



Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Investment timing and length choice for a rail transit line under demand uncertainty

Qianwen Guo^{a,e}, Shumin Chen^b, Yanshuo Sun^{c,*}, Paul Schonfeld^d^a Department of Civil and Environmental Engineering, Florida International University, 10555 W Flagler Street, EC 3627, Miami, FL 33174, USA^b School of Management, Guangdong University of Technology, No. 161, Yinglong Road, Guangzhou 510520, China^c Department of Industrial and Manufacturing Engineering, FAMU-FSU College of Engineering, Florida State University, 2525 Pottsdamer Street, Tallahassee, FL 32310, USA^d Department of Civil and Environmental Engineering, University of Maryland, College Park, 1173 Glenn Martin Hall, College Park, MD 20742, USA^e Department of Civil and Environmental Engineering, FAMU-FSU College of Engineering, Florida State University, 2525 Pottsdamer Street, Tallahassee, FL 32310, United States

ARTICLE INFO

Keywords:

Transit mode evolution

Feeder-and-trunk

Rail transit

Investment timing

Demand shocks

Uncertainty modeling

ABSTRACT

This paper is motivated by the inadequate treatment of uncertainty in public transit infrastructure investments. As observed in the past, rare but dramatic events can heavily disrupt public transit system operations and negatively affect the transit riders. For example, the COVID-19 pandemic caused an 80%-90% transit demand decline in March 2020 in the U.S. However, the existing transit infrastructure planning studies have not modeled such sudden demand shocks. We thus improve the modeling realism of uncertain transit demand by formulating demand evolution as a jump-diffusion process, which is a combination of continuous-time Brownian motion and a discrete counting process, namely Poisson process, and present analytical optimization models for the development of a rail transit line under such uncertainty. We jointly optimize two related decisions, namely the timing for introducing rail transit to a commuter corridor and length choice for the rail line. We refute a misconception that investment in a project should always start immediately if a positive cost saving over the planning horizon is expected. We also find that investment timing and sizing decisions are closely related and behave quite differently for the same change in some parameters, such as the infrastructure construction period. The developed modeling and analysis framework should be transferable to other civil infrastructure development and investment problems under uncertainty.

1. Introduction

The effectiveness of long-term plans for most transportation infrastructure projects (e.g., airport runways, railway terminals, and highway networks) largely depends on how well future uncertainty is addressed in the planning process (De Neufville and Scholtes, 2011). For example, in an airport master plan, planners need to design and schedule airport development projects to ensure the airport capacity can meet anticipated demand on both airside and landside facilities in the next few decades (Sun and Schonfeld, 2015). At the same time, design flexibility should be built into a master plan to enhance an airport's ability to adapt to unpredictable changes in regulations, markets, and technologies (Sun and Schonfeld, 2017). Similarly, public transportation system planners must make

* Corresponding author.

E-mail address: y.sun@eng.famu.fsu.edu (Y. Sun).

strategic decisions regarding when the current transit service should be upgraded, and when new infrastructure should be built to satisfy the travel needs of future transit riders. For instance, when demand is relatively low in a corridor, bus transit remains a cost-effective choice, while light rail, which can increase capacity and level of service, may be introduced when demand grows to a threshold at which it becomes more effective than bus transit. The transit mode evolution problem is concerned with the optimal selection of transit modes over time (Sun et al., 2017). The main motivation for this paper is the inadequate treatment of demand uncertainty in optimizing transit mode selection on the tactical level.

A few researchers, e.g., Saphores and Boarnet (2006) and Li et al. (2015), have studied a closely related problem, namely the public transit investment problem, considering demand uncertainty. In those studies, demand uncertainty was modelled with a standard geometric Brownian motion (GBM) and a demand trigger was derived to model the replacement of bus transit by rail transit. Although a few key factors, such as the time-inconsistent preferences (Guo et al., 2018b) or the ambiguity attitudes (Gao and Driouchi, 2013) of the planning authority, have been investigated, previous studies (e.g., Gao and Driouchi (2013), Li et al. (2015), and Guo et al. (2018b)) did not fully characterize demand uncertainty with a continuous-time stochastic process, i.e., GBM. Not only does the future demand evolve continuously, but it also jumps sometimes. A case in point is the dramatic transit demand reduction observed in the early months of the COVID-19 pandemic. A survey conducted by Liu et al. (2020) indicated that on average the transit ridership dropped in the U.S. by an average of 80% in the span of a few weeks, due to various disease control measures, such as social distancing and working from home. Such sudden and rare changes can be modeled with a low-intensity Poisson process (Chen et al., 2019), which has been adopted in related areas, such as modeling spikes in electricity price (Seifert and Uhrig-Homburg, 2007; Benth et al., 2012). Nonetheless, the sudden demand changes, also called jumps, have not been considered in the public transit investment literature. Considering the catastrophic consequences of possible extreme events in the planning horizon of transit systems, the effect of those demand jumps arising from those sudden events on the long-term transit planning process must be incorporated. In addition, a major difference between the transit mode evolution problem (Sun et al., 2017) and the transit investment problem (Li et al., 2015) is that the latter problem is essentially a two-period problem, which means it seeks to find the threshold between bus-only and rail-only services. Conversely, the transit mode evolution problem explores the change of transit modes over multiple periods. The feeder-and-trunk service, which is a hybrid mode of bus and rail, is shown to be advantageous in the evolution process of transit modes by Sun et al. (2017) while it has not been considered in transit investment studies such as Guo et al. (2018b). Considering this hybrid mode is nontrivial because it implies another important decision, namely the length of rail line in various periods. Then, simultaneous optimization of the investment timing and rail length becomes a necessity. Considering the above identified research gaps, we seek to improve the transit mode evolution literature by making the following methodological contributions. This study is the first known one to model transit demand jumps due to sudden events with a jump-diffusion process and propose a stochastic model for the development of a rail transit line under such uncertainty. It also concurrently optimizes rail investment timing and sizing decisions, which are essentially inseparable considering the trade-off between rail capital investments and operational savings.

To achieve the stated contributions, we first developed two analytical optimization models for bus-only and hybrid (feeder-and-trunk) services in a commuter corridor. As closed-form solutions could be found, the hourly system cost became a simple function of demand density only. Next, we formulated a dynamic optimization problem where the time for introducing rail transit and the rail length were both optimized, for a given demand growth curve. Then, demand evolution was governed by a jump-diffusion process over an infinite planning horizon, consisting of both gradual and sudden changes; a stochastic optimization model was thus formulated to minimize the expected aggregated cost over the infinite horizon. Leveraging the optimal stopping theory, we derived the optimal rail investment timing and presented a numerical solution of rail length. Through numerical studies, we arrived at a few important conclusions. For instance, it is not always optimal to start an infrastructure investment as soon as a cumulative cost saving is expected over the planning horizon. We also confirm the close interrelation between investment timing and sizing decisions, which cannot be separately analyzed or optimized, due to the complex tradeoff of capital costs and operational savings. In addition, when some parameters (such as demand volatility) vary, the optimal demand threshold and rail length change in the same direction; those two decisions may change in different directions if some other parameters (such as jump intensity) vary.

The rest of this paper is organized as follows. Section 2 identifies the main research gaps in the transit mode evolution as well as transit investment studies. Section 3 defines static optimization models for two types of transit services, which are solved analytically, and further introduces a deterministic dynamic model. After introducing a new stochastic process for modeling demand uncertainty, we propose a stochastic dynamic optimization model and present analytical solutions in Section 4. In Section 5, we analyze results from deterministic and stochastic analyses, as well as the sensitivity of results to some key parameters. Finally, we draw conclusions and suggest future research directions in Section 6.

2. Literature review

As we seek to optimize the development of a rail transit line over time considering uncertain demand shocks, we organize relevant studies into different groups, depending on whether temporal variations are considered and whether demand uncertainty is incorporated. We begin with a review in Section 2.1 of those studies in which demand is static and deterministic; we next review studies on the dynamic selection of transit modes over time without any demand uncertainty in Section 2.2; in Section 2.3, a number of studies which consider demand uncertainty over a planning horizon are reviewed, although they simplify the evolution of a rail transit line and do not yet incorporate the effect of transit demand shocks. In Section 2.4, we expand the scope of our review to cover transportation problems outside public transit. An additional major research gap is identified, namely very few real-option studies have jointly considered timing and sizing decisions, which are both essential in the rail transit line development.

2.1. Static models for public transit mode comparison and selection

In practice, there is a wide spectrum of public transit services ranging from demand-responsive and flexible bus to light and heavy rail transit, each of which is designed to effectively accommodate some distinct demand characteristics. Transportation researchers have presented many comparisons of various types of public transit modes. Note that in some studies (Parajuli and Wirasinghe, 2001; Sivakumaran et al., 2014), transit modes are also referred to as transit technologies, such as bus transit vs rail transit.

Even though comparisons of public transit modes were conducted many decades ago, such as in Vuchic (1976), a few representative studies published more recently are reviewed here. Tirachini et al. (2010) developed analytical models for the comparison of a few public transit modes, namely bus, bus rapid transit, light rail, and heavy rail, under different optimization objectives. Moccia and Laporte (2016) extended Tirachini et al. (2010) by considering additional operational parameters, such as optimal stop spacing. Kim and Schonfeld (2013) developed bus operational models for different service types (flexible vs fixed) and fleets (uniform or mixed vehicle sizes). They also compared those service alternatives under different demand scenarios. Zhang et al. (2018) compared park & ride with on-demand bus considering a transportation corridor. They concluded that park & ride should serve low-demand areas while the latter can cover high-demand areas.

In addition to the analytical models reviewed above, some researchers also conducted empirical analyses of different public transit modes. For instance, Casello et al. (2014) compared economic costs and emissions for bus rapid transit and light rail with a case study of Waterloo, Ontario in Canada. Puchalsky (2005) also compared those two modes from an environmental perspective only (i.e., emissions).

The above review suggests that many studies have been devoted to the comparison of different public transit modes when demand is given and static.

2.2. Public transit mode selection over time

With models reviewed in Section 2.1, we can select an optimal transit mode for given demand inputs at any given time. However, studying the selection of optimal transit modes over time, also called the evolution of transit modes, is not a straightforward extension of a static mode comparison or selection model, primarily because of the transition from one mode to another is not “frictionless.” In other words, major infrastructure upgrades and thus substantial capital investments are needed, which complicates the evolution of transit modes over time.

Sun et al. (2017) considered a commuter corridor and investigated when rail transit should be introduced to partially replace bus transit and how the rail length should be optimized over time. As an extension of Cheng and Schonfeld (2015), where demand was assumed to be invariant with fare or service quality, Sun et al. (2018) focused on the development of a single rail transit line or the optimal extension of a rail line, with a full consideration of demand elasticity. Further, Wu and Schonfeld (2022) considered bi-directional extensions of a rail line in multiple stages. While maximizing net present value over a long analysis period, Wu and Schonfeld (2022) assumed that the rail line extensions determined through a genetic algorithm should be implemented as soon as sufficient construction funds became available for the next stage. This assumption implied that, in the absence of a budget limit, a rail line should be immediately extended if a positive net present value is expected. While this assumption can greatly simplify the rail extension scheduling decisions, it may not always hold, as shown later in Section 5.2 of this paper. The extension timing indeed depends on the tradeoff between infrastructure capital investment and operational savings when infrastructure becomes in place.

Chang and Schonfeld (1991) derived a demand density threshold at which two optimized service types (conventional bus and subscription bus) had the same average cost per user trip. When the demand density was below the critical value or threshold, subscription bus was preferred. However, this does not imply that whenever the demand exceeds the threshold, conventional bus should replace subscription bus, when a transition or switching cost is considered. In a similar setting, Kim and Schonfeld (2013) derived thresholds during daily demand cycles at which large buses and small buses would lead to the same cost. When demand exceeded the thresholds in various service regions, larger buses were used; otherwise, smaller ones were used. Similarly, this assumption will no longer hold when there is a positive transition cost. For instance, Guo et al. (2018a) studied the optimal switching between fixed-route transit and flexible-route transit. Because service switching is not costless, Guo et al. (2018a) showed that it was suboptimal to immediately switch to another service type when a demand threshold was reached.

Therefore, the literature on the evolution of public transit modes is underdeveloped and a few key issues (such as how to determine a trigger demand for introducing a new transit service or mode) should be fully examined and addressed. Another major research gap is that the effect of uncertainty on rail transit investment has not been fully examined, to be elaborated below.

2.3. Rail transit investment under uncertainty

When the future demand growth is known with certainty over a planning period, the right timing for introducing a new mode (such as rail transit) is when a critical demand threshold is reached. For example, Chen et al. (2015) derived a critical population density of two transit modes by setting their corresponding benefits to be equal. As a deterministic demand growth function was given in Chen et al. (2015), finding the right investment timing was simply determining when the demand reached the threshold. When population growth was uncertain, Li et al. (2015) employed real options to determine the population trigger. Gao and Driouchi (2013) also studied the rail transit investment problem while considering population uncertainty and the ambiguity attitudes (e.g., pessimism) of the planning authority. Guo et al. (2018b) enhanced those two studies by considering the time-inconsistent preferences of the planning authority with a quasi-hyperbolic discount function.

A distinction of the above studies from [Sun et al. \(2017\)](#) is that they addressed the effect of demand uncertainty on a long-term planning problem. However, the common shortcoming in those studies is that they oversimplified the evolution process of transit modes. They essentially considered only two options or choices, such as bus-only at present and rail-only at a future time. Nonetheless, as pointed out by [Sun et al. \(2017\)](#) and [Sun et al. \(2018\)](#), rail transit may expand gradually over time and may not extend to the city boundary when the distribution of demand over space is not uniform. Specifically, the feeder-and-trunk service (rail combined with bus) is advantageous over the rail-only service even in the long-term. Clearly, the existing studies on rail transit investment under uncertainty should be improved by considering this important layout. In this case, an important decision on the rail line length in various periods emerges. Additionally, uncertainty of demand in the rail transit investment literature is modelled with a GBM, which cannot capture demand jumps or shocks caused by sudden events.

Although demand jumps have not been modeled in the rail transit investment literature, they have been considered in a few other related infrastructure investment problems, to be reviewed next.

2.4. Applications of real option analysis to transportation project investments

Real options approaches have been widely used to study the investment in irreversible infrastructure projects, since the pioneering work on investment under uncertainty by [Dixit et al. \(1994\)](#). [Martins et al. \(2015\)](#) provided a comprehensive survey of real options studies in various infrastructure systems, such as power plants, hospitals, water supply systems and transportation infrastructures. We will next review some representative real options analyses in non-rail transit investments.

