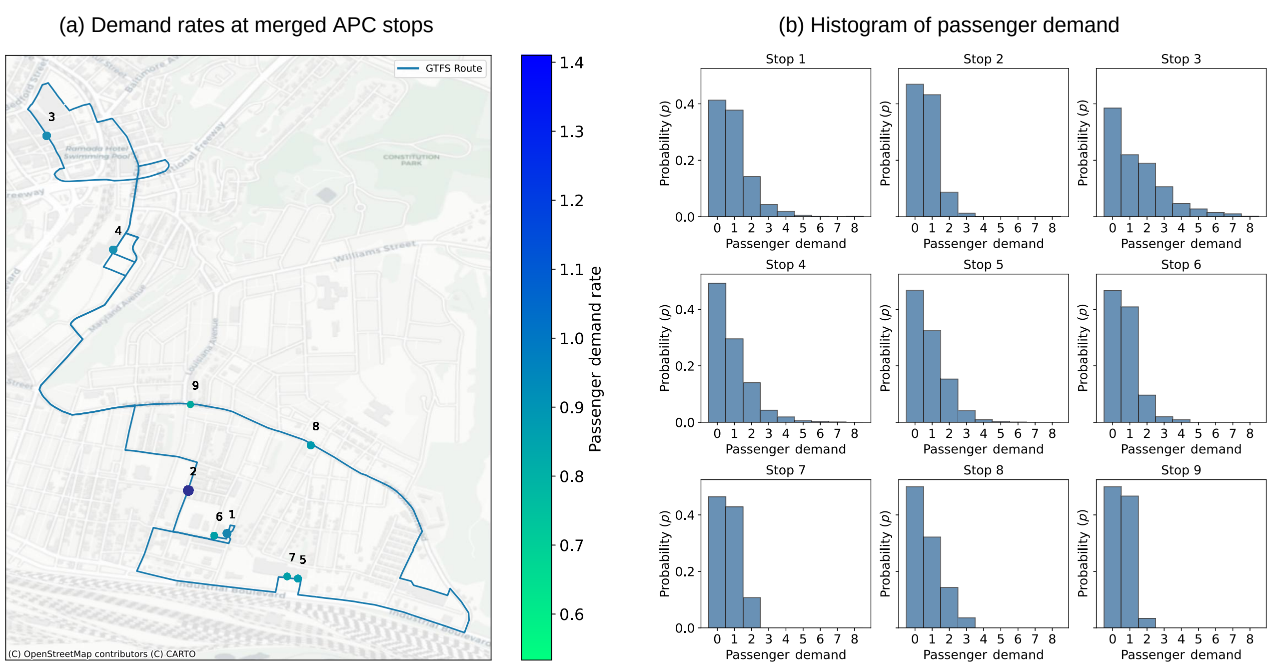
6.1. Minimizing Walking Distance with Fixed Demand

In scenarios with fixed demand, we introduce new stops into the route, maintaining the same demand levels despite increased mobility. Consequently, we allocate all generated passengers to the nearest stops, ensuring comprehensive service coverage by the bus. In this specific scenario, our focus has been on adding new stops to minimize the expected walking distance, as outlined in Section 4.1. The outcomes of these route optimizations using Algorithm 3 are discussed further below.

Figures 10–12 demonstrate the effects of introducing a semiflexible transit system on the number of stops, the implementation of shortcuts, and their impact on walking distances across different amounts of advance notice for the improved routes, namely, the Blue, Red, and Yellow Lines. It is evident that extending the advance notice enables the bus system to bypass more stops and integrate more shortcuts. For example, when the advance notice is extended to 30 minutes, the number of stops on the Blue and Red Lines increases by approximately 160%, leading to a substantial decrease in passenger walking distance by about 27%. In contrast, the Yellow Line shows a lesser increase, adding only four stops, primarily because of its straight layout, which limits the opportunities for taking shortcuts when stops are bypassed.

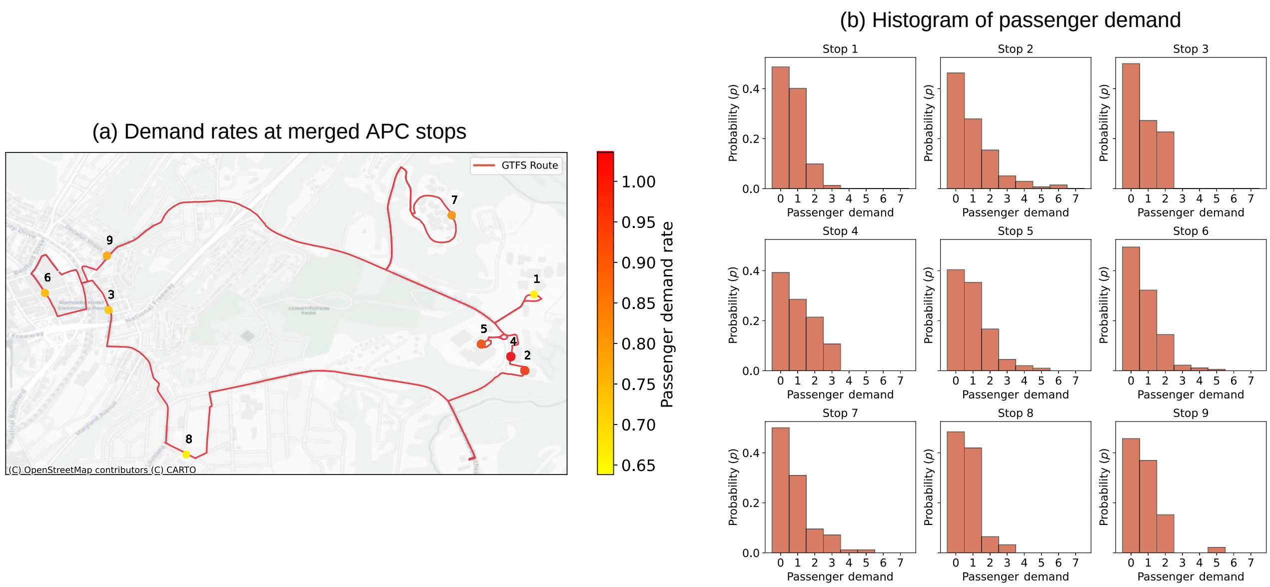
Tables 3–5 present a detailed comparison between the fixed and semiflexible bus systems operating on these lines, considering various amounts of advance notice while maintaining constant demand. For the Blue and Red Lines, the analysis indicates that the semiflexible bus systems, especially with longer advance notice (20 and 30 minutes), offer significant benefits over fixed systems in reducing walking distances. Another benefit for

Figure 7. (Color online) Stop Locations on the Blue Line Based on APC and GTFS Data



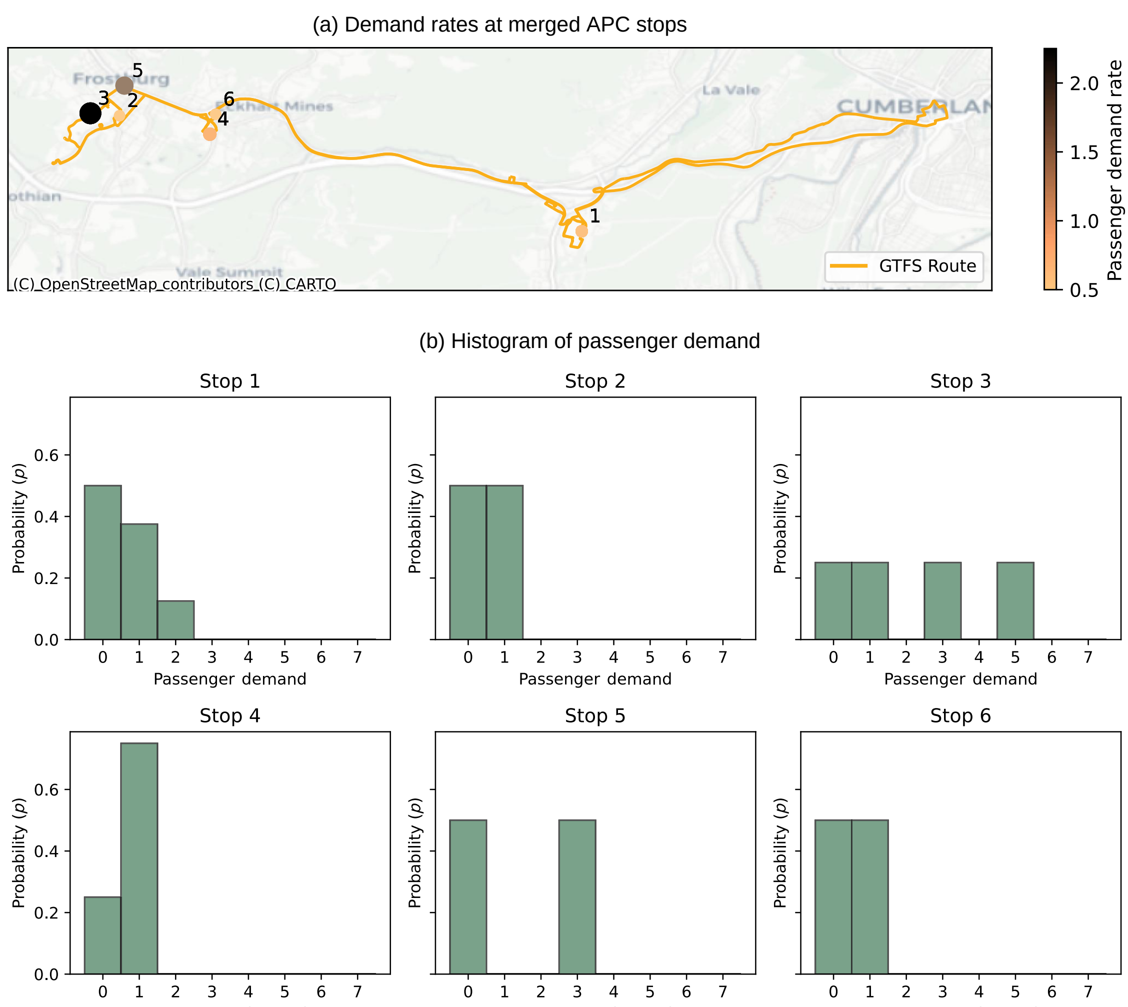
Note. APC stops that are within the buffer of GTFS stops are utilized for adjusting passenger demand rates.