[Couto et al. \(2015\)](#) considered a High-Speed Rail (HSR) project and derived the demand threshold for implementing the HSR project. They did not consider any sizing decisions, such as how long the HSR line should be. [Balliauw and Onghena \(2020\)](#) optimized the capacity investment decision of a private airport under demand uncertainty. They modeled the aircraft movements with a GBM and selected profit maximization as the optimization objective, given the private nature of the considered airport. [Table 1](#) presents a comparison of some representative studies. We find that most existing studies considered the timing decision only, i.e., deciding when a project should be implemented, while neglecting the sizing decision, e.g., optimizing the magnitude of an increment. In most cases, a GBM was used to model demand or traffic uncertainty, which cannot incorporate the dramatic change caused by sudden events. In addition, the major infrastructure improvements in practice may take years to complete, while the investment project was typically assumed to be implemented instantly in the literature.

While other infrastructure investment studies are helpful, especially because they have offered useful insights into how realistic demand uncertainty should be modelled, the transit mode evolution problem is distinct, as reviewed in [Section 2.3](#). Instead of only two scenarios, bus-only vs rail-only, there exists a hybrid of two options, namely the feeder-and-trunk service, which may dominate rail-only services in the long run. The sizing decision is thus an indispensable part of the mode evolution optimization and becomes inseparable from the investment timing decision. In addition, many existing studies modeled a rail line or highway segment as a single point without considering the spatial distribution of demand and the impact on transportation facility operations. For instance, [Couto et al. \(2015\)](#) studied the annual line-level demand for a high-speed rail without modeling any demand patterns at the station level. In other words, the spatial structure of a rail line is reduced to a single node or atom. Conversely, it is essential to model a reasonably realistic setting of transit demand distribution and service layout in the transit mode evolution problem.

In summary, the transit mode evolution literature should be advanced by systematically evaluating the impact of future uncertainty, especially regarding rare but dramatic changes. We thus develop analytical optimization models extending [Sun et al. \(2017\)](#) by modeling realistic demand uncertainty with a jump-diffusion process and jointly determining investment timing and sizing decisions.

3. Deterministic model

3.1. Model setup

Throughout this paper, we analyze a commuter corridor that stretches from the central business district (CBD) to the city boundary,

Table 1
Real options analyses for other transportation infrastructure investments.

Study	Project type	Objective function	Source of uncertainty	Sizing decision	Jumps/shocks	Construction time
Saphores and Boarnet (2006)	A highway corridor	Max. expected utilities of residents	Urban population	No	No	Yes
Friesz et al. (2008)	Transportation network	Max. expected net trip value	Trip cost	No	No	No
Galera and Soliño (2010)	A highway corridor	Max. expected cash flows	Traffic volume	No	No	No
Chow and Regan (2011a, 2011b)	Highway network	Max. option value	Travel demand	No	No	No
Couto et al. (2015)	High-speed rail transport	Max. expected project value	Annual demand	No	Yes	No
Balliauw and Onghena (2020)	An airport roadway	Max. profit	Demand	Yes	No	No

shown in Fig. 1. As assumed in other public transit studies (Li et al., 2015; Sun et al., 2017; Guo et al., 2018b), residents living in this corridor travel to a common destination, i.e., CBD, for employment. The travel demand density $q(x)$ at location $x \in (0, A]$ is given by a linear function:

$$q(x) = \theta - \frac{\theta}{A}x, \quad (1)$$

where θ represents the demand density at the CBD, and A is the distance between the CBD and city boundary, i.e., the corridor length. As $\frac{1}{A}$ determines how quickly the demand density at location x drops as x increases, the demand density diminishes to zero when $x = A$, i.e., at the end of the corridor.

Fig. 1 shows the co-existence of two transit modes, namely rail and bus transit, each of which covers a portion of the corridor. Specifically, the interval $[0, L]$ is covered by rail and the remaining of the corridor $[L, A]$ is served by bus transit. Residents located in the bus service area first take bus and then transfer to rail through the transfer point located at L . Prior to the co-existence of rail and bus, namely the feeder-and-trunk service (called interchangeably in this paper the “hybrid” service), the entirety of the commuter corridor is covered by bus.

3.2. Cost functions

We start by defining the generalized travel cost function $W(x)$ for a resident at x as follows:

$$W(x) = \begin{cases} \varphi \frac{x}{V_R} + \frac{e_w}{2}h_R + f_R, & 0 < x \leq L, \\ \varphi \left(\frac{x-L}{V_B} + \frac{L}{V_R} \right) + \frac{e_w}{2}h_R + \frac{e_w}{2}h_B + f_B, & L < x \leq A. \end{cases} \quad (2)$$

If a resident lives between the CBD and transfer point, the generalized travel cost of commuting by rail consists of in-vehicle travel cost $\varphi \frac{x}{V_R}$, waiting cost $\frac{e_w}{2}h_R$, and a fixed cost f_R , which includes fare and access cost. φ is the value of in-vehicle time, e_w is the value of waiting time, V_R is the average travel speed of rail, and h_R is headway of rail services. The subscript “R” represents rail. Residents beyond the transfer point face additional waiting time. Similarly, V_B is the average travel speed of bus, h_B is bus headway, and f_B is the fixed cost for transferring passengers, which includes access and transfer costs. The subscript “B” represents bus.

The total user cost can be aggregated as follows:

$$C_u = \int_0^A W(x)q(x)dx. \quad (3)$$

The operating cost is the hourly cost per vehicle (bus or train) multiplied by fleet size. The fleet size for each mode can be computed as round-trip time divided by headway. The total operating cost is thus:

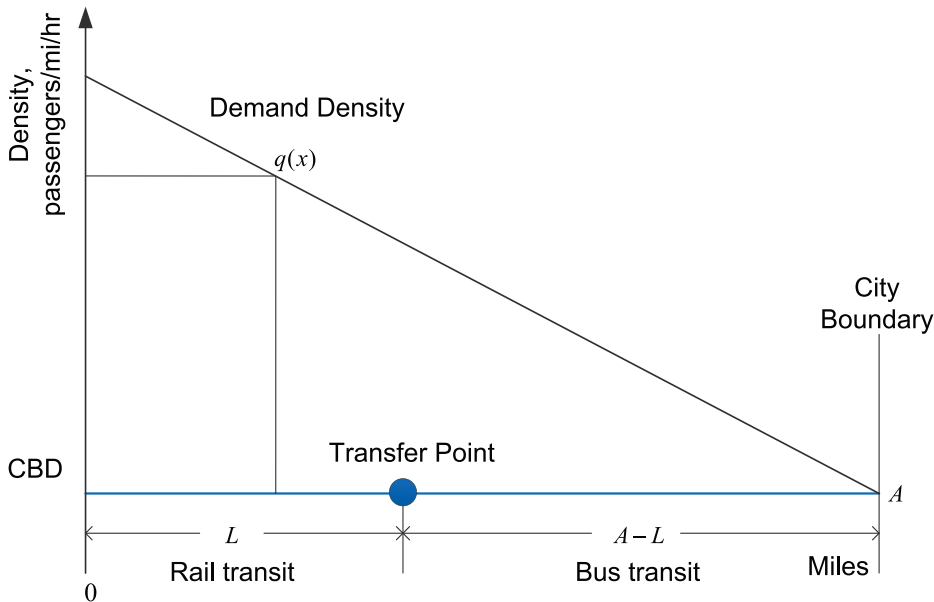


Fig. 1. Illustration of the demand density function.

$$C_p = \frac{2b_R}{V_R h_R} L + \frac{2b_B}{V_B h_B} (A - L), \quad (4)$$

where b_R and b_B are the hourly costs per train and per bus, respectively. b_R is a linear function of vehicle capacity (i.e., numbers of seats) for rail:

$$b_R = \alpha_R + \beta_R k_R. \quad (5)$$

Similarly, b_B is written as follows:

$$b_B = \alpha_B + \beta_B k_B. \quad (6)$$

In Eqs. (5) and (6), α is the fixed hourly cost, β is the marginal cost, and k is the vehicle capacity. The subscript indicates the mode (“bus” B or “rail” R).

3.3. Static model for a single mode

Before rail is introduced and becomes operational, the corridor is served by bus transit entirely. A static model for optimizing bus operations assuming time-invariant demand is presented as follows:

$$\min_{\{h_B, k_B > 0\}} C_B = \int_0^A \left(\varphi \frac{x}{V_B} + \frac{e_w}{2} h_B + f_B \right) q(x) dx + 2 \frac{(\alpha_B + \beta_B k_B) A}{V_B h_B}, \quad (7)$$

s.t.

$$h_B \int_0^A \left(\theta - \frac{\theta}{A} x \right) dx \leq k_B \delta_B. \quad (8)$$

where δ_B is the allowable load factor for bus.

We seek to minimize the objective function Eq. (7), which is the sum of user cost and operating cost, also called the *system cost* in this paper, because both users and the supplier (or operator) are considered. There are two decision variables, namely bus headway h_B and bus capacity k_B , both of which should be positive. The bus capacity constraint Eq. (8) is binding at optimality, because any additional bus capacity beyond what is sufficient to cover the travel demand accumulated during one headway would unnecessarily worsen the objective function and thus should be avoided in the first place.

After solving the static problem Eqs. (7) and (8) analytically in Appendix A, we write the bus system cost C_B as a simple function of θ of the following form:

$$C_B = e_1 \theta + e_2 \theta^{\frac{1}{2}}, \quad (9)$$

where $e_1 = \left(\frac{\varphi}{6V_B} + \frac{\beta_B}{\delta_B V_B} \right) A^2 + \frac{f_B}{2} A$, and $e_2 = \sqrt{\frac{2\alpha_B e_w}{V_B}} A$. Note that both e_1 and e_2 depend on known parameters, e.g., bus travel speed and operating costs, rather than demand density θ .

3.4. Static model for hybrid modes (feeder-and-trunk)

When a rail transit line partially covers the corridor, residents beyond the transfer point make multimodal transit trips. We thus formulate another static optimization model when rail and bus jointly serve the residents. The optimization objective Eq. (10) now includes the total user cost for non-transferring riders (rail only) and transferring riders (bus + rail, or hybrid), as well as the total cost of operating trains and buses. Decision variables include train and bus headways (h_R and h_B) as well as capacities (k_R and k_B). As in the static problem for a single mode, there are vehicle capacity constraints Eqs. (11) and (12), for rail and bus, respectively. The complete formulation is presented as follows:

$$\min_{\{h_R, h_B, k_R, k_B > 0\}} C_{B+R} = \int_0^L \left(\varphi \frac{x}{V_R} + \frac{e_w}{2} h_R + f_R \right) q(x) dx + \int_L^A \left(\varphi \left(\frac{x-L}{V_B} + \frac{L}{V_R} \right) + \frac{e_w}{2} h_R + \frac{e_w}{2} h_B + f_B \right) q(x) dx + \frac{2(\alpha_R + \beta_R k_R)}{V_R h_R} L + \frac{2(\alpha_B + \beta_B k_B)}{V_B h_B} (A - L) \quad (10)$$

s.t.

$$h_R \int_0^A \left(\theta - \frac{\theta}{A} x \right) dx \leq k_R \delta_R, \quad (11)$$

$$h_B \int_L^A \left(\theta - \frac{\theta}{A} x \right) dx \leq k_B \delta_B. \quad (12)$$

In Eq. (11), δ_R is the maximum load factor for rail. Note that in this setting, rail and bus operations are independent and uncoordinated, i.e., they can have different headways (Sun et al., 2017).

We analytically solve the static problem Eqs. (10)–(12) in Appendix B and re-write the objective Eq. (10) as:

$$\min_{\{A>L>0\}} C_{B+R} = e_3 \theta + e_4 \theta^{\frac{1}{2}}, \quad (13)$$

where $e_3 = \varphi L \left(\frac{1}{V_B} - \frac{1}{V_R} \right) \left(\frac{L^2}{3A} - \frac{L}{2} - \frac{(A-L)^2}{2A} \right) + \frac{\varphi A^2}{6V_B} + f_R \left(L - \frac{L^2}{2A} \right) + f_B \frac{(A-L)^2}{2A} + \frac{\beta_R AL}{V_R \delta_R} + \frac{\beta_B (A-L)^3}{\delta_B V_B A}$, and $e_4 = \sqrt{\frac{2e_w \alpha_R LA}{V_R}} + \sqrt{\frac{2e_w \alpha_B (A-L)^3}{V_B A}}$.

In Appendix C, we further derive the second-order derivatives and prove the convexity of Eq. (10) in h_R and h_B .

The objective function Eq. (10) has been simplified as a simple function of θ , namely Eq. (13), which consists of a linear term $e_3 \theta$ and a square root term $e_4 \sqrt{\theta}$. The largest exponent in Eq. (13) is only 1, which is expected. This is because the generalized travel cost function $W(x)$ is a quadratic function of $\sqrt{\frac{1}{\theta}}$ and the demand density function $q(x)$ is linear in θ . The aggregated user cost only has a linear term and a square root term. Similarly, the supply cost has only linear and square root terms.

As in other microeconomic models for public transit operations, some simplifications are made in this paper mainly to preserve analytical tractability. For instance, a few notable differences between rail and bus transit, such as service reliability (van Oort, 2014; van Oort, 2016), in-vehicle crowding, and level of comfort, should be considered in determining an appropriate transit mode, but these are not yet incorporated in the current static models, because otherwise the resulting models become analytically intractable. If analytical tractability is no longer pursued, additional operational details and complex network structures can be considered, which will yield numerical relations between the total system cost and transit demand.

3.5. Deterministic dynamic model

In a dynamic setting, the demand density at the CBD at time t is given by:

$$\theta_t = \theta_0 e^{\eta t}. \quad (14)$$

where θ_0 is the demand density when $t = 0$ and η is the demand growth rate.

If only bus transit is available over the whole planning horizon $[0, +\infty]$, the cumulative system cost is computed as the integral of the discounted system cost over time, as follows:

$$T_B = \int_0^{+\infty} N_p C_B(\theta_t) e^{-kt} dt. \quad (15)$$

Here N_p is the number of equivalent operating hours per year. k is the discount factor in exponential discounting, which is a very commonly used discount function in decision-making over continuous time. Although an infinite planning horizon $[0, +\infty]$ is assumed, by varying the discount factor k , the infinite horizon technically reduces to a finite one, as e^{-kt} is virtually zero when t is sufficiently large.