Figure 8. (Color online) Stop Locations on the Red Line Based on APC and GTFS Data



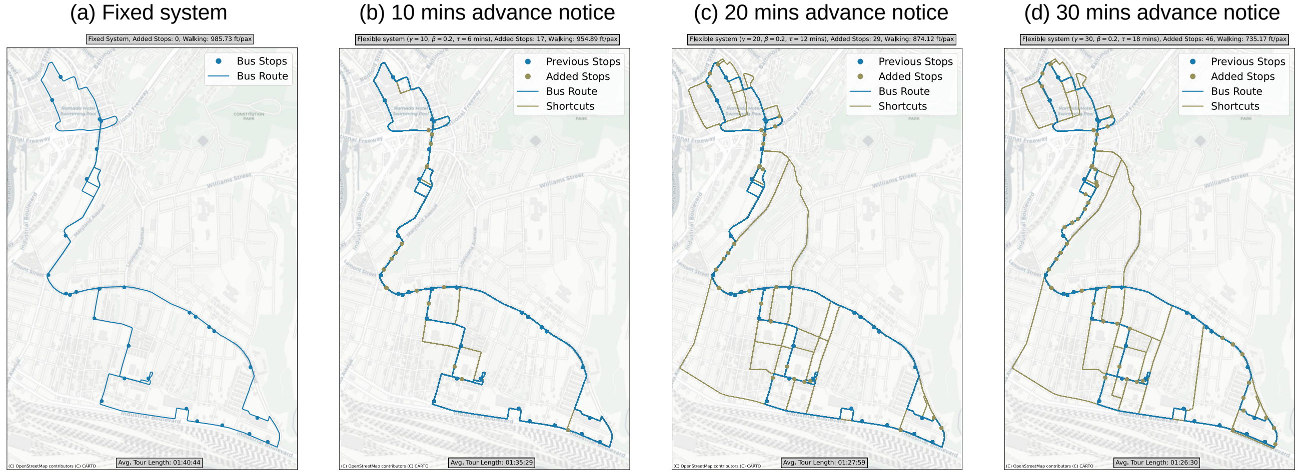
Note. APC stops that are within the buffer of GTFS stops are utilized for adjusting passenger demand rates.

Figure 9. (Color online) Stop Locations on the Yellow Line Based on APC and GTFS Data



Note. APC stops that are within the buffer of GTFS stops are utilized for adjusting passenger demand rates.