Next, we analyze the scenario where rail transit is introduced at time τ and becomes operational after a known construction period Δ . Before $\tau + \Delta$, there is bus service only; after $\tau + \Delta$, rail and bus transit jointly serve the corridor. The cumulative system cost is expressed as follows:

$$T_{B+R}(\tau, L) = \int_0^{\tau+\Delta} N_p C_B(\theta_t) e^{-kt} dt + \int_{\tau+\Delta}^{+\infty} N_p C_{B+R}(\theta_t) e^{-kt} dt + (K_0 + K_1 L) e^{-k\tau}. \quad (16)$$

$C_{B+R}(\theta_t)$ is the total system cost for the hybrid operations after rail transit becomes operational, namely after $\tau + \Delta$. $K_0 + K_1 L$ is the capital cost of rail transit, which depends on rail length L . K_0 is the fixed capital cost of rail, and K_1 is the marginal capital cost per mile.

When demand grows in a deterministic manner, a deterministic dynamic problem is thus described as follows. We need to decide whether to introduce rail transit to the commuter corridor. If so, when should the rail line be constructed and how long should it be? This is the so-called investment timing and length choice problem. The optimization objective is to minimize the minimum of T_B and $T_{B+R}(\tau, L)$. T_B does not depend on any decision variables, while $T_{B+R}(\tau, L)$ depends on the investment timing τ and rail length L . Formally, this optimization problem is written as follows:

$$\min_{\{\tau, L>0\}} \min \{T_B, T_{B+R}(\tau, L)\}. \quad (17)$$

For any given demand growth over the planning horizon, T_B can be computed easily because it is a constant. Considering this, we can

first minimize $T_B + R(\tau, L)$ by optimizing two decision variables τ and L , which can be done numerically. Once the optimal value of $T_B + R(\tau, L)$ is obtained, it can be compared with the constant T_B to find the optimal planning decision regarding rail investments.

The essence of the investment timing and length choice problem lies in the tradeoff between the capital investment for constructing a rail line and the system cost savings (operating and user cost savings) brought by hybrid operations compared to bus operations only. The rail capital cost, and hence the system cost savings, depend on the investment timing as well as the rail length.

4. Stochastic model

4.1. Demand uncertainty modeling

Conventionally, a continuous-time random process, such as geometric Brownian motion (GBM), is employed to model uncertain demand variations due to evolving macroeconomic factors or automobile ownership, among other possible factors. Unfortunately, such a process is unable to incorporate sudden and extreme changes in demand, also known as *jumps*, such as a drastic change caused by a terrorist attack, global financial crisis, or public health crisis. For example, Fig. 2 shows how the bus and subway ridership plummeted as the COVID-19 pandemic hit New York City (NYC), according to the New York's Open Data Portal (NY Open Data, 2022). Such rare events causing sudden disruptions to demand should by no means be neglected and unfortunately, cannot be modeled with GBM. Instead, they can be modeled with a discrete counting process, namely Poisson process. In an extended planning horizon, such as 30 or 50 years, the number of jumps follows a Poisson process. The magnitude of a jump, upward or downward, can be a constant or follow a different distribution. Therefore, in this paper, we use a diffusion-jump process, which is a combination of GBM and Poisson process to realistically model demand evolution.

The jump-diffusion process under consideration is given by the following stochastic differential equation:

$$\frac{d\theta_t}{\theta_t} = \eta dt + \sigma dw(t) + d \sum_{i=1}^{N_t} U_i. \quad (18)$$

Here, η is the demand growth rate as defined in Section 3.5 (also called drift rate), dt is an infinitesimal time increment, and σ is the volatility rate. The drift rate η captures the expected demand change while the volatility rate models the extent to which demand randomly changes. $dw(t)$ is an increment of a standard Wiener process (diffusion, which is also called Brownian motion). By definition, $dw(t) = \varepsilon_t \sqrt{t}$, where ε_t is a random variable that follows the standard normal distribution with a mean of 0 and a standard deviation of 1. In the jump component $d \sum_{i=1}^{N_t} U_i$, N_t follows a Poisson process with intensity λ , which should be relatively small. U_i is the jump size

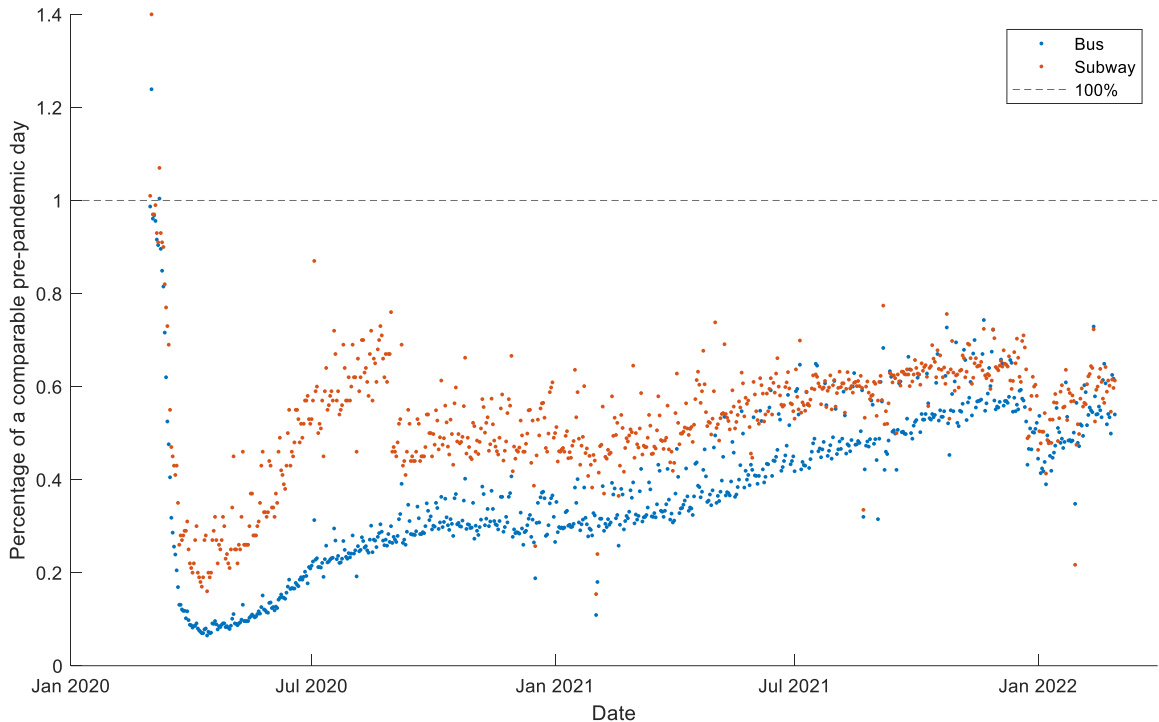


Fig. 2. Public transit demand during the COVID-19 pandemic in New York City.

(percentage change, e.g., a 10% decrease) for the i th jump, $i = 1, 2, \dots, N_t$, which is independent of GBM. U_i can be a constant or random variable, which will be discussed further in Section 4.5. Thus, the jump component in Eq. (18), i.e., $d \sum_{i=1}^{N_t} U_i$, becomes U_i if a jump occurs, with probability λdt ; $d \sum_{i=1}^{N_t} U_i$ is 0, otherwise, with probability $(1 - \lambda dt)$.

Eq. (18) indicates that over a small interval dt , if a jump does not occur, the jump-diffusion process reduces to GBM; otherwise, one or a few jumps occur.

As shown in Appendix D, the jump-diffusion differential equation, namely Eq. (18), has an analytical solution, as follows:

$$\theta_t = \theta_0 e^{\left(\eta - \frac{\sigma^2}{2}\right)t + \sigma w(t)} \prod_{i=1}^{N_t} (1 + U_i). \quad (19)$$

θ_0 is the demand density at the CBD when time is 0; N_t represents the total number of jumps occurred until time t .

Given Eq. (19), the expectation of θ_t^γ for any given value of t , where γ is a positive real number, can be derived as Eq. (20) as shown in Appendix E.

$$\mathbf{E}[\theta_t^\gamma] = \theta_0^\gamma e^{\left[\gamma\eta + \frac{\gamma^2\sigma^2}{2}(\gamma-1) + \lambda(\mathbf{E}[(1+U)^\gamma] - 1)\right]t}. \quad (20)$$

U is a random variable that has the same distribution as U_i , $i = 1, 2, \dots$. It is clear that all parameters η , σ , and λ in Eq. (18) affect the expectation of θ_t^γ . Eq. (20) will be used later in Section 4.3.

4.2. Parameter estimation for a stochastic process

It is worth noting that researchers have developed various methods for estimating the parameters for the stochastic process considered in this paper. For instance, Croghan et al. (2017) estimated the drift and volatility rate for a GBM using historical oil price data and assessed how well the oil price fit the GBM. Ramezani and Zeng (2007) employed maximum likelihood estimation to estimate a jump-diffusion model and compared its fitness with other stochastic processes, such as GBM, with daily data on stock returns. Yu (2007) also presented closed-form likelihood approximations for a jump-diffusion process. Therefore, similar approaches can be employed to derive key parameters from empirical transit demand data, which is beyond the scope of this paper.

4.3. Optimization objective formulation

Depending on whether rail transit is introduced at a future time, we consider two scenarios: not built (Scenario 0) and built (Scenario 1). For any given demand θ_0 , the cost difference between Scenario 0 and Scenario 1, is simply the difference between $C_B(\theta_0)$ and $C_{B+R}(\theta_0)$, both of which depend on θ_0 . As θ_t is stochastic, we seek to maximize the expectation of the cumulative cost difference over the planning horizon, namely to maximize $\mathbf{E}[T_B(\theta_0) - T_{B+R}(\theta_0)]$ or minimize $\mathbf{E}[T_{B+R}(\theta_0)]$, because $\mathbf{E}[T_B(\theta_0)]$ is a constant. Although there are two decision variables, we next focus on how to optimize the investment timing τ , while keeping the rail length choice L fixed at $\bar{L}$.

Noting the definitions of T_B and T_{B+R} in Eqs. (15) and (16), respectively, we write the expected cumulative cost difference $F(\tau, \bar{L})$ as:

$$F(\tau, \bar{L}) = \mathbf{E} \left[\int_0^{+\infty} N_p C_B(\theta_t) e^{-kt} dt - (K_0 + K_1 L) e^{-k\tau} - \int_0^{\tau+\Delta} N_p C_B(\theta_t) e^{-kt} dt - \int_{\tau+\Delta}^{+\infty} N_p C_{B+R}(\theta_t) e^{-kt} dt \right], \quad (21)$$

By further noting Eqs. (9) and (13), we rewrite Eq. (21) as:

$$F(\tau, \bar{L}) = \mathbf{E} \left[\int_{\tau+\Delta}^{+\infty} \left(\alpha_1 \theta_t^{\frac{1}{2}} + \alpha_2 \theta_t \right) e^{-kt} dt \right] - (K_0 + K_1 L) e^{-k\tau}, \quad (22)$$

where $\alpha_1 = N_p(e_2 - e_4)$, and $\alpha_2 = N_p(e_1 - e_3)$. e_1 to e_4 have been specified in Section 3.

It is quite clear that $F(\tau, \bar{L})$ is the expected value of the cumulative system cost savings minus rail capital cost.

By noting that the rail capital cost is a one-time investment, to transform the cumulative cost difference $F(\tau, \bar{L})$ into an integral, we re-write the one-time cost as an integral:

$$(K_0 + K_1 L) e^{-k(\tau+\Delta)} e^{k\Delta} = (K_0 + K_1 L) e^{k\Delta} \int_{\tau+\Delta}^{+\infty} -de^{-kt}, \quad (23)$$

Then, after defining $\alpha_0 = -(K_0 + K_1 L)ke^{-k\Delta}$, Eq. (22) is rearranged as an integration of a discounted cost function from $\tau + \Delta$ to the end of the planning horizon:

$$F(\tau, \bar{L}) = \mathbf{E} \left[\int_{\tau+\Delta}^{+\infty} \left(\alpha_0 + \alpha_1 \theta_t^{\frac{1}{2}} + \alpha_2 \theta_t \right) e^{-kt} dt \right]. \quad (24)$$

The lower limit of the integral in Eq. (24) can be changed to τ , i.e., the investment timing, by defining another the variable of integration $s = t - \Delta$. Finally, Eq. (24) is written as:

$$F(\tau, \bar{L}) = \mathbf{E} \left[\int_{\tau}^{+\infty} \left(\alpha_0 + \alpha_1 \theta_{s+\Delta}^{\frac{1}{2}} + \alpha_2 \theta_{s+\Delta} \right) e^{-k(s+\Delta)} ds \right] \quad (25)$$

4.4. Demand threshold and optimal timing

For the given optimization objective function Eq. (25), we seek to find a demand threshold at a time (called optimal investment timing) that maximizes $F(\tau, \bar{L})$, for any fixed rail length $\bar{L}$. Formally, we let θ^* be the *optimal trigger value* at which rail construction starts. The optimal investment timing is thus $\tau = \min\{t: \theta_t \geq \theta^*\}$. In other words, we launch the rail construction project when demand density reaches θ^* for the first time. The optimization objective Eq. (25) becomes:

$$F(\theta^*) = \int_0^{+\infty} \mathbf{E}_{\theta^*} \left[\alpha_0 + \alpha_1 \theta_{s+\Delta}^{\frac{1}{2}} + \alpha_2 \theta_{s+\Delta} \right] e^{-k(s+\Delta)} ds \quad (26)$$

in which the conditional expectation $\mathbf{E}_{\theta^*}[\cdot]$ indicates that $\theta_0 = \theta^*$. According to Eq. (20), by letting $\gamma = j/2$, $\theta_0 = \theta^*$ in Eq. (20), we obtain:

$$\mathbf{E}_{\theta^*} \left[\theta_{s+\Delta}^{\frac{j}{2}} \right] = \theta^{*\frac{j}{2}} e^{\omega_j(s+\Delta)}, \quad (27)$$

where coefficient $\omega_j = \frac{j}{2}\eta + \frac{\sigma^2}{2} \frac{j}{2} \left(\frac{j}{2} - 1 \right) + \lambda \left(\mathbf{E} \left[(1+U)^{\frac{j}{2}} \right] - 1 \right)$, $j = 1, 2$.

Thus, Eq. (26) can be further rewritten as:

$$F(\theta^*) = \int_0^{+\infty} \left[\alpha_1 \theta^{*\frac{1}{2}} e^{(\omega_1-k)(s+\Delta)} + \alpha_2 \theta^* e^{(\omega_2-k)(s+\Delta)} + \alpha_0 e^{-k(s+\Delta)} \right] ds. \quad (28)$$

We simplify Eq. (28) as:

$$F(\theta^*) = l_0 + l_1 \theta^{*\frac{1}{2}} + l_2 \theta^*, \quad (29)$$

after defining $l_0 = \frac{\alpha_0}{k} e^{-k\Delta}$, and $l_j = \frac{\alpha_j e^{(\omega_j-k)\Delta}}{k-\omega_j}$, $j = 1, 2$. Let $V(\theta)$ denote the option function of investing in rail transit. Following the classic theory of optimal stopping (Chow et al., 1971), we let $\{\theta > 0: V(\theta) = F(\theta)\}$ be the set where it is optimal to invest in rail transit, also called stopping region; let $\{\theta > 0: V(\theta) > F(\theta)\}$ be the set where it is not optimal to invest, also called continuation region.

Then, before investment, the Bellman equation (Dixit et al., 1994) is given by:

$$kV(\theta(t))dt = \mathbf{E}_t[dV(\theta(t))], \quad (30)$$

which suggests that the total return of an investment $kV(\theta(t))dt$ equals the expected rate of capital appreciation (an increase in the project value) over a small interval dt .

Applying Ito's lemma, V satisfies the following ordinary differential equation:

$$\frac{1}{2}\sigma^2\theta^2 V''(\theta) + \eta\theta V'(\theta) + \lambda E[V((1+U)\theta) - V(\theta)] - kV(\theta) = 0. \quad (31)$$

Noting that Eq. (31) is a Cauchy–Euler equation, we know that the solution of Eq. (31) is as follows:

$$V(\theta) = a_1(\theta)^{b_1}, \quad (32)$$

where a_1 and b_1 are positive constants to be determined. Constant b_1 satisfies the non-linear equation: $\varphi(b_1) = 0$, where

$$\varphi(b_1) = \frac{1}{2}\sigma^2 b_1(b_1 - 1) + \eta b + \lambda E[(1+U)^{b_1}] - (\lambda + k). \quad (33)$$

Note that $\varphi(1) = \eta + \lambda E[U] - k$ and $\lim_{b_1 \rightarrow \infty} \varphi(b_1) = \infty$. If

$$\eta + \lambda E[U] < k, \quad (34)$$

then there exists $b_1 > 1$ such that $\varphi(b_1) = 0$.

Constant a_1 is determined by using the boundary conditions, namely the value-matching and smooth-pasting conditions:

$$V(\theta^*) = F(\theta^*), \quad (35)$$

$$V'(\theta^*) = F'(\theta^*). \quad (36)$$

The value-matching condition Eq. (35) states when the demand trigger θ^* is reached, the value of the investment opportunity $V(\theta^*)$ is the same as the net present value of the project $F(\theta^*)$. The smooth-pasting condition Eq. (36) ensures the option function $V(\cdot)$ is maximized.

Then, the following conditions are satisfied:

$$a_1 \theta^{*b_1} = l_0 + l_1 \theta^{*\frac{1}{2}} + l_2 \theta^*, \quad (37)$$

$$a_1 b_1 \theta^{*b_1-1} = l_1 \frac{1}{2} \theta^{*\frac{1}{2}} + l_2. \quad (38)$$

By comparing Eqs. (37) and (38), we can obtain the analytical solution of the optimal demand density threshold from the following equation:

$$a_1 = \left(l_0 + l_1 \theta^{*\frac{1}{2}} + l_2 \theta^* \right) \theta^{*-b}. \quad (39)$$

Finally, the optimal demand density at CBD at time t is:

$$\theta^{\frac{1}{2}} = \frac{l_1(1 - 2b_1) + \sqrt{(l_1(2b_1 - 1))^2 - 16l_0l_2b_1(b_1 - 1)}}{4l_2(b_1 - 1)}. \quad (40)$$

For each fixed rail length $\bar{L}$, we can follow the above procedure to find the analytical solution of the trigger value θ^* . We can then conduct a numerical search on a given interval for L to find the optimal rail length, because finding an analytical solution of L is challenging given the complexity of Eq. (25).

4.5. Alternative demand jump processes

In Section 4.1, we stated the jump magnitude (percentage change) U as either a constant or a random variable; in Section 4.4, we did not further specify its type of distribution, because the demand trigger can be derived similarly. Next, we analyze the results under

Table 3

Values of some parameters under different jump processes.

Scenario (a): No jumps	l_1	$\frac{\alpha_1 e^{\left(\frac{1}{2}\eta - \frac{\sigma^2}{8} - k\right)\Delta}}{k - \left(\frac{1}{2}\eta - \frac{\sigma^2}{8}\right)}$
	l_2	$\frac{\alpha_2 e^{(\eta-k)\Delta}}{k - \eta}$
	b_1	$\frac{1}{2}\sigma^2 b_1(b_1 - 1) + \eta b_1 - k = 0$
Scenario (b): Constant jumps	l_1	$\frac{\alpha_1 e^{\left(\frac{1}{2}\eta - \frac{\sigma^2}{8} + \lambda\left((1+U)^{\frac{1}{2}} - 1\right) - k\right)\Delta}}{k - \left(\frac{1}{2}\eta - \frac{\sigma^2}{8} + \lambda\left((1+U)^{\frac{1}{2}} - 1\right)\right)}$
	l_2	$\frac{\alpha_2 e^{(\eta+\lambda U-k)\Delta}}{k - (\eta + \lambda U)}$
	b_1	$\frac{1}{2}\sigma^2 b_1(b_1 - 1) + \eta b_1 + \lambda(1+U)^{b_1} - (\lambda + k) = 0$
Scenario (c): Stochastic jumps	l_1	$\frac{\alpha_1 e^{\left(\frac{1}{2}\eta - \frac{\sigma^2}{8} + \lambda\left(\frac{E[U]}{2} - \frac{E[U]^2}{8} + \frac{E[U]^3}{16} - \frac{5E[U]^4}{128} + \frac{7E[U]^5}{256} \dots\right) - k\right)\Delta}}{k - \left(\frac{1}{2}\eta - \frac{\sigma^2}{8} + \lambda\left(\frac{E[U]}{2} - \frac{E[U]^2}{8} + \frac{E[U]^3}{16} - \frac{5E[U]^4}{128} + \frac{7E[U]^5}{256} \dots\right)\right)}$
	l_2	$\frac{\alpha_2 e^{(\eta+\lambda E[U]-k)\Delta}}{k - (\eta + \lambda E[U])}$
	b_1	$\frac{1}{2}\sigma^2 b_1(b_1 - 1) + \eta b_1 + \lambda\left(1 + b_1 E[U] + \frac{E[U]^2}{2} b_1(b_1 - 1) + \frac{E[U]^3}{6} b_1(b_1 - 1)(b_1 - 2) + \dots\right) - (\lambda + k) = 0$

each of three scenarios for U .

Note that a jump may occur in either direction, namely upward (positive U) or downward (negative U). In Scenario (a), we set $U = 0$, meaning no jumps are considered. Scenario (a) can also be achieved by setting the intensity of the Poisson process to be zero, namely $\lambda = 0$. In Scenario (b), the magnitude of each jump is the same constant, such as a 10% drop. In Scenario (c), the magnitude of the jump U follows an exponential distribution, whose probability density function is as follows:

$$f(x; g) = \begin{cases} ge^{-gx} & x \geq 0, \\ 0 & x < 0. \end{cases} \quad (41)$$

where g is the rate parameter for the exponential distribution with $1/g$ denoting the expected magnitude of each jump. For example, when $g = 10$, the expected value of U is 10%.

Eq. (40) applies in each scenario, while determinations of parameters l_1 , l_2 , and b_1 differ across different scenarios. Specifically, Table 3 shows the values of l_1 , l_2 , and b_1 under each scenario.

5. Model applications

5.1. Numerical inputs

We next conduct extensive numerical studies to demonstrate the usefulness of the analytical methods and highlight some important insights into rail transit investment timing and sizing decisions. Unless stated explicitly otherwise, the input parameters and baseline values in Table 2 are adopted largely from Sun et al. (2017) and Guo et al. (2018b). Other values of some key parameters are used in sensitivity analyses in Section 5.4.

5.2. Deterministic analyses

We evaluate the operating advantage of the hybrid service over bus only. It is understandable that as demand density increases, the hybrid service become advantageous as evidenced by a lower system cost per hour; however, this relative advantage depends on the configuration of the hybrid service, specifically the length of the rail line. We first consider an efficient hybrid configuration in which the rail length is 35 miles. As shown in Fig. 3, when demand density is relatively low, e.g., θ is slightly below 10 pax/hr./mi., bus-only has a smaller hourly system cost; when $\theta > 10$ pax/hr./mi. for example, the hourly system cost of the hybrid service is lower than that of bus-only, namely $C_{B+R} < C_B$ when $L = 35$ mi. This finding does not hold when the rail length is quite limited, such as $L = 2$ mi., regardless of demand density θ . In other words, when $L = 2$ mi., the hybrid service is always inferior to bus only.

Fig. 4 further shows the system cost breakdown for bus-only and hybrid services. Clearly, bus-only is less advantageous mainly because its operating cost increases more quickly as demand density increases, even though the user cost is higher under hybrid than under bus-only operations. In this case, the configuration of the hybrid service is said to be “efficient” in the sense that it can dominate bus-only operations. Fig. 5 shows that when rail length is limited, the configuration becomes “inefficient”, because the operating cost under hybrid operations is on par with that of bus-only while its users’ cost is higher. Under hybrid services, residents living beyond the transfer point must change from bus to rail; the additional waiting for trains at the transfer point explains why the users’ cost is higher than in bus-only operations where no transfers are needed. It should also be noted that rail and bus have quite different operating cost

Table 2
Baseline values.

Notation	Definitions	Baseline values	Unit
A	Corridor length	50	mi.
φ	Value of in-vehicle travel time	10	\$/hr.
e_w	Value of waiting time	40	\$/hr.
θ_0	CBD demand density at time 0	15	pax/hr./mi.
V_R	Rail average travel speed	40	mph
V_B	Bus average travel speed	30	mph
α_R	Rail fixed hourly cost	1000	\$/hr.
α_B	Bus fixed hourly cost	120	\$/hr.
β_R	Rail variable hourly cost	3.0	\$/hr./seat
β_B	Bus variable hourly cost	6.0	\$/hr./seat
δ_R	Rail maximum load factor	1.2	–
δ_B	Bus maximum load factor	1.1	–
f_R	Fixed user cost – Rail riders	1.0	\$
f_B	Fixed user cost – Bus riders	1.0	\$
N_p	Operation hours per year	3600	–
K_0	Fixed capital cost of rail	100	Million \$
K_1	Marginal capital cost of rail	1.0	Million \$/mi.
k	Annual discount rate	0.05	–
η	Annual demand growth rate	0.01	–
Δ	Number of years of construction	5.0	yrs.

parameters, such as fixed and variable operating costs, which also contribute to the cost difference between bus-only and hybrid services.

We next show the optimal solutions to the deterministic dynamic problem defined in Section 3.5. The initial demand density at the CBD θ_0 , demand growth rate η , as well as other parameters are specified in Table 2. Fig. 6 shows how the cumulative cost of the hybrid service $T_B + R(\tau, L)$ varies with the investment timing τ and rail length L . The optimal solutions, namely $\tau^* = 18$ and $L^* = 27$, are also highlighted in Fig. 7. It is worth noting that the cumulative cost of the hybrid operations, i.e., optimization objective, is relatively flat around the optimum, especially along the axis for investment timing. This implies that although an optimum investment timing exists, the objective is not very sensitive to an investment timing change as long as the deviation from the optimum timing is not substantial.

Fig. 7 shows the annualized system cost for each investment scenario (no built: bus only throughout the planning horizon, and built: rail introduced in Year 18), along with the one-time rail investment cost when rail construction starts. All costs shown in Fig. 7 are not yet discounted. Before Year 23 when rail becomes operational, the annualized system cost $N_p C_B$ is the same for both scenarios; after Year 23 (i.e., $18+5$), the operating advantage of the hybrid service becomes evident, as $N_p C_{B+R} < N_p C_B$. If we kept the optimal L^* but change the optimal timing to 0, we would have observed a different system cost curve for the hybrid service from time 0 to Year 23, as represented by a dashed curve labeled “Hybrid - annualized system (virtual).” Clearly, even at the beginning of the planning horizon, namely at time 0, operating hybrid services saves money, namely $N_p C_{B+R} < N_p C_B$. However, the rail construction does not start until Year 18. This is because of the important trade-off between the operating savings due to the introduction of rail and the one-time capital cost of rail construction, to be elaborated next.

Fig. 8 plots the cumulative cost saving $T_B - T_B + R(\tau, L)$ at each possible investment timing. Clearly, if the rail investment occurs at time 0 and the rail length is optimizable, a positive cost saving is observed immediately, namely $T_B > T_B + R(\tau, L)$. However, $T_B - T_B + R(\tau, L)$ reaches the maximum much later, at $\tau^* = 18$. This occurs because when the decrease in the rail capital cost accounting for future cost discounting exceeds the decrease in the cumulative system cost saving, investment should be postponed. When the rail investment is delayed on purpose, $T_B - T_B + R(\tau, L)$ can become negative, which means investing in rail is no longer worthwhile. This occurs because although the rail capital cost is discounted more heavily, the opportunity to take advantage of the operating savings due to the introduction of rail also dramatically diminishes. In this case, introducing rail cannot yield a net cost saving. In other words, a rail line can no longer be justified.

Those above results are important because they directly refute the misconception that rail should be introduced immediately after hybrid services become more operationally advantageous to bus only (namely $C_{B+R} < C_B$) or the net cumulative cost saving becomes positive (namely $T_B + R(\tau, L) < T_B$). Note that in a related study Chen et al. (2015), it was assumed that as long as positive cost savings are expected, a new transit mode should be introduced. This limited assumption neglected the possibility that postponing a project may possibly yield more savings or a higher NPV. To summarize, investing when a project generates a positive cash flow or results in an overall cost saving over the planning horizon is not necessarily optimal.