Figure 10. (Color online) Comparison Between the Current Fixed Transit System and the Proposed Semiflexible System with Varying Advance Notice and Fixed Demand for the Blue Line

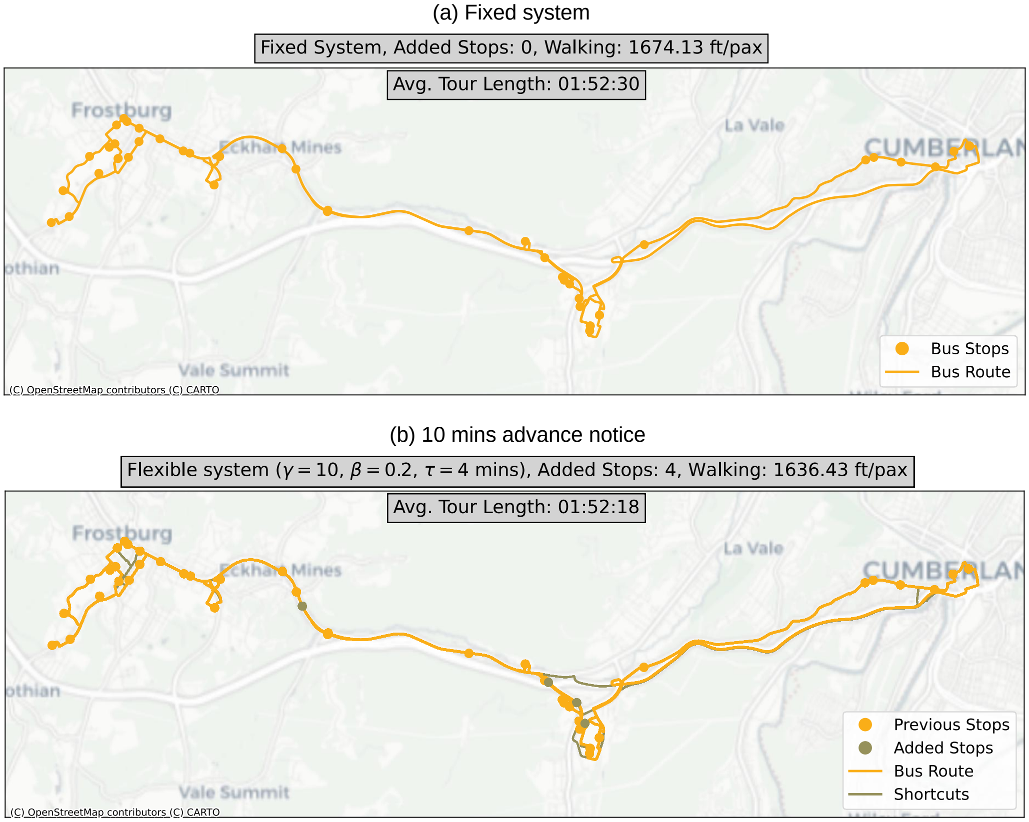


Note. A simulation of the presented systems can be found at https://youtu.be/hK4zpgkyNxg.

Figure 11. (Color online) Comparison Between the Current Fixed Transit System and the Proposed Semiflexible System with Varying Advance Notice and Fixed Demand for the Red Line

Note. A simulation of the presented systems can be found at https://youtu.be/lsX3fRUfWMU.

Figure 12. (Color online) Comparison Between the Current Fixed Transit System and the Proposed Semiflexible System with Advance Notice and Fixed Demand for the Yellow Line



Note. A simulation of the presented systems can be found at https://youtu.be/QRX9j_1ZtS4.

Table 3. Fixed vs. Semiflexible Bus System (Fixed Demand) for the Blue Line with Varying Notice

		Semiflex system		
Performance indicator	Fixed system	γ = 10 min	γ = 20 min	γ = 30 min
Stops added	—	17	29	46
Lookup period (τ^*)	—	6	12	18
Mean tour length (min)	100.73	97.07	88.07	86.72
SD tour length (min)	—	3.09	6.22	9.31
Walking (ft/passenger)	983.47	932.10	862.04	719.81

Note. SD, standard deviation.

Table 4. Fixed vs. Semiflexible Bus System (Fixed Demand) for the Red Line with Varying Notice

		Semiflex system		
Performance indicator	Fixed system	γ = 10 min	γ = 20 min	γ = 30 min
Stops added	—	15	25	43
Lookup period (τ^*)	—	5	11	15
Mean tour length (min)	119.43	122.02	106.95	101.67
SD tour length (min)	—	3.74	6.88	11.61
Walking (ft/passenger)	1,423.52	1,348.15	1,263.13	1,015.83

Note. SD, standard deviation.