Therefore, the trade-off between system cost savings and rail investment must be fully examined when the investment timing decision is made. As rail length affects the rail capital cost and largely determines the operational advantage of hybrid services, a

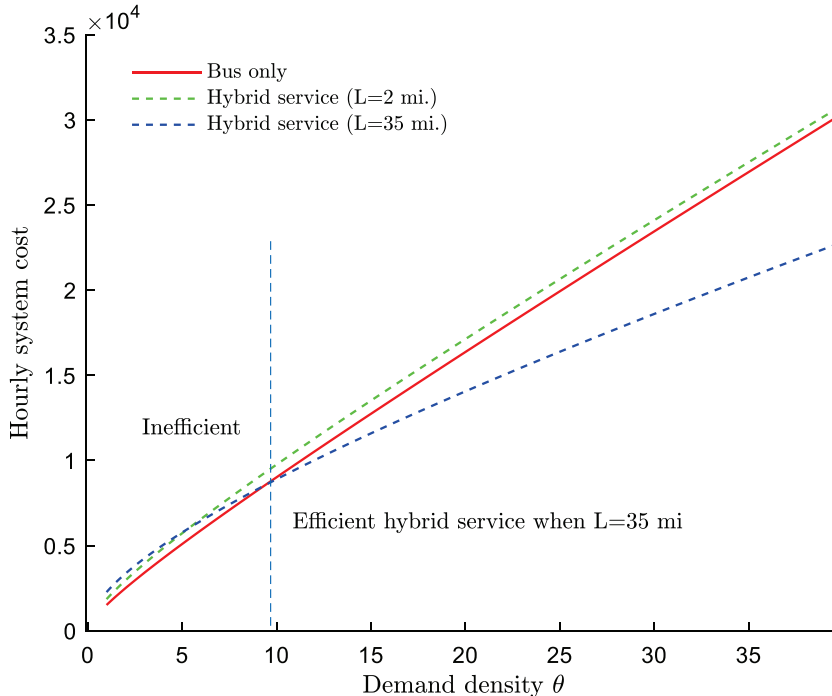


Fig. 3. Hourly system cost comparisons.

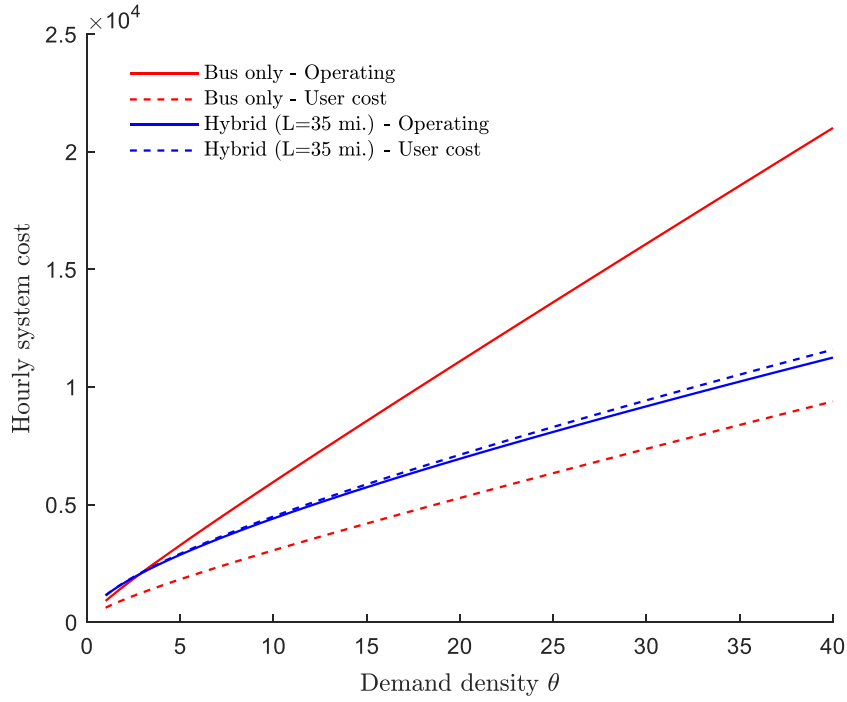


Fig. 4. Hourly system cost breakdown when rail length (i.e., 35 miles) is sufficient.

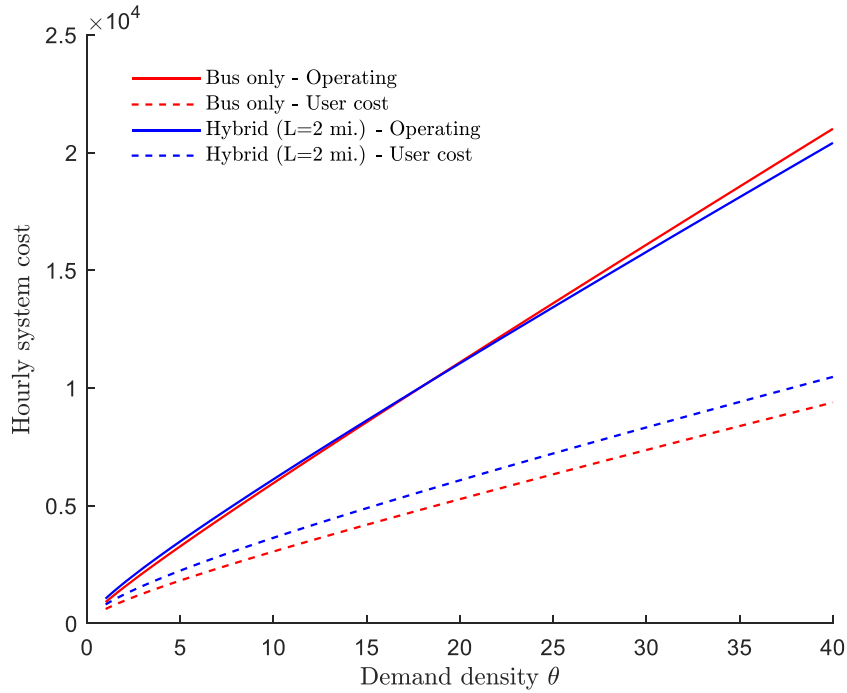


Fig. 5. Hourly system cost breakdown when rail length (i.e., 2 miles) is insufficient.

similar trade-off analysis should be made. The above analyses thus indicate that the rail transit investment timing and length choice problem is nontrivial even when demand uncertainty is not considered.

We next explore the impact of the rail capital cost parameters on various cumulative costs, namely T_B , T_R , and $T_{B+R}(\tau, L)$, where T_R represents the cumulative cost if rail transit covers the entire corridor throughout the planning horizon. Clearly, regardless of the fixed

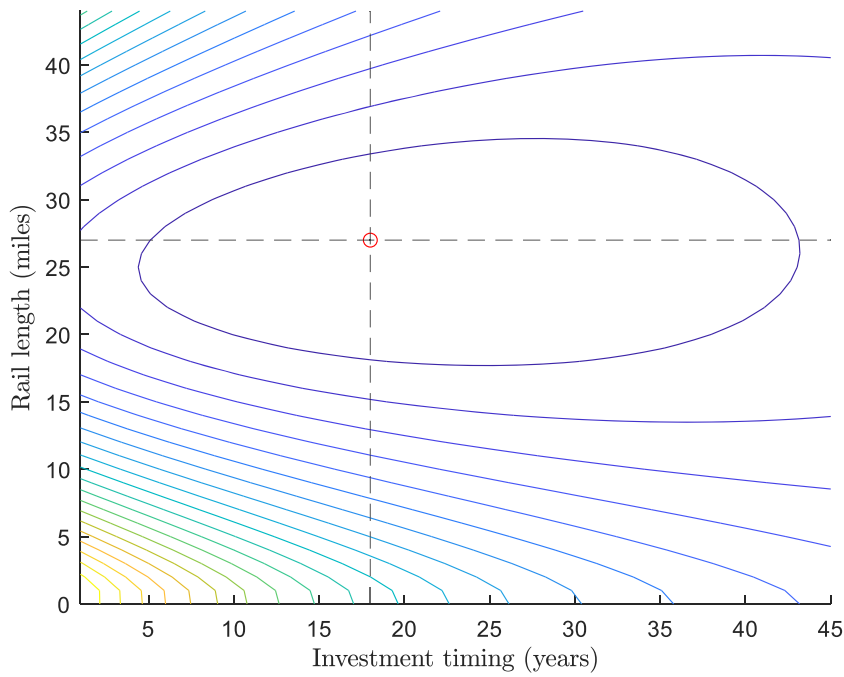


Fig. 6. Cumulative cost of hybrid operations as a function of investment timing and sizing.

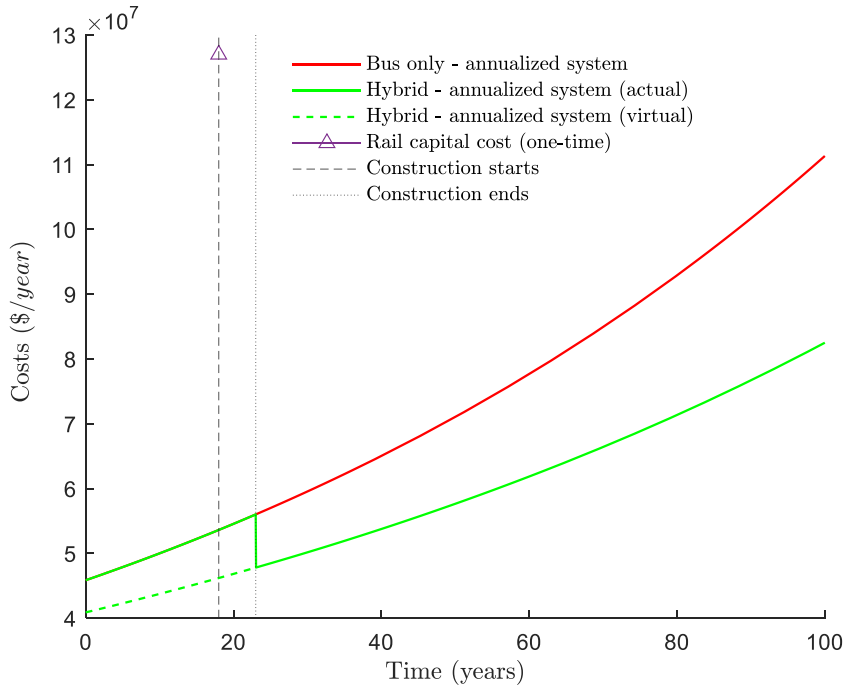


Fig. 7. Various costs over time when optimal timing and length choice are selected.

capital cost parameter K_0 , having rail serve the entire corridor from time 0 is not a good decision, because its cumulative cost T_R is the highest, as shown in Fig. 9. T_R grows at the same rate as K_0 as expected, because all the capital cost is assumed to occur at time 0. The hourly or annualized system cost of rail-only operations does not rely on K_0 . T_B does not vary with K_0 , because bus-only operations do not involve any rail investment. We can observe from Fig. 9 that as K_0 increases, the gap between $T_{B+R}(\tau, L)$ and T_B approaches zero, after which $T_{B+R}(\tau, L)$ exceeds T_B , which means when the fixed capital cost of rail becomes overly large, the hybrid service will not be introduced. Similarly, Fig. 10 shows that as the variable capital cost of rail K_1 increases from 2 to 12 million while keeping K_0 at 10

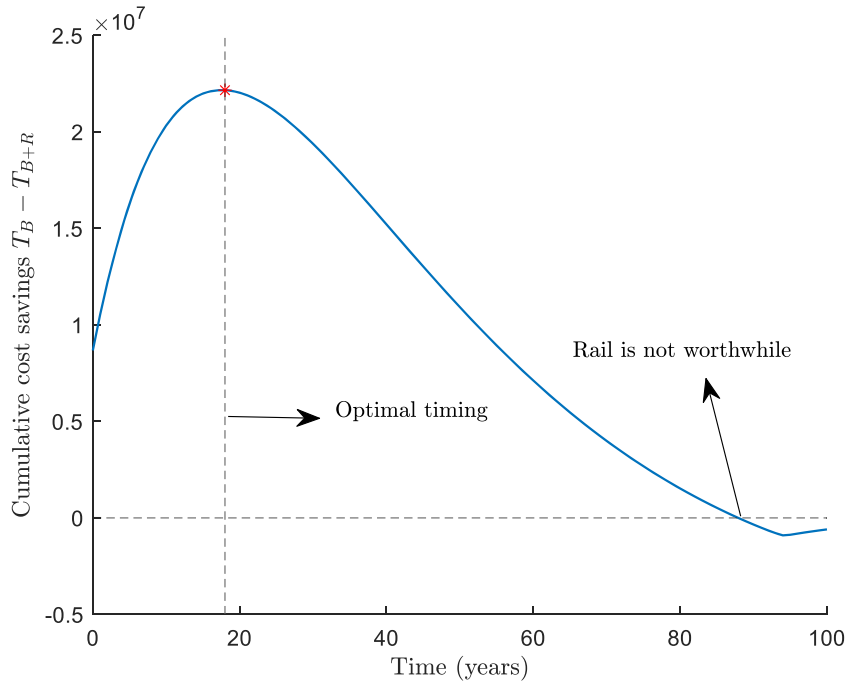


Fig. 8. Relation between cumulative cost savings and investment timing.

million, the advantage of hybrid over bus-only gradually diminishes.

In the above analyses, the baseline value of in-vehicle travel time φ is fixed at \$10 per hour. We now consider a time-varying $\varphi_t = \varphi e^{\bar{\eta}t}$, where $\bar{\eta}$ is the growth rate for the value of time. Fig. 11 shows how the optimal investment timing and rail length vary with the growth rate. Specifically, as $\bar{\eta}$ increases, the rail line should be introduced earlier, and it should be longer.

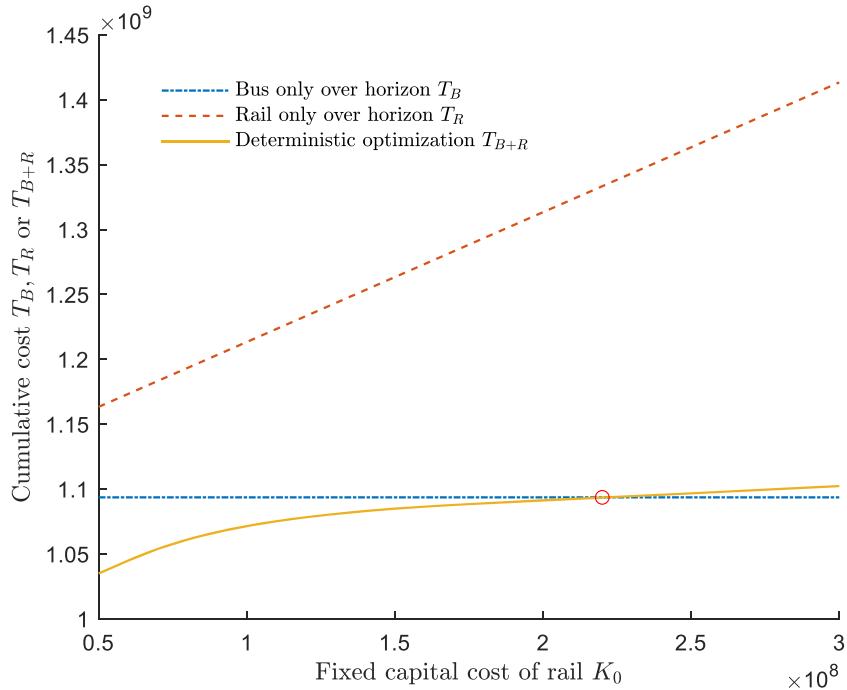


Fig. 9. Comparisons of cumulative costs under various fixed capital cost parameter values.

5.3. Stochastic analyses

We next consider uncertain future demand with a volatility rate $\sigma = 0.1$. The second jump process, namely Scenario (b) defined in Section 4.4, is considered, with a constant magnitude for each downward jump, i.e., $U_i = 10\%$, $i = 1, \dots, N_t$. The intensity of the underlying Poisson process is $\lambda = 0.1$. The values of all other parameters are the same as those in Section 5.2. For instance, at time 0, the demand density is fixed at $\theta_0 = 15$ pax/hr./mi.

Following the approach presented in Section 4.4 and fixing the rail length at $\bar{L} = 20$, we can analytically derive the optimal demand density threshold from the value-matching and smooth-pasting conditions Eqs. (35) and (36), as shown in Fig. 12. As defined in Section 4.4, when demand density is below the demand density threshold, i.e., in the continuation region, the best strategy is to defer the investment. Otherwise, investment should be implemented when demand density reaches the stopping region. As a comparison, we reduce the magnitude of a jump to zero, namely $U_i = 0$, which means no demand jumps are considered. Consequentially, the derived demand density threshold decreases from 21.1 pax/hr./mi. to 18.5 pax/hr./mi., representing a premature investment timing if downward demand jumps are neglected.

We next show in Fig. 13 how the expectation of the cumulative cost difference, defined in Eq. (25), can be maximized by varying the rail length. As we can find the optimal demand density threshold and the resulting optimization objective for each given rail length, we can numerically find the optimal rail length (i.e., 27 mi.) that maximizes the expected cumulative cost difference, and the corresponding optimal demand density threshold (i.e., 20.3 pax/hr./mi.). Fig. 14 shows how the optimal demand density threshold θ^* depends on the rail length L . Their interrelations can be described as follows: (1) when the rail length is insufficient (e.g., less than 2 mi.), θ^* is very large, implying a delayed investment timing, due to the inefficient hybrid service layout; (2) when the rail length is excessive (close to the corridor length, e.g., 48 mi.), θ^* is also considerably large, because a substantial rail investment cost should be justified only by sufficiently high demand density; (3) when the rail configuration is efficient, determined by an appropriate rail length, θ^* is relatively small, meaning that such an efficient hybrid service layout should be introduced relatively early.

5.4. Sensitivity analyses

We next conduct sensitivity analyses of a few key parameters while keeping the same type of demand jump process as in Section 5.3, namely constant jumps. Fig. 14 shows that as volatility rate σ increases, both the optimal demand density threshold θ^* and rail length L^* increase. The policy implication is that increasing demand uncertainty favors a later investment timing and a larger investment size. However, when the volatility rate is fixed, we find that downward demand jumps have different impacts on those decisions. Specifically, Fig. 14 shows that downward jumps lead to an increased demand threshold but a decreased rail length in comparison with no jumps (dashed lines in Fig. 14).

Fig. 15 shows that as the demand growth rate η grows, the demand threshold decreases while the rail length increases. The different impacts of downward demand jumps on the two decisions are observed again. Without jumps, the demand threshold is lower while the rail line is longer.

Fig. 16 shows the effects of construction period Δ and annual discount rate k on the two decision variables. For a given construction period, as the discount rate increases, the demand threshold increases. For a fixed discount rate, as the construction duration increases, the demand threshold decreases. Unlike the investment timing, the investment size (rail length) is not very sensitive to the construction period. As the construction period varies from 3 to 7, the variation in the optimal rail length stays within 0.3 miles, which is very limited. The optimal rail length is indeed sensitive to the annual discount rate.

It is interesting to observe in Fig. 17 that as the jump magnitude, the demand threshold monotonically increases, while the rail length first decreases and then increases.

5.5. Effect of the jump process

We finally consider the last jump process, namely Scenario (c) where the magnitude of a jump follows an exponential distribution. We set the rate parameter for the exponential distribution to be 10, which implies the expected jump magnitude U is 0.1. Fig. 18 shows how the optimal demand threshold and rail length vary with the intensity of the underlying Poisson process. As the intensity λ grows, the optimal demand threshold increases while the optimal rail length decreases. Compared with constant jumps, the demand density threshold is higher, and the rail length is also higher under exponential jumps.

6. Conclusions

Analytical models are presented for simultaneously determining when a rail transit line should be constructed and how long it should be, with a realistic formulation for long-term demand evolutions, namely a jump-diffusion process consisting of a standard geometric Brownian motion and a jump component driven by a Poisson process. We seek to maximize the expected cumulative cost difference between operating buses throughout and introducing rail services at a certain future time in a commuter corridor. Analytical solutions for the investment timing under uncertainty have been derived under various scenarios for the jump process.

When demand uncertainty is neglected, we refute a major misconception. Contrary to the common wisdom in the transit infrastructure planning literature (Chen et al., 2015), we find that even when demand is assumed to be deterministic it is not necessarily optimal to implement an investment project immediately when a positive cash flow occurs or a cumulative cost saving over the entire

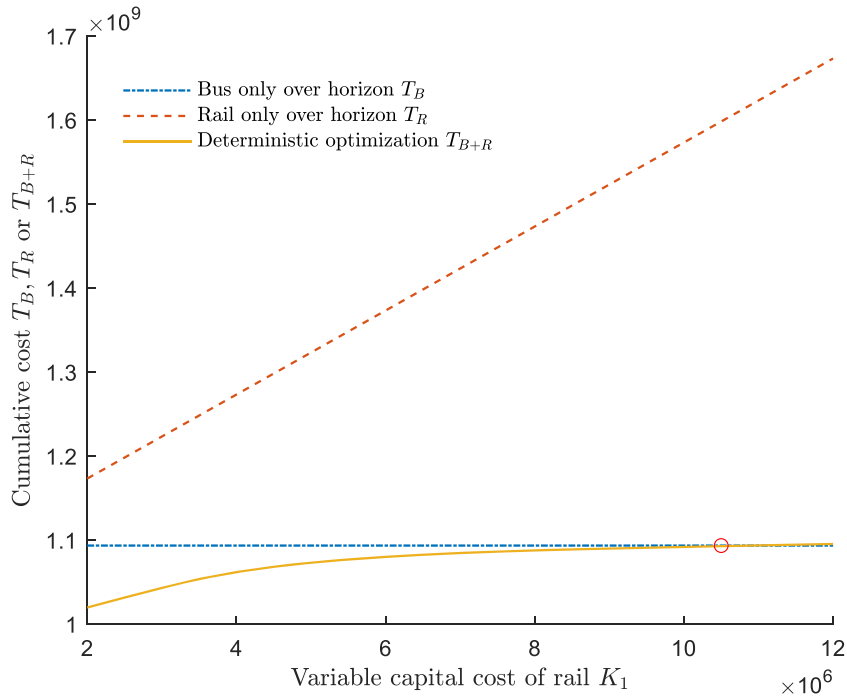


Fig. 10. Comparisons of cumulative costs under various variable capital cost parameter values.

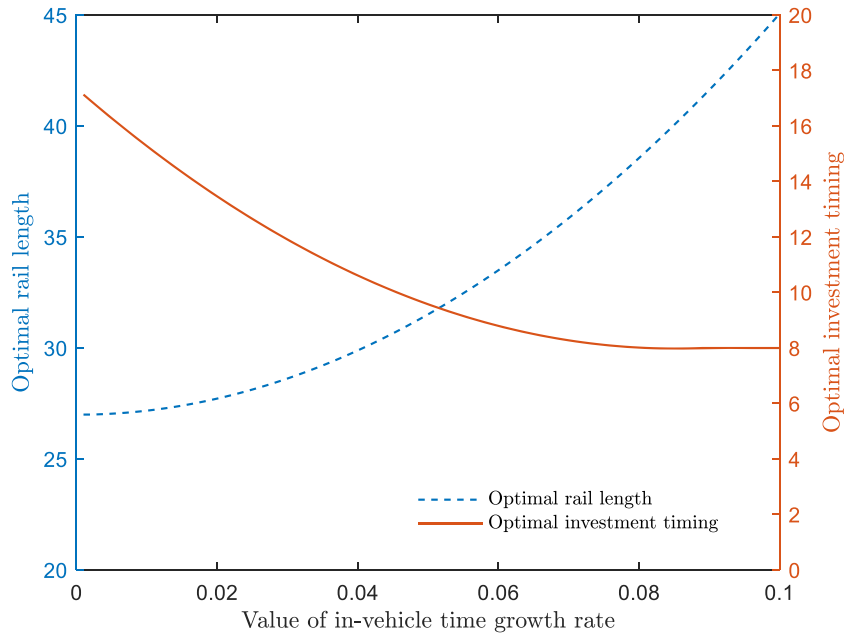


Fig. 11. Effect of time-varying value of time on planning decisions.

planning horizon is expected. Note that in a deterministic study, [Chen et al. \(2015\)](#) concluded that as long as a certain demand threshold is reached leading to an indifference between two transit modes, the new transit mode should be implemented, which is suboptimal. [Li et al. \(2015\)](#) also described a cost benefit analysis based on the net present value. It was stated that a new transit mode should be introduced if and only if the net improvement in the project value (measured by social welfare) becomes positive, which is also suboptimal even under the deterministic setting.

Several other important findings are derived through extensive numerical studies considering demand uncertainty, as follows:

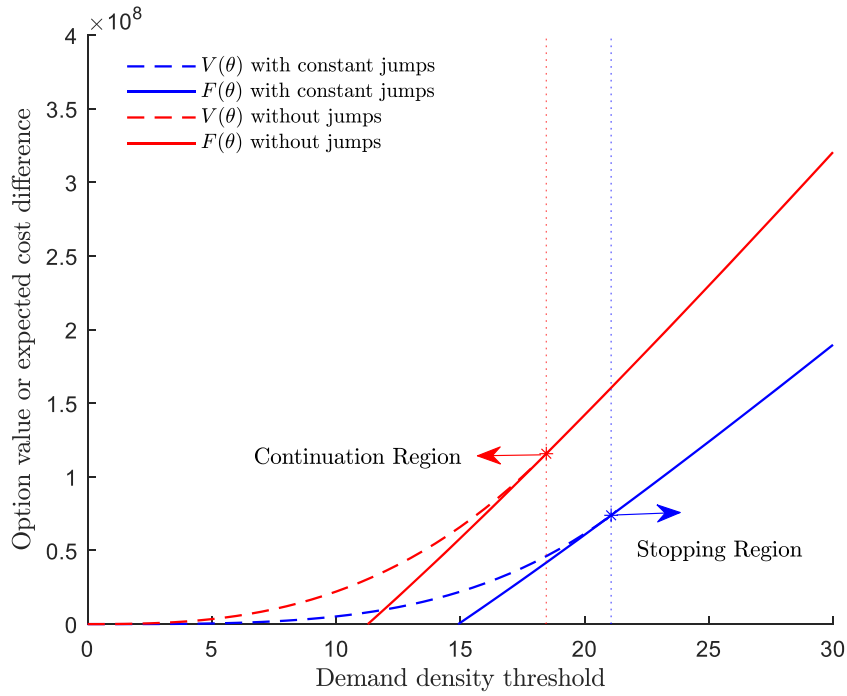


Fig. 12. Deriving demand density threshold under uncertainty for a fixed rail length.

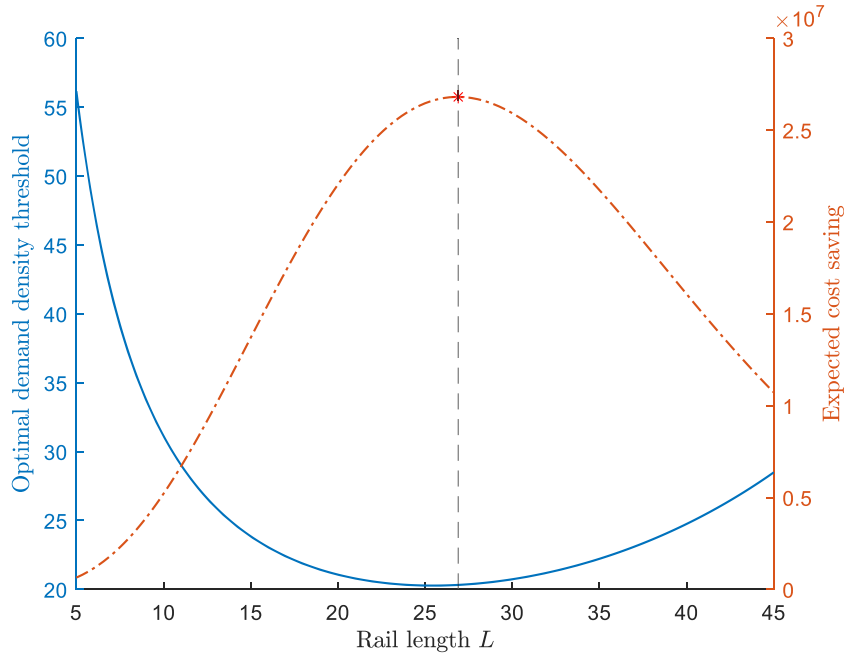


Fig. 13. Concurrent optimization of rail length and demand threshold under uncertainty.

- (1) Under demand uncertainty, two decisions (investment timing and sizing) are interrelated due to the complex tradeoff of capital investments and operational savings: depending on the range, as the rail length grows, the optimal demand threshold may increase or decrease.
- (2) While the demand threshold clearly varies with values of important parameters, such as construction period and discount rate, the optimal rail length can be relatively insensitive to similar changes, such as the infrastructure construction time.
- (3) Both the demand threshold and rail length increase as the demand volatility increases.