Table 5. Fixed vs. Semiflexible Bus System (Fixed Demand) for the Yellow Line with Varying Notice

		Semiflex system		
Performance indicator	Fixed system	γ = 10 min	γ = 20 min	γ = 30 min
Stops added	—	4	4	4
Lookup period (τ^*)	—	4	9	13
Mean tour length (min)	112.50	112.30	111.68	110.14
SD tour length (min)	—	4.59	8.51	12.28
Walking (ft/passenger)	1,674.13	1,636.43	1,636.43	1,636.43

Note. SD, standard deviation.

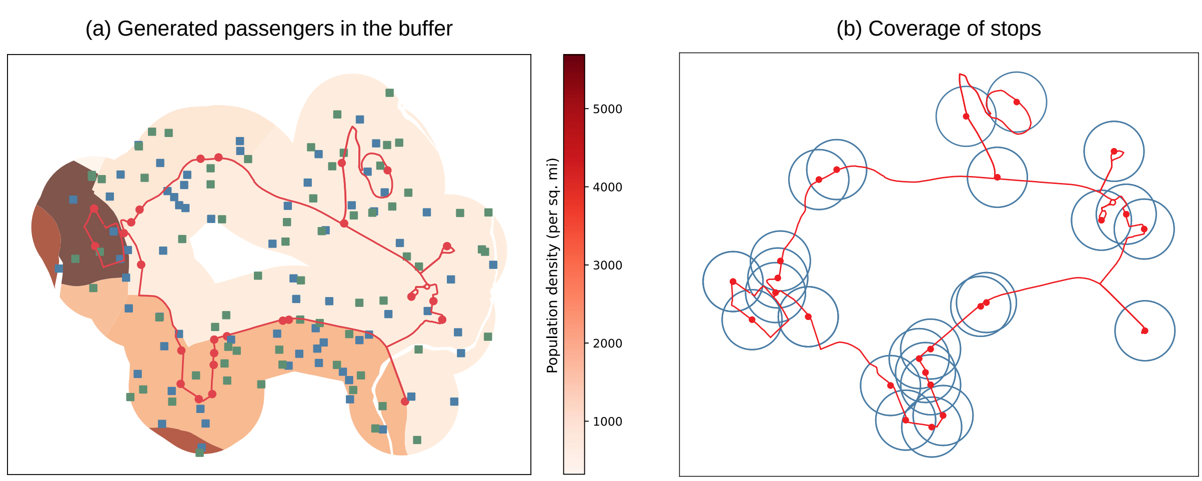
the passengers is that the passengers will spend less time on the trips because of shorter walking distances and reduced bus tour durations. Notably, the average tour duration in the semiflexible system is about 15–20 minutes shorter than in the fixed system when the advance notice is extended to 30 minutes. Although there is more slack in the tour duration, which might suggest the possibility of adding more stops, the constraints on acceptable variance in arrival times at stops limit the addition of further stops.

6.2. Maximizing Service Coverage with Increasing Demand

In this scenario, we assume that demand will increase because of increased coverage from added stops; thus, new stops are allocated to maximize demand coverage. Specifically, passenger origins and destinations are randomly generated based on the population within a 0.25 square-mile area surrounding the bus route, as depicted in the buffer area shown in Figure 13(a). Once generated, passengers are assigned to their nearest bus stops, considering a maximum walking distance threshold of 0.125 miles, visualized as circles in Figure 13(b). Passengers whose origins or destinations fall outside this threshold are deemed unable to use the bus. Excluding those beyond the walking distance threshold, we aggregate the origins and destinations within the threshold circle to determine the number of boarding and alighting passengers at each stop. Given that each additional stop covers more demand in this scenario, we optimized the routes to maximize service coverage as outlined in Algorithm 3.

After optimizing the routes, the number of stops in the Blue and Red Lines is increased by about 38% when the advance notice is increased to 30 minutes, resulting in 47% and 62% increases in the population coverage, respectively. In contrast, only two stops could be added to the Yellow Line, and the resulting improvement was not as significant as for the Blue and Red Lines; nonetheless, they still manage to increase the population coverage by 15%, which is even achievable using a 10-minute advance notice. It is noteworthy that the number of stops added in the increasing demand scenario is less than that of the fixed demand scenario. Because the addition of new stops is attracting more demand in this scenario, the bus is not able to skip as many stops, which results in a lesser number of stops than in the fixed demand scenario. The high-level summary and detailed analysis of the results are provided in Online Appendix L.