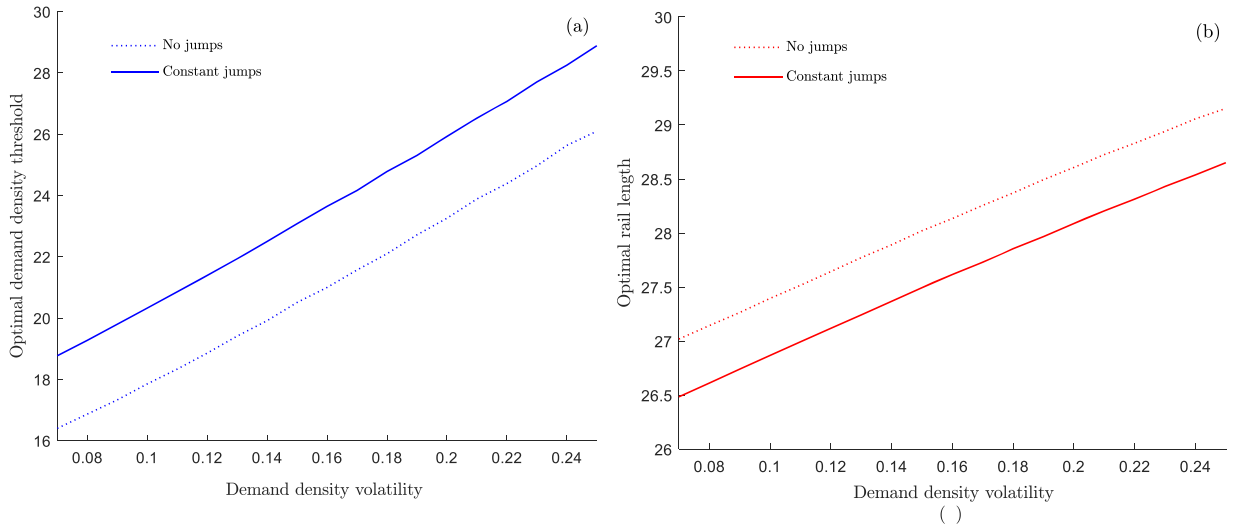


Fig. 14. Effect of volatility rate on optimal demand density threshold and rail length.

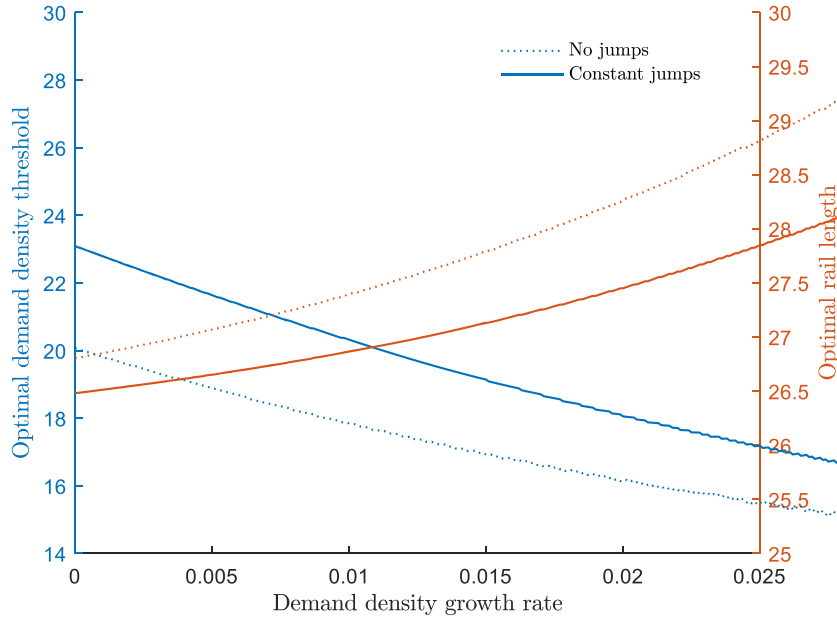


Fig. 15. Effect of demand growth rate on optimal demand density threshold and rail length.

- (4) As the value of time grows in the future, the optimum rail length should be longer, and the investment timing should be earlier.
- (5) When downward demand jumps are considered, a larger demand threshold for investment is derived, implying a delayed investment timing. In contrast, the optimal rail length decreases rather than increases.

In this study, we focus on the mobility benefits brought by rail transit, such as increased travel speed, while neglecting other benefits of investments in rail transit. For instance, [Litman \(2007\)](#) indicated that good rail transit projects can stimulate compact urban development, reduce car-dependency, and promote environmental sustainability. Further including such benefits may favor earlier rail investments. In addition, we focus on the development of a single line, while a transit line may be connected to other lines. In such a case, the length of a line may depend on the structure and scale of the rest of the network. A much more complex method would then be needed.

Although some new insights and important findings are achieved with the proposed analytical models, we can strengthen the current methods and enhance the relevant analyses as follows:

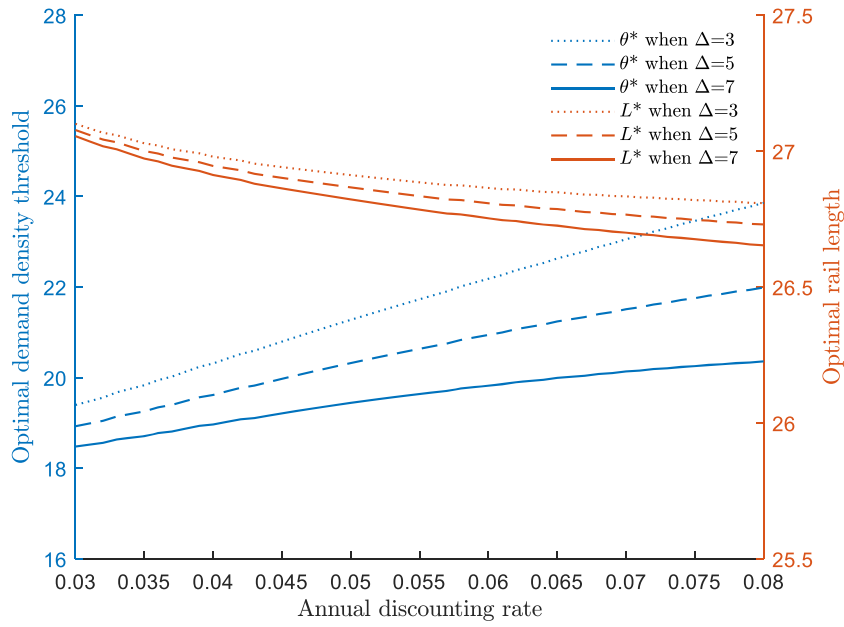


Fig. 16. Sensitivity analyses of construction period and annual discount rate.

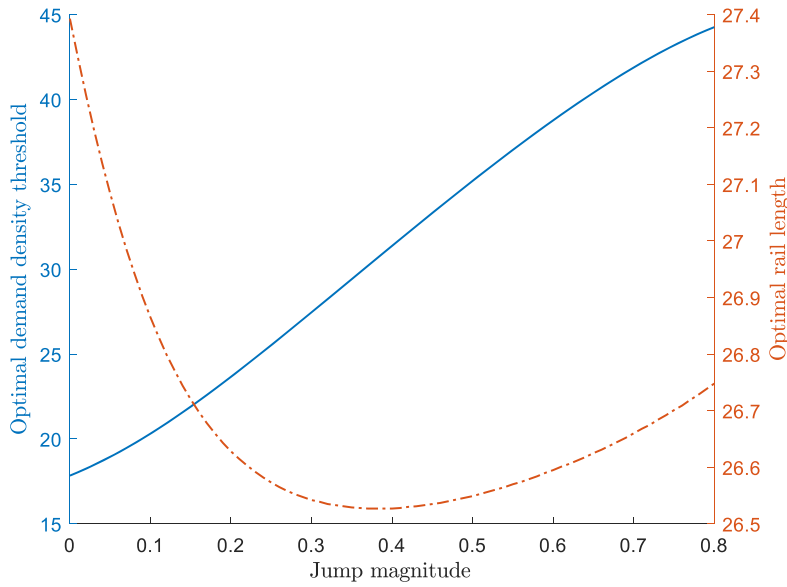


Fig. 17. Effect of jump magnitude on optimal demand density threshold and rail length.

- i Only a single source of uncertainty, namely demand density, is considered in the current analysis. Additional sources of uncertainty, such as uncertain capital cost parameters, uncertain construction periods, budgets and riders' value of time, should be considered.
- ii As the public transit planning authority is implicitly assumed to be risk-neutral, in a future study a mean-variance formulation (Sun and Schonfeld, 2016) can be developed incorporating certain risk metrics.

CRedit authorship contribution statement

Qianwen Guo: Conceptualization, Methodology, Investigation, Visualization. **Shumin Chen:** Methodology, Formal analysis, Validation. **Yanshuo Sun:** Conceptualization, Validation, Writing – original draft, Writing – review & editing. **Paul Schonfeld:** Writing – review & editing, Supervision.

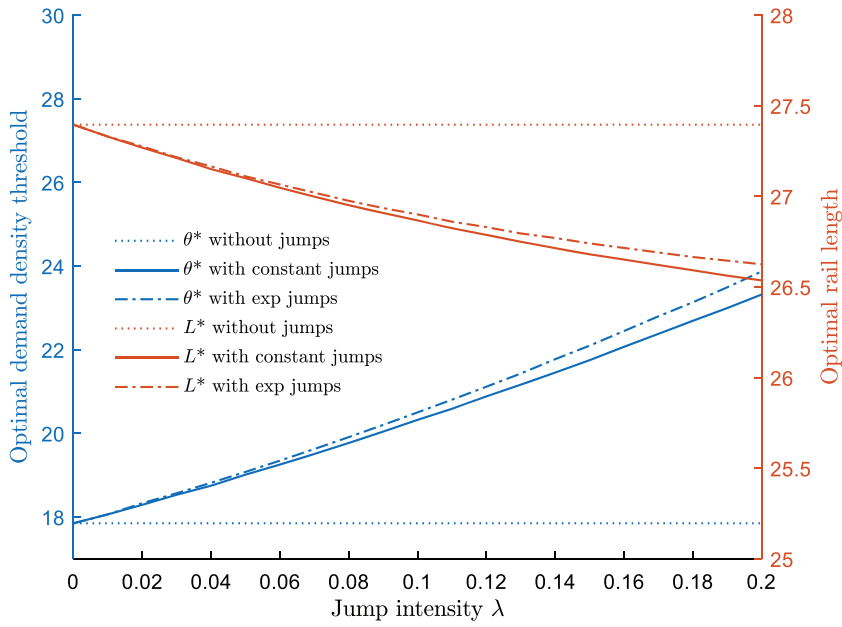


Fig. 18. Effect of the intensity of the Poisson jump process on two decisions.

Data availability

All data are included in the paper.

Acknowledgements

Comments from anonymous reviewers have led to significant improvements to this paper. The authors gratefully acknowledge this assistance. Chen is supported by the National Natural Science Foundation of China (Grant No. 72171056, and 72071215). The corresponding author is partially supported by National Science Foundation (Grant Nos. 2200384, 2100745, and 2055347).

Appendix A. Analytical Solution of Static Problem Eqs. (7) and (8)

The optimal headway h_B^* can be found by setting the first-order partial derivative of C_B with respect to headway h_B to be zero, which is:

$$h_B^* = 2\sqrt{\frac{2\alpha_B}{V_B e_w \theta}}. \quad (A1)$$

Because of the binding vehicle capacity constraint Eq. (8), the optimal vehicle capacity k_B^* can be found:

$$k_B^* = \frac{1}{\delta_B} \sqrt{\frac{2\alpha_B \theta A^2}{V_B e_w}}. \quad (A2)$$

Substituting the optimal headway h_B^* and vehicle capacity k_B^* into the optimization function Eq. (7) yields:

$$\min_{\{h_B, k_B > 0\}} C_B = \int_0^A \left(\varphi \frac{x}{V_B} + \frac{e_w 2\sqrt{\frac{2\alpha_B}{V_B e_w \theta}}}{2} + f_B \right) \left(\theta - \frac{\theta}{A} x \right) dx + \frac{\left(\alpha_B + \beta_B \frac{1}{\delta_B} \sqrt{\frac{2\alpha_B \theta A^2}{V_B e_w}} \right) A}{V_B \sqrt{\frac{2\alpha_B}{V_B e_w \theta}}}. \quad (A3)$$

Eq. (A3) can be rearranged as Eq. (9).

Appendix B. Analytical Solution of Static Problem Eqs. (10)–(12)

To find the optimal headways, we set the first-order partial derivatives of the objective Eq. (10) with respect to h_R and h_B to be zero, as follows:

$$\frac{\partial C_{B+R}}{\partial h_R} = \frac{e_w}{2} \int_0^L \left(\theta - \frac{\theta}{A} x \right) dx + \frac{e_w}{2} \int_L^A \left(\theta - \frac{\theta}{A} x \right) dx - \frac{2L\alpha_R}{V_R h_R^2} = 0, \quad (B1)$$

$$\frac{\partial C_{B+R}}{\partial h_B} = \frac{e_w}{2} \int_L^A \left(\theta - \frac{\theta}{A} x \right) dx - \frac{2(A-L)\alpha_B}{V_B h_B^2} = 0. \quad (B2)$$

As the headways for bus and train must be positive, from the first-order conditions, namely Eqs. (B1) and (B2), we obtain:

$$h_R^* = 2 \sqrt{\frac{2L\alpha_R}{V_R e_w \theta A}}, \quad (B3)$$

$$h_B^* = 2 \sqrt{\frac{(A-L)\alpha_B}{V_B e_w \theta \left(\frac{1}{2}A - L + \frac{L^2}{2A} \right)}}. \quad (B4)$$

After substituting Eqs. (B3) and (B4) into the vehicle capacity constraints Eqs. (11) and (12), we obtain the optimal vehicle capacities for bus and train as follows:

$$k_R^* = \frac{1}{\delta_R} \sqrt{\frac{2L\alpha_R \theta A}{V_R e_w}}, \quad (B5)$$

$$k_B^* = \frac{2}{\delta_B} \sqrt{\frac{(A-L)\alpha_B \theta \left(\frac{1}{2}A - L + \frac{L^2}{2A} \right)}{V_B e_w}}. \quad (B6)$$

Since the closed-form solutions for headways and vehicle capacities are available, they can be inserted into the objective Eq. (10), which becomes a function of rail length L and demand density θ , expressed as follows:

$$\begin{aligned} \min_{\{A>L>0\}} C_{B+R} = & \int_0^L \left(\varphi \frac{x}{V_R} + \sqrt{\frac{2e_w L \alpha_R}{V_R \theta A}} + f_R \right) q(x) dx + \int_L^A \left(\varphi \left(\frac{x-L}{V_B} + \frac{L}{V_R} \right) + \sqrt{\frac{2e_w L \alpha_R}{V_R \theta A}} - \frac{1}{\frac{L}{\sqrt{2A}} - \sqrt{\frac{A}{2}}} \sqrt{\frac{e_w (A-L)\alpha_B}{V_B \theta}} + f_B \right) q(x) dx \\ & + \sqrt{\frac{\theta A e_w L \alpha_R}{2V_R}} + \frac{\beta_R \theta A L}{V_R \delta_R} - \left(\frac{L}{\sqrt{2A}} - \sqrt{\frac{A}{2}} \right) \sqrt{\frac{\theta e_w (A-L)\alpha_B}{V_B}} + \frac{2\beta_B \theta (A-L) \left(\frac{1}{2}A - L + \frac{L^2}{2A} \right)}{\delta_B V_B}. \end{aligned} \quad (B7)$$

The seemingly complex objective Eq. (B7) can be simplified as Eq. (13).

Appendix C. Convexity proof

Here we prove the Hessian matrix of Eq. (10) is positive definite, which further means Eq. (10) is convex in h_R and h_B . Any local minimum of Eq. (10) is guaranteed to be a global minimum.

Given the first-order derivatives in Eqs. (B1) and (B2), we derive the second-order partial derivatives of C_{B+R} with respect to h_R and h_B as follows:

$$\frac{\partial^2 C_{B+R}}{\partial h_R^2} = \frac{4L\alpha_R}{V_R h_R^3}, \quad (C1)$$

$$\frac{\partial^2 C_{B+R}}{\partial h_R \partial h_B} = 0, \quad (C2)$$

$$\frac{\partial^2 C_{B+R}}{\partial h_B \partial h_R} = 0, \quad (C3)$$

$$\frac{\partial^2 C_{B+R}}{\partial h_B^2} = \frac{4(A-L)\alpha_B}{V_B h_B^3}. \quad (C4)$$

The Hessian matrix is then written as follows:

$$\Delta^2 C_{B+R}(h_R, h_B) = \begin{pmatrix} \frac{4(A-L)\alpha_B}{V_B h_B^3} & 0 \\ 0 & \frac{4L\alpha_R}{V_R h_R^3} \end{pmatrix}. \quad (C5)$$

Its determinant is always positive, as $h_R > 0$, $h_B > 0$, and $A > L$.

$$\frac{16(A-L)L\alpha_{BaR}}{V_R V_B h_B^3 h_R^3} > 0. \quad (C6)$$

Note that in the feeder-and-trunk service, rail cannot cover the entirety of the corridor, namely L is strictly less than A . Otherwise, it is the rail-only service, which should be separately analyzed.

Therefore, the Hessian matrix has been proved to be positive definite.

Appendix D. Analytical Solution of Stochastic Demand Density

According to [Oksendal and Sulem \(2005\)](#), [Eq. \(18\)](#) can be rewritten as the following Levy process:

$$\frac{d\theta_t}{\theta_t} = \eta dt + \sigma dw(t) + U dN(t). \quad (D1)$$

In [Eq. \(D1\)](#), U has the same distribution as U_i . $U dN(t)$ denotes a possible jump at time t . In other words, $U dN(t) = d \sum_{i=1}^{N_t} U_i$. By Ito's formula ([Oksendal and Sulem, 2005](#)), we obtain:

$$d \ln \theta_t = \frac{\theta_t}{\theta_t} [\eta dt + \sigma dw(t)] - \frac{1}{2} \theta_t^2 \sigma^2 \theta_t^{-2} dt + [\ln(\theta_t + \theta_t U) - \ln(\theta_t)] dN(t). \quad (D2)$$

[Eq. \(D2\)](#) can be further rearranged as:

$$d \ln \theta_t = \left(\eta - \frac{1}{2} \sigma^2 \right) dt + \sigma dw(t) + \ln(1 + U) dN(t). \quad (D3)$$

Therefore, we can obtain:

$$\ln \theta_t = \ln \theta_0 + \left(\eta - \frac{1}{2} \sigma^2 \right) t + \sigma w(t) + \ln(1 + U) N(t). \quad (D4)$$

The dynamic demand density θ_t is thus given by the following equation:

$$\theta_t = \theta_0 e^{\left(\eta - \frac{1}{2} \sigma^2 \right) t + \sigma w(t) + \ln(1 + U) N(t)}. \quad (D5)$$

[Eq. \(D5\)](#) can be further rearranged as [Eq. \(19\)](#).

Appendix E. Expectation of Stochastic Demand Density

Let $J(u)$ denote the cumulative probability distribution function of U (or U_i). [Eq. \(19\)](#) can be further derived as follows according to Ito's lemma ([Oksendal and Sulem, 2005](#)).

$$\begin{aligned} d\theta_t^\gamma &= \gamma \theta_t^{\gamma-1} \theta_t (\eta dt + \sigma dw(t)) + \frac{1}{2} \gamma(\gamma-1) \theta_t^{\gamma-2} \sigma^2 \theta_t^2 dt + \int_0^\infty [(\theta_t + u \theta_t)^\gamma - \theta_t^\gamma] \lambda dJ(u) dt + [(\theta_t + U \theta_t)^\gamma - \theta_t^\gamma] dN(t) \\ &\quad - \int_0^\infty [(\theta_t + u \theta_t)^\gamma - \theta_t^\gamma] \lambda dJ(u) dt, \end{aligned} \quad (E1)$$

where the fourth and the fifth terms of [Eq. \(E1\)](#) are martingales with zero mean. Thus, taking expectation of both sides of the [Eq. \(E1\)](#) leads to:

$$dE(\theta_t^\gamma) = E(\theta_t^\gamma) \left[\gamma \eta + \frac{1}{2} \gamma(\gamma-1) \sigma^2 + \int_0^\infty [(1+u)^\gamma - 1] \lambda dJ(u) \right] dt \quad (E2)$$

[Eq. \(E2\)](#) can be further rearranged as:

$$dE(\theta_t^\gamma) = E(\theta_t^\gamma) \left[\gamma \eta + \frac{1}{2} \gamma(\gamma-1) \sigma^2 + \lambda E((1+U)^\gamma - 1) dt \right]. \quad (E3)$$

Thus, $E(\theta_t^y)$ can be determined by solving Eq. (E3) with initial value θ_0^y :

$$E(\theta_t^y) = \theta_0^y e^{\left[\gamma \eta + \frac{\sigma^2}{2} \gamma (\gamma - 1) + \lambda (E(1+U)^{\gamma} - 1) \right] t}. \quad (\text{E4})$$

References

- Balliauw, M., Onghena, E., 2020. Expanding airport capacity of cities under uncertainty: strategies to mitigate congestion. *J. Air Transp. Manag.* 84, 101791.
- Benth, F.E., Kiesel, R., Nazarova, A., 2012. A critical empirical study of three electricity spot price models. *Energy Economics* 34 (5), 1589–1616.
- Casello, J.M., Lewis, G.M., Yeung, K., Santiago-Rodríguez, D., 2014. A transit technology selection model. *J. Public Transp.* 17 (4), 5.
- Chang, S.K., Schonfeld, P.M., 1991. Optimization models for comparing conventional and subscription bus feeder services. *Transp. Sci.* 25 (4), 281–298.
- Chen, B., Blanchet, J., Rhee, C.H., Zwart, B., 2019. Efficient rare-event simulation for multiple jump events in regularly varying random walks and compound Poisson processes. *Math. Oper. Res.* 44 (3), 919–942.
- Chen, Y.J., Li, Z.C., Lam, W.H., 2015. Modeling transit technology selection in a linear transportation corridor. *J. Adv. Transp.* 49 (1), 48–72.
- Cheng, W.C., Schonfeld, P., 2015. A method for optimizing the phased development of rail transit lines. *Urban Rail Transit* 1 (4), 227–237.
- Chow, Y.S., Robbins, H., Siegmund, D., 1971. *Great Expectations: The Theory of Optimal Stopping*. Boston: Houghton Mifflin.
- Chow, J.Y., Regan, A.C., 2011a. Network-based real option models. *Transp. Res. Part B Methodol.* 45 (4), 682–695.
- Chow, J.Y., Regan, A.C., 2011b. Real option pricing of network design investments. *Transp. Sci.* 45 (1), 50–63.
- Couto, G., Nunes, C., Pimentel, P., 2015. High-speed rail transport valuation and conjecture shocks. *Eur. J. Financ.* 21 (10–11), 791–805.
- Croghan, J., Jackman, J., Min, K.J., 2017. Estimation of geometric Brownian motion parameters for oil price analysis. In: *Proceedings of the IIE Annual Conference*. Institute of Industrial and Systems Engineers (IIE), pp. 1858–1863.
- De Neufville, R., Scholtes, S., 2011. *Flexibility in Engineering Design*. MIT Press.
- Dixit, R.K., Dixit, A.K., Pindyck, R.S., 1994. *Investment Under Uncertainty*. Princeton University Press.
- Friesz, T.L., Mookherjee, R., Yao, T., 2008. Securitizing congestion: the congestion call option. *Transp. Res. Part B Methodol.* 42 (5), 407–437.
- Galera, A.L.L., Soliño, A.S., 2010. A real options approach for the valuation of highway concessions. *Transp. Sci.* 44 (3), 416–427.
- Gao, Y., Diouchi, T., 2013. Incorporating Knightian uncertainty into real options analysis: using multiple-priors in the case of rail transit investment. *Transp. Res. Part B Methodol.* 55, 23–40.
- Guo, Q.W., Chen, S., Schonfeld, P., Li, Z., 2018b. How time-inconsistent preferences affect investment timing for rail transit. *Transp. Res. Part B Methodol.* 118, 172–192.
- Guo, Q.W., Chow, J.Y., Schonfeld, P., 2018a. Stochastic dynamic switching in fixed and flexible transit services as market entry-exit real options. *Transp. Res. Part C Emerg. Technol.* 94, 288–306.
- Kim, M.E., Schonfeld, P., 2013. Integrating bus services with mixed fleets. *Transp. Res. Part B Methodol.* 55, 227–244.
- Li, Z.C., Guo, Q.W., Lam, W.H., Wong, S.C., 2015. Transit technology investment and selection under urban population volatility: a real option perspective. *Transp. Res. Part B Methodol.* 78, 318–340.
- Litman, T., 2007. Evaluating rail transit benefits: a comment. *Transp. Policy* 14 (1), 94–97.
- Liu, L., Miller, H.J., Scheff, J., 2020. The impacts of COVID-19 pandemic on public transit demand in the United States. *PLoS ONE* 15 (11), e0242476.
- Martins, J., Marques, R.C., Cruz, C.O., 2015. Real options in infrastructure: revisiting the literature. *J. Infrastruct. Syst.* 21 (1), 04014026.
- Moccia, L., Laporte, G., 2016. Improved models for technology choice in a transit corridor with fixed demand. *Transp. Res. Part B Methodol.* 83, 245–270.
- NY Open Data. (2022). MTA daily ridership data: beginning 2020, <https://data.ny.gov/Transportation/MTA-Daily-Ridership-Data-Beginning-2020/vxuj-8kew> (accessed on August 1, 2022).
- Øksendal, B., Sulem, A., 2005. *Stochastic control of jump diffusions*. Applied Stochastic Control of Jump Diffusions. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-26441-8_3. Universitext.
- Parajuli, P.M., Wirasinghe, S.C., 2001. A line haul transit technology selection model. *Transp. Plan. Technol.* 24 (4), 271–308.
- Puchalsky, C.M., 2005. Comparison of emissions from light rail transit and bus rapid transit. *Transp. Res. Rec.* 1927 (1), 31–37.
- Ramezani, C.A., Zeng, Y., 2007. Maximum likelihood estimation of the double exponential jump-diffusion process. *Ann. Financ.* 3 (4), 487–507.
- Saphores, J.D.M., Boarnet, M.G., 2006. Uncertainty and the timing of an urban congestion relief investment: the no-land case. *J. Urban Econ.* 59 (2), 189–208.
- Seifert, J., Uhrig-Homburg, M., 2007. Modelling jumps in electricity prices: theory and empirical evidence. *Rev. Deriv. Res.* 10 (1), 59–85.
- Sivakumaran, K., Li, Y., Cassidy, M., Madanat, S., 2014. Access and the choice of transit technology. *Transp. Res. Part A Policy Pract.* 59, 204–221.
- Sun, Y., Guo, Q., Schonfeld, P., Li, Z., 2017. Evolution of public transit modes in a commuter corridor. *Transp. Res. Part C Emerg. Technol.* 75, 84–102.
- Sun, Y., Schonfeld, P., 2015. Stochastic capacity expansion models for airport facilities. *Transp. Res. Part B Methodol.* 80, 1–18.
- Sun, Y., Schonfeld, P., 2016. Holding decisions for correlated vehicle arrivals at intermodal freight transfer terminals. *Transp. Res. Part B Methodol.* 90, 218–240.
- Sun, Y., Schonfeld, P.M., 2017. Coordinated airport facility development under uncertainty. *Transp. Res. Rec.* 2603 (1), 78–88.
- Sun, Y., Schonfeld, P., Guo, Q., 2018. Optimal extension of rail transit lines. *Int. J. Sustain. Transp.* 12 (10), 753–769.
- Tirachini, A., Hensher, D.A., Jara-Díaz, S.R., 2010. Restating modal investment priority with an improved model for public transport analysis. *Transp. Res. Part E Logist. Transp. Rev.* 46 (6), 1148–1168.
- van Oort, N., 2014. Incorporating service reliability in public transport design and performance requirements: international survey results and recommendations. *Res. Transp. Econ.* 48, 92–100.
- van Oort, N., 2016. Incorporating enhanced service reliability of public transport in cost-benefit analyses. *Public Transp.* 8, 143–160.
- Vuchic, V.R., 1976. Comparative analysis and selection of transit modes. *Transp. Res. Rec.* 559, 51–62.
- Wu, F., Schonfeld, P., 2022. Optimized two-directional phased development of a rail transit line. *Transp. Res. Part B Methodol.* 155, 424–447.
- Yu, J., 2007. Closed-form likelihood approximation and estimation of jump-diffusions with an application to the realignment risk of the Chinese Yuan. *J. Econom.* 141 (2), 1245–1280.
- Zhang, J., Wang, D.Z., Meng, M., 2018. Which service is better on a linear travel corridor: park & ride or on-demand public bus? *Transp. Res. Part A Policy Pract.* 118, 803–818.