Figure 13. (Color online) Sample of Passengers Generated on the Red Line



7. Conclusions

We introduced a novel semiflexible transit system in which passengers submit advance notices of their intended stops, allowing buses to bypass stops without passenger demand by taking shortcuts. This strategy facilitated the addition of more stops, thus enhancing mobility while maintaining operational efficiency. Specifically, we developed an optimization model aimed at maximizing the number of stops while ensuring that the bus tour duration remains comparable to conventional transit systems and reliably honoring advance passenger requests, both with a high probability. This model enabled theoretical analysis of the problem, as well as efficient exploration of system design trade-offs and assessment of the impacts of various parameters on the number of stops and user experience. The theoretical analyses also motivated the application of heuristics and the evaluation of the derived constraints within a Bayesian learning framework to address more intricate scenarios.

To demonstrate the real-world application of our semiflexible routing strategy, we developed and applied simulation-based models to various bus routes in Allegany County, Maryland, United States. Our analysis, covering a range of route configurations, revealed how the strategy's advantages varied depending on each route's specific characteristics. We observed considerable improvements in routes with underutilized stops and configurations conducive to shortcuts. In scenarios with fixed and increasing passenger demand, our simulation-based approach led to increases in the number of stops by approximately 160% and 38%, respectively, when the average probability of skipping existing stops was greater than 45% and had layouts conducive to substantial shortcuts. By evaluating the performance of the semiflexible transit system on these routes, we have provided insights into its effectiveness and feasibility, demonstrating its potential to enhance bus routes with underutilized stops.

Acknowledgments

Funding support from the National Science Foundation is gratefully acknowledged, but it does not imply endorsement of the findings.

References

Aktaş D, Sörensen K, Vansteenwegen P (2023) A variable neighborhood search algorithm for a public bus line with a demand-responsive operation during peak hours. *Transportation Planning Tech.* 46(5):615–652.

Aldaihani M, Dessouky MM (2003) Hybrid scheduling methods for paratransit operations. *Comput. Indust. Engrg.* 45(1):75–96.

Aldaihani MM, Quadrifoglio L, Dessouky MM, Hall R (2004) Network design for a grid hybrid transit service. *Transportation Res. Part A: Policy Practice* 38(7):511–530.

Cayford R, Yim Y (2004) Personalized demand-responsive transit service. Technical report, University of California, Berkeley.

Chandra S, Quadrifoglio L (2013) A model for estimating the optimal cycle length of demand responsive feeder transit services. *Transportation Res. Part B: Methodological* 51(2013):1–16.

Chang SK, Schonfeld PM (1991) Optimization models for comparing conventional and subscription bus feeder services. *Transportation Sci.* 25(4):281–298.

Chen PW, Nie YM (2017) Analysis of an idealized system of demand adaptive paired-line hybrid transit. *Transportation Res. Part B: Methodological* 102(2017):38–54.

Chen Y, Ryzhov IO (2020) Consistency analysis of sequential learning under approximate Bayesian inference. *Oper. Res.* 68(1):295–307.

Chen X, Hellinga B, Chang C, Fu L (2015) Optimization of headways with stop-skipping control: A case study of bus rapid transit system. *J. Advanced Transportation* 49(3):385–401.

Chen Y, Zhang Q, Li M, Cai W (2022) Sequential selection for accelerated life testing via approximate Bayesian inference. *Naval Res. Logist.* 69(2):336–351.

Cortés CE, Pagès L, Jayakrishnan R (2005) Microsimulation of flexible transit system designs in realistic urban networks. *Transportation Res. Record* 1923(1):153–163.

Crainic TG, Malucelli F, Nonato M, Guertin F (2005) Meta-heuristics for a class of demand-responsive transit systems. *INFORMS J. Comput.* 17(1):10–24.

Daganzo CF (1984) Checkpoint dial-a-ride systems. *Transportation Res. Part B: Methodological* 18(4–5):315–327.

Davison L, Enoch M, Ryley T, Quddus M, Wang C (2014) A survey of demand responsive transport in Great Britain. *Transport Policy* 31(2014):47–54.

DeGroot MH (1970) *Optimal Statistical Decisions* (John Wiley and Sons, Hoboken, NJ).

Errico F, Crainic TG, Malucelli F, Nonato M (2013) A survey on planning semi-flexible transit systems: Methodological issues and a unifying framework. *Transportation Res. Part C: Emerging Tech.* 36(2013):324–338.

Errico F, Crainic TG, Malucelli F, Nonato M (2021) The single-line design problem for demand-adaptive transit systems: A modeling framework and decomposition approach for the stationary-demand case. *Transportation Sci.* 55(6):1300–1321.

Hadas Y, Ceder A (2008) Multiagent approach for public transit system based on flexible routes. *Transportation Res. Record* 2063(1):89–96.

Hadas Y, Ceder AA (2010) Optimal coordination of public-transit vehicles using operational tactics examined by simulation. *Transportation Res. Part C: Emerging Tech.* 18(6):879–895.

Koch KR (2007) *Introduction to Bayesian Statistics* (Springer, Berlin).

Koffman D (2004) *Operational Experiences with Flexible Transit Services* (Transportation Research Board, Washington, DC).

Li X, Quadrifoglio L (2010) Feeder transit services: Choosing between fixed and demand responsive policy. *Transportation Res. Part C: Emerging Tech.* 18(5):770–780.

Li X, Liu W, Qiao J, Li Y, Hu J (2023) An enhanced semi-flexible transit service with introducing meeting points. *Networks Spatial Econom.* 23(2023):487–527.

Liu Z, Yan Y, Qu X, Zhang Y (2013) Bus stop-skipping scheme with random travel time. *Transportation Res. Part C: Emerging Tech.* 35(2013):46–56.

Nourbakhsh SM, Ouyang Y (2012) A structured flexible transit system for low demand areas. *Transportation Res. Part B: Methodological* 46(1):204–216.

Osama A, Mahdy H, Kandil K, Elhabiby M (2016) Monitoring the impact of unscheduled stops and unscheduled route deviations on bus service using automatic vehicle location system: A case study. *Adv. Transportation Stud.* 39(2016):37–48.

Pratelli A, Schoen F (2001) A mathematical programming model for the bus deviation route problem. *J. Oper. Res. Soc.* 52(5):494–502.

Pratelli A, Lupi M, Farina A, Pratelli C (2018) Comparing route deviation bus operation with respect to dial-a-ride service for a low-demand residential area. *Seventh Internat. Conf. Data Anal.* (International Academy, Research, and Industry Association), 141–148.

Qiu F, Shen J, Zhang X, An C (2015) Demi-flexible operating policies to promote the performance of public transit in low-demand areas. *Transportation Res. Part A: Policy Practice* 80(2015):215–230.

Quadrifoglio L, Li X (2009) A methodology to derive the critical demand density for designing and operating feeder transit services. *Transportation Res. Part B: Methodological* 43(10): 922–935.

Quadrifoglio L, Hall RW, Dessouky MM (2006) Performance and design of mobility allowance shuttle transit services: Bounds on the maximum longitudinal velocity. *Transportation Sci.* 40(3):351–363.

Rahman MH, Chen S, Sun Y, Siddiqui MIY, Mohebbi M, Marković N (2023) Integrating dial-a-ride with transportation network companies for cost efficiency: A Maryland case study. *Transportation Res. Part E: Logist. Transportation Rev.* 175(2023): 103140.

Shapiro A, Dentcheva D, Ruszczynski A (2021) *Lectures on Stochastic Programming: Modeling and Theory* (Society for Industrial and Applied Mathematics, Philadelphia).

Silva DF, Vinel A, Kirkici B (2022) On-demand public transit: A Markovian continuous approximation model. *Transportation Sci.* 56(3):704–724.

Sipetas C, Gonzales EJ (2021) Continuous approximation model for hybrid flexible transit systems with low demand density. *Transportation Res. Record* 2675(8):198–214.

Sun Y, Guo Q, Schonfeld P, Li Z (2017) Evolution of public transit modes in a commuter corridor. *Transportation Res. Part C: Emerging Tech.* 75(2017):84–102.

Zhang L, Huang J, Liu Z, Vu HL (2021) An agent-based model for real-time bus stop-skipping and holding schemes. *Transportmetrica A: Transport Sci.* 17(4):615–647.

Zheng Y, Li W, Qiu F (2018) A methodology for choosing between route deviation and point deviation policies for flexible transit services. *J. Advanced Transportation* 2018(1):6292410